Understanding the Effect of the Combination of Navigation Tools in Learning Spatial Knowledge

Sanorita Dey
sdey4@illinois.edu
University of Illinois at Urbana-Champaign

Wai-Tat Fu
wfu@illinois.edu
University of Illinois at Urbana-Champaign

Karrie Karahalios
kkarahal@illinois.edu
University of Illinois at Urbana-Champaign

ABSTRACT

Spatial knowledge about the environment often helps people accomplish their navigation and wayfinding tasks more efficiently. Off-the-shelf mobile navigation applications often focus on guiding people to go between two locations, ignoring the importance of learning spatial knowledge. Drawing on theories and findings from the area of learning spatial knowledge, we investigated how the background reference frames (RF) and navigational cues can be combined in navigation applications to help people acquire better spatial (route and survey) knowledge. We conducted two user studies, where participants used our custom-designed applications to navigate in an indoor location. We found that having more navigational cues in a navigation application does not always assist users in acquiring better spatial knowledge; rather, these cues can be distracting in some specific setups. Users can acquire better spatial knowledge only when the navigational cues complement each other in the interface design. We discussed the implications of designing navigation interfaces that can assist users in learning spatial knowledge by combining navigational elements in a complimentary way.

CCS CONCEPTS

• **Human-centered computing**; • **Human computer interaction**; • **HCI design and evaluation methods**; • **User studies**;

KEYWORDS

Spatial learning, interface design, navigational cues, frame of references, allocentric and egocentric encoding

ACM Reference Format:
Sanorita Dey, Wai-Tat Fu, and Karrie Karahalios. 2019. Understanding the Effect of the Combination of Navigation Tools in Learning Spatial Knowledge. In *Symposium on Spatial User Interaction (SUI '19), October 19–20, 2019, New Orleans, LA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3357251.3357582

1 INTRODUCTION

Imagine a typical urban search and rescue mission that takes place in a large building such as an international airport. During such emergencies, rescue teams will benefit from any information that can help them search for critical people, objects, or locations in the environment, and to formulate plans to navigate in the environment as efficiently as possible. In addition to floorplans, rescue teams will benefit most from information from people who work or visit the environment regularly, as they can provide critical information such as landmark locations, environmental structures, spatial patterns of where people go and what they do at different times of the day, etc. Spatial knowledge, in the broad sense, refers to this kind of knowledge about the environment that is directly or indirectly related to the spatial structures of the environment.

One may think that spatial knowledge can be acquired automatically with repeated visits. However, existing literature in spatial cognition suggests that this assumption is not always correct. Even after years of experiences, people surprisingly do not gain accurate spatial knowledge of the environment (although they may know how to go between specific locations). Prior work found that people often get lost in hospitals, libraries, conference centers, or shopping malls even after multiple visits [2, 20, 27]. For example, Peponis et al. [20] found that hospital patients were reluctant to leave their rooms for fear that they would not find their way back. These findings suggest that technologies that help people to acquire spatial knowledge are not only useful, but in many cases, also necessary and critical.

Although the importance of acquiring spatial knowledge is a well-established concept among the researchers, in practice, navigation applications primarily focus on how to guide people to go from one location to another using turn-by-turn instructions. A recent study by Dey et al. [4] showed that using only turn-by-turn directions in regular navigation interfaces cannot assist people in learning spatial knowledge incidentally. This is because, with turn-by-turn directions, users need to pay less attention to the navigation tasks. Thus, by design, it did not encourage users to learn spatial knowledge incidentally. However, relative location updates motivated users to actively process spacial information, which resulted in better learning of spatial knowledge. Since navigational cues such as relative location updates helped people learn incidental spatial knowledge, we asked – do people learn a greater degree of spatial knowledge when more than one navigational cues are presented simultaneously in a single

---

SUI '19, October 19–20, 2019, New Orleans, LA, USA

ACM ISBN 978-1-4503-6975-6/19/10...$15.00
https://doi.org/10.1145/3357251.3357582

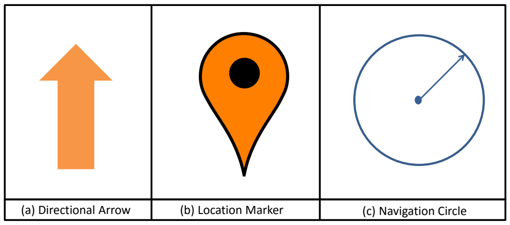


Figure 1: Three navigational cues used to design our interfaces. The *directional arrow* was used for both map-based and video-based interfaces. The *location marker* was used as a relative location update cue for the map-based interface. The *navigation circle* was used as a relative location update cue for the video-based interface.

navigation application? Does the effect of the navigational cues depend on the type of the interface (map vs video)?

In this paper, we extended Dey et al.'s [4] work to examine the effect of adding multiple navigational cues on incidental learning of spatial knowledge. To this end, we designed two new navigation interfaces combining more than one navigational cues (turn-by-turn direction and relative location update) for both map-based and video-based interfaces (the map-based and video-based interfaces were designed by Dey et al. [4] as the reference frame for navigation applications). We used the survey and route knowledge tests to measure the acquired spatial knowledge of the participants. We found that the idea of combining the navigational cues did not always help users to learn the spatial knowledge better; rather, the effect depended on the background reference frames (map and video). The combination of navigational cues helped users to learn spatial knowledge better for the map-based interface but did not perform similarly for the video-based interface.

A follow-up interview with participants revealed the need for combining the background map- and video-based interfaces along with the navigational cues. We varied the size of the map- and video-based interfaces in two separate versions (a smaller map in a larger video-based interface and vice-versa). We aim to understand how the relative size difference between these two reference frames affects users' spatial learning process. We found that a larger video-based interface with a smaller map was the most effective interface design for gaining spatial knowledge. Based on our findings, we discuss the implications of our findings in designing navigation applications that will encourage incidental learning of spatial knowledge. Before describing our results in detail, we will discuss relevant backgrounds and theories on spatial knowledge next, which guided our interface design.

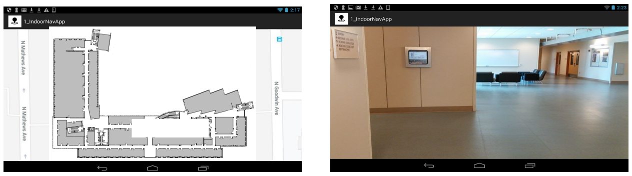
2 RELATED WORK

Development of navigational systems has been inspired and informed by systematic studies on how humans and animals navigate in their environment [8]. For example, in "The Image of the City", Kevin Lynch [17] described five elements for navigating successfully in a city: path, edges, districts, nodes, and landmarks. Some of these elements were later used to build navigation devices. Wiesman [27] did a classic study that showed that even when the navigator was familiar with the environment, up to 40% of the participants reported that they were lost in certain buildings. Darken et al. [3] proposed a toolset for navigation in the virtual environment based on the real-world analogy. In our work, we explored the tools primarily used in mobile navigation applications and investigated how these tools can contribute in learning the mental model of the surrounding environment through incidental learning.

One of the early candidates of the map UI is the hand-drawn sketch maps. Gell [7] explained how a map could include images on their coordinates. Wright et al. [29] compared the map to the wall signs and how these two features affected navigation in indoor space. Thorndyke et al. [25] showed how the spatial knowledge acquired differed when participants navigated using a map representation of an environment. Their study showed that although the map helped people learn the relative distances between spatial objects, having navigated physically in the environment with the map significantly helped them to acquire a better spatial knowledge of the surrounding. Richardson et al. [22] extended this concept and found that people had more problems learning from the augmented virtual environment compared to navigating in the real environment, suggesting that the map combined with the experience of navigation was more useful for the acquisition of spatial knowledge.

Another essential feature for acquiring spatial knowledge is whether people can effectively utilize unique landmarks and cues in the environment [25]. Vinson et al. [26] provided detailed guidelines about designing and placing landmarks in the environment to assist navigation. With the massive improvement of mobile technology, Aslan et al. [1] showed that mobile assisted navigation helped users to learn route knowledge but failed to support learning of survey information.

Recent studies have found that learning of spatial knowledge is not necessarily an automatic process [11, 18]. For example, Henson et al. [11] found that learning of survey knowledge from navigational experiences required explicit encoding and integration of past navigation experiences. When participants navigated using a physical map, acquisition of route knowledge required precise attention and encoding of environmental cues [19, 28]. These findings suggest that navigational cues that focus solely on helping people to go from one location to another may not always lead to good spatial knowledge. Paradoxically, it is possible that interfaces that require (or encourage) participants to transform and integrate their navigational experiences (from route to



(a) Map-based Reference Frame (b) Video-based Reference Frame

Figure 2: The figure at the top shows the map-based reference frame, whereas the figure at the bottom shows the video-based reference frame. We grayed out the room numbers on the map-based RF to avoid distractions since participants were not allowed to enter into any of these rooms during the user study.

survey knowledge and vice versa) will lead to better spatial knowledge.

Although there has been a recent work [4] showing the possibility of using navigational cues to design assistive navigation applications for learning spatial knowledge, there is still a need to explore how different navigational cues can be combined in a single interface to nudge the user to learn spatial knowledge incidentally. A systematic study on whether and how different combination of navigational elements can encourage learning of different types of spatial knowledge will be important and useful.

3 SPATIAL REPRESENTATION, REFERENCE FRAMES, AND NAVIGATIONAL CUES

The spatial location of an object in the environment can be represented with reference to two fundamental classes of coordinate frames: egocentric and allocentric [12, 15]. In the egocentric representation, the location of objects is encoded with reference to the body of the observer or, more specifically, to the body parts, such as the head or trunk of the observer. Egocentric representations of objects may be used for the organization of goal-directed movements, such as reaching a target or avoiding a dangerous stimulus. In the allocentric representation, by contrast, objects are primarily represented with reference to their configurational properties, such as the relationships among their different components and different objects in the environment. Allocentric representation is widely used for path planning or remote processing of spatial information.

The goal of our work is to understand how off-the-shelf navigation applications can assist people in learning these representations incidentally through regular navigation activities. To learn egocentric and allocentric representations of an environment, users need to actively process this information so that they can retrieve them when required. The task of active processing can be achieved with the following two encoding tasks: egocentric encoding and allocentric encoding. These encoding tasks allow users to store spatial knowledge about their environment in the offline system instead of the online system of our memory. It is important to store spatial knowledge in the offline system because prior work showed that although an online system allows users to be continuously connected to the environment, knowledge stored in the online system is highly fragmentary, parse, and not suitable for long-term information retrieval [23]. We aim to design our interfaces in a way with the help of reference frames and navigational cues so that they can allow users to store egocentric and allocentric encoding incidentally in their offline memory. We hypothesized that when the background reference frames (map or video) and combination of the navigational cues (directional arrow and relative location updates) would complement each other, they will enable users to perform better egocentric and allocentric encoding. Thus, users will have a better opportunity to gain spatial knowledge incidentally. However, when the RF and the navigational cues would not complement each other, just having more than one navigational cues would not assist users in gaining more spatial knowledge. Here, the term reference frame means a structure that represents the locations of entities in the space [13] whereas the term navigational cues mean the tools that are placed on top of a reference frame to guide users during navigation.

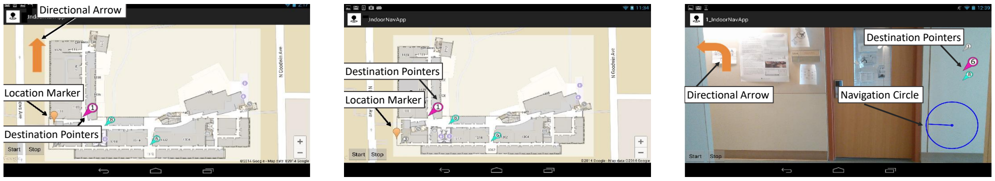
To test our hypothesis, we used the same set of design elements (reference frames and navigational cues) as proposed by Dey et al. [4] to compare our interfaces with their best-performing interface designs. We considered two types of reference frames (RF): 1) map-based RF and 2) video-based RF (See Figure 2). Also, we considered two types of navigational cues: 1) directional arrow and 2) relative location update (location marker for the map-based interfaces and navigation circle for the video-based interfaces) (See Figure 1). By combining these two types of RFs and navigational cues, we designed the following two new navigation interfaces: 1) map-based interface with both directional arrow and location marker (MWAM), and 2) video-based interface with both directional arrow and navigation circle (VWAC). We compared these two interfaces with the two best-performing interfaces reported by Dey et al. [4] where only one navigational cue was used for each interface design: 1) map-based interface with location marker (MWM) and 2) video-based interface with navigation circle (VWC).

4 STAGE 1: IMPACT OF COMBINING NAVIGATIONAL CUES

Here, we briefly explained the design of the four navigation interfaces. Next, we described the procedure followed for the user study. Finally, we described the results of the user study.

4.1 Interface Design

4.1.1 Map-based interface with both directional arrow and location marker (MWAM). In this new design, we used the map-based FOR as the background of the interface where we showed the floorplan of the entire floor of a building. On top of this background FOR, we added two navigational cues: 1)



(a) Map-based Interface with directional arrow and location marker **(b) Map-based Interface with location marker** **(c) Video-based Interface with Directional Arrow and Navigation Circle** **(d) Video-based Interface with Navigation Circle**

Figure 3: Map-based interfaces with (a) a directional arrow and a location marker (b) a location marker. Video-based interface with (c) a directional arrow and a navigation circle, (d) a navigation circle. We used the directional arrow to provide turn-by-turn directions to assist users in their navigation tasks. The location marker and the navigation circle were updated after every step to provide relative location updates suitable for the background reference frame. The map shown in map-based interfaces was the simple version of the original floor plan where we grayed out all the rooms since participants did not have access to any rooms. They could only access the hallways during the user study.

Table 1: List of interfaces considered for our user study. The third column lists the acronyms for all the interfaces (for example, MWAM: Map with Arrow and location Marker). The fourth and the fifth column of this table shows our hypothesis whether a specific interface design will assist users to perform allocentric encoding (column 3) and egocentric encoding (column 4) or not.

Serial	Interface Design	Acronym	Allocentric Encoding	Egocentric Encoding
1	**M**ap-based interface **W**ith both directional **A**rrow and location **M**arker	MWAM	Yes	Yes
2	**M**ap-based interface **W**ith location **M**arker	MWM	No	Yes
3	**V**ideo-based interface **W**ith both directional **A**rrow and navigation **C**ircle	VWAC	No	Yes
4	**V**ideo-based interface **W**ith navigation **C**ircle	VWC	No	Yes

directional arrow on the top-left corner and 2) the location marker (Figure 3a). The directional arrow provided the turn-by-turn instructions during the navigation tasks, whereas the location marker showed the exact location of the user on the floorplan at every step.

4.1.2 Map-based interface with only location marker (MWM). This was the best performing map-based interface developed by Dey et al. [4](Figure 3b). We reproduced it to compare with our new interface design (described above). The only difference between the MWAM and MWM interface designs was that MWAM presented both directional arrow and location marker simultaneously, but MWM presented only location marker to assist users to reach their destinations.

4.1.3 Video-based interface with both directional arrow and Navigation Circle (VWAC). In this video-based interface design, we used the live video feed (captured through the camera of the mobile device) of the environment as the background of the application. On top of this background FOR, we again added two navigational cues: 1) directional arrow on the top-left corner and 2) the navigation circle (relative location update suitable for video-based interfaces) (Figure 3c). Similar to the map-based interface, the directional arrow again provided the turn-by-turn instructions, whereas the navigation circle pointed to the direction of the destination relative to the current position of the user. We updated the direction of the navigation circle after every step taken by the user.

4.1.4 Video-based interface with only navigation circle (VWC). Finally, we recreated this interface design as this was the best performing video-based interface reported by Dey et al. [4] (Figure 3d). Unlike the VWAC interface, this VWC interface did not use the directional arrow to provide turn-by-turn instructions to the user.

In the map-based interface, we hypothesized that the background interface and the navigation tools would work coherently. Since the map-based interface presented a complete floorplan of the environment, we expected that the directional arrow of the map-based would assist users to perform allocentric encoding of their surrounding which would allow users to perform complete path planning during their navigation tasks. On the other hand, the location marker would help users orient themselves in the environment relative to their own positions, which would help them perform egocentric encoding. In contrast to the map-based interface, for the video-based interface, we hypothesized that the background interface and the navigation tools might not complement each other. Since the video-based interface did not present a complete representation of the environment, we expected that the directional arrow would not particularly help users to perform allocentric encoding of their environment. Instead, the instructions provided by the directional arrow would, most likely, be stored in the online memory system temporarily. Although we expected that the navigation circle would still help users to perform egocentric encoding of the environment, the combination of the directional arrow and the navigation circle in the video-based interface would not be more effective than video-based interface with only navigation circle for gaining spatial knowledge. Table 1 lists all these four interfaces and our hypothesis how each of these interfaces will assist users to perform better egocentric and allocentric encoding of their surrounding environment.

In summary, we hypothesized the following:

(1) Map-based interface with both directional arrow and location marker will assist users to gain better spatial knowledge compared to the map-based interface with only location marker.
(2) Video-based interface with both directional arrow and navigation circle will not assist users in gaining better spatial knowledge compared to the video-based interface with only navigation circle.

4.2 Method

We conducted a user study to test our hypothesis. We recreated a similar navigation application as Dey et al. [4] described for their user study using the dead-reckoning technique based on accelerometer and gyroscope readings. We used a Samsung Tablet A (8 inches) to deploy our navigation application with an option of using four different types of interface designs, as we proposed earlier. Since we used the same algorithm in the background for all interface designs, the accuracy of the navigation application remained the same for participants across all interfaces. We conducted all our user studies on a single floor of an academic building. Figure 2a shows the floorplan of the location where we performed all the experiments.

4.3 Participants

Figure 4: The figure shows the path that each participant had to traverse to complete an assisted navigation task. There were three destinations on this path that each participant had to reach sequentially to complete the task. The destinations are marked here with their corresponding sequence numbers.

We posted flyers in community restaurants and libraries and sent emails to local communities to recruit participants. We recruited 62 participants for our user study. Two participants could not complete the experiment due to time constraint. Finally, we had 27 male and 33 female participants between 18 and 60 years of age, M = 26.17 ($\sigma = 9.32$) from a university town who successfully completed all the experiments. 66% of the participants were Caucasian, 17% were Asian, 12% were African-American, the remaining 5% were Hispanic. Participants were nearly equally distributed for four interface designs. All participants were familiar with using navigation applications in mobile devices. Participants were recruited only when they reported through an introductory online survey that they had never been to the building used for the experiments. We randomly assigned each participant to one of the four interface designs. For each interface, we assigned 15 participants. Participants received $10/hour for their participation.

4.4 Study Procedure

Our user study included six stages: 1) pre-experiment spatial ability test surveys, 2) a preview of a randomly assigned interface design of the navigation application, 3) a set of four assisted navigation tasks and incremental survey knowledge tests, 4) one integrated survey knowledge test, 5) a set of three unassisted navigation tasks and route knowledge tests, and 6) a semi-structured interview. Stages 1 - 5 are replicates of Dey et al.s [4] study procedure. We added the semi-structured interview section to assess the observation of the participants using our proposed interface designs compared to the previously reported best-performing interface. To measure the performance of the participants, we used survey and route knowledge tests.

4.4.1 Pre-experiment spatial ability test surveys. Before conducting our user study, we measured each participant's mental rotation, object manipulation and perspective-taking spatial abilities as these are important for processing spatial knowledge. Each participant completed the perspective-taking and spatial orientation test [10, 14] and self-reported measurements of environmental abilities by "sense of direction (SOD)" scale [9]. We also asked all of the participants to complete the Kit of Factor-Referenced Cognitive Test VZ-2 [6] to assess their spatial visualization ability. No significant difference was found in these pre-tests among the participants in each group.

4.4.2 Preview of a randomly assigned interface design. We randomly assigned one interface design to each participant. One member from our research team spent ten minutes to explain the purpose of the study, the basic functionality of our navigation application, and the utility of the FOR and navigational cues of the assigned interface design.

4.4.3 Assisted navigation tasks and incremental survey knowledge tests. First, we asked each participant to complete four assisted navigation tasks using their assigned interface design. During each task, participants had to reach three different destinations sequentially with the help of their navigation application. When a participant successfully reached a destination, the application automatically sent a notification and asked the participant to find the next designated destination. When a participant arrived at the third destination, the application notified that the task was successfully completed. Figure 4 shows the complete path of one of the assisted navigation tasks on the floorplan. Next, the experimenter asked the participant to complete two incremental survey knowledge tests related to the navigation task.

We reused the orientation test and path recall test as proposed by Thorndyke et al. [25]. In the orientation test, the experimenter asked participants to mark the locations of the three destinations that they had to reach during their assisted navigation task. We measured the performance of the participants by calculating the difference between the ground truth and the participants' reported location as the orientation error. In the path recall test, the experimenter asked participants to draw the entire path that they took to complete their assisted task. We compared the reported path with the original path (recorded by the navigation app) to calculate the number of turns and corridors reported by mistake by the participants. The orientation and path recall tests together allowed us to measure the incremental survey knowledge gained by the participants while using our interface designs.

4.4.4 Integrated survey knowledge test. Once participants completed all four assisted navigation tasks, the experimenter asked participants to complete an integrated survey knowledge test (floor plan recall test) to test their overall survey knowledge acquired from all four navigation tasks. In this test, participants drew the floor plan of the building. Since participants were allowed to walk only through the hallways of a specific floor during their assigned tasks, they were asked to draw only the hallways where they indicated as many landmark objects as possible. Two independent coders rated these sketches on a 5-point Likert-type scale [16] using two criteria: (a) to what degree the sketch resembled the original floorplan and (b) to what degree potential landmark locations were represented in the correct position. A Cohens kappa test showed the inter-rater agreement as 0.82. So we took the average score of each sketch for our analysis.

4.4.5 Unassisted navigation tasks and route knowledge tests. In this stage, each participant completed three unassisted navigation tasks. The goal of these unassisted navigation tasks was to reach to a location shown in a picture without any external assistance. During each unassisted task, the experimenter showed one of these pictures to each participant. First, participants had to tell the experimenter whether they could recognize that place. We called this the location recognition test. Next, they had to reach to that location by themselves without any external help. We called this the unassisted navigation test. We replicated the location recognition test and the unassisted navigation test from Dey et al. [4] to measure the route knowledge gained by participants.

Table 2: Mean and standard deviation of the measurements in the survey and route knowledge tests for all four groups of participants in stage 1. Asterisk(*) denotes statistical significance (p<.05)

Tasks/ Interfaces	Map with Arrow and Marker, M (SD)	Map with Marker, M(SD)	Video with Arrow and Circle, M (SD)	Video with Circle, M (SD)
Orientation Test	22.04* (10.11)	31.33 (8.22)	35.67 (9. 31)	29.01 (11.23)
Path Recall Test	1.95* (0.79)	3.79 (1.43)	3.33 (1.39)	3.18 (1.56)
Floor Plan Recall Test	3.85* (0.66)	3.75* (0.97)	2.29 (1.12)	2.44 (1.18)
Location Recognition Test	0.79 (0.51)	0.72 (0.48)	1.50 (0.39)	1.86* (0.37)
Unassisted Navigation Test	1.92 (0.87)	1.82 (0.81)	2.40 (0.62)	2.68* (0.58)

4.4.6 Semi-structured interview. Finally, the experimenter conducted a semi-structured interview with each participant. The interviews took 15-20 minutes on average to complete. We mainly focused on the following questions during the interview: 1) What is your overall experience of using this interface design for an indoor navigation app? 2) What features of the interface did you find useful? 3) What was the most challenging aspect of using the application? 4) How will you like to change the design of the interface to make it more user-friendly?

4.5 Results

For our analysis, we had five dependent variables from the survey and route knowledge tests: 1) orientation test, 2) path recall test, 3) floor plan recall test, 4) location recognition test, and 5) unassisted navigation test and two independent variables: 1) FOR and 2) navigational cues. Since we had more than one dependent variables, we performed two-way MANOVA test (2 (FOR) X 2 (types of navigational cues)) with post-hoc Tukey analysis.

We found statistically significant interaction effect between FOR and types of navigational cues on the combined dependent variables, F(2, 55) = 18.38, $p < 0.01$, Wilks' $\lambda = 0.29$, partial $\eta^2 = 0.76$. To further examine the dependent variables separately, we observed the univariate interaction effect. We found that there was a statistically significant interaction effect between RF and type of navigation cues for orientation test, F(1, 56) = 19.18, $p < 0.01$, partial $\eta^2 = 0.77$, path recall test, F(1, 56) = 15.04, $p < 0.01$, partial $\eta^2 = 0.62$, the location recognition test, F(1, 56) = 5.78, $p = 0.02$, partial $\eta^2 = 0.30$, and unassisted navigation test, F(1, 56) = 5.02, $p = 0.02$, partial $\eta^2 = 0.34$. No statistically significant interaction effect was found for the floor plan recall test, F(1, 56) = 3.01, $p = 0.23$, partial $\eta^2 = 0.16$. Next, we discussed the post-hoc analysis for all survey and route knowledge tests individually. Table 2 shows the mean and standard deviation of the measurements of the participants in all the navigation tasks.

4.5.1 Orientation Test. Since the univariate interaction effect of the orientation test was statistically significant, we examined the main effect for the independent variables. We found that for the map-based interfaces, participants using the MWAM interface performed significantly better than those who used the MWM interface, F(1, 56) = 8.16, $p < 0.01$, $\eta^2 = 0.42$. We also found that participants using the map-based MWAM interface performed significantly better than participants using the video-based VWAC interface F(1, 56) = 5.28, p = 0.02, $\eta^2 = 0.31$. No other effect was statistically significant.

4.5.2 Path Recall Test. Similar to the orientation test, we also found a univariate interaction effect significant for the path recall test. So we examined the simple main effect for the independent variables for the path recall test too. We found that for the map-based interfaces, participants using the MWAM interface performed significantly better than those who used the MWM interface, F(1, 56) = 7.01, $p < 0.01$, $\eta^2 = 0.38$. No other effect was statistically significant.

4.5.3 Floor Plan Recall Test. We did not find the univariate interaction effect significant for the floor plan recall test. So we focused on examining the main effect directly. We found that overall participants using map-based interfaces performed significantly better in the floor plan recall test than those using the video-based interfaces, F(1, 56) = 6.45, $p < 0.01$, $\eta^2 = 0.33$. Further post-hoc analysis showed that participants using the MWAM interface performed significantly better than both versions of the video-based interfaces. We observed a similar trend for participants using the map-based MWM interface as well.

4.5.4 Location Recognition Test. Since the univariate interaction effect of the location recognition test was statistically significant, we examined the main effect for the independent variables. We found that participants using the video-based VWC interface performed significantly better than those using map-based MWM interface, F(1, 56) = 4.85, p = 0.02, $\eta^2 = 0.09$ and those using map-based MWAM interface, F(1, 56) = 4.84, p = 0.02, $\eta^2 = 0.09$. No other effect was statistically significant.

4.5.5 Unassisted Navigation Test. Similar to the location recognition test, we examined the main effect for the independent variables for the unassisted navigation test. We again found that participants using the video-based VWC interface performed significantly better than those using map-based MWM interface, F(1, 56) = 4.32, p = 0.03, $\eta^2 = 0.07$ and those using map-based MWAM interface, F(1, 56) = 4.24, p = 0.03, $\eta^2 = 0.06$. No other effect was statistically significant.

4.5.6 Summary of the survey and route knowledge tests. Our results showed that participants using the map-based interfaces performed significantly better in survey knowledge tests whereas in route knowledge tests, participants using the video-based interfaces performed better. Given, participants of the map-based interfaces could directly access the floorplan of the building, this finding was consistent with our expectation. In addition, we found that combining the directional arrow and location marker for the map-based interface helped users to perform better in gaining incremental survey knowledge (orientation and path recall tests). This finding is consistent with our first hypothesis. However, the strategy of combining the directional arrow and the navigation circle for the video-based interface did not work the same way. Participants of the VWAC interface performed no better or worse than those who used the VWC interface in the incremental survey and route knowledge tests, which is consistent with our second hypothesis. However, participants using the VWAC interface performed worse than those who used the VWC interface in the route knowledge tests. This finding was not consistent with our second hypothesis. We were surprised to observe that combining more navigational cues in a single interface can even make an interface design less effective than before. To understand why participants using the VWAC interface performed worse than their counterpart using the VWC interface, we analyzed participants' interview responses.

4.5.7 Summary of the semi-structured interview. Analyzing the interview responses of the participants for the VWAC interface, we found that since both the arrow and the navigation circle were providing egocentric information about the surrounding environment, having those two cues together did not assist participants during their navigation tasks. Instead, participants (N = 9) found them distracting at the beginning of their navigation task and eventually, majority of the participants (N=11) using this interface decided to focus on only one cue of their choice and ignored the other one. We believe that the act of ignoring one cue put additional stress on the participants, and that might be the reason why they performed poorly in the route knowledge tests.

Moreover, participants using the map-based interface (N = 18) felt that following the app for the navigation tasks did not allow them to look around for natural landmarks. The application indirectly grabbed most of their attention during the navigation tasks, and thus, they performed poorly during the route knowledge tests. Participants of the map-based interface particularly struggled to recognize one location where we showed the picture of an elevator door in a dark corridor. During the semi-structured interview, participants of the map-based interfaces (N = 12) explained that they were surprised during the location recognition test to know that there was a second elevator in that floor and they failed to notice such an important landmark during their assisted navigation tasks.

On the other hand, participants of the video-based interfaces could never access the map of the environment. They felt that without any reference of the floorplan, it was challenging for them (N = 14) to complete the survey knowledge tests. Participants of the video-based interface (N=14) felt that by default, they expected to see a floorplan of the environment in a navigation application. Since we did not provide any map of the environment, they struggled to create a mental map of the environment. Prior work showed that creating a mental map of the environment is critical to perform day-to-day navigational activities with ease [5, 17, 21, 24]. Majority of the participants of the video interfaces (N = 22) felt that users of the video-based interfaces should have access to a floorplan or a schematic map of the environment for the better mental encoding of their surrounding. Similarly, almost half of the participants of the map-based interfaces (N=14) wanted to add a feature in the app itself that would allow them to find and remember the landmarks of the environment naturally for future navigational activities. Based on the observation from this user study and the semi-structured interviews, we designed a pair of new interface designs by combining both map-based and video-based RFs and conducted a second user study to explore the effectiveness of those new designs.

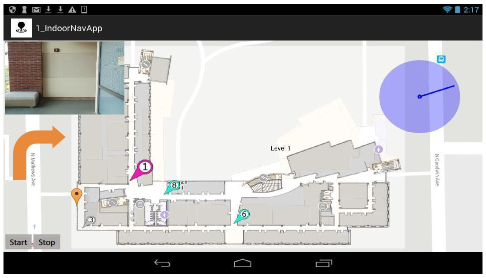
5 STAGE 2: DESIGNING INTERFACES COMBINING MAP- AND VIDEO-BASED RFS

The observations from the first user study motivated us to design interfaces combining both map- and video-based RFs. We also considered the effect of navigational cues in our design (depending on the RF). Here, we discussed the two versions of the interfaces that we designed in stage 2.

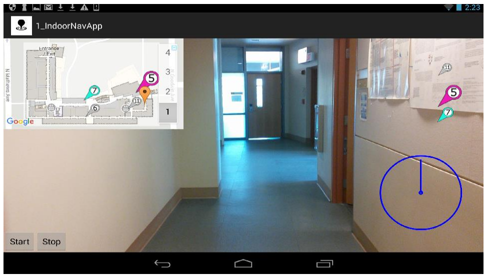
5.1 Interface Design

5.1.1 Map-based interface with a video window, a directional arrow, a navigation circle, and a location marker. Similar to our previous map-based interfaces, we displayed the floorplan of the entire floor as the background of this interface. On the top-left corner of the interface, we showed the live camera feed of the surrounding environment (as we did for the video-based interfaces). We hypothesized that this camera feed would allow users to notice the landmarks of the surrounding environment even when they would focus their attention on the navigation app. Additionally, we included all three navigational cues (1) the directional arrow, 2) the navigation circle, and 3) the location marker) since we found that including multiple cues in the map-based interface did not negatively affect users to acquire survey and route knowledge. Unlike other map-based interfaces, this design provided users an egocentric point of view. Fig 5a shows a screenshot of this interface.

5.1.2 Video-based interface with map window, navigation circle, and location marker. Like other video-based interfaces, we showed the live camera feed of the environment as the primary background of this interface. On the top-left corner, we included the floorplan of the building, which showed the current location of the user on the floorplan with the location marker. Additionally, we added the navigation circle to provide real-time direction update of the destination with respect to the current position of the user. We did not use the directional arrow in this design since we found from study 1 that using directional arrow and navigation circle together can distract users from gaining spatial knowledge. In this interface design, users received the allocentric perspective of the surrounding environment, which was missing in all other video-based interfaces. Fig 5b shows a screenshot of this interface.



(a) Map-based Interface with a window to show the video feed



(b) Video-based Interface with a window to show the map

Figure 5: (a) A map-based interface with a directional arrow, a location marker, a navigation circle, and a window to show a video feed of the surrounding environment. (b) A video-based interface with a navigation circle, a window to show the floorplan of the location along with a location marker indicating the current location of the user on the map.

We conducted the second user study to compare the effectiveness of these two interface designs with the best performing designs so far. We considered the map-based interface with both directional arrow and location marker (MWAM) and video-based interface with only navigation circle (VWC) to compare with our new interface designs.

5.2 Stage 2: User Study Procedure

We followed the same procedure as we did in stage 1 to conduct the user study. The only difference was the interface designs that we compared in this stage. We compared the following interface designs in this stage: 1) map-based interface with video window, directional arrow, navigation circle, and location marker (new), 2) video-based interface with map window, navigation circle, and location marker (new), 3) map-based interface with both directional arrow and location marker (existing), and 4) video-based interface with only navigation circle (existing).

Table 3: Mean and standard deviation of the measurements in the survey and route knowledge tests for all four groups of participants of stage 2. Asterisk(*) denotes statistical significance (p<.05)

Tasks/ Interfaces	Map with Video Window, M (SD)	Map with Arrow and Marker, M (SD)	Video with Map Window, M (SD)	Video with Circle, M (SD)
Orientation Test	19.82* (11.33)	21.89 (9.54)	21.11* (10.33)	27.98 (8.11)
Path Recall Test	1.97* (0.87)	1.98 (1.31)	2.01* (1.44)	3.26 (1.29)
Floor Plan Recall Test	3.97 (0.78)	3.76 (1.05)	4.02* (1.20)	2.39 (1.02)
Location Recognition Test	1.33 (0.66)	0.82 (0.54)	1.85* (0.48)	1.97* (0.41)
Unassisted Navigation Test	2.72* (0.69)	1.84 (0.92)	2.81* (0.85)	2.65* (0.89)

5.3 Participants

We again posted flyers in community restaurants and libraries and sent emails to local communities to recruit participants. In total, we recruited 61 participants. One participant failed to complete the study due to physical challenges. 60 participants (36 female) between 18 and 60 years of age limit, M = 29.27 ($\sigma = 7.98$) completed the user study. 61% of the participants were Caucasian, 22% were Asian, 10% were African-American, the remaining 7% were Hispanic and Native American. Participants were nearly equally distributed across all four interface designs. All participants were familiar with using navigation applications during their daily activities, and none of them visited the building before where we conducted the user study. We again randomly assigned each participant to an interface design. Since we wanted to maintain the assumption of independent sampling, we did not reuse the test results of the participants who used existing interfaces in study 1. Instead, we recruited new participants for the existing interface designs as well. Each participant received $10/hour for their participation.

5.4 Results

Before conducting our user study, we measured each participant's mental rotation, object manipulation, and perspective-taking spatial abilities as we did in study 1. No significant difference was found in these pre-tests among the participants in each group.

Similar to stage 1, we had five dependent variables from survey and route knowledge tests: 1) orientation test, 2) path recall test, 3) floor plan recall test, 4) location recognition test, and 5) unassisted navigation test and two independent variables: 1) RF and 2) navigational cues. Since we had more than one dependent variables, we performed two-way MANOVA test (2 (RF) X 2 (types of navigational cues)) with post-hoc Tukey analysis.

We found statistically significant interaction effect between RF and types of navigational cues on the combined dependent variables, F(2, 55) = 11.57, $p < 0.01$, Wilks' $\lambda = 0.21$, partial $\eta^2 = 0.55$. To further investigate each dependent variable separately, we observed the univariate interaction effect. We found that there was a statistically significant interaction effect between RF and type of navigation cues for the orientation test, F(1, 56) = 9.52, $p < 0.01$, partial $\eta^2 = 0.54$, the path recall test, F(1, 56) = 4.31, p = 0.03, partial $\eta^2 = 0.12$, the floor plan recall test F(1, 56) = 4.74, p = 0.03, partial $\eta^2 = 0.13$, and the location recognition test, F(1, 56) = 4.16, p = 0.04, partial $\eta^2 = 0.07$. No statistically significant interaction effect was found for the unassisted navigation test, F(1, 56) = 2.36, p = 0.27, partial $\eta^2 = 0.02$. Next, we discussed the post-hoc analysis for all the spatial knowledge tests individually. Table 3 shows the mean and standard deviation of the measurements of the participants in all the navigation tasks.

5.4.1 Orientation Test. Since the univariate interaction effect of the orientation test was statistically significant, we examined the main effect for the independent variables. We found that participants of the map-based interfaces with video window performed significantly better than those who used the VWC interface, F(1, 56) = 5.79, p = 0.02, $\eta^2 = 0.31$. We also found that participants using the video-based interface with map window performed significantly better than participants using the VWC interface, F(1, 56) = 5.41, p = 0.02, $\eta^2 = 0.29$. No other effect was statistically significant.

5.4.2 Path Recall Test. Similar to the orientation test, we also found a univariate interaction effect significant for the path recall test. So we proceed to examine the simple main effect for the independent variables for the path recall test. We found that participants using the map-based interfaces with video window performed significantly better than those who used the VWC interface, F(1, 56) = 4.85, p = 0.03, $\eta^2 = 0.22$. We also found that participants using the video-based interfaces with map window performed significantly better than those who used the VWC interface, F(1, 56) = 4.86, p = 0.03, $\eta^2 = 0.22$. No other effect was statistically significant.

5.4.3 Floor Plan Recall Test. Here again, we examined the simple main effect for the independent variables. We found that participants using the video-based interface with map window performed significantly better in the floor plan recall test than those who used the VWC interface, F(1, 56) = 4.21, p = 0.03, $\eta^2 = 0.16$. No other effect was statistically significant.

5.4.4 Location Recognition Test. Since the univariate interaction effect of the location recognition test was statistically significant, we examined the main effect for the independent variables. We found that participants using the VWC interface performed significantly better than those using the MWAM interface, F(1, 56) = 4.78, p = 0.02, $\eta^2 = 0.08$. We also found that participants using the video-based interface with map window performed significantly better than those using the MWAM interface, F(1, 56) = 4.69, p = 0.02, $\eta^2 = 0.08$. Finally, we found that participants using the video-based interface with map window performed significantly better than those using the map-based interface with video window, F(1, 56) = 3.56, p = 0.04, $\eta^2 = 0.04$. No other effect was statistically significant.

5.4.5 Unassisted Navigation Test. Finally, we examined the main effect of the independent variables for the unassisted navigation test. We found that participants using the VWC interface performed significantly better than those using the MWAM interface, F(1, 56) = 4.19, p = 0.03, $\eta^2 = 0.06$. We also found that participants using the video-based interface with map window performed significantly better than those using the MWAM interface, F(1, 56) = 3.85, p = 0.04, $\eta^2 = 0.04$. Finally, we found that participants using the map-based interface with video window performed significantly better than those using the MWAM interface, F(1, 56) = 3.90, p = 0.04, $\eta^2 = 0.05$. No other effect was statistically significant.

5.4.6 Stage 2: Summary of the survey and Route Knowledge Tests. The findings of the user study in stage 2 showed that combining the map- and video-based interfaces in a single interface design helped users acquire better survey and route knowledge. Adding the map window in the video-based interface provided an allocentric perspective of the environment and thus allowed them to perform better in the orientation, path recall, and floor plan recall test. Similarly, including the video window in the map-based interface allowed users to look around the environment even when they were following the app for the navigation task. Therefore, users performed significantly better in the location recognition and unassisted navigation tests than before when the map-based interface did not have any camera feed. Moreover, we could not find any significant differences between the two new interface designs in general except for the location recognition test. Only for this route knowledge test, participants of the video-based interface with map window performed significantly better than those using the map-based interface with the video window. To understand the reasoning behind this finding, we analyzed the interview responses of the participants.

We found that although participants liked both of these new interface designs, some participants (N=5) of the map-based interface with video window found that checking the camera feed in the small video window was hard, especially in those cases when the corridor was dark. On the contrary, checking the floorplan through the small window was not difficult for the participants of the video-based interface since the floorplan was a static feature compared to a live camera feed.

6 DISCUSSION AND CONCLUSION

Prior work [4] showed that interface designs equipped with certain RFs and navigational cues could encourage people to process spatial knowledge about their surrounding environment actively. Active processing of spatial information can help people learn spatial knowledge incidentally during regular navigation tasks. The primary goal of our study was to understand whether having more navigational cues in a single interface can make the incidental learning process more effective. We found that having more navigational cues do not always help users learn spatial knowledge better. It can even work counter-intuitively and can distract users during their navigation activities. Interface designs creating such distractions can hamper the natural process of incidental learning of spatial knowledge. To better facilitate the process of spatial learning, interface designers should combine the navigational cues and RFs in a way so that they can complement each other.

Designers can also optimize the interface design of navigation applications based on the environment and the nature of the navigation activities. For example, in a busy shopping mall or during the rush hour of the day, the navigation interface can present minimum, more directive navigational cues to users to keep the task of reaching to the destination simple. However, during the late afternoon or over the weekend, more cognitively demanding cues can be presented to allow users to process spatial information more actively, which will facilitate incidental spatial learning. Further investigation needs to be conducted to understand how the dynamic modification of the interface design of the navigation applications can affect users' experience of using such navigation applications during regular wayfinding tasks.

In our interface designs at stage 2, we kept a small window at the top left corner of the interface to include the secondary RF. During the semi-structured interview, some participants (N=4) suggested dividing the screen half-way to include the secondary RF. Other participants suggested presenting the secondary RF on demand instead of always showing it on the screen (N=3). For example, P31 suggested offering the secondary RF through a specific gesture such as a double-tap to keep the screen less cluttered. She mentioned that a less cluttered interface design would be more suitable for a smaller mobile device such as a smartphone. The relative proportion of the RFs can be explored further in the future to design more effective interfaces for navigation applications.

In our work, we represented allocentric RF using a static map of the location, whereas the egocentric RF was presented using the live camera feed. In practice, an interactive map may also provide an egocentric representation of the surrounding. For example, off-the-shelf GPS applications generally use an egocentric interactive map to provide directional guidance to drivers. On the other hand, a top-down live camera feed can present an allocentric representation of the environment too. The interface designs used in our current study are not sufficient to understand the impact of various forms of maps and camera feeds in acquiring spatial knowledge. Further investigations need to be conducted to understand the broader effects of RFs (both maps and videos) in learning spatial knowledge.

To conclude, our work shows that there are many possibilities to design interfaces of navigation applications that can help users to acquire spatial knowledge incidentally through day-to-day navigation activities. However, interface designers need to proceed with caution since a specific combination of navigation cues may not facilitate the learning process. We believe that navigation interface designers need to focus more on how human beings encode and create a mental model of their environment. Perhaps that understanding will allow them to find out effective combinations of design elements for interfaces of smart navigation applications.

REFERENCES

[1] Ilhan Aslan, Maximilian Schwalm, Jörg Baus, Antonio Krüger, and Tim Schwartz. 2006. Acquisition of spatial knowledge in location aware mobile pedestrian navigation systems. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM, 105–108.

[2] Laura A Carlson, Christoph Hölscher, Thomas F Shipley, and Ruth Conroy Dalton. 2010. Getting lost in buildings. *Current Directions in Psychological Science* 19, 5 (2010), 284–289.

[3] Rudy P Darken and John L Sibert. 1993. A toolset for navigation in virtual environments. In *ACM symposium on User interface software and technology*. Citeseer, 157–165.

[4] Sanorita Dey, Karrie Karahalios, and Wai-Tat Fu. 2018. Getting There and Beyond: Incidental Learning of Spatial Knowledge with Turn-by-Turn Directions and Location Updates in Navigation Interfaces. In *Proceedings of the Symposium on Spatial User Interaction (SUI 2018)*. ACM, New York, NY, 100–110.

[5] Roger M Downs and David Stea. 1977. *Maps in minds: Reflections on cognitive mapping*. HarperCollins Publishers.

[6] Ruth B Ekstrom, Diran Dermen, and Harry Horace Harman. 1976. *Manual for kit of factor-referenced cognitive tests*. Vol. 102. Educational testing service Princeton, NJ.

[7] Alfred Gell. 1985. How to read a map: remarks on the practical logic of navigation. *Man* (1985), 271–286.

[8] Reginald G Golledge. 1999. *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU Press.

[9] Mary Hegarty, Anthony E Richardson, Daniel R Montello, Kristin Lovelace, and Ilavanil Subbiah. 2002. Development of a self-report measure of environmental spatial ability. *Intelligence* 30, 5 (2002), 425–447.

[10] Mary Hegarty and David Waller. 2004. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* 32, 2 (2004), 175–191.

[11] Agnes Henson, Hanspeter A Mallot, Heinrich H Bülthoff, and Tobias Meilinger. 2011. When do we integrate spatial information acquired by walking through environmental spaces?. In *33rd Annual Meeting of the Cognitive Science Society (CogSci 2011)*. Cognitive Science Society, 2764–2769.

[12] Ian P Howard and William B Templeton. 1966. Human spatial orientation. (1966).

[13] Roberta L Klatzky. 1998. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition*. Springer, 1–17.

[14] Maria Kozhevnikov and Mary Hegarty. 2001. A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition* 29, 5 (2001), 745–756.

[15] F Lacquaniti. 1997. Frames of reference in sensorimotor coordination in: Handbook of Neuropsychology Boller F. Grafman J.(Eds) Vol. 11 pp. 27–64.

[16] K Lohmann. 2011. The use of sketch maps as a basis for measures of spatial knowledge. In *Understanding and Processing Sketch Maps. In: Proceedings of the COSIT 2011 workshop. AKA Verlag, Heidelberg.*

[17] Kevin Lynch. 1960. *The image of the city*. Vol. 11. MIT press.

[18] Tobias Meilinger, Alain Berthoz, and Jan M Wiener. 2011. The integration of spatial information across different viewpoints. *Memory & cognition* 39, 6 (2011), 1042–1054.

[19] Avi Parush, Shir Ahuvia, and Ido Erev. 2007. Degradation in spatial knowledge acquisition when using automatic navigation systems. In *Spatial information theory*. Springer, 238–254.

[20] John Peponis, Craig Zimring, and Yoon Kyung Choi. 1990. Finding the building in wayfinding. *Environment and behavior* 22, 5 (1990), 555–590.

[21] Gould Peter and White Rodney. 1974. Mental maps.

[22] Anthony E Richardson, Daniel R Montello, and Mary Hegarty. 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & cognition* 27, 4 (1999), 741–750.

[23] Daniel J Simons and Christopher F Chabris. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *perception* 28, 9 (1999), 1059–1074.

[24] Robert Sommer and Susan Aitkens. 1982. Mental mapping of two supermarkets. *Journal of Consumer Research* 9, 2 (1982), 211–215.

[25] Perry W Thorndyke and Barbara Hayes-Roth. 1982. Differences in spatial knowledge acquired from maps and navigation. *Cognitive psychology* 14, 4 (1982), 560–589.

[26] Norman G Vinson. 1999. Design guidelines for landmarks to support navigation in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 278–285.

[27] Jerry Weisman. 1981. Evaluating architectural legibility wayfinding in the built environment. *Environment and behavior* 13, 2 (1981), 189–204.

[28] Katharine S Willis, Christoph Hölscher, Gregor Wilbertz, and Chao Li. 2009. A comparison of spatial knowledge acquisition with maps and mobile maps. *Computers, Environment and Urban Systems* 33, 2 (2009), 100–110.

[29] Patricia Wright, Audrey J Hull, and Ann Lickorish. 1993. Navigating in a hospital outpatients' department: The merits of maps and wall signs. *Journal of Architectural and Planning research* (1993).