Available online at www.sciencedirect.com
ScienceDirect
Cognition 104 (2007) 231–253
COGNITION
www.elsevier.com/locate/COGNIT

Learning your way around town: How virtual taxicab drivers learn to use both layout and landmark information ☆, ☆☆

Ehren L. Newman, Jeremy B. Caplan, Matthew P. Kirschen, Igor O. Korolev, Robert Sekuler, Michael J. Kahana *

Brandeis University, Waltham, MA, 02254-9110, USA

Received 22 May 2006; accepted 23 May 2006

Abstract

By having subjects drive a virtual taxicab through a computer-rendered town, we examined how landmark and layout information interact during spatial navigation. Subject-drivers searched for passengers, and then attempted to take the most efficient route to the requested destinations (one of several target stores). Experiment 1 demonstrated that subjects rapidly learn to find direct paths from random pickup locations to target stores. Experiment 2 varied the degree to which landmark and layout cues were preserved across two successively learned towns. When spatial layout was preserved, transfer was low if only target stores were altered, and high if both target stores and surrounding buildings were altered, even though in the latter case all local views were changed. This suggests that subjects can rapidly acquire a survey representation based on the spatial layout of the town and independent of local views, but that subjects will rely on local views when present, and are harmed when associations between previously learned landmarks are disrupted. We propose that spatial navigation reflects a

☆ This manuscript was accepted under the editorship of Jacques Mehler.
☆☆ This research was funded by National Institutes of Health Grants MH55687, MH61975 and MH68404. We thank Ben Burack for creating our virtual reality engine and Greg Schwartz for technical assistance with the OpenGL Utility Toolkit libraries.

* Corresponding author. Present address: Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA.
E-mail address: kahana@psych.upenn.edu (M.J. Kahana).

doi:10.1016/j.cognition.2006.05.013

hierarchical system in which either layout or landmark information is sufficient for orienting and wayfinding; however, when these types of cues conflict, landmarks are preferentially used.

Keywords: Spatial memory; Navigation; Virtual reality; Landmark; Layout; Spatial learning

1. Introduction

The cities, neighborhoods, and buildings in which we live are rich in spatial structure, and the ability to orient within that structure is crucial for effective navigation. The opportunity to move through an environment allows people to integrate various routes into a *cognitive map* – a mental model of objects' spatial configuration that permits navigation along optimal paths between arbitrary pairs of points (Tolman, 1948). Previous research has pointed to environmental landmarks (salient objects) and environmental layout (geometrical and topological properties of spaces, also known as survey knowledge) as distinct means by which people orient in environments (Kaplan, 1976; Lynch, 1960; Montello, 1998; Trowbridge, 1913; Trullier, Wiener, Berthoz, & Meyer, 1997). Evidence points to these types of spatial knowledge, along with route knowledge, operating together and even learned at the same stages of experience and development (Montello, 1998; Peponis, Zimring, & Choi, 1990) rather than in distinct stages of learning or development, as previously thought (Hart & Moore, 1973; Siegel & White, 1975). However, there is little understanding of whether and how landmark and layout information are integrated when both are available. One formidable obstacle to understanding the integration of landmark and layout information is the fact that in natural environments, the two types of cues are usually correlated, although the nature of that correlation varies from one environment to another. As a result, special tools are needed to examine the two separately and in combination.

1.1. Layout information

Multiple sources of prior research point to the importance of layout information in orienting and navigation. Loomis, Lippa, Klatzky, and Golledge (2002) showed that blindfolded subjects have little difficulty navigating simple paths of a few segments to a remembered target location. Thus, in such extreme cases, in which landmarks are not available, subjects can still perform navigation tasks involving few path segments. Fujita, Klatzky, Loomis, and Golledge (1993) showed how a simple model could account for the pattern of errors in these blindfolded wayfinding data if systematic variability was introduced at encoding but not retrieval. Given that blindfolded subjects lack visual feedback about the progress of their navigation, their mode of navigation is likely to differ in important ways from the behavior of sighted navigators, who can draw on both layout and landmark information.

Investigating a much simpler type of layout information, Hermer and Spelke (1994) found that young children (ages 1.5–2 years old) will use layout information and not landmark information when orienting in a very simple environment – a rectangular room. This contrasts with the performance of adults (ages 17–26 years old) who utilize both landmark and layout information. However, when required to verbally shadow a tape-recorded passage while performing the reorientation task, adults ceased to use landmark information to reorient (Hermer-Vazquez, Spelke, & Katsnelson, 1999).

In a more formalized approach to analyzing the relationship between layout and wayfinding behavior, Hillier, Burdett, Peponis, and Penn (1987) introduced a graph theoretic spatial syntax, focusing on environmental features that included axial lines, connectivity and lines of sight. Haq and Zimring (2003) showed that this spatial syntax influences wayfinding behavior. In a related approach, Kim and Penn (2004) reported that the way subjects drew sketch maps of their neighborhood reflected biases toward axial lines with high numbers of branch points; a similar conclusion was drawn by Kuipers, Tecuci, and Stankiewicz (2003) in matching robotic behavior to human wayfinding behavior. Similarly, Peponis et al. (1990) found that subjects made especially frequent use of spaces through which many locations were accessible, including at times when the target location was unknown. Trullier et al. (1997) summarize how various models have made use of layout information to drive wayfinding behavior.

1.2. Landmark information

Other studies have focused on the role of landmarks in navigation. Examining navigation in virtual towns, Mallot and Gillner (2000) showed that subjects learned associations between landmarks and the direction of a turn at an intersection. Switching these landmarks (with other landmarks from the environment) impaired subjects' performance. Thus, landmark information clearly has value in orienting and navigation. However, one could conceptualize different ways in which landmark cues could place a subject within different types of spatial maps. Kaplan (1976) proposed that multiple landmarks are associated together based on their contiguity, and this information could be used to orient to a reference direction as well as to identify choice points, especially in routes. Mou and McNamara (2002) and Shelton and McNamara (2001) suggested that subjects learn landmark-to-landmark associations, and use those to orient during navigation. This was further supported by Mou, Zhang, and McNamara (2004) in cued recall of spatial configurations learned from verbal narratives. Interestingly, McNamara and colleagues refer to subjects as using clusters of landmarks as intrinsic frames of reference, thus blurring the distinction between landmark and layout information; this could also be said of the Mallot and Gillner (2000) findings. Both Benhamou, Bovet, and Poucet (1995) and Schölkopf and Mallot (1995) implemented the notion of landmark-to-landmark associations in a model which learns to associate views with one another, which are, in turn, linked with a separate reference direction. Alternatively, Couclelis, Golledge, Gale, and Tobler (1987) suggested that subjects use landmarks as anchor points for orienting.

1.3. Layout and landmark information

Although the work summarized above establishes the importance of both landmarks and layout information as orienting cues, these studies have yet to address a crucial question: when both types of information are available, how do they interact to control wayfinding? Whereas specific models have implemented algorithms that make use of both landmark and layout information (cf. Trullier et al., 1997), little empirical evidence exists to constrain them. We addressed this question using virtual environments. Virtual reality technology affords the opportunity to tailor virtual environments for research, and to measure navigation behavior in great detail. The capacity to generate environments whose characteristics can be controlled and manipulated, and the ability to create multiple environments that differ along specific dimensions, are crucial for studying the interacting roles of landmark and layout cues during navigation. Under realistic conditions with natural environments, these two sources of information may be correlated to varying degrees. Further, spatial representations generated through active navigation of virtual environments are comparable to those formed in navigation of real environments (Péruch & Gaunet, 1998; Péruch, May, & Wartenberg, 1997; Tlauka & Wilson, 1996; Tong, Mariin, & Frost, 1995; Whitmer, Bailey, Knerr, & Parsons, 1996). Thus, findings generated from virtual reality experiments has substantial relevance to real-world navigation, while allowing considerable experimental control.

We designed a taxi-driver task to test how subjects navigate relative to landmark and layout cues. Subjects maneuvered a virtual taxi through computer-generated, virtual, 3-D-rendered towns using the arrow keys on a standard computer keyboard.[1] Within these towns, subjects searched for passengers ("searching" phase) and then delivered them to target stores that were located at different locations in the environment ("goal-seeking" phase). Subjects earned virtual money for each successful delivery, and expended virtual money as a function of the distance traveled. A display of current earnings provided explicit continuous feedback about success, which mimicked a common condition of everyday navigation. Subjects sought to maximize their earnings by delivering passengers via the most efficient routes possible.

In Experiment 1, subjects navigated two virtual towns. The aim was to test whether subjects could learn to reduce their delivery path lengths as they learned the structure of the town. This first experiment was meant to test our overall method; the subsequent experiment was designed to separate the contributions of landmark and layout information during navigation.

2. Experiment 1

2.1. Methods

2.1.1. Subjects

Thirty Brandeis undergraduates (15 male and 15 female) participated for pay or course credit. An experimental session lasted approximately one hour.

[1] For these experiments, we wrote an environment creation and navigation program using C++ in conjunction with the OpenGL Utility Toolkit. Our software may be downloaded from http://memory.psych.upenn.edu.

2.2. Construction of virtual towns

The experiment used three specially designed towns (see Figs. 1 and 2). Each had a unique road layout and a unique set of five stores. Defining the width of a road as one *unit*, the size of the entire town was 8×10 units, for a total area of 80 square units. Of these, 44 square units were covered by roads and could be traversed freely. In most of the town, gray featureless walls rose on either side of the roadway, giving

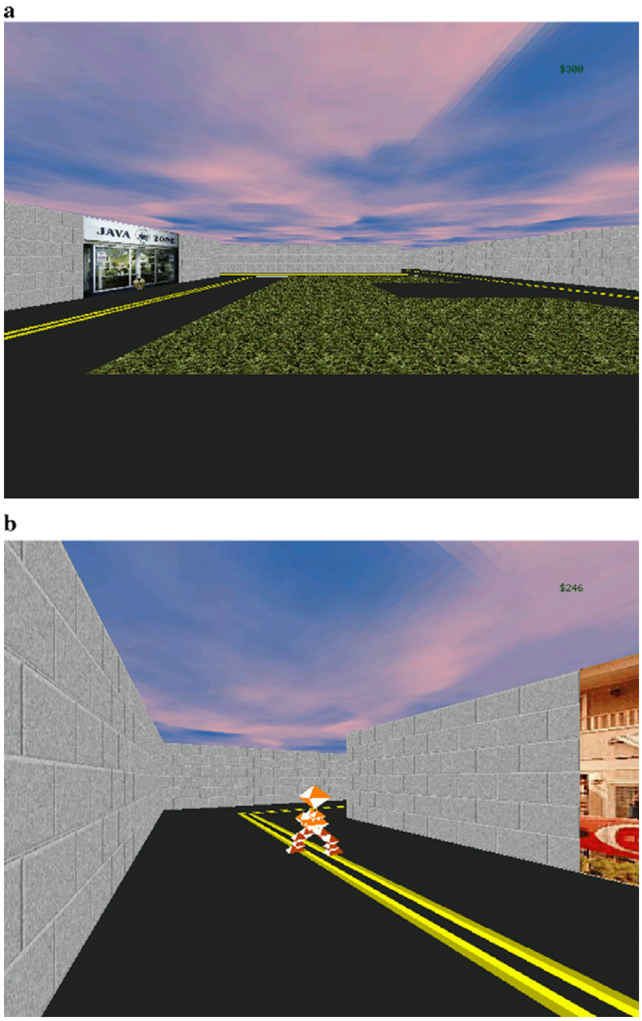

Fig. 1. Screenshots of the virtual towns used in Experiment 1. (a) This image shows road, grass and wall blocks. Movement is restricted to the road blocks. The store fronts were placed in the walls. To make a delivery to a store subjects had to drive close to the store. (b) This image shows a passenger waiting to be picked up. To pick up a passenger subjects had to drive close to the passenger.

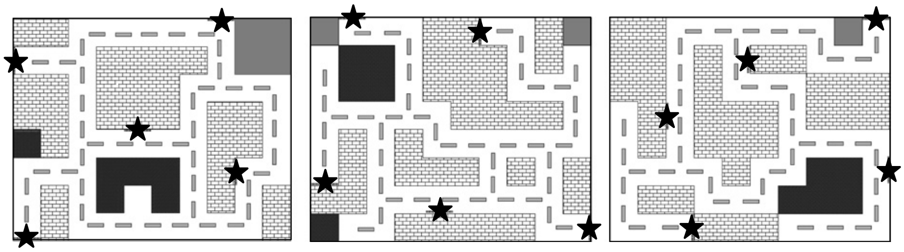

Fig. 2. Blueprints of the three towns used in Experiment 1. The roads are marked by dashed lines. Water and grass are marked by gray and black, respectively. Solid walls are marked with brick texture. The store locations are marked by stars.

the environment a maze-like appearance (Fig. 1). Small ponds and/or lawns occupied several locations in each town. Fig. 2 shows the blueprints of each of the three towns. The stores were constructed by mapping unique photos of real-world stores onto the virtual walls at different locations. Storefronts were 1 unit long and 0.5 units tall. Passengers were represented by diamond-shape humanoid figures (see Fig. 1).

2.2.1. Navigation

Subjects navigated from a first-person point of view (field of view: $106° \times 90°$ in 640×480 pixel mode). Subjects controlled their movement using the four arrow keys on a standard computer keyboard. The ↑ and ↓ keys (*translation* keys) allowed subjects to travel forward and backward, respectively. The ← and → keys (*rotation* keys) caused the cab to rotate in place (counterclockwise or clockwise, respectively). When both a translation key and a rotation key were pressed together the resultant movement was an arcing turn (linear sum of translation and rotation). Movement started immediately when a key was depressed, continuing at a constant velocity until the key was released. The view was refreshed every 40 ms, producing the appearance of smooth movement, which is important for route learning (Kirschen, Kahana, Sekuler, & Burack, 2000). The turning rate was 75 °/s; a full rotation took 4.8 s to complete. The driving speed was 1.6 units/s. Acceleration was instantaneous as soon as a keypress was detected.

During a search phase, a single would-be passenger was placed in the town at a location chosen randomly subject to the constraints that the target storefront was (a) not visible to the driver from the passenger pickup location, and (b) at least five units away. When the passenger was picked up, a text screen instructed the subject to take the passenger to a specific target store. After reading the text, the subject pressed the ENTER key to return to the virtual town. Pickup or delivery occurred when the taxi came within 0.20 units of the passenger or the target storefront, respectively.

As soon as a passenger had been successfully delivered, a text screen informed the driver of this success, urging the driver to search for another passenger. The driver pressed ENTER to return to the virtual town. Upon a successful delivery, the value was incremented by \$20. Subjects were charged \$1 for every 10 s of movement and charged \$1 for any continuous period of standing still longer than 10 s. This reward schedule was designed to favor direct paths over short navigation times. The subject's total current earnings were shown in the upper right-hand corner of the screen (Fig. 1) at all times.

2.2.2. Procedure

Each driver began by navigating through a practice town (not one of the three towns shown in Fig. 2). Practice consisted of one delivery to each of 5 stores (none of which were used in the test towns) placed along the four walls of an empty room.

After this practice phase, subjects navigated two of the three towns shown in Fig. 2. The assignment of towns was counterbalanced so that across subjects each town appeared an equal number of times as the first and second town navigated. Within each town, subjects delivered a total of 25 passengers, with 5 passengers requesting delivery to each of the 5 target stores. Order of deliveries was blocked into *sets* such that (a) subjects delivered to each store once before any store was visited again, and (b) no store was a target twice in a row. On average, about five minutes intervened between successive deliveries to the same target store.

2.2.3. Results

The combination of randomly placed passengers and randomly chosen goal locations encouraged the driver to find flexible, efficient routes for each delivery; thus, each delivery was a novel test of the subject's survey knowledge. We hypothesized that delivery path length would decrease with increased exposure to a given town, and would show evidence of transfer between towns to the extent that those towns were similar.

Fig. 3 shows path length as a function of delivery number for the first and second navigated towns. There is a clear decrease in path length with increased number of deliveries in a town. For both towns, learning is rapid. We confirmed this with a two-factor analysis of variance (ANOVA) on First vs. Second town [2] $\times$ delivery number [25]. As expected, the main effect of delivery number was significant, $F(24,696) = 5.19$, $\mathrm{MSE} = 6038$, $p < 0.001$. Confirming the apparent equality of the learning curves between the two towns, the main effect of town, first versus second, was not significant, $F(1,29) = 0.001$, n.s. The interaction term also failed to approach significance, $F(24,696) = 0.976$, n.s. The magnitude of learning was considerable, with delivery path length decreasing nearly to the optimal value,[2] which was approximately half the initial value.

2.2.4. Discussion

With experience in picking up and dropping off passengers at random locations, subjects' ability to find near optimal paths improved dramatically. This improvement suggests that subjects successfully formed a survey representation of each town. The lack of transfer between Towns 1 and 2 suggests that subjects learned town-specific

[2] We calculated the approximate most efficient path length by computing the mean number of blocks from passenger pickup to storefront along the shortest route to each store in all three towns.

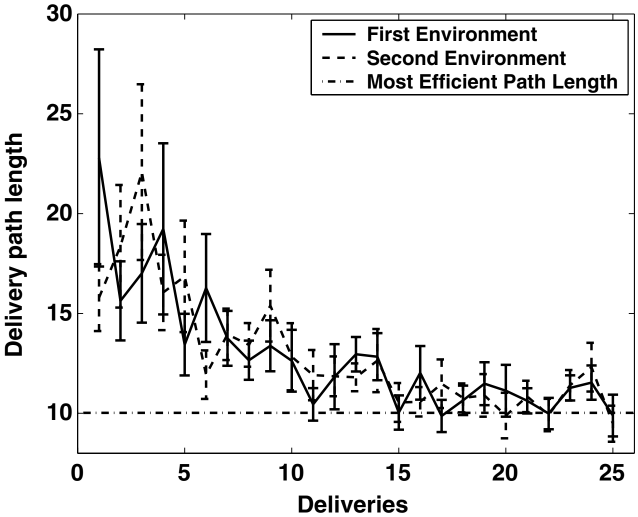

Fig. 3. Learning curves for the first and second towns. The length of the delivery paths decreases with each additional visit to the stores. Error bars represent ±1 SEM. The horizontal dot-dashed line denotes the average most efficient path length (see text).

information rather than general features of the task. Also, by allowing subjects to experience a practice town prior to actual testing, learning-to-learn effects (Keppel, Postman, & Zavortnik, 1968) may have been minimized.

3. Experiment 2

Experiment 1 demonstrated that our general method can be used to quantify navigational learning. With each successive delivery, subjects made use of increasingly more efficient paths from random pickup points to target stores. However, because Towns 1 and 2 differed along several dimensions, it was not possible to determine what sources of information would be important for transfer. For Experiment 2, we modified the design of our virtual towns so that we could systematically vary the overlap of landmark and layout information across towns.

Each town in Experiment 2 comprised a 5×5 block grid, where each block contained a single building (see Figs. 4 and 5). Certain buildings were potential destinations for passengers; we call such targets *stores*. The remaining buildings were never targets, but provided visual context information (from here on, we use "building" to refer to non-targets and "stores" to refer to targets). We considered the stores and buildings the landmarks in these towns, and their locations relative to each other and the global shape of the environment as the layout.

To measure the separate effects of landmarks and layout information on spatial navigation, we created three experimental and two control conditions. All five conditions used the same layout and landmarks in the second town. Each of the

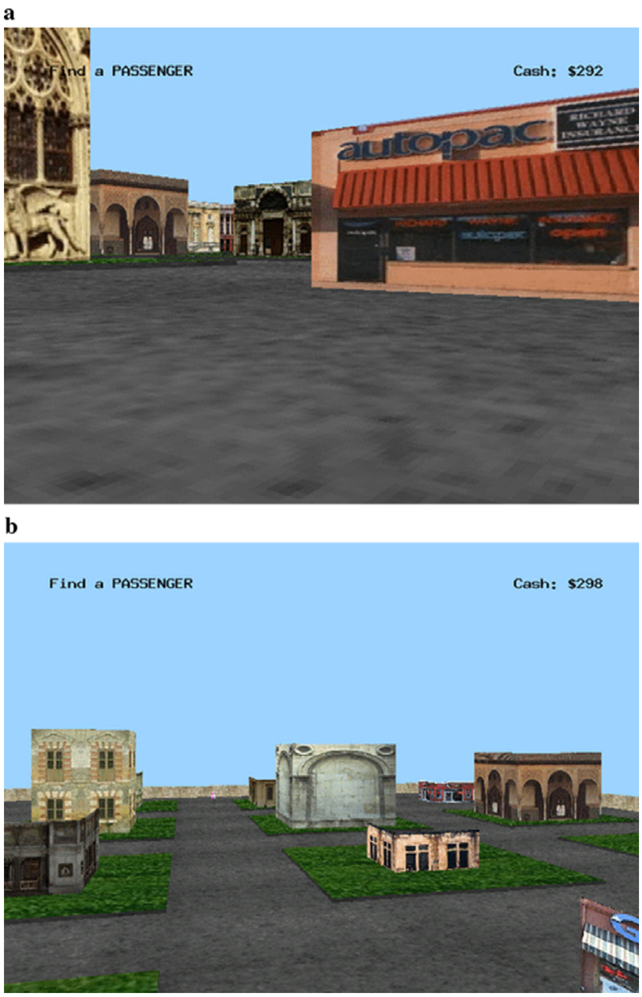

Fig. 4. Images of the virtual towns used in Experiment 2. (a) A snapshot from the subject's perspective during virtual navigation in Experiment 2. Subjects only saw first-person views. (b) An aerial view (not used in the experiment) to illustrate the grid structure of the roads.

experimental conditions replaced some set of the landmarks in the first town with novel landmarks while the layout of the target locations remain unchanged. In one control condition, we made two unique towns by replacing all the landmarks (stores and buildings) and changing the relative locations of stores and buildings within the towns. In the other control condition the two towns were identical. These manipulations allowed us to test whether subjects could navigate based on previously learned landmark information or layout information. The manipulations also made it possible to examine subjects' use of these two types of information when they conflicted.

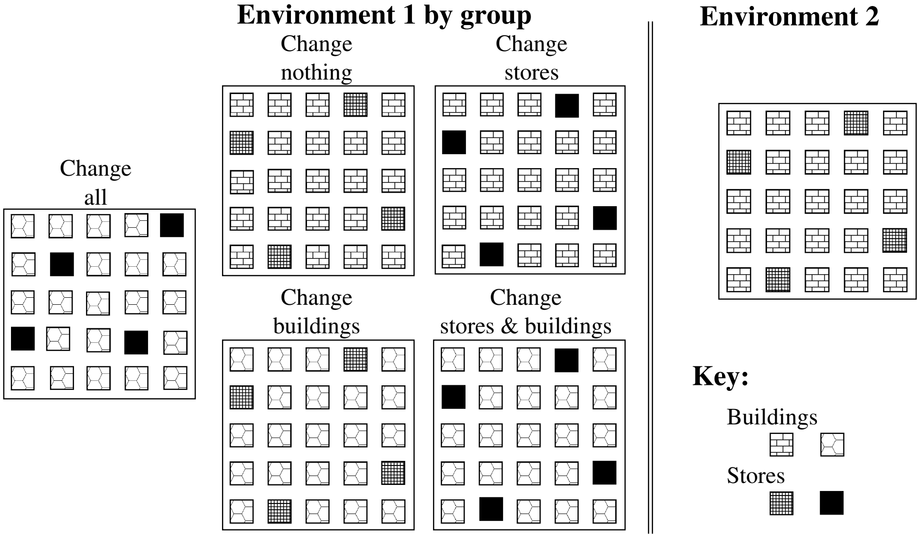

Fig. 5. The design of the towns used in Experiment 2. Each set of five subjects had the same Town 2, but Town 2 was randomly generated for each such set of five subjects. Within a set of five subjects, how Towns 1 and 2 differed from each other varied as a function of transfer group. Depicted are all the Town 1 blueprints for a given, example Town 2 blueprint. The towns contain 4 goal objects (stores) and 21 contextual objects (buildings).

3.1. Methods

3.1.1. Subjects

One hundred and thirty Brandeis undergraduates (66 males and 64 females) participated for monetary compensation plus a performance-based bonus. Subjects were randomly assigned to each of the five transfer conditions.

3.1.2. Construction of virtual towns

The virtual towns were laid out on a 5×5 regular grid of roads. The regularity of the layout facilitated automatic generation of towns, randomizing the sets of stores and buildings used, and their locations. This randomization reduced the kind of town-specific variability associated with the small number of unique towns used in Experiment 1 (compare Figs. 2 and 5).

Defining the width of a road as one *unit* in our virtual world, the size of the entire town was 16×16 units. Of the 256 total square units, 100 square units consisted of 25 equally sized blocks, each containing one structure (building or store). Blocks were separated from each other and the outer wall by 1-unit-wide roads. Each town contained 21 buildings and 4 stores (Fig. 5).

Each building occupied approximately one square unit centered within a block. Buildings could vary slightly in the area of their base, and could vary substantially in their heights. As shown in Fig. 4, each building had a unique façade mapped onto all four walls, and a lawn separating the building from the road. These non-target structures provided a rich visual context for our virtual towns.

The stores were of uniform shape and size, with a single storefront image mapped onto all four sides. Each store occupied a $0.7 \times 0.7 \times 0.35$ unit rectangular cube centered within a block. Unlike buildings, stores were not surrounded by a lawn, but by paved roads on which subjects were able to drive.

The outer boundary of the town had the image of a stone wall mapped along its length. No other heterogeneous visual information could be seen beyond this boundary.

3.1.3. Navigation

As in Experiment 1, the virtual taxi allowed subjects to navigate from a first person point of view using the four arrow keys. Now, however, the field of view was made somewhat narrower ($56° \times 44°$ in 640×480 pixel mode) and subjects were no longer permitted to make arcing turns – if more than one key were pressed, only the most recent key press would apply, and a key had to be released before it again would have an effect. This constraint separated turning and translating behaviors in time. Turning rate was 20 °/s such that a full rotation took 18 s. Driving speed was constant at 1.17 *units/s* and the view was refreshed every 30 ms. These rates of movement, which were slower than those used in Experiment 1, were chosen to reduce the choppiness experienced during navigation of the larger towns and to ensure that subjects could comfortably track the turn.

During the search phase, a single would-be passenger was placed in the town. On each delivery, the location of the passenger was chosen randomly subject to the constraint that the location was not within line of sight of either the current or previous target store. As in Experiment 1, this randomization tested subjects' survey representation by requiring them to find novel short routes for each delivery. When the passenger was picked up, a text screen instructed the subject to take the passenger to a specific target store. Subjects pressed ENTER to return to the virtual town. As in Experiment 1, pick up or delivery of passengers occurred when the taxi came within 0.20 units of the passenger or the storefront.

As soon as the passenger was delivered, a text screen told the subject that they were successful and to look for another passenger. Subjects were rewarded $50 virtual cash for each delivery and were docked $1 for every 10 s spent moving, turning or standing still, but there was a restriction that a maximum of $1 could be docked for any continuous period of standing still. Their earnings were continuously displayed in the upper right corner of the screen. In the upper left, a short description of their current goal was shown (e.g., "*Find a passenger*" or "*Find the Java Zone*"). We added this feature because several subjects in Experiment 1 reported that they occasionally forgot the identity of their destination target.

3.1.4. Transfer design

All subjects were tested in two different towns. The similarity between the two towns was varied across five groups of subjects. The first towns could differ from each other and the second town along three dimensions: the identity of the goal structures (stores), the identity of the contextual structures (buildings) and the set of locations of the goals. Group *change-all*, which served as a control condition for minimal transfer, changed all three dimensions, that is, the appearance of stores, buildings and store locations. Group *change-none*, which served as a control condition for maximal transfer, varied none of the three dimensions, keeping all information constant across the two towns. The remaining three groups, which represent our key experimental manipulations, either varied only the stores, Group *change-stores;* only the buildings, Group *change-buildings*; or both stores and buildings, Group *change-stores and buildings*. Table 1 illustrates the full set of manipulations across the 5 groups.

To promote the generalizability of our results, we created a unique pair of towns (having their own stores, buildings and layouts) for each subject in a given transfer group. One hundred pairs of towns (one for each of the first hundred of the subjects) were yoked across groups such that one subject from each group navigated an identical second town (for a total of 20 distinct second towns), and a first town that differed only along the dimensions being manipulated for that group. Fig. 5 illustrates an example of the towns given to one set of subjects across the transfer groups. When we began to see interesting effects in Groups *change-stores* and *change-stores and buildings*, we added 15 pairs of subjects in those two groups, yoked to each other, in an effort to firm up those effects.

The rationale for the transfer groups is as follows. The layout hypothesis suggests that the landmarks in a town are not crucial for orienting within that town. To test this hypothesis, we created Group *change-stores and buildings*, in which we altered all landmark information (i.e., the appearance of all stores and buildings; this included altering the names of stores and the shapes of the buildings) while preserving the configuration of target vs. contextual landmarks (stores versus buildings). If subjects can rely on layout alone to orient, then this condition should show relatively high levels of transfer. If, on the other hand, people do rely on landmarks this group ought to show little to no transfer.

We also wanted to test how subjects would respond if target or contextual landmarks were altered while layout information was preserved between the two towns. For that reason, we included two additional transfer groups, Group *change-stores*

Table 1
The information (see text) that differed between Towns 1 and 2 for each of the groups in Experiment 2

Experimental group	Goal objects	Contextual objects	Set of goal locations
change-none	Same	Same	Same
change-stores	Diff.	Same	Same
change-buildings	Same	Diff.	Same
change-stores and buildings	Diff.	Diff.	Same
change-all	Diff.	Diff.	Diff.

In Group *change-none*, the first and second towns were identical. In Group *change-all*, all objects and the set of goal object locations were changed (only the grid structure was preserved). The three remaining groups had the same target layout in both towns, however, either the buildings and/or stores were different between towns. Diff. – Different.

and Group *change-buildings*, for which we altered the appearance of either the stores or the buildings, respectively. Changing a subset of the landmarks present in the town would disrupt landmark-to-landmark associations. If subjects favor layout over landmark information, Groups *change-stores* and *change-buildings* should show as much transfer as Group *change-stores and buildings*. However, if subjects favor landmark over layout information, these groups may show less transfer than Group *change-stores and buildings*.

3.1.5. Procedure

Before encountering the test towns, subjects completed two different practice tasks. In the first task, subjects delivered four passengers, one to each store, in a practice town. The practice town was a 3×3 grid with four stores, one in each corner block. These stores were not used in the main task. The other blocks were covered with grass, which restricted movement to the paved areas without obstructing subjects' views of the town. Navigating this small practice environment familiarized subjects with the controls of the taxi and with the method for picking up and delivering passengers.

In the second practice task, a subject viewed the images of all eight stores that would be encountered later in both towns. Below each image, the store's name was displayed; these names were later used by passengers to communicate where they wanted to go. Subjects looked at each picture and read its name aloud. The list was presented four times, each time in a new random order. This practice task was designed to familiarize subjects with the appearance of the stores before entering the towns.

The experimenter remained in the testing room during both practice tasks, and answered any questions unrelated to strategy. Once the practice tasks were completed the experimenter left the room and subjects began with the first test town. To complete this town subjects picked up and delivered 20 passengers, five to each of the four stores. Between the two test towns subjects did the second practice task (viewing and naming store fronts) again with the experimenter in the room. The second test town also had 20 passengers.

3.2. Results

Our performance measure was *excess path length,* defined as the difference between the length of the subject's delivery path and the city block distance ($\Delta X + \Delta Y$) between the pickup and delivery points. This measure[3] removes passenger-specific variability from our delivery distance data. Fig. 6a shows the learning curve for subjects in the first town. A two-way ANOVA on delivery number [20] $\times$ transfer group [5] confirmed that the effect of delivery number on excess path length was significant

[3] The qualitative pattern of results did not change when we used Euclidean distance. Note that the city block distance measure of optimal path length allows for negative excess path length as it does not account for subjects' ability to take shortened, curved paths through intersections. The Euclidean measure, on the other hand, overestimates subjects' ability to do so.

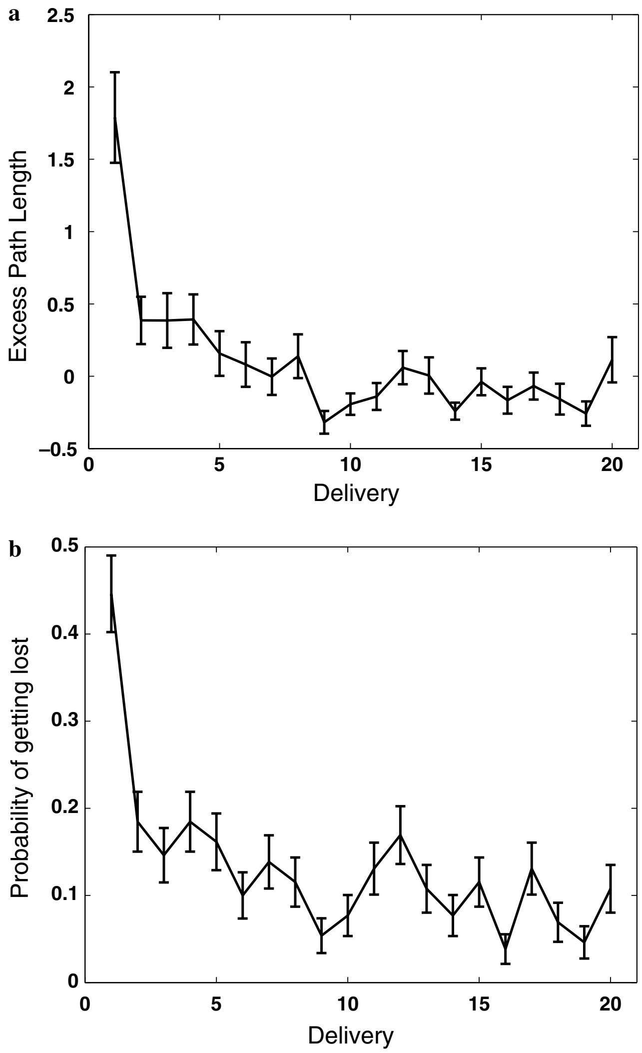

Fig. 6. Learning curves for all subjects in the first town of Experiment 2. (a) Learning as measured in excess path length (actual path length minus minimum path length; see text for note on negative values). (b) Learning as measured by the probability of getting lost. Error bars represent ±1 SEM.

($F(19, 2500) = 15.59$, $\text{MSE} = 7646$, $p < 0.001$). Because none of our manipulations should have affected performance in the first town we neither expected nor observed any significant main effect of group on the first town, $F(4, 2700) = 1.70$, n.s. nor an interaction, $F(76, 2700) = 1.19$, n.s.

Subjects in the second town also show a significant main effect of delivery number on excess path length ($F(19, 2500) = 2.65$, $\text{MSE} = 794.9$, $p < 0.001$). Because the similarity between Towns 1 and 2 varied across the five groups, we expected differences in

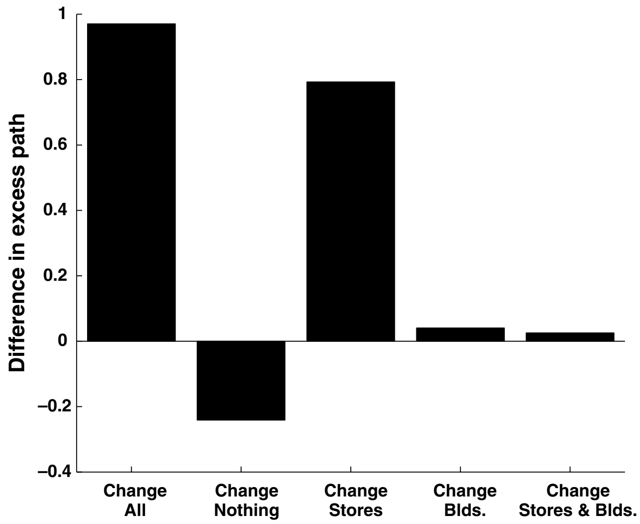

Fig. 7. Average increase in excess pathlength from the end of the first town to the beginning of the second town in Experiment 2. Positive increase in excess pathlength for the *change-all* and *change-stores* groups indicates that subjects did not transfer knowledge from the first town. The zero or near zero difference in excess pathlength for the *change-nothing, change-buildings* and *change-stores and buildings* conditions indicates maximum transfer between towns.

performance as a function of transfer condition. This was confirmed by a significant main effect of transfer group ($F(4,2500) = 4.83$, $\mathrm{MSE} = 1446$, $p < 0.001$). The interaction between delivery and transfer group was also significant ($F(76,2500) = 1.28$, $\mathrm{MSE} = 385$, $p < 0.05$).

Because subjects mastered the first town within the first few deliveries (Fig. 6), one would expect to see the strongest evidence for transfer on the first delivery in the second town. We therefore computed an index of transfer, which was defined as the difference between excess path length on delivery 1 of the second town and the asymptotic excess path length[4] of the first town across the five transfer groups. Fig. 7 shows the mean values of this index for each condition. The large bars in conditions *change-all* and *change-stores,* indicate poor transfer. Both of these conditions show an average index of roughly one indicating that they navigated one block further than the optimal pathlength. Because the environment was five blocks wide this is equivalent to traveling across a fifth of the environment. The near-zero values seen by both the *change-nothing* and *change-buildings* conditions indicates that these subjects were able to use a path that closely matched the optimal pathlength. In post-hoc comparisons, we found that the *change-all* condition was not significantly different from the *change-stores* condition and both were

[4] We defined the asymptotic excess path length to be the average of the last 12 deliveries (3 to each of the 4 stores) for each subject.

significantly greater than the *change-nothing* ($p < 0.05$) and *change-buildings and stores* ($p < 0.05$) conditions, and showed trends toward being significantly greater than the *change-buildings* condition ($p < 0.1$). All other comparisons were non-significant.

Fig. 8 shows the distributions of the index of transfer for each of the five conditions. One can clearly see that for condition *change-none,* all subjects showed an index value less than one block. For the other conditions, some subjects had larger values.

Upon inspecting these distributions, it appears that subjects either navigated quite directly to the first target store in the second environment, or could not easily orient toward the target store, and took an inefficient path (see tails of the distributions). We refer to the latter condition as the subject "getting lost" within the town. To separate these types of paths, we used condition *change-none* as the standard. We selected a threshold, θ_{lost} to be the 95th percentile of the distribution from *change-none*, which was 0.70 blocks and it plotted on the histograms in Fig. 8 as a vertical grey dashed line. For each condition, we could then compute an estimate of the number of subjects who got lost along their path, N_{lost}, defined as the number of subjects with transfer index exceeding θ_{lost}. First, note that when we replot the learning curve from the first town in this measure, the overall shape of the learning curve remain unchanged (Fig. 6b). The values of N_{lost} were: *change-none*: 1/20; *change-stores*: 12/35; *change-buildings*: 3/20; *change-stores and buildings*: 4/35; *change-all*: 5/20. We tested whether these observed rates differed from one another. All lost rates were greater than *change-none (change-stores:* $\chi^2(1) = 60.5$, $p < 0.01$; *change-buildings*: $\chi^2(1) = 4.0$, $p < 0.05$; *change-all*: $\chi^2 = 16.0$, $p < 0.01$) with the exception of *change-stores and buildings*, which nonetheless showed a trend toward significance ($\chi^2(1) = 2.9$, $p = 0.09$). None of the conditions showed significantly lower lost rates than *change-all (change-stores*: $\chi^2 = 0.3$; *change-buildings*: $\chi^2 = 1.3$; *change-stores and buildings*: $\chi^2 = 0.11$, $p > 0.1$) except for condition *change-none* ($\chi^2 = 16.0$, $p < 0.01$). Condition *change-stores and buildings* showed a trend toward lower lost rate than *change-all* ($\chi^2(1) = 3.2$, $p = 0.07$). Finally, condition *change-stores* had greater lost rate than conditions *change-stores and buildings* ($\chi^2(1) = 16.0$, $p < 0.01$) and *change-buildings* ($\chi^2(1) = 8.7$, $p < 0.01$).

It is possible that although Groups *change-all* and *change-stores* show less transfer than the other transfer groups, they still show *some* transfer (including learning-to-learn effects). We thus compared the path length for the initial delivery made in the first and in the second town. They were not significantly different for Group *change-all* ($t(34) = -1.18$, $p > 0.1$, two-tailed, paired samples) and Group *change-stores* ($t(19) = -145$, $p > 0.1$). Thus, not only do Groups *change-all* and *change-stores* show less transfer than the other groups, they show no significant transfer between towns, reminiscent of the subjects' performance in Experiment 1.

Remarkably, subjects in Group *change-stores and buildings* showed maximal transfer even when all store fronts were novel. This suggests that subjects had seen the store locations prior to the first passenger pickup. To rule out the possibility that the performance of group *change-stores and buildings* was due to more target exposure than other groups we used a one-factor ANOVA to compare store exposure

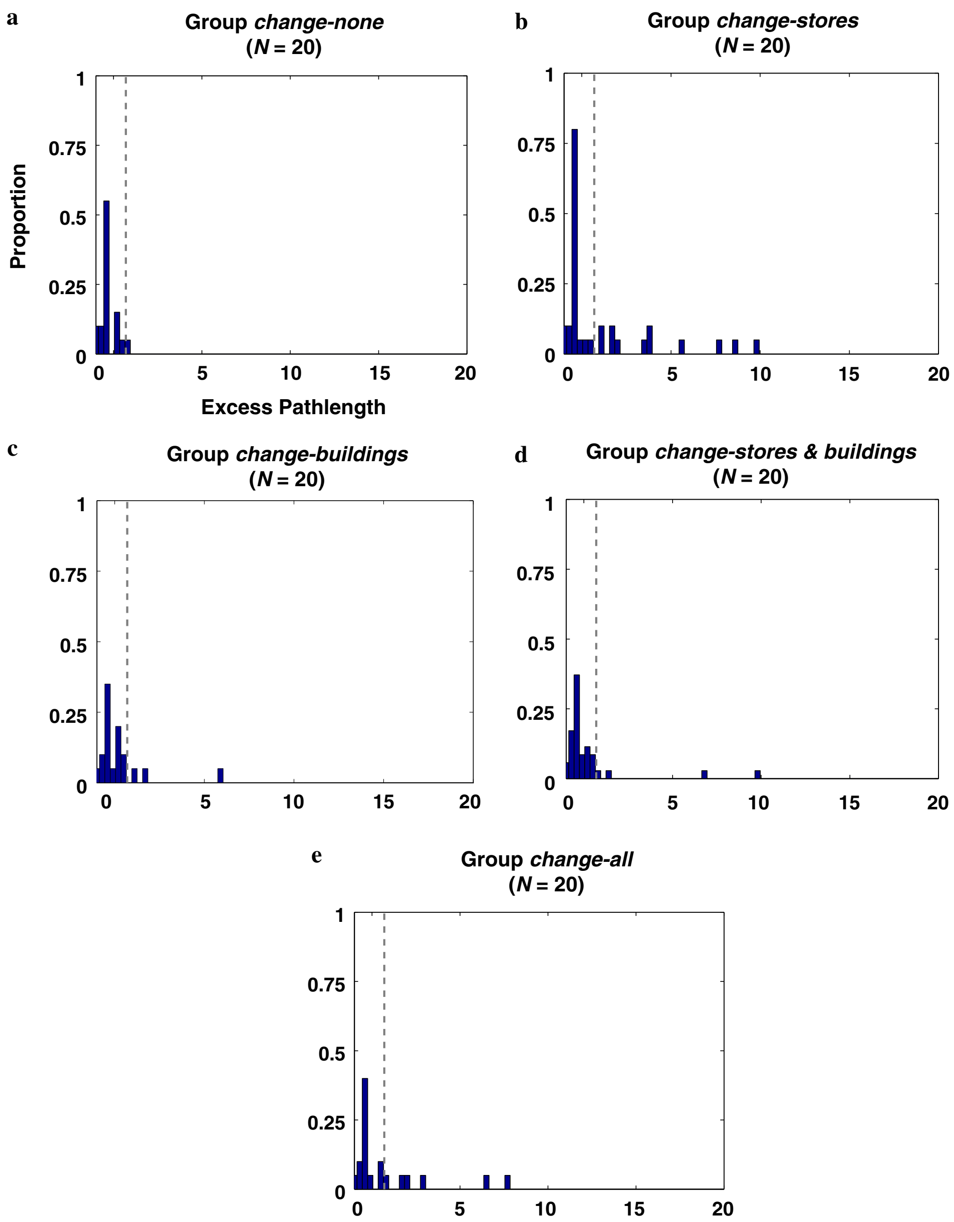

Fig. 8. Distributions of the index of transfer, excess pathlength on the first delivery path of the second town, for each transfer condition. The grey dashed line marks the threshold used to determine whether subjects navigated directly to the target or indirectly.

across the five groups. The result indicate that the store exposure did not differ across groups ($F(4,125) = 0.23$, n.s.), showing that differences in store exposure cannot be responsible for the sizable groups differences we found.

Finally, it should be pointed out that during even the first search path subjects experience considerable exposure to both stores and buildings. In quantitative terms, subjects had 2.9 times as much store viewing during their first search than during their first goal-seeking path. A store was considered to have been viewed when a store occupied at least one 300th of the horizon. To compute the amount of store viewing, we summed the area of the horizon filled by stores over the entire search or goal-seeking path. Given that quite a large proportion of the learning in Town 1 occurred between the first and second goal-seeking path, it is likely that subjects were learning much of their spatial representation during the first search path. The same is likely to hold for Town 2. Thus, subjects in all groups would have had a good view of Town 2 during their first search path. This amount of exposure, combined with the constant layout pattern in all groups other than Group *change-all,* could have easily enabled subjects in these groups to orient effectively enough to seek a short path to the first target store.

3.3. Discussion

As in Experiment 1, while subjects navigated their first town, they developed a survey representation of the town, reducing excess path length from random pick-up points to targets in that town (Fig. 6). Of special interest was performance on the first delivery in the second town. As shown in Fig. 8, subjects exhibited near-perfect transfer in all but Groups *change-all* and *change-stores*. The relatively high level of transfer in Group *change-stores and buildings* (comparable to the amount level of transfer in Group *change-none*, the condition with identical Towns 1 and 2) suggests that people are capable of orienting based on pure layout information.

4. General discussion

We asked whether subjects could navigate on the basis of landmark and layout cues, and how these two types of cues might have interacted to drive behavior. We measured transfer of training between virtual towns that differed along specific dimensions. Subjects learned to navigate in these towns as they played a virtual taxi-driver game, repeatedly picking up passengers from random locations and attempting to deliver them along the shortest possible route to designated target stores.

Subjects' ability to find novel shortest routes within our virtual towns suggests that they were not simply memorizing learned paths, but rather forming some higher-order survey representation of the environment, as first argued by Tolman (1948). Experiment 2 used a transfer methodology to probe the nature of this survey representation. Specifically, we varied the set of goal objects (stores; e.g., coffee shop vs. clothing store), the set of contextual objects (non-goal buildings) and the set of locations of goal vs. contextual objects. As expected, subjects showed near-perfect transfer of training when two subsequently learned towns (Town 1 and Town 2) were identical, and showed minimal transfer when they varied in every way (Experiment 1) or in all ways except the general layout of the town (Experiment 2, Group *change-all*). However, when Town 2 consisted of novel stores and surrounding buildings, but used the same sets of store and building locations as Town 1 (Group *change-stores and buildings)*, transfer of training was just as high as in the condition where the two towns were identical. Even without available landmark cues, the preserved spatial layout and set of goal locations enabled subjects to learn the new targets within a single pick-up and delivery.

However, what happens when layout information is preserved, but landmark information is perturbed? In Group *change-stores*, target stores were altered, while the buildings were unchanged. The low level of transfer obtained in this condition suggests that the disrupted landmark information impeded subjects' ability to orient to learned layout information alone. This disorientation might have resulted from familiar buildings (which subjects may have used as contextual landmarks to orient to) cueing target stores from the previously learned environment. In Group *change-buildings*, in which we found high transfer of learning, the buildings were altered, while the layout and target stores were preserved. The lack of interference from previously learned landmark information is not surprising given that only the small number of target stores would have been familiar, and those were the targets themselves. Thus, this condition may have in effect approximated condition *change-stores and buildings.*

An alternative explanation for the lack of transfer in the *change-stores* condition might be that the high degree of visual similarity between the first and second towns kept subjects from noticing that the town had changed. Thus these subjects might have paid relatively less attention to the town during their initial search for a passenger. This hypothesis suggests that experience within the environment, or more specifically with the changed stores, would have a large impact upon the degree of transfer. In several follow-up analyses, we found convergent evidence that make this interpretation implausible, and at the very least, could not account for the substantial differences among transfer groups. First, the distance traveled to find the first passenger did not differ between the *change-stores* condition and the other conditions (8.33 [$SD = 5.51$] for change-stores vs. 8.10 [$SD = 5.07$] for the other conditions), and that in each condition this search would have afforded the opportunity to see all stores. Thus, subjects in the *change-stores* condition had an equivalent amount of experience with both the environment and the stores.

To further test this alternate explanation we examined the relationship between store experience and transfer. If subjects increase their attention to encode the new information only after they have encountered a changed store then the attention hypothesis predicts that shorter times to encounter a changed store would lead to shorter excess path length on the first delivery. However, we failed to detect any significant effect of time to encounter a changed store and excess path length ($r(35) = -0.113, p > 0.5$). We also asked whether the distance traversed during the initial passenger search (regardless of when or whether a store was encountered) influenced transfer in the change-stores condition, as would be expected if transfer largely reflects variation in learning during this first search. Here, too, we failed to observe a significant correlation ($r(35) = -0.16$, $p > 0.1$). Therefore, if the alternate account is relevant, it is subtle, and is unlikely to account for the large levels of negative transfer we observed.

4.1. Positional versus associative representations of space

Spatial cognition researchers have proposed different structures for the spatial representation learned through navigation. In positional coding models, it is assumed that subjects learning the positions of landmarks within the environment (e.g., Devlin, 1976; Gouteux & Spelke, 2001; Hart & Moore, 1973; Hermer & Spelke, 1996; Trowbridge, 1913; Wang & Spelke, 2000). In associative models, in contrast, it is assumed that subjects learn direct relationships among landmarks, which could account for several distortions that have been observed empirically (e.g., Kaplan, 1976; Mallot, Franz, Schölkopf, & Bülthoff, 1997; Schölkopf & Mallot, 1995).

Although subjects' ability to orient purely on the basis of layout information (Group *change-stores and buildings)* might be taken to suggest that pure positional information is the principal basis for navigational spatial memory, our findings in the other transfer conditions suggest a richer underlying structure. When Town 2 changed the set of goal objects but maintained the same set of contextual objects (Group *change-stores*), subjects exhibited little or no transfer. This indicates that the buildings in the second town interfered with subjects' ability to cue their survey representation. This may have been due to subjects retrieving context-target object associations formed during navigation of the first town. The retrieved associations would have conflicted with the context target object associations in the second town, causing subjects to become disoriented.

When the identity and location of target objects was preserved but the visual context (surrounding buildings) was changed (Group *change-buildings*), any interfering associations would have been irrelevant to goal-seeking, which was the subject's task. This pattern of results supports the notion that, in addition to positional information, subjects learn associations among landmarks, and that this information can disrupt subjects' ability to orient based on layout information alone.

4.2. Relevance to specific navigational memory models

Benhamou et al. (1995) and Schölkopf and Mallot (1995) suggested that human survey representations are based on directional associations among neighboring local views. These models, as well as any theoretical accounts that rely on direct landmark-to-landmark associations (e.g., Kaplan, 1976; Mou & McNamara, 2002; Mou et al., 2004; Shelton & McNamara, 2001), could account for the high level of transfer in the identical transfer town, as well as the lack of transfer in Experiment 1 and in the *change-all* condition in Experiment 2. Additionally, it would account for the low level of transfer in the condition that varied the identities of the target stores due to interference from previously learned building–store associations. However, these models would incorrectly predict little transfer when all surrounding buildings were varied (in the *change-buildings* condition), because that manipulation would disrupt all nearest-neighbor associations. Finally, such models would have great difficulty accounting for the high level of transfer between towns that maintained their spatial layouts and set of goal locations while changing the identities of the target and contextual objects – a manipulation that altered all of the local views in the environment.

Our finding that subjects can transfer spatial layout information when the identities of the objects are drastically altered is consistent with Hart and Moore (1973) and the work of Spelke's group (Gouteux & Spelke, 2001; Hermer & Spelke, 1996; Wang & Spelke, 2000). These groups found that people often depend upon spatial location information rather than upon other properties of environments. Such a view is consistent with our finding that subjects rapidly learned a transfer town with novel stores and buildings but an identical spatial layout. However, it fails to explain the slow learning of a transfer town with novel stores but identical buildings and spatial layout. This latter result seems to call for an associative account in which target stores and their surrounding context are somehow linked, and mapped onto an abstract representation of the layout of the town as a whole. Alternatively, the survey representation may rely on goal objects as anchors, along the lines of Couclelis et al. (1987).

It appears that the only way to reconcile these incomplete sets of accounts of our data is to create composite models, in which both landmark and layout information are learned in parallel (Montello, 1998; Peponis et al., 1990) and used as orienting cues, but, under some conditions, a distracting set of cues can undermine use of reliable cues.

4.3. Some notes on methodology

The methodology used in this work allowed us to explore the interaction between landmark and layout information during spatial navigation. As noted in the introduction, virtual reality affords the ability to explore navigation in a way that is rarely possible in real environments. For example, in the *change-stores* condition we were able to simulate a large scale turnover of the commercial property in a neighborhood and explore the effects of this change on subjects' ability to orient in what should have been a familiar environment. Alternately, the *change-stores and buildings* condition could be thought of as exploring peoples' ability to orient on different floors of a building between which, the appearance of all the landmarks differed, but the layout would remain the same. Such manipulations have the advantage of giving unique insight into the navigational representations people use.

Our specific methodology, however, required a large number of subjects to obtain reliable estimates of transfer. Had we foreseen how easily and rapidly subjects would learn the regular towns used in Experiment 2, we could have overcome this limitation in our methodology. For example, we might have increased the complexity of the environment by enlarging our environment, adding irregular street layouts, and/or using fog to diminish visibility and reduce the availability of remote cues. In our design, subjects also received a great deal of experience with the environment during the initial search for the passenger. Starting the subject with a passenger in clear view might have enabled us to better evaluate transfer between towns. In designing such an experiment one walks the tight rope of experimental control versus realism and generalizability. Future work should move in both directions, pursuing both more realistic paradigms as well as ones that are more tightly controlled.

In summary, the results of our experiments further specify the abilities of subjects to orient to a learned spatial representation. First, subjects can orient to layout information alone; second, when landmark and layout information conflict, subjects orient preferentially to landmark information, which can interfere with effective orienting based on layout.

References

Benhamou, S., Bovet, P., & Poucet, B. (1995). A model for place navigation in mammals. *Journal of Theoretical Biology, 173*, 163–178.

Couclelis, H., Golledge, G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology, 7*, 99–122.

Devlin, A. S. (1976). The "small town" cognitive map: Adjusting to a new environment. In G. T. Moore & R. G. Golledge (Eds.), *Environmental knowing: Theories, research, and methods* (pp. 58–66). Stroudsberg, Pennsylvania: Dowden, Hutchinson & Ross.

Fujita, N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1993). The encoding-error model of pathway completion without vision. *Geographical Analysis, 25*(4), 295–314.

Gouteux, S., & Spelke, E. (2001). Children's use of geometry and landmarks to reorient in an open space. *Cognition, 81*, 119–148.

Haq, S., & Zimring, C. (2003). Just down the road a piece: the development of topological knowledge of building layouts. *Environment & Behavior, 35*(1), 132–160.

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition. In R. M. Downs & D. Stea (Eds.), *Image and environment*. Chicago: Aldine.

Hermer, L., & Spelke, E. (1996). Modularity and development: the case of spatial reorientation. *Cognition, 61*, 195–232.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*, 57–59.

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cognitive Psychology, 39*(1), 3–36.

Hillier, B., Burdett, R., Peponis, J., & Penn, A. (1987). Life: or, does architecture determine anything? *Architecture and Behavior/Architecture et Comportment, 3*, 233–250.

Kaplan, S. (1976). Adaptation, structure, and knowledge. In G. T. Moore & R. G. Golledge (Eds.), *Environmental knowing: Theories, research, and methods* (pp. 32–45). Stroudsberg, Pennsylvania: Dowden, Hutchinson & Ross.

Keppel, G., Postman, L., & Zavortnik, B. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior, 7*, 790–796.

Kim, Y. O., & Penn, A. (2004). Linking the spatial syntax of cognitive maps to the spatial syntax of the environment. *Environment & Behavior, 36*(4), 483–504.

Kirschen, M. P., Kahana, M. J., Sekuler, R., & Burack, B. (2000). Optic flow helps humans learn to navigate through synthetic environments. *Perception, 29*, 801–818.

Kuipers, B., Tecuci, D. G., & Stankiewicz, B. J. (2003). The skeleton in the cognitive map: a computational and empirical exploration. *Environment & Behavior, 35*(1), 81–106.

Loomis, J. M., Lippa, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(2), 335–345.

Lynch, K. (1960). The image of the city. In (pp. 123–139). Cambridge, Massachusetts: The Technology Press & Harvard University Press.

Mallot, H. A., Franz, M., Schölkopf, B., & Bülthoff, H. H. (1997). The view-graph approach to visual navigation and spatial memory. In Proceedings of the 7th International Conference on Artificial Neural Networks. Lausanne.

Mallot, H. A., & Gillner, S. (2000). Route navigation without place recognition: what is recognised in recognition-triggered responses? *Perception, 29*, 43–55.

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

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 162–170.

Mou, W., Zhang, K., & McNamara, T. P. (2004). Frames of reference in spatial memories acquired from language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 171–180.

Peponis, J., Zimring, C., & Choi, Y. K. (1990). Finding the building in wayfinding. *Environment & Behavior, 22*(5), 555–590.

Péruch, P., & Gaunet, F. (1998). Virtual environments as a promising tool for investigation human spatial cognition. *Current Psychology of Cognition, 17*, 445–463.

Péruch, P., May, M., & Wartenberg, F. (1997). Homing in virtual environments: effects of field of view and path layout. *Perception, 26*, 301–311.

Schölkopf, B., & Mallot, H. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior, 3*, 311–348.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43*, 274–310.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large scale environments. In H. W. Reese (Ed.), *Advances in child development and behavior* (pp. 9–55). New York: Academic Press.

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

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

Tong, F. H., Mariin, S. G., & Frost, B. J. (1995). Visual-motor integration and spatial representation on a visual virtual environment. *Investigative Opthalmology & Visual Science, 36*, S1679.

Trowbridge, C. C. (1913). On fundamental methods of orientation and "imaginary maps. *Science, 38*, 888–897.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J.-A. (1997). Biologically based artificial navigation systems: review and prospects. *Progress in Neurobiology, 51*(5), 483–544.

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition, 77*, 215–250.

Whitmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real-world places: transfer of route knowledge. *International Journal of Human–Computer Studies, 45*, 413–428.