Acquisition of Structural Versus Object Landmark Knowledge

Brian J. Stankiewicz and Amy A. Kalia
University of Texas at Austin

Three experiments investigated the acquisition and retention of structural and object landmarks in virtual indoor environments. The experiments investigated the rate of acquisition and memory retention for hallway structure (structural landmarks) and pictures (object landmarks). The experiments investigated the rate of acquisition, the role of information content, and memory retention of this information when participants were trained and tested in novel virtual indoor environments. The results from these experiments suggest that (a) even initially, participants are biased toward encoding building structure over object landmarks; (b) participants are sensitive to the information content of landmarks and will allocate memory resources to landmarks that are more informative; and (c) information about these landmarks is retained even after a 1-year delay.

Keywords: spatial navigation, spatial cognition, virtual reality, spatial memory, knowledge acquisition

Humans, and many other animals, possess the remarkable ability to engage in goal-directed spatial navigation behavior in large-scale spaces. That is, they can find their way from one location within a familiar environment to another predetermined and unobservable location. This behavior can be accomplished without reference to external (noncognitive) directions or representations of the space (e.g., a map) by using an internal representation of the large-scale space. The internal representation of a large-scale space is typically referred to as a cognitive map (Tolman, 1948). Most adults carry around within them hundreds, if not thousands, of these maps, which allow them to easily navigate in the building in which they work, among buildings on a campus, and/or within multiple cities that they have lived in or visited.

There is little debate that humans possess the ability to generate a cognitive map. What is typically debated is what is made explicit within the cognitive map and how this spatial information is acquired (Gillner & Mallot, 1998; Ruddle, Payne, & Jones, 1997; Siegel & White, 1975; Thorndyke & Hayes-Roth, 1982; Tversky, 1993). Several theories claim that landmarks play a role in the development of a cognitive map, but it is not clear which features of the environment are used as landmarks and why these features are useful when navigating. This article investigates this issue by defining two types of landmarks: structural landmarks and object landmarks. Structural landmarks are defined as geometric features of a layout that can serve as visual cues, such as dead-ends and T-junctions, and object landmarks are objects in the environment that are independent of its structure, such as pictures on the walls.

With respect to these landmarks, the current experiments investigated three primary questions concerning the acquisition and retention of landmark knowledge. First, do people acquire knowledge about one type of landmark faster than knowledge about the other? Second, is the memory retained (following a delay) the same for these two types of landmarks? And finally, how does the information content of these landmarks affect encoding?

To investigate the issue of acquisition, we measured participants’ landmark knowledge during initial learning of an environment to see if the two types of visual cues are treated differently. Second, to investigate the retention of landmark knowledge, we tested participants after delays of 1 day, 7 days, 30 days, and 1 year (Experiment 1) after the initial training. The retention rate demonstrates the resiliency of the memory representations for these different types of landmarks in addition to the effect of information content on retention. We also manipulated the information content of these landmarks (i.e., how well participants could theoretically localize themselves simply by observing a particular landmark) and measured how the information content of a landmark affected the acquisition and retention of the landmark knowledge (Experiment 2).

ACQUISITION OF SPATIAL KNOWLEDGE

In the development of a cognitive map of a novel environment, visual cues can be used and translated into a mental reference that can be reused at a later time. Several theories have suggested ways in which visual information, primarily landmarks, can be used to learn an environment. One of the first spatial acquisition theories, proposed by Siegel and White (1975), states that learning of the structure of an environment is based on remembering key landmarks; from this, route knowledge develops and ultimately a “map-like” survey representation of the space is generated. Having a survey representation implies having some understanding of the topography of the environment as well as the spatial relations between locations. Subsequent research suggested that route
knowledge may not necessarily be an intermediary stage between landmark and survey knowledge (Moar & Carleton, 1982). It is also possible that landmarks themselves are the basis for spatial representations of environments. According to the anchor-point hypothesis, salient cues are used to "anchor" each region of a space into an organized, even hierarchical, mental map (Couclelis, Golledge, Gale, & Tober, 1987). It is interesting to note that developmental work has shown that children as young as 8 years old are sensitive to the landmarks that are on a route versus the landmarks that are not observable on a route (Cornell, Heth, & Alberts, 1994; Cornell, Heth, & Rowat, 1992), suggesting that the strategy of encoding landmarks is a fundamental function in way-finding.

Alternatively, the view-graph approach suggests that spatial representations consist of a series of views, or snapshots, of the environment, each associated with specific actions and goals. According to this theory, there is no need for a survey representation of space. Instead, one can simply associate a specific view with a specific action. The model moves through the environment by recognizing a view and looking up what action to generate given the view. The model then continues the action until it recognizes another view and then generates the action associated with that view. According to the authors (Mallot & Gillner, 2000; Schölkopf & Mallot, 1995), linking of views to actions will ultimately lead the model to its goal. Recent work by Gillner and Mallot (1998) in which they manipulated the positions of landmarks following training showed that participants' navigation performances were the poorest when the new positions of the landmarks offered conflicting action choices (e.g., turn left vs. right). This finding suggests that movements may be associated with specific object configuration.

It is also important to note that not all theories of spatial learning explicitly involve the encoding of object landmarks. Kuipers (2000, 2001; Kuipers & Byun, 1991) has developed a robot navigation algorithm with a hierarchy of representations. This algorithm, called the spatial semantic hierarchy, represents space as a collection of paths that intersect at particular places. The model represents this space at three levels of abstraction: local control laws (e.g., move down the hallway), topological structure (Path A and Path B intersect at Place C), and finally a metrical level. Object landmark knowledge is not encoded within this hierarchy. Instead, the model relies on place recognition, which is dependent on the local structure of the building. Recent research by Kuipers, Tecuci, and Stankiewicz (2003) has shown evidence for the topological structure proposed by this model.

Defining Landmarks

The term landmark seems to have as many shades of meaning as does the term cognitive map. According to Lynch (1960), a landmark is any element in an environment that can function as a reference point. Siegel and White (1975) defined landmarks as “unique configurations of perceptual events (patterns) . . . [that] identify a specific geographical location” (p. 23). These definitions highlight several important aspects of landmarks. First, landmarks are distinctive features in the environment, a property that allows them to be more easily remembered than other features. Second, landmarks serve as mental reference points for various locations in an environment, which emphasizes the relational nature of cognitive maps. Consequently, landmarks are sometimes considered to be spatial cues that are associated with target locations or behavioral responses, or they are specifically defined as spatial reference points that organize mental representations according to their relations with other elements (Presson & Montello, 1988).

Because landmarks can be virtually any distinctive feature in an environment, selection of them can be somewhat subjective given a sufficiently complex space and depending on the needs of the navigator. Yet in the laboratory, where the participant’s experience of an environment can be manipulated using a virtual system, several methods for the selection of landmarks have been discovered, indicating that the process of incorporating significant external features into a mental representation is not merely random. It has been suggested that landmarks are selected according to their perceptual salience, the frequency of visits to a particular object (Presson, 1987), or the number of locations it is associated with (Sadalla, 1988). The way landmarks are selected also influences how the resulting cognitive map is organized (Hardwick, Worrallidge, & Rinalducci, 1983).

Landmark Properties

Models that propose the use of landmarks as a form of knowledge implicitly assume three properties of a landmark. First, a landmark should be persistent. That is, a given landmark at time $t$ should be present when the navigator returns to that location at some later time. Second, a landmark needs to be perceptually salient. That is, on a navigator’s return to a particular state (position and orientation), the landmark needs to be detectable and identifiable. A subtle landmark may be easily missed. Finally, the landmark needs to be informative. That is, the landmark should provide information about the navigator’s position within the environment and/or what action the navigator should take when observing the landmark. Some landmarks may be more common (e.g., a door in an office building) and, therefore, provide less information about the observer’s position when they are observed than do landmarks that are less common (e.g., a drinking fountain). To better elucidate the role of these properties in identifying a landmark, we have provided some simple examples of landmarks in Table 1.

We hypothesize that when humans are acquiring knowledge about a large-scale space, they are evaluating the landmarks on the basis of these three properties. That is, they want to store landmarks that are persistent, perceptually salient, and informative. For example, Ruddle et al. (1997) showed evidence that participants used object landmarks (statues presented within a virtual environment) when they were made available. In these experiments, the object landmarks were highly informative because there was only one example of each object within the environment. Thus, storage of the landmark and its spatial location was highly informative regarding one’s location. It is interesting to note that when the statues were not meaningful but were like “abstract art,” participants had a more difficult time using them for navigation. We suspect that participants were having a difficult time differentiating between the different statues, and thus, these abstract statues were less informative than the meaningful statues.
Studying Landmark Acquisition and Retention

The three experiments described here compared how well landmarks are learned when their informativeness is manipulated. Informativeness is defined as the number of states (positions and orientations) that an observer can be in given that they observe a specific landmark (structural vs. object). The more unique a landmark is within an environment (i.e., fewer instances within a specific environment), the more informative it is. Using a college campus as an example, a clock tower or a fountain is typically more useful in determining one’s location on the campus than is, perhaps, an ordinary tree.

In these experiments, two types of perceptual information were distinguished: structural landmarks and object landmarks. Certain structural characteristics of an environment, such as long corridors or T-junctions, provide visual cues that are useful for navigation, thus serving the function of landmarks. For example, if the hallway layout of a building contains only one T-junction, participants can make a movement decision on the basis of this unique perceptual information without knowing the overall configuration of the building. The term structural landmarks refers to this information embedded in the structure of the building’s topology.

Although learning structural landmarks is closely linked to developing a survey representation of an environment, a distinction between the two should be made. Whereas encoding structural landmarks implies learning the topology of a space, having a survey representation allows for inference of spatial relations between locations that are not necessarily made perceptually explicit during navigation. Therefore, the definition of structural landmarks, for the current experiments, refers only to a type of visual cue for learning layout topology—it does not refer to the development of a survey representation.

Object landmarks are visual objects within the environment that are independent of its structure, such as paintings on the wall or drinking fountains. These are the visual cues that are typically associated with the term landmark. Structural and object landmarks closely correspond to geometric and featural cues described in studies of animal and human orienting behavior (Cheng, 1986; Hermer & Spelke, 1996).

One of the unique aspects of the current research is that it investigated the role of information in the acquisition of these different types of landmarks. Landmarks are important because they allow a navigator to localize him- or herself and/or decide what action to generate when he or she sees the landmark (e.g., take a left at the drinking fountain). Contemplating the role of landmarks makes one thing intuitively obvious—object landmarks typically provide more information concerning the spatial coordinates of the navigator than do structural landmarks. Knowledge that you are at a T-junction may not be very informative in certain environments (i.e., those in which there are many T-junctions). However, observation of a conference poster on someone’s office door may be unique, and it thus provides a large amount of information about where one is located within an environment. The current experiments recognized this inequality in different landmarks and investigated how manipulating their informativeness affects whether landmarks are actually remembered. In Experiment 1, the information content of the structural and the object landmarks was equivalent, and in Experiment 2, we increased the information content of the object landmarks.

Table 1

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Persistent</th>
<th>Perceptually salient</th>
<th>Informative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student sitting in hallway&lt;sup&gt;a&lt;/sup&gt;</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Small crack in paint on wall</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Presence of an office door in an office building</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Conference poster on an office door</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<sup>a</sup> It is assumed that the hallway is void of other students.

Retention of Landmarks

In this study, retention rate refers to two aspects of the learning curve for landmarks: acquisition and decay. In the current experiments, memories for structural and object landmarks were tracked over time, first to determine whether one type of landmark is learned earlier or at the same rate as the other and then to observe if knowledge of the landmarks is remembered equivalently over periods of delay.

The rationale for tracking acquisition of landmark knowledge derives from the results of previous studies that explored the use of object landmarks in learned environments. Although it is intuitive that object landmarks are used in navigation through an environment, it has been difficult to confirm this empirically. Several studies have compared navigational performance in well-learned environments with and without highly informative object landmarks and have shown no difference in the acquisition of survey knowledge (Ruddle et al., 1997; Wilson, Foreman, & Tlauka, 1997). Nevertheless, participants reported seeing and using the object landmarks. Also, more salient objects decreased travel distance without affecting overall survey knowledge (Ruddle et al., 1997). When other strategies were suppressed during testing, participants successfully used object landmarks to navigate through an environment (Tlauka & Wilson, 1994). These findings suggest that meaningful object landmark information is encoded while an environment is being learned even if it is not used during navigation.
Some studies have attempted to demonstrate that object landmarks are used within a view-based framework (Mallot & Gillner, 1998). Accordingly, navigation is a process of comparing an observed image with stored views of landmark configurations. This matching is then associated with a specific action or a specific state (position and orientation) within the environment. However, support for this theory does not necessarily provide evidence for the use of object landmarks, because each "view" most likely contains structural landmarks in addition to object landmarks. Mallot and Gillner even remarked that most participants did not notice when object landmarks had been shifted from their original locations in three of the four conditions.

The second reason for exploring retention rate for structural and object landmarks is to determine how well this information is retained after periods of delay. It is a common task to find one's way around a place that is not visited on a regular basis. For example, a student may have to find a specific room in a building on campus to attend a class once a week, or a person may revisit a city after several years and have to navigate to a friend's house. In such cases in which an environment is not visited frequently, what visual information about the environment is retained to allow for successful navigation? This study addressed this question by testing structural and object landmark knowledge after several time delays: 1 day, 1 week, 1 month, and 1 year after initial learning of a building layout. Previous studies and Siegel and White's (1975) hypothesis suggest that object landmarks may not be useful once the topology of an environment is learned. If this is the case, perhaps it is not necessary to retain object landmark information over long periods of time. As for structural landmarks, if the ultimate goal of learning a space is to develop an understanding of the topography of the layout, then structural landmark knowledge should be retained. Thus, we predicted that object landmark knowledge would decline over time, whereas structural landmark knowledge would be more resilient over time.

SUMMARY

In this article, two types of visual information, structural landmarks and object landmarks, are defined as visual cues that can be used by an observer learning to navigate novel indoor environments. The experiments reported below explored two main issues regarding the use of these landmarks. First, we investigated whether landmark informativeness influences the acquisition and retention of landmark knowledge. Second, we tested the rate at which landmark information is acquired and retained over time delays. These issues were investigated in a series of three experiments. Experiment 1 established the retention rate for structural and object landmarks when both have equal informativeness. Experiment 2 looked at how retention was affected when object landmarks are more informative than structural landmarks. Finally, in Experiment 3, we controlled for the nonindependence, or predictability, of the structural landmarks within the environment.

EXPERIMENT 1

Experiment 1A

In Experiment 1A, we investigated memories for structural versus object landmarks over varying delay periods when the two landmark types were equally informative. According to Siegel and White (1975), participants should initially learn the landmarks (in this study, referred to as object landmarks), followed by the structural properties of a layout (i.e., structural landmarks) as they develop a survey representation of the environment. However, when these landmarks are equally informative, we predict that participants will be biased toward those landmarks that have other advantageous properties described previously, such as temporal persistence and/or salience (the probability of accurately detecting and identifying a landmark while navigating; see the Landmark Properties section in the introduction).

Method

Participants

Eight participants (4 male, 4 female) were tested in Experiment 1. Participants were 7 undergraduate and graduate students and 1 professor (one of the authors, B.J.S.) from the University of Texas at Austin. Participation was voluntary. All participants, with the exception of B.J.S., were monetarily compensated for their time at a rate of $8.00/hr.

Materials

This experiment used randomly generated indoor environments viewed with desktop virtual reality. The environments were built on a grid system so that all intersecting hallways were perpendicular. Figure 1 is a map illustration of the type of environment used. One hallway unit corresponds to a line connecting two dots on the grid, and each connected dot corresponds to a position in the environment. Each environment consisted of 40 hallway units. To move through the environment, participants made keypresses using the number pad of a computer keyboard. The 8 corresponded to moving forward 1 hallway unit, the 6 rotated the participant clockwise by 90°, and the 4 rotated the participant counterclockwise by 90°. When participants made an action, the computer generated the appropriate collection of images to give the optic flow of rotating or translating within the environment (a translation took approximately 1 s, and a rotation took approximately 750
ms). The environments also contained seven target positions labeled with audio clips that played when participants walked over one of these locations (“Position 1,” “Position 2,” etc). Participants were tested in one of two different environments to ensure that the results could be generalized to different environments and that any effects were not due to idiosyncratic properties of a particular environment.

**Structural landmarks.** Structural landmarks were defined as the specific hallway structures in the environments. For each hallway juncture (i.e., in Figure 1, any dot on the grid connected to another dot by a hallway unit), there were four possible combinations of structures: walls to the left and right, hallways to the left and right, a wall on the left and a hallway on the right, and a hallway on the left and a wall on the right. Figure 2 depicts a view in a sample environment as it was seen by participants. The structural landmarks of the immediate juncture are hallways on the left and right (i.e., hall-left and hall-right).

**Object landmarks.** Every hallway unit in the environment contained two object landmarks in the form of pictures on the walls. There was one picture on the left wall and one on the right wall. In Experiment 1A, each object landmark could be one of two possible pictures (an apple or a fish). Therefore, the environment contained 80 randomly distributed pictures: 40 pictures of the apple and 40 pictures of the fish. In Figure 2, the object landmarks are a picture of the fish on the left wall and a picture of the apple on the right wall.

**Procedure**

Participants engaged in a training phase and a testing phase. During the training phase, participants started from a start state in the environment and were allowed to freely explore for 100 forward actions. During this training phase, participants were instructed to explore the environment and learn as much about the environment as possible. They were also told that there were seven locations within the environment that would serve as target locations. These target locations were specified by auditory signals that were given when the participant “walked” over the location. The computer generated these auditory signals that said “Position X,” where X was the target number (1–7). Participants were informed that during the testing phase, they would be given a series of target locations to navigate to. They were also told that they would be queried about the hallways and pictures in the environment.

In the testing phase, participants started from the same start state and were given a target position by the computer. The instruction came in the form of a computer-generated voice that said, “Go to Position X” (where X was 1–7). Participants then began to navigate to the target position. After 3–5 translations (the number of translations was randomly selected by the computer), the participant was tested about his or her knowledge of either the structural or the object landmark. The testing was completed by covering the next intersection (see Figure 3) or covering the pictures in the hallway corridor (see Figure 4). On a structural landmark trial, the participants’ task was to indicate what structural landmarks were behind the covers (hall vs. wall), and in the object landmark condition, their task was to indicate what object landmarks were behind the covers (fish vs. apple).

After participants answered the structural or object landmark query, the landmarks were uncovered and the participants continued navigating to the target location. When a participant reached the target location, the computer stated “Position X, go to position Y.” X was the current target location number, and Y was a new target location that was randomly selected (excluding the current target location). A single testing phase ended after 20 queries of landmark knowledge (10 structural and 10 object, in random order). This was followed by another training phase.

The entire experiment on the 1st day consisted of 10 blocks, each block containing one training phase followed by a testing phase. After the 1st day, participants participated in a series of 10 test blocks 1 day, 1 week, and 1 month after the initial study. Each testing phase consisted of 50 queries of landmark knowledge (25 structural and 25 object), instead of just 20 as on the initial day of testing. The goal of this paradigm was to plot the accuracy of the two types of landmark knowledge with respect to the amount of exposure to the environment and the amount of time after initial acquisition of spatial knowledge.

**Results**

A two-alternative forced choice (2AFC) $d'$ analysis was completed on the participants’ data. The mean performance and standard errors of the means are presented in Figure 5. A 2 (landmark type: structural vs. object) $\times$ 4 (delay: 0, 1, 7, or 30 days) univariate analysis of variance (ANOVA) was computed for data compiled across participants. Results indicated that participants had greater accuracy identifying structural landmarks than object landmarks and that there was an effect of delay. Figure 5 depicts the accuracy of structural and object landmark knowledge as a function of delay across participants. There was a significant main effect of landmark type, $F(1, 6) = 14.275, p < .01$, with structural landmarks being remembered more accurately than object landmarks (structures: $M = 1.956, SE = 0.185$; objects: $M = 1.153, SE = 0.120$). The compiled data also showed a significant effect of delay, $F(3, 18) = 11.544, p < .001$, but no Delay $\times$ Landmark Type interaction.
The primary dependent measure in Experiment 1A was a participant’s ability to remember either structural or object landmarks. Although this provides us with an insight into the information encoded about a large-scale space, it does not provide any information about the participant’s ability to navigate through this space. Ultimately, we are interested in understanding how knowledge acquisition leads to better navigation performance. To this end, we computed a measure of participants’ abilities to navigate between the seven target locations in the two environments. More specifically, we measured how efficiently participants traveled between each of the target locations. It should be pointed out that participants were never instructed to take the shortest routes; however, we suspect that participants may have naturally adopted this strategy.

For each participant, we computed their navigation efficiency (NE). This was done by first computing the shortest distance between all possible combinations of the seven target locations ($\text{MinDist}_{\text{TargPair}}$). For each session we computed which target pairs were given to the participant, and we computed the number of actions it would have taken the participant to complete these target pairs if he or she had taken the shortest routes between these target pairs (the sum in the numerator of Equation 1). We divided this by the total distance actually traveled in the session to give us a navigation efficiency measure:

$$NE = \frac{\sum_{\text{TargPair} \in \text{TargetsReached}} \text{MinDist}_{\text{TargPair}}}{\text{TotalDistance}}.$$

As $NE$ approaches 1.0, participants are getting closer to taking the shortest route between two target locations. As the number approaches 0.0, participants are wandering more and more through the environment.

Figure 6 illustrates the average navigation performance as a function of delay period. The results show that participants were least efficient at navigating on the training day, and their efficiencies improved as a function of delay. The results in Figure 6 are remarkably similar to those in Figure 5. In fact, the correlation between navigation efficiency and structural memory was significant, $r(30) = .840, p < .01$, and the correlation between navigation efficiency and object memory was also significant, $r(30) = .726, p < .01$. These results show a clear relationship between memory for landmarks and navigation abilities.
Initial Learning

According to Siegel and White (1975), participants should initially learn a set of landmarks, followed by routes between those landmarks, and finally a survey representation of the space. Figure 5 shows that even on the 1st day, participants encoded structural landmarks better than object landmarks. However, it is possible that in this 1st day of training and testing, participants initially learned the object landmarks and then eventually, through experience, learned the structural landmarks. To investigate this issue, we measured the participant’s ability to identify the structural and object landmarks in each block of trials in the 1st day. We used d’ analysis to compare accuracy of structural and object landmark knowledge in the first 10 sessions conducted on Day 1 of testing (see Figure 7). A 2 (landmark type: structural vs. object) × 10 (Sessions 1–10) ANOVA was calculated for the first 10 sessions conducted on Day 1 of testing. Results indicated that there was a significant effect of landmark type, F(1, 140) = 8.473, p = .004, with greater accuracy for structural landmarks (M = 1.40, SE = 0.115) than for object landmarks (M = 0.943, SE = 0.102).1

Experiment 1B

We were presented with a unique opportunity to run participants following a 1-year delay. Three of the participants (1 of them being B.J.S.) were available for data collection approximately 1 year after the original study. We were interested in understanding how resilient the memory was for the two different types of landmarks as a function of a long delay period.

Method

The procedure for Experiment 1B was the same as that for Experiment 1A. In Experiment 1A, participants ran in one of two different environments, whereas in Experiment 1B they ran in both of these environments. The environment participants had learned in Experiment 1A was the familiar environment, and the second, unexplored environment was the novel environment. Participants ran in the familiar and novel environments in random order. In each environment, they were given an initial exploration period of 100 forward actions followed by 10 test blocks of 50 landmark queries (25 structural and 25 object). The query procedure was identical to that used in Experiment 1A.

Results

We used d’ analysis to compare accuracy of structural and object landmark knowledge following a 1-year delay. Average d’ performance for structural and object landmarks in both familiar (environments that participants ran in previously) and novel environments are illustrated in Figure 8. To provide a relative measurement, we also plotted the average performance for the 3 participants in the 30-day delay condition from Experiment 1A.

We computed paired t-tests to determine the effect of delay on memory. There was a significant difference in performance between the familiar and novel environments in the 1-year delay condition for structural landmarks, t(2) = 3.3400, p < .05, and for object landmarks, t(2) = 6.6921, p < .05. There was no significant difference between the 30-day delay and the 1-year delay for structural landmarks, t(2) = 1.5218, p = .2675, or the object landmarks, t(2) = 1.3056, p = .3056. The lack of an effect between the 30-day and 1-year delays may be attributable to the small sample size. However, we were able to detect a reliable effect between the familiar and novel conditions for structural and object landmark accuracy.

Discussion

The purpose of this experiment was to understand the baseline memory for structural and object landmarks when the information

1 It should be noted that due to errors during data collection, some participants were queried a greater number of times than indicated in the procedure description. This analysis was done using only the first 10 queries for both structural and object landmarks in each block.
Participants were still learning the environment. The second surprising finding was the resiliency of memory for the landmarks. There was very little decrement in memory after a 30-day delay or even after a 1-year delay for a subset of participants. This result suggests that once a spatial representation is established, it remains relatively stable over a long delay. The long-term retention of this spatial memory is obviously very adaptive for the spatial navigation system, because often navigators need to reference their cognitive map for a specific environment after a significant delay (months or even years).

We believe that the acquisition of building structure is similar to a participant attempting to extract a survey or topological representation of an environment. Specifically, the present results are consistent with a system that is attempting to extract the global topology of the large-scale space from the local topology. Under conditions in which the object landmarks provide similar information about the participant’s location as the local topology, we find that participants are biased toward remembering the local topology over the object landmarks. These results are somewhat inconsistent with the strong interpretation of Siegel and White (1975) that stresses the acquisition of object landmark knowledge before route and survey knowledge.

These results are consistent with the type of topological representations proposed by Kuipers (2000, 2001) and Kuipers and Byun (1991). Kuiper’s spatial semantic hierarchy argues that right from the beginning, one should focus on extracting a topological representation of a space that is composed of the structure of the building. In fact, according to this approach, even place recognition is based on the local structure of the building. It should be pointed out that the spatial semantic hierarchy is a theory of spatial knowledge acquisition and use that is motivated by human behavior but has been implemented in robots. It is a well-known fact that adding computer vision to an autonomous robot is a difficult task, and thus the lack of object landmark identification in this theory might be a result of this limitation. However, it should be noted that this theory has been used to model and explain human navigation through complex spaces similar to those used in the present study (Kuipers et al., 2003).

**EXPERIMENT 2**

Experiment 1 showed a bias for remembering structural landmarks over object landmarks when the two types of landmarks were equally informative. Experiment 2 investigated the effect of increasing the information content of the object landmarks relative to the structural landmarks. To do this, we increased the variability of the landmarks from two different pictures (one of a fish and one of an apple) to eight different pictures. By increasing the variability of landmarks, we increased the information content of the object landmarks relative to the structural landmarks.

Again, we define the information content of a landmark as its ability to reduce an observer’s uncertainty about his or her current state (position and orientation) in the environment. To demonstrate this, we computed the conditional entropy (uncertainty) for the environments used in Experiments 1 and 2. The equation to compute the conditional entropy is as follows:

\[
H_{\text{State}}(LM) = \sum_{lm \in LM} p(lm) \sum_{s \in \text{State}} p(s|lm) \log \left( \frac{1}{p(s|lm)} \right)
\]  

(2)
The conditional entropy computes the degree of uncertainty \( (H) \) about what state the observer is in \( (\text{State}) \) given the local landmarks \( (\text{LM}) \). In Experiment 1, simply identifying the local structural landmarks or the local object landmarks provided the observer with the same amount of uncertainty (about 5 bits or, on average, about 32 different states in the environment).\(^2\) In Experiment 2, we increased the number of different object landmarks to eight, which increased their information content. This is shown by a lower entropy value for the object landmarks than for the structural landmarks. Identifying the local structural landmarks still generates about 5 bits of uncertainty (32 out of 160 states), whereas identifying the local object landmarks reduces the uncertainty to 1 bit (2 states). A plot of the average conditional entropy for Experiments 1 and 2 is shown in Figure 9.

We hypothesize that with the information content of the object landmarks increased, participants will be more likely to remember these landmarks. That is, according to Table 1, participants will be more likely to encode the object landmarks in Experiment 2 than they were the object landmarks in Experiment 1. Although there should be an increase in memory for the object landmarks in Experiment 2 over Experiment 1, it is not clear whether such an increase in information would overcome the prior bias for encoding hallway structure.

**Participants**

Six participants (3 male, 3 female) were tested in Experiment 2. Participants were 5 undergraduate or graduate students at the University of Texas at Austin and 1 professor (B.J.S.). Participation was voluntary. All participants, with the exception of B.J.S., were monetarily compensated for their time at a rate of $8.00/hr. Two participants were tested in both Experiment 1 and Experiment 2 using a different environment for each experiment.

**Materials**

This experiment used two randomly generated indoor environments similar to those used in Experiment 1. These environments were viewed in the same way as those in Experiment 1. The primary difference between the two environments used in Experiment 2 and Experiment 1 was the collection of images used as the object landmarks. In the current experiment, the objects were the following pictures: an apple, a fish, a butterfly, a flower, a bird, a shoe, a phone, and a tiger.

**Procedure**

Experiment 2 used virtually the same procedure as Experiment 1. The only difference was that in the object landmark queries, participants entered numbers from 0 to 7 that corresponded with each type of object. Participants were given a reference sheet that had the object landmark’s name and its corresponding response.

**Results**

We computed \( d' \) analyses on the data. For the structural objects, we computed a 2AFC analysis, and for the object landmarks, we computed an 8AFC analysis. The mean performance and standard errors of the mean are presented in Figure 10. A 2 (landmark type: structural vs. object) \( \times 4 \) (delay: 0, 1, 7, or 30 days) ANOVA was computed for data compiled across participants and across days of testing. Unlike in Experiment 1, there was no significant effect of landmark type, \( F(1, 5) = 2.860, p = .152 \), but there was still a significant effect of delay, \( F(3, 15) = 9.624, p < .001 \).

**Discussion**

Experiment 2 investigated participants’ memory for structural and object landmarks when the information content of the object landmarks was greater than the information content of the structural landmarks. Information content was defined as the amount of state uncertainty (the number of locations that an observer could be in if they identified the local landmarks) given correct identification of the structural landmarks versus the object landmarks. As shown in Figure 9, the conditional entropy for identifying the object landmarks leaves the participant with approximately 1 bit of uncertainty (on average, 2 states within the environment). By contrast, by identifying the structural landmarks, the participant has 5 bits of uncertainty (or an average of about 32 states within the environment). We predicted that if participants are sensitive to the information content of the landmarks, their memory for the

\[^2\] This analysis was completed only on the immediate landmarks (those immediately in front of the observer). We designed the environments in Experiment 1 so that if the participant considers landmarks beyond the immediate set of landmarks, this behavior will reduce the conditional entropy for both the object and structural landmarks at roughly the same rate.
object landmarks should improve compared with Experiment 1. We found that there was a significant effect of delay (as found in Experiment 1) but no effect of landmark type across participants. In this experiment, participants were more sensitive to remembering object landmarks than they were in Experiment 1.

Figure 11 shows the difference in performance between Experiment 2 and Experiment 1 ($d'$ [Experiment 2] − $d'$ [Experiment 1]). This illustration reveals an interesting pattern of effects. The memory for the object landmarks increases in Experiment 2 relative to Experiment 1. At the same time, the memory for the structural landmarks decreases (as shown by the negative difference). This result suggests a specific memory or attentional limitation when participants are acquiring this landmark knowledge. That is, when the information content of the object landmarks is greater than that of the structural landmarks, participants seem to attend to these landmarks more at the expense of the structural landmarks. With this said, one needs to be careful about drawing these conclusions. The participants used in Experiment 1 were different from those used in Experiment 2, and thus, they may have used slightly different strategies.

The environments used in Experiment 2 were more similar to those used in previous experiments investigating the use of landmarks in human spatial navigation. That is, the objects provide a large amount of information about the navigator’s position within the environment. Yet even under these conditions, we did not find a bias in initially remembering object landmarks over structural landmarks. According to a strong interpretation of Siegel and White’s (1975) hypothesis, we should have found that participants were initially biased toward remembering the object landmarks (perhaps in the Delay 0 and 1 conditions) but that with experience, they eventually learned the structure of the environment (through the acquisition of routes and/or a survey representation).

One might question why the memory for the object landmarks in Experiment 2 was not greater than that for the structural landmarks. We hypothesized that information content is a factor in whether a participant chooses to store a landmark in memory. Given that the object landmarks had more information content, one might have predicted that the participants would have been more likely to store those landmarks than the structural landmarks in memory. Our claim is not that information content is the sole factor in deciding what landmarks to store in memory, but it is one of the factors. Other factors include the persistence and the perceptual salience of the landmark (see Table 1).

We speculate that participants may have a bias for remembering the structural landmarks over the object landmarks even when the structural landmarks have less information content for one of two reasons. First, in most environments, structural landmarks are more persistent than object landmarks. Thus, there is an inherent advantage for remembering these landmarks. Second, participants are attempting to extract a topological representation of the space even from the beginning. The structural landmarks provide critical information for generating that representation; thus, participants concentrate on acquiring that knowledge.

**EXPERIMENT 3**

Experiment 1 showed a clear effect of landmark type, and Experiment 2 showed that even when the information content of the object landmarks was greater than that of the structural landmarks, participants remembered both equally well. However, these two types of landmarks differed not only in their form (structure vs. object) but also in their level of independence. This independence–dependence is illustrated in a simple environment in Figure 12. The upper portion of the figure illustrates four different structural landmarks (A, B, C, and D), and the lower
portion of the figure illustrates four different object landmarks. As can be seen from the figure, by knowing that Structural Landmark C is a hallway, one can infer that the Structural Landmark D must also be a hallway. That is, if you know the type of structural landmark at C, you can infer the structural landmark at D. However, for the object landmarks, E (for example) provides no information about the identity of F.

Experiment 3 investigated the role of this nonindependence on structural versus object landmarks. To investigate this issue, we used two new randomly generated layouts containing the same two object landmarks used in Experiment 1. However, in Experiment 3, the object landmarks were arranged such that they had the same interdependence as the structural landmarks. In other words, the object landmarks on adjacent walls in adjacent hallways (see Object Landmarks E and F in Figure 12) were identical. Thus, knowing the identity of one object landmark perfectly predicts the identity of the corresponding object landmark in the adjacent hallway.

**Method**

**Participants**

Four participants (2 male, 2 female) were tested in Experiment 3. Participants were undergraduate or graduate students at the University of Texas at Austin. Participation was voluntary. All participants were monetarily compensated for their time at a rate of $8.00/hr.

**Materials**

This experiment used randomly generated indoor environments similar to those used in Experiment 1. These environments were viewed in the same way as those in Experiment 1. In Experiment 3, there were two different images used for the pictures: a fish and an apple. These images were placed in a structured format so that the image in one hallway was the same on the adjacent wall in the adjacent hallway (when there was an adjacent wall).

**Procedure**

Experiment 3 used the same procedure as Experiments 1 and 2. In the current experiment, participants were told that the pictures on one hallway would be the same as the pictures on the adjacent hallway. Also, participants ran for only 2 consecutive days.

**Results**

We computed a 2AFC $d'$ analysis for the structural and object landmarks. The means and standard errors of the means are presented in Figure 13. A paired $t$ test revealed a significant difference between participants’ memory for object landmarks ($M = 0.6293$, $SE = 0.120$) and their memory for structural landmarks ($M = 1.2270$, $SE = 0.166$), $t(3) = 7.330$, $p < .05$.

**Discussion**

Experiment 3 investigated whether participants’ superior memory for structural landmarks over object landmarks was due to the fact that the structural landmarks were not independent. To investigate this issue, we generated virtual environments in which the object landmarks had the same dependence as the structural landmarks. Even with this dependence, structural landmarks were
remembered better than object landmarks. The results from Experiment 3 suggest that the superior memory for structural landmarks in Experiments 1 and 2, even when the object landmarks were more informative, cannot be attributed to the dependence of the landmarks across hallways.

GENERAL DISCUSSION

The goal of these experiments was to understand what type of landmark information is encoded when an observer is learning to navigate a novel environment. More specifically, these experiments investigated how the encoding of structural and object landmarks differ when their informativeness is manipulated and the rate at which this information is acquired and retained. In Experiment 1, both structural and object landmarks had equal information in terms of informing participants about their current state. In this experiment, we found that participants remembered the structural landmarks better than the object landmarks. In Experiment 2, we wanted to determine whether participants were sensitive to the information content of the landmarks. When the information content of the object landmarks was greater than that of the structural landmarks, participants’ memory for the two types of landmarks was the same. Further analysis of this data, comparing performance in Experiments 1 and 2, showed that there was some evidence of a “push–pull” system in which attempting to encode one type of landmark seemed to have a negative affect on encoding the other type of landmark. Experiment 3 showed that the bias for remembering structural landmarks over object landmarks in Experiment 1 was not a result of the nonindependence of these landmarks within the environments.

In Experiments 1 and 2, the acquisition of structural and object knowledge was tracked as participants were initially learning the unfamiliar environments during the first day of testing. This procedure allowed for testing of Siegel and White’s (1975) proposal that object landmarks are learned before structural information. In both experiments, participants remembered structural landmarks more accurately than object landmarks, although the distinction was not as great in Experiment 2. Even at this early stage in learning, these two types of visual information are treated differently, and there is a clear bias toward learning structural landmarks.

To compare the retention of structural and object landmark knowledge over time, we tested participants on this information 1 day, 1 week, and 1 month after the initial day of training in Experiments 1 and 2 (and a subset of participants were tested approximately 1 year later). Although an effect of delay was found, this result seems to indicate that participants continued learning structural and object landmarks after a period of delay rather than forgetting the information that they learned previously, contrary to previous predictions. Furthermore, participants seemed to retain structural and object landmark knowledge equally well following a 1-year delay. This seems to be a testament to the resiliency of the cognitive map as a construction for storing information about an environment—that is, once this information is encoded, it does not seem to be lost. Furthermore, observers are able to retain both the topological structure of an environment and the object landmarks within the environment.

Factors Influencing Landmark Memory

One question that arises from these results is why participants are inclined to learn structural landmarks better than they are object landmarks. It seems clear that different sources of visual information are treated differently when one is learning about an environment. As shown in Figure 9, knowing the identity of a specific pair of structural landmarks and object landmarks in Experiment 1 left a participant with a similar amount of state uncertainty. If one assumes that the observer is storing information that will reduce state uncertainty, then there was no advantage to encoding one type of landmark over the other. Despite this, participants still remembered structural landmarks better than they did object landmarks. In Experiment 2, in which object landmarks were more informative than structural landmarks (see Figure 9), participants’ memory for object landmarks was about the same as their memory for the structural landmarks.

These results are inconsistent with theories that propose that spatial navigation is accomplished by storing images of the environment. For example, the view-graph hypothesis argues that images or views are stored in memory that are associated with specific actions (Mallot, Franz, Schölkopf, & Bülthoff, 1997; Mallot & Gilmer, 2000; Schölkopf & Mallot, 1995). Because these models use noninterpreted images, there should be no bias for remembering one type of landmark over another. In fact, memory for one type of landmark, according to these theories, should be perfectly correlated with the other type of landmark since they both appear in a view. This was not what we found in the current experiments.

The greater accuracy of structural knowledge can also be a result of how participants directly interact with the structure of the environment when making movement decisions. Previous studies have shown that landmarks at decision points (locations where heading changes are made) are remembered better than those that are remembered at other states (Aginsky et al., 1997). It is possible that the interactions with the corridors reinforced memory for these landmarks. This would explain the increase in memory for the structural landmarks over the object landmarks in Experiment 1 but not the lack of an advantage for the structural landmarks in Experiment 2.

These experiments were premised on the hypothesis that there are three primary factors influencing the storage of a landmark into long-term memory: temporal persistence, perceptual salience, and informativeness. That is, when learning about an environment, participants are evaluating landmarks on the basis of these three factors to decide what to store in memory. Landmarks that are highly informative are stored in memory, whereas those that are not as informative may not be stored in memory. Furthermore, potential landmarks that are not salient or not persistent may not be stored in memory either. It is possible that over time, these landmarks may eventually be stored in memory. We hypothesize that, initially, landmarks that are salient, persistent, and informative will be stored before those that are nonpersistent, noninformative, and less informative.

The stated hypothesis is different but not entirely inconsistent with that proposed by Siegel and White (1975). For example, given an environment that has highly informative (perhaps unique) object landmarks, the current proposal predicts that those landmarks would be stored in memory before the structural landmarks.
These highly informative landmarks may allow for successful navigation with minimal memory storage. However, the memory may be highly dependent on the set of start and goal positions and not robust for navigating between start–goal positions that have not been traveled before. As suggested by Siegel and White, participants may then begin to generate a more robust representation by extracting a survey representation or a topological representation (Kuiipers, 2000, 2001; Kuiipers & Byun, 1991). To this end, the cognitive system may then begin to encode and attend to the structural landmarks.

Summary and Conclusions

In summary, there are three primary findings from the present experiments. First, all other things being equal, humans have a natural bias for remembering structural landmarks over object landmarks. Second, participants are sensitive to the information content of a landmark. Third, the memory representations acquired for large-scale space are highly resilient. When a landmark provides more information about the observer’s state in the environment, it is more likely to be stored in memory. However, information content is not the only variable that is used to evaluate whether a specific landmark will be stored in memory. Other factors might include perceptual salience and temporal persistence. These results suggest that participants are interpreting the images that they are observing to infer the properties of specific landmarks within the environment.

References


Received July 19, 2005
Revision received March 6, 2006
Accepted March 12, 2006