Optic flow helps humans learn to navigate through synthetic environments

Matthew P Kirschen, Michael J Kahana¶, Robert Sekuler, Benjamin Burack
Center for Complex Systems, Mailstop 013, Brandeis University, Waltham, MA 02454-9110, USA;
e-mail: kahana@brandeis.edu

Abstract. Self-movement through an environment generates optic flow, a potential source of heading information. But it is not certain that optic flow is sufficient to support navigation, particularly navigation along complex, multi-legged paths. To address this question, we studied human participants who navigated synthetic environments with and without salient optic flow. Participants used a keyboard to control realistic simulation of self-movement through computer-rendered, synthetic environments. Because these environments comprised series of identically textured virtual corridors and intersections, participants had to build up some mental representation of the environment in order to perform. The impact of optic flow on learning was examined in two experiments. In experiment 1, participants learned to navigate multiple T-junction mazes with and without accompanying optic flow. Optic flow promoted faster learning, mainly by preventing disorientation and backtracking in the maze. In experiment 2, participants found their way around a virtual city-block environment, experiencing two different kinds of optic flow as they went. By varying the rate at which the display was updated, we created optic flow that was either fluid or choppy. Here, fluid optic flow (as compared with choppy optic flow) enabled participants to locate a remembered target position more accurately. When other cues are unavailable, optic flow can be a significant aid in wayfinding. Among other things, optic flow can facilitate *path integration*, which involves updating a mental representation of *place* by combining the trajectories of previously travelled paths.

1 Introduction

After a hiatus of decades, interest in spatial memory and wayfinding has returned to the scientific forefront. The work on animal wayfinding and its physiological substrates (eg O'Keefe and Nadel 1978; Samsonovich and McNaughton 1997) has in recent years led to an explosion of behavioral and functional neuroimaging studies of spatial learning and wayfinding in humans (Aguirre et al 1996; Berthoz 1997; Kahana et al 1999; Maguire et al 1996, 1997, 1998). Various complex navigation tasks have been used in these studies in an attempt to identify brain structures involved in spatial learning and wayfinding.

In a landmark paper some forty years ago, Gibson (1958) pointed out that an animal's locomotion through the environment could be guided by visual information that results from the animal's movements. This visual information, which is the ensemble of optical velocities that surround the moving animal, is called *optic flow*. The characteristics of optic flow are related to the speed and direction of locomotion, as well as to the visual properties of the environment such as texture gradients. Although optic flow normally accompanies locomotion, it is unclear whether optic flow actually benefits spatial learning and wayfinding. In particular, doubts have been raised about the influence of optic flow in the absence of associated vestibular cues (eg Berthoz 1997; Klatzky et al 1998). For example, Berthoz (1997) found that participants who experienced a single passive motion through the environment, in the absence of vision (and consequently optic flow), were later able to reproduce the extent and acceleration/deceleration profile of that movement. Furthermore, optic flow alone may not be sufficient to ensure proper spatial/directional coding (Klatzky et al 1998). Note, however, that these studies focused

¶ Author to whom correspondence and requests for reprints should be addressed.

on relatively simple responses, requiring just a single movement, such as one turn. This leaves open the possibility that optic flow alone could still facilitate wayfinding, but along more complex paths, perhaps by knitting together a series of individual, coordinated movements.

To directly assess the contribution of optic flow to the formation of spatial representations, we created a spatial-memory task in which participants learned their way through complex spatial environments to a target location and were then tested for their memory of that location. We hypothesized that the presence of salient optic flow would facilitate spatial learning, even with no accompanying vestibular activation.

1.1 Learning to navigate in a computer-generated synthetic environment

Our computer-generated virtual 3-D environments were created in the Open GL graphics language.[1] Open GL rendered textured walls, floors, and ceilings of the environment as realistic images on a large high-resolution monitor.[2] Participants navigated virtual environments using the ↑, ↓, ←, and → keys on a standard computer keyboard.

Similar virtual-reality simulations have been used effectively in tests of spatial learning. For example, Witmer et al (1996) demonstrated that virtual-reality simulations enabled participants to learn a complex route through an office building. In a subsequent test, in which a real office building was explored, people who had learned the virtual environment performed just as well as people who learned the same route either by physically exploring the building or by encoding the route purely symbolically.

To explore the effects of optic flow on spatial learning and navigation, we devised paradigms that imposed differing constraints on the paths that participants could take. In the first paradigm, participants learned a single path in each of several different computer-rendered, multiple T-junction mazes. Except for the visually rich synthetic environment, this task resembled a constrained sequence-learning paradigm (eg Kahana and Jacobs 2000). In this task, we examined the effects of optic flow and visual landmarks on spatial learning, using acquisition curves and inter-response times (IRTs) as dependent measures.

In the second paradigm, we removed the constraints of the multiple T-junction design, allowing participants to move freely through a large city-block environment. We also shifted the emphasis from mastering a single path (or one series of turns), to finding the most direct path to a studied target position. This enabled us to examine performance in terms of accuracy of participants' memory for the target position, and in terms of the route they took from the starting point to the target.

1.2 Multiple T-junction mazes

Studies of maze learning fueled much early work on learning theory and comparative psychology. Most often in these studies animals explored mazes constructed of multiple T-junctions to find a reward (Miles 1928; Stone and Nyswander 1927). Similar studies were conducted on human participants with finger, stylus, or full-body mazes. Some mazes were of elaborate design, fashioned after famed garden mazes like the Hampton Court maze (Perrin 1914). However elaborate the overall design of the maze, its methodological strength as a research instrument came from the simplicity and

---

[1] Open GL is a language for describing virtual environments. Within Open GL, one first defines a virtual environment in terms of geometric primitives. Once defined, the virtual environment can be viewed from any position or sequence of positions. For these experiments, we wrote a maze creation and navigation program. Our program supports a number of features including texture mapping, file operations, sound, and simple geometric shapes. An editor accompanying the program allows visualized construction of the mazes. The editor generates multiple versions of each maze so that a given maze can be seen with and without additional graphics (arrows or landmarks). Our software may be downloaded from http://fechner.ccs.brandeis.edu/maze
[2] We used an NEC multisync monitor with a display approximately 46.3 deg across and 34.2 deg vertically. The Open GL program operated at a resolution of 640 by 480 pixels.

uniformity of its choice points—the T-junctions. Because every junction was identical to every other one, the maze afforded little environmental support from external cues, forcing participants to fall back on some internal representation of the maze layout or on the remembered sequence of turns.

There are several different modes that participants might use in learning to navigate through a maze. For example, participants might encode their own actions into a sequence, forming a symbolic representation of some kind (ie memorizing a series of binary responses). Alternatively, participants might generate a spatial representation of the layout of the environment. The latter mode brings the maze learning task closer to studies of wayfinding in complex environments; the former mode connects maze learning with symbolic sequence learning.

According to current theories of serial learning (eg Burgess and Hitch 1999), successful retrieval requires that some representation of positional context be linked to each symbolic element in the sequence. This is necessary, for example, in order to learn single lists in which individual elements are repeated (eg Kahana and Jacobs 2000), or to learn multiple lists that share subsequences of items (Chance and Kahana 1997). Within a maze learning task, some mental representation of the geometry of the maze may provide participants with something akin to positional context. Alternatively, the presence of distinctive visual landmarks at junctions could make positional context even more explicit. Of course, these alternative modes of learning are not mutually exclusive.

By manipulating visual properties of the environment through which our participants attempted to find their way, we sought to identify conditions that promote formation and use of various alternative kinds of representations. To correctly traverse a multiple T-junction maze, it is sufficient to learn a series of left or right turns. What makes our task different from a symbolic learning task is the experience of movement through space. In a virtual environment, one could manipulate this experience by either including or excluding the optic flow that would normally be associated with movement through the environment. We expected that the presence of movement-dependent visual information (optic flow) would make it easier for participants to build up an accurate spatial representation of the environment, but that the absence of this information would make such a representation hard to construct.

2 Experiment 1: Navigating multiple T-junction mazes
For use in this experiment, we generated a library of all possible mazes with eight and twelve junctions. All mazes were created by linking a series of T-junctions in various combinations, creating a large number of mazes, each with a different spatial layout. The sequence of successive turns in any maze was constrained to foreclose the possibility of three consecutive turns in the same direction, either left or right. This constraint results from the geometry of identical T-junctions: after three successive turns in the same direction, the maze runs into itself. We added the further constraint that no sequence of two turns could be repeated more than three times. For example, if a randomly generated maze included the subsequence right – left – right – left – right – left the maze was rejected. This eliminated mazes that could be coded very simply, as a series of alternations of turns. Figure 1a shows the layout of one twelve-junction maze. The start and finish of the maze are indicated by a plus sign and a star, respectively.

The appearance of all junctions and corridors was identical. Figure 1b shows a detailed aerial diagram of a maze junction. The side walls were textured with a brick pattern and the ceiling and floor were textured with a series of parallel and perpendicular lines. The mean luminance of a corridor was approximately 19.7 cd m^{-2}. Transparent barriers prevented participants from moving forward into a dead end after a wrong turn. These barriers were meant to minimize participants' disorientation.
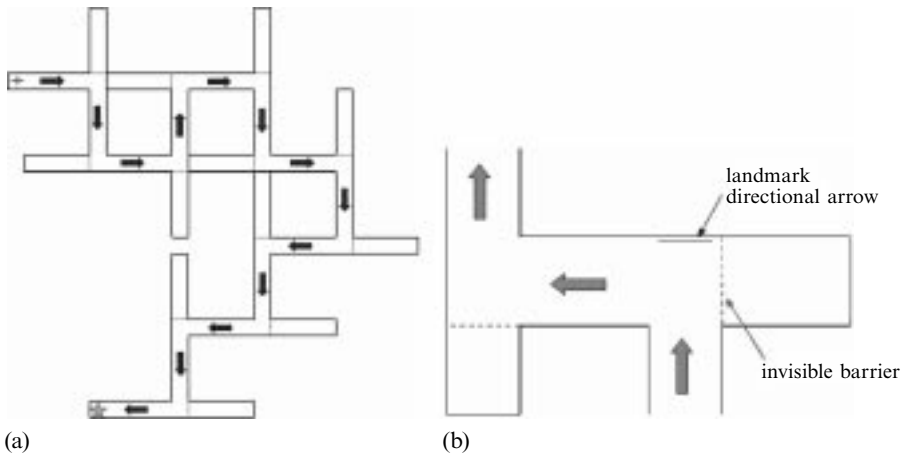
Figure 1. Example layout of a twelve-junction maze (a) and enlarged diagram of a sample T-junction (b). In (a), the cross indicates the start of the maze, and the star marks the end of the maze. The length of each corridor was four times its width. In both panels, the arrows show the correct path through the maze, and the dotted lines represent invisible barriers that prevent a participant from entering a dead-end corridor.

2.1 Maze navigation

Participants' possible movements through our T-junction mazes were highly constrained. To navigate a maze, participants had to first move forward into a T-junction intersection, then turn either right or left, and proceed forward again into the next junction. This sequence of forward and then either left or right was repeated until the target was reached. To facilitate the analysis of IRTs and to maintain a direct link between keypresses and movements in the maze, our program ensured that a single keypress—no matter of what duration—would produce just a single move.[3] Although these constraints sacrificed some of the participants' freedom to move ad lib through a maze, the constraints equated the data format across all trials on which participants navigated mazes without error. This, in turn, facilitated the analysis of IRTs as a function of output position.

2.2 Landmarks

To introduce landmarks into our virtual mazes, we placed highly visible, colored geometric shapes on junction walls. These landmarks were drawn from the sixteen combinations of four colors (red, green, blue, and yellow) and four shapes (circle, square, star, and plus sign). To promote good visibility, each landmark appeared in the center of a black square. Viewed from the start of a corridor, landmarks averaged 1.83 deg across, which allowed easy discrimination among landmarks on the basis of either color or shape. Viewed from the inside of a junction, landmarks averaged 20.4 deg across.

We tested the impact of landmarks on maze learning under three different conditions: (i) no-landmark, (ii) constant-landmark, and (iii) distinct-landmark. In the distinct-landmark condition, each junction had a different, randomly chosen landmark (subject to the constraint that consecutive landmarks could not have the same shape or the same color). This condition produced a consistent relationship, within any maze, between a landmark and the correct turn at that landmark. In principle, participants could learn and use this consistent relationship, although for each maze the relationship would have to be learned anew. In the constant-landmark condition, a single, randomly chosen landmark appeared in all of the junctions in a given maze; a new, randomly chosen landmark was used for different mazes. Such landmarks

[3] This restriction was enforced by clearing the keyboard buffer after each keypress.

provided no information about the layout of the maze, ie participants could not make correct choices in the maze by relying on landmarks alone.[(4)]

2.3 *Manipulation of optic flow*

We manipulated participants' experience of optic flow as they traversed the maze corridors. Our hypothesis was that, under some conditions at least, optic flow would influence participants' learning their way through a maze.

Movement, however fast or slow, through any textured environment, produces directional changes over time in the retinal image, which satisfies a minimal, literal definition of optic flow. Hypothetical conditions in which successive visual samples were separated by a large interval, say a year, would qualify as flow-producing, despite the fact that here flow demands temporal integration that lies outside the capability of the visual system. In this paper, though, we choose to hold the term *optic flow* to a higher standard, reserving it for visual conditions that produce perceived motion, not mere succession. By this definition, motion sampled at too low a rate is not considered to produce optic flow. We use the term no-motion to describe such conditions.

Consistent with this definition, the two conditions of optic flow in this experiment are described as *motion* and *no-motion*. For both conditions, a single press of the ↑ key was adequate to traverse the length of a corridor. Likewise, pushing the ← or → key turned the participant's viewpoint through 90° to the left or right, respectively. Pressing any key (including the ↓ key) while traversing a corridor or turning a corner had no effect on movement. The participant's eye height was 60% of the wall height, as measured from the floor.

In the motion condition, a single keypress caused the program to move the participant's view along the corridor incrementally, pausing briefly after each spatial increment. This produced a compelling experience of self-motion as participants traversed a corridor. Figure 2 shows eight views that a participant would see when traversing a corridor with motion, distinct landmarks, and directional arrows (see section 3). Panel (a) shows the view at the start of the corridor. After pressing the ↑ key, the participant's view along

(a)   (b)   (c)   (d)

Traversing down a corridor in motion condition

Approaching a junction   At a decision point   View after a single →
                                                                                       keypress
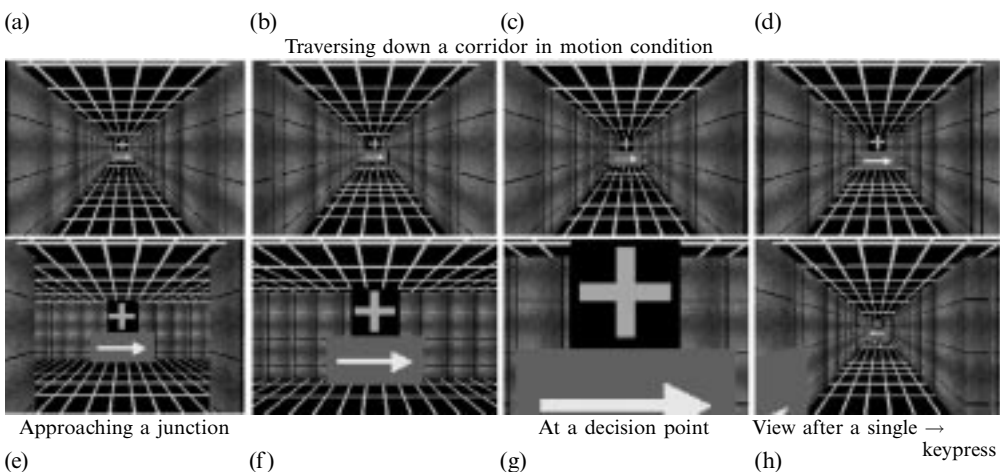(e)   (f)   (g)   (h)

Figure 2. Eight successive views that a participant sees while traversing a maze in the study mode. Panels (a)–(f) show incremental views as a participant traverses a corridor towards a T-junction. Panel (g) shows a participant's view from inside the junction and panel (h) is after a single → keypress.

[(4)] Because all maze corridors look identical in the no-landmark condition, a participant may turn the wrong way and think that she/he can proceed forward along the blind corridor. Pressing the forward key at this point will indicate that the participant is facing the wrong way. In the constant-landmark condition, a participant who turns the wrong way will not see a landmark at the end of the blind corridor. This indicates that the participant is facing a blind corridor and must turn back to face the correct direction.

the corridor was updated eight times, at equal intervals. This update rate of about 16 Hz was sufficiently high to produce an experience of smooth motion. The time taken to traverse the corridor was approximately 500 ms. Successive panels of figure 2 show six views that a participant would see. The final panel, figure 2h, shows the view after a 90° turn to the right. In both conditions, turning either left or right was instantaneous, ie there was no delay between pressing the ← or → key and the participant's next view in the maze. As a result, participants got immediate feedback from turning because their view would change from seeing a wall (eg figure 2g) to looking down a corridor (figure 2h). In the no-motion condition, a single keypress moved the participant's view directly to the next junction, after a delay that was approximately equal to that in the motion condition. In this way, we ensured that the time taken to traverse a corridor was comparable in both the motion and no-motion conditions.[5]

3 Method
Participants traversed each maze successively in two different modes, study and test. In either mode, on a single trial, participants navigated a maze from start to finish. In the study mode, arrows on the wall of each T-junction indicated the correct turn at that junction. In the test mode, these arrows were absent. In the constant-landmark and distinct-landmark conditions, landmarks appeared on junction walls throughout both study and test trials. Participants began by navigating a maze twice in the study mode. They then performed test trials, one after another, until a criterion of three successive perfect trials had been achieved. A perfect trial was defined as one in which the maze was navigated from start to finish with no unnecessary keystrokes; that is, the studied pathway was reproduced perfectly.

3.1 *Participants*
Forty-eight Brandeis students (mean age 19.7 years) participated for payment. Three different groups of sixteen participants each (eight males, eight females) were randomly assigned to no-landmark, constant-landmark, and distinct-landmark conditions. Optic flow (motion versus no-motion) was manipulated within participants as described below.

3.2 *Procedure*
Participants first completed a practice phase consisting of four eight-junction mazes. These mazes were navigated twice in the study mode and then to a criterion of two consecutive perfect trials in the test mode. After this practice phase, participants learned twelve different twelve-junction mazes (six in the motion condition and six in the no-motion condition) to a criterion of three consecutive perfect trials. These mazes were also navigated twice in the study mode. If neither criterion was met after twelve test trials, the computer automatically advanced the participant to the next maze. This criterion prevented differential loss of participants across the three between-participant conditions. All sixteen mazes (four practice + twelve experimental) were presented in succession during a single session lasting approximately 1 h.[6]

Within each landmark group, motion and no-motion conditions alternated every two mazes. Half of the participants in each group navigated the alternating pattern of mazes with the motion condition first while the other half traversed the same ordered set of mazes with the no-motion condition first. Maze layouts were counterbalanced across the different orders.

[5] In order to equate the timing in the motion and no-motion conditions, we adopted a very fast rate of motion. On the assumption that participants' views are from an average eye height of 1.75 m, the equivalent rate of movement is greater than 25 m s^{-1}.
[6] At the conclusion of all mazes, we asked participants to complete a survey, rating how effective they thought landmarks and the optic flow had been in facilitating maze learning. None of the participants' ratings correlated with performance measures.

4 Results

We examined several dependent measures, including total time in maze, number of wrong turns at maze junctions, and IRTs as a function of maze junction for each participant's final three perfect trials. All results for this experiment were consistent with findings from two pilot studies we conducted while improving our experimental methods. The pilot studies were designed to compare maze learning in the motion and no-motion conditions. In the first pilot study, participants navigated mazes with no landmarks or with distinct landmarks. In the second pilot study, participants navigated mazes with either constant landmarks or distinct landmarks. Unlike experiment 1, landmark conditions were manipulated within participants in the pilot studies. In the second pilot study we also examined the effect of training and found that the advantage in the motion condition was significant only early in training. For this reason, we restrict our analyses of the current experiment to the first six of the twelve presented mazes.

Figure 3 shows time in maze as a function of degree of learning (vincentized trial number).[7] The three panels show vincentized learning curves for (a) the no-landmark, (b) constant-landmark, and (c) distinct-landmark conditions. As can be seen, participants took longer to traverse mazes in the no-motion condition, and this effect was especially pronounced in the no-landmark condition. Consistent with this observation, an analysis of variance (ANOVA) on time in maze yielded a significant main effect of motion ($F_{1,45} = 16.46$, MSE $= 32.96$, $p < 0.01$), a significant motion $\times$ landmark interaction ($F_{2,45} = 3.84$, MSE $= 32.96$, $p = 0.029$), and a significant motion $\times$ landmark $\times$ degree of learning interaction ($F_{10,225} = 2.11$, MSE $= 9.23$, $p = 0.024$). A marginally significant main effect of landmark ($F_{2,45} = 3.15$, MSE $= 186.07$, $p < 0.10$) was carried by the no-landmark, no-motion condition.

On certain trials some participants became disoriented and spent an unusually long time wandering through the maze. A characteristic feature of these trials is that at some point in the maze, participants turned around and started going backwards towards the start of the maze. We used this feature, retracing one's path, to classify trials as ones in which a participant was lost in maze. Figure 4 shows learning curves
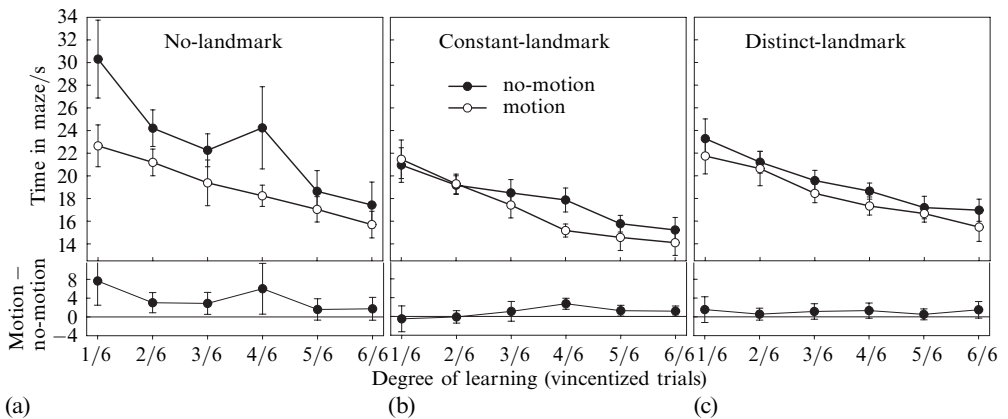

Figure 3. Total time in maze as a function of degree of learning for experiment 1 with the lost-in-maze values included. Filled circles mark the data from the no-motion conditions, and open circles mark the data from the motion conditions for (a) no-landmark, (b) constant-landmark, and (c) distinct-landmark conditions, respectively. Within each panel, learning curves are plotted in the upper graph, and the difference between motion and no-motion conditions is plotted in the lower graph. Error bars reflect 95% confidence intervals adjusted for between participant variability (Loftus and Masson 1994).

[7] Because participants took a variable number of trials to reach criterion, the number of trials was vincentized into six levels (Vincent 1912). This procedure was done separately for each maze a participant navigated.
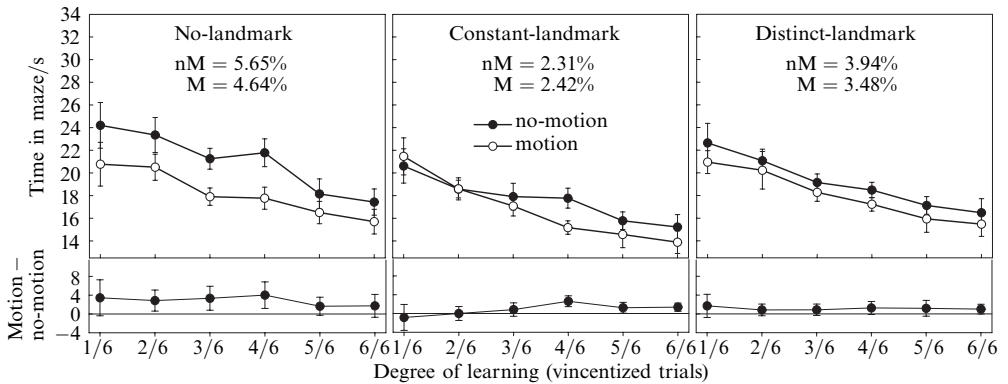
Figure 4. Total time in maze as a function of degree of learning in experiment 1 with data from lost-in-maze trials excluded from the calculations. In each panel, the percentages of lost-in-maze trials are indicated by separate numbers for the motion (M) and no-motion (nM) conditions. Participants are considered lost in maze on a given trial if they backtracked (or retraced their path) at any point while traversing the maze.

in each of the three landmark conditions with these lost-in-maze trials excluded. Although the general patterns of these results are quite similar to those shown in figure 3, the exclusion of lost-in-maze trials attenuated the motion effect in the no-landmark condition. This reduced the motion × landmark interaction to below the significance threshold ($F_{2, 45} = 2.21$, MSE = 24.96, $p = 0.12$), but preserved the significant main effect of motion ($F_{1, 45} = 14.94$, MSE = 24.96, $p < 0.01$).

We also examined the effect of motion and landmark conditions on participants' error rates, where an error is defined as a wrong turn at a maze junction. Note that by convention, even if a participant backtracks through the maze, the maximum number of errors cannot exceed the number of junctions, ie multiple errors at the same junction count as one error. Unlike time in maze, error rates were not statistically different across motion and no-motion conditions (lost-in-maze trials included, $F_{1, 45} = 1.87$, MSE = 4.12, ns; and lost-in-maze trials excluded, $F_{1, 45} = 1.93$, MSE = 2.63, ns). In addition, neither the main effect of landmark nor the motion × landmark interaction effect reached significance (either with lost-in-maze data included or excluded).

Figure 5 shows IRT serial-position curves for the final three perfect trials (note that the number of learning trials is not the same for the motion and no-motion conditions). An ANOVA on IRTs between maze junctions yielded a significant main effect of output position ($F_{11, 484} = 12.39$, MSE = 0.039, $p < 0.01$). Although there appears to be slight differences in IRTs across landmark and motion conditions, none of these effects approached significance—landmark ($F_{2, 44} = 1.18$, MSE = 0.44, ns) and motion ($F_{1, 44} = 2.06$, MSE = 0.043, ns)—nor did they interact with output position (all interactions, $F < 1$).
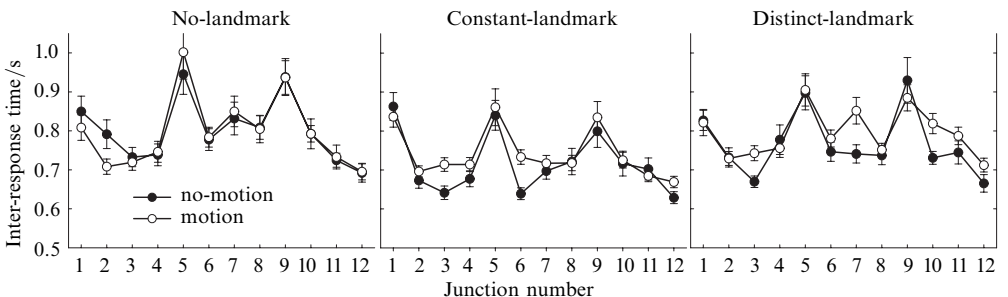

Figure 5. Inter-response times (between junctions) for the final three perfect trials in experiment 1.

The phenomenon of response bursting (eg Kahana and Jacobs 2000)—a single slow response preceding each succession of fast responses—is clearly apparent in all conditions. Because this response pattern persists despite averaging the data over mazes of different spatial structures, it can be concluded that this IRT pattern is not dictated by maze structure, but rather by human cognition. This temporal output pattern is characteristic of verbal sequence memory. The pattern of IRTs suggests that participants code several elements of the sequence as a single higher-order unit, or chunk (eg Johnson 1972; Kahana and Jacobs 2000; Martin and Noreen 1974; Terrace 2000). Because it takes time to retrieve each chunk, IRTs are relatively short within a chunk, but are relatively slow at boundaries between successive chunks. Although this account is somewhat circular (the properties that define a chunk are also the properties that chunking is designed to explain), the IRT data in our maze task and in studies of symbolic lists are strikingly similar.

Although the effects of optic flow suggest that participants are using some type of spatial information to orient themselves, the IRT data and participants' self-reports suggest that the maze is ultimately learned as a series of left and right turns. The spatial dimension of the task re-asserts itself when participants make an error and need to know where they are in the sequence. In this case, either motion or landmarks may provide orienting information and prevent the participant from becoming lost in maze.

5 Discussion

The effect of optic flow on learning (in this experiment and in two pilot studies with the same paradigm) proved to be fragile: moderate familiarity with the task caused the effect to disappear. Because of the highly constrained nature of the multiple T-junction maze, participants could effectively learn the maze in a purely symbolic mode, as a series of turns. Nonetheless, even with T-junction mazes, we found significant effects of optic flow, especially when landmarks were absent. One can imagine other kinds of spatial navigation experiments, using the same basic methodology, that would promote greater reliance on spatial rather than symbolic coding. As a step in that direction, we created a spatial-memory task that was free from the physical constraints of the T-junction maze and, therefore, was less likely to facilitate symbolic encoding of the environment.

6 Experiment 2: Remembering targets in a city-block environment

In experiment 1, participants did not need a spatial representation of the environment in order to complete a multiple T-junction maze. A memorized sequence of left and right turns would suffice. In designing experiment 2 we had two main objectives: first, we aimed to develop a spatial learning task that would be more resistant to symbolic coding; and, second, we sought to clarify the role of optic flow in spatial learning. Our plan was to refine the manipulation of optic flow, varying the degree of optic flow rather than its presence or absence.

7 Method

Instead of constraining participants' movement to the one path defined by the corridors, turns, and dead ends of a maze, our new paradigm allowed participants to move in any direction through a city-block environment. Because the goal was to encourage participants to form a spatial representation of the environment, we asked participants to traverse *different* paths from a start position to a target location. It was our hope that, by approaching the target along several different routes in succession, participants would form a more accurate spatial representation of the target location within the environment. In addition, we allowed participants to move multiple city-blocks in one direction by holding down a key. This feature was intended to discourage participants from counting city-blocks as they traversed the environment.

Several measures were taken to eliminate visual cues from the city-block environment. The size of the environment was greatly increased because a pilot study indicated that participants were using perimeter walls as visual landmarks to help locate the target. The textures on the floor, ceiling, and walls were modified to create a homogeneous environment of random-luminance squares. Additionally, this new texture made the corridors of the environment endless, making it impossible to use the perimeter walls as distal cues to the target location.

The city-block environment consisted of 25 north–south and 25 east–west corridors. All walls, ceiling, and floor were covered with low-contrast squares of random luminance (Michelson contrast 7%, mean luminance 75 cd m^{-2}), creating an environment that was uniformly patterned. The environment was navigated from an egocentric perspective with the four arrow-keys on a standard computer keyboard. The navigator's eye height was 60% of the wall height.

7.1 Manipulation of optic flow

For this experiment we improved the optic-flow manipulation in several important ways. First, rather than comparing optic flow with no optic flow, we varied the salience of the optic flow. Second, we increased realism by providing appropriate optic flow when participants made turns in the environments; in the previous experiment, a turn produced no such flow. Third, we slowed the speed at which participants traversed corridors by a factor of 7, bringing the rate of movement through the environment into line with rates used in other studies (eg Warren 1998). This decrease in speed also gave participants a more realistic sense of the virtual distances they were traversing. Finally, we were able to precisely equate the rate of the optic flow in the fluid and choppy conditions, ensuring that the time taken to traverse a corridor in both conditions was exactly 2.5 s.

The two optic-flow conditions in this experiment will be described as *fluid* and *choppy*. The mean rate of optic flow along the corridors for both of these conditions was two eye heights per second.[8] In the fluid condition, the display was updated 50 times during movement along a corridor; in the choppy condition, the display was updated just twice during the same distance. This 25-fold difference in the rate at which the display was updated produced distinctly different impressions of self-motion. Pauses of different lengths between successive display updates equated the time to move down a corridor in the two conditions. Optic flow was also present during turning, with the display being updated 15 times during a turn in the fluid condition and just twice during the choppy condition.

As in experiment 1, a single press of the ↑ or ↓ key was necessary to traverse an entire city-block. Likewise, pushing the ← or → key resulted in rotating through 90° in the appropriate direction. After a single keypress, the program moved the participant's view along the city-block corridor incrementally, with a brief pause after each spatial increment. Pressing any key while traversing a corridor or turning a corner had no effect on movement. However, multiple blocks could be traversed by holding down the ↑ or ↓ key. Likewise, holding down the ← or → key resulted in multiple turns in the appropriate direction.

The upper three panels, (a)–(c), in figure 6 show a participant's view while traversing through two city-blocks (requiring two ↑ keypresses) in the study mode. The forward arrows on the floor and the suspended left arrow direct the participant along the study route to the target position. Panel (d) shows the participant's view after a single ← keypress and panel (e) shows the view as the participant approaches the target location (indicated by a diamond symbol). After navigating along directed paths to the target,

[8] The measure of eye heights per second is the standard way to describe speed of movement in virtual environments. Since it is difficult to map conventional units (such as meters) onto virtual units, eye heights per second relates rate of movement to the height of the navigator in the optic-flow field. Two eye heights per second is approximately equal to a brisk walk.
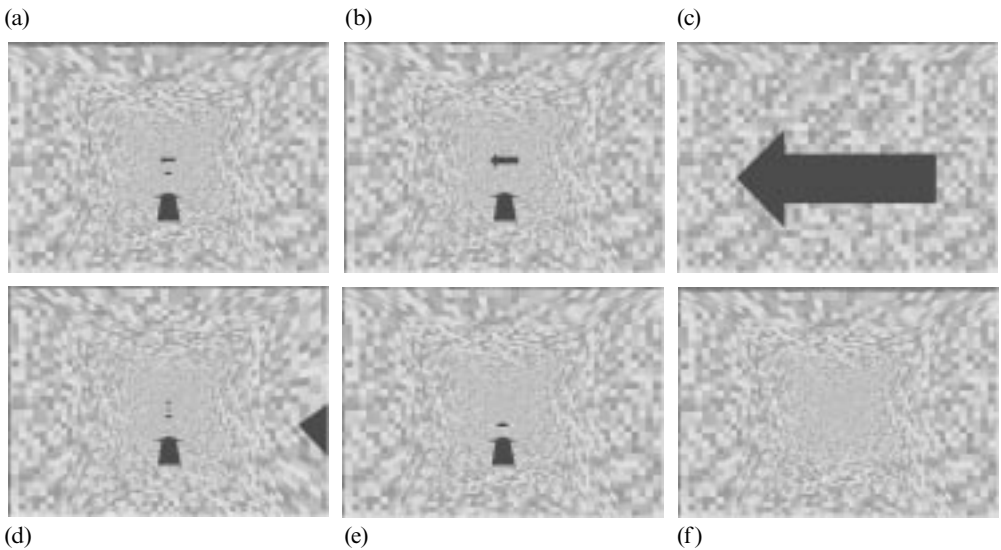
Figure 6. Views from within the city-block environment. See the text for a complete description of the individual panels.

participants attempt to find the target without any directional arrows. A view of the environment without these arrows (test mode) is shown in panel (f).

7.2 *Targets and routes*
Participants learned the location of a target in the environment by following a series of four different routes, designated by arrows (as in experiment 1). Each route from the center of the environment (the starting location) to a target position comprised five turns, and covered a total of 23 city-blocks. The lengths of the straightaways between turns varied between one and five city-blocks (uniform distribution). Routes that intersected themselves were excluded. All legitimate routes between the starting location and a target position were created, and four were chosen at random. These same four routes to a given target position were traversed by all participants in all conditions. Because experiment 1 indicated that the presence of visual cues masked the effect of optic flow, no visual landmarks were present along the route to aid participants in navigation.

Figure 7a shows the eight different target positions used in this experiment. The participant's starting location is indicated by a triangle, and the starting orientation is facing the top of the page or north. Target positions, represented by crosses, were distributed throughout the environment, with two targets lying in each of the four quadrants. Figures 7b – 7e are aerial diagrams of the four arrow routes leading participants to one of the target locations.

7.3 *Participants*
Sixty-four students (thirty-two male and thirty-two female, mean age 21.3 years) participated for either monetary compensation or course credit. The experiment was conducted in a single sitting of approximately 1.25 h.

7.4 *Procedure*
Before beginning the experiment, participants were introduced to both the fluid and the choppy optic-flow conditions. Participants were given 45 s to locate a randomly placed blue diamond within the environment, without the aid of any arrows. This allowed participants to become familiar with both the layout of the environment and the movement associated with each arrow-key. The program stopped after 45 s if the target had not been found. This was repeated for both optic-flow conditions.
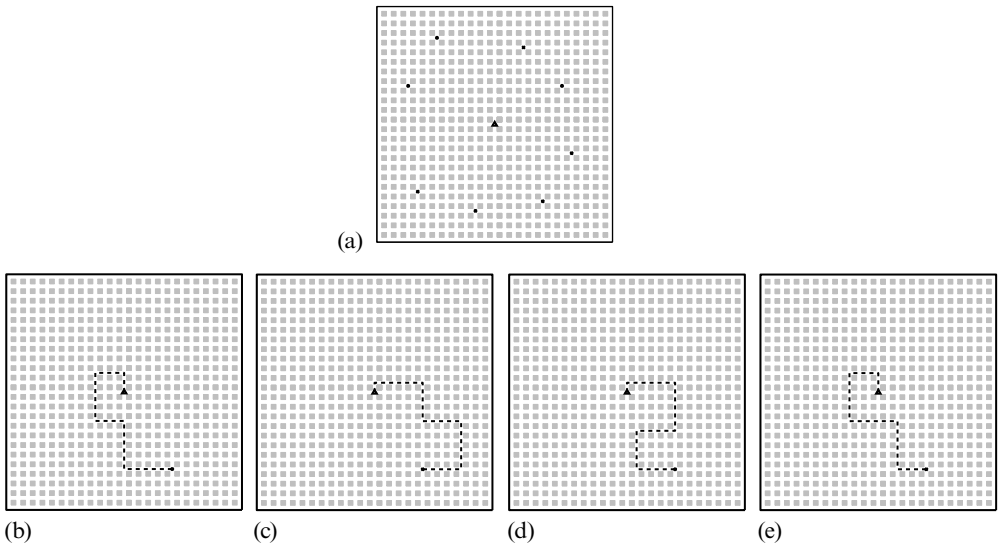
Figure 7. (a) The placement of eight different target locations (represented by small dots) in the city-block environment. The triangle in the center indicates the starting position and direction. (b) – (e) Illustration of the four study routes that a participant would navigate to learn the location of the target position.

The experiment consisted of one practice target and eight experimental targets, all presented in a study/test paradigm. Participants first learned the target position by following a series of four different arrow paths (study mode). All paths began in the center of the environment (facing north), and the target position was indicated by a blue diamond placed on the floor of the final junction.

Participants were instructed to remember the location of the target with respect to their starting location. In addition, special emphasis was given to the importance of not deviating from the designated path. In the test phase, neither the arrows nor the blue diamond were visible. Beginning at the same start location, participants were to navigate directly to where they remembered the blue diamond to be and press an exit key. After this single traversal, participants were asked to make a confidence judgment (on a scale from 1 to 7) based on their performance. Participants were then given feedback as to their approximate distance from the actual target.[9]

All participants completed trials with each of the eight target positions. Four of these trials were done with fluid optic flow, and four with choppy optic flow.[10] The order of the target positions, and the assignment of target positions to the two optic-flow conditions were counterbalanced across participants. For each target position, the order of the four study paths was randomized. The practice target was navigated in the fluid optic-flow condition.

[9] Feedback was based on the number of city-blocks the participant landed from the actual target. If the participant remembered the exact target location, the computer displayed the message, "Congratulations, you landed directly on the target!!". If the subject did not land precisely on the target location, but did land within 3 city-blocks, the message read, "You landed very close to the target position." Between 4 and 7 blocks away the message read, "You landed in the vicinity of the target position." And if the participant landed more than 8 blocks from the target position, the message read, "You landed quite a way from the target position."

[10] At the conclusion of the experiment, participants were given a survey, similar to the one in experiment 1. Participants were asked to judge whether the optic-flow manipulation affected their performance. They were also asked to rate the quality of their sense of direction. None of the participants' judgments correlated with performance measures.

8 Results

Overall, participants performed well, landing on the exact target position on 37% of all trials and navigating to within two city-blocks of the target on 54% of all trials. Of trials (37% of the total) on which participants landed exactly on the target location, 58% were traversed with fluid optic flow, whereas only 42% were traversed with choppy optic flow ($\chi^2 = 4.7$, $p < 0.05$). Optic flow also affects the average Euclidean distance between the remembered target positions and the actual target positions. An ANOVA on this error measure yielded a significant main effect for fluidity of optic flow ($F_{1,63} = 9.96$, MSE $= 5.32$, $p < 0.01$). Additionally, optic flow affects the total time spent navigating the environment ($F_{1,63} = 20.07$, MSE $= 111.94$, $p < 0.01$). However, the number of keystrokes used was not significant ($F < 1$) between the two optic-flow conditions. Table 1 reports the means and standard errors for these three measures. In addition to these objective measures, participants' confidence judgments indicated that they were well aware of the accuracy with which they navigated to a target position. Higher-rated confidence correlated strongly with proximity to target ($r = 0.47$, $p < 0.01$, two-tailed).

Figure 8 shows a histogram of the number of keystrokes taken to reach the remembered target location for all trials. The choppy optic-flow condition is represented by the solid gray line and the fluid optic-flow condition by the broken black line.

Table 1. The average error (Euclidean distance measure of how far a participant landed from the actual target), the average total time spent in the environment, and the average number of keystrokes a participant used to reach the remembered target, for fluid and choppy optic-flow conditions. Standard errors are shown in parentheses.

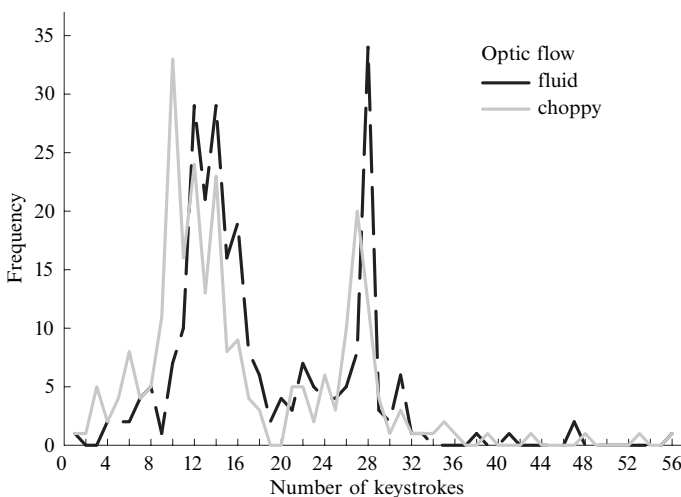| Performance measure	Optic flow	
	choppy	fluid
Average error	4.5 (0.3)	3.2 (0.3)
Total time/s	55.0 (1.9)	46.6 (1.5)
Keystrokes	18.2 (0.5)	18.3 (0.5)

Figure 8. Histogram of the number of keystrokes used in reaching the remembered target location. Choppy optic flow is represented by the solid gray line and fluid optic flow by the broken black line. There were 28 keystrokes in all arrow routes leading participants to the target location in the learning phase and it took an average of 12.5 keystrokes to reach targets when using an L-path strategy (see text for details).

Both distributions are bimodal, with the first mode concentrated around $10-15$ keystrokes and the second centered around $27-28$ keystrokes. As will become apparent, these distributions are useful for identifying the strategies participants used to reach the remembered location of the targets.

Table 2 reports the average error for the fluid and choppy optic-flow conditions at each target position. Because each participant navigated to each target position once, either with fluid or choppy optic flow, the interaction between optic flow and target position could only be assessed by computing independent sample t-tests for each target position. Although only one target position exhibited a statistically significant advantage of fluid optic flow, there was a strong trend for target locations situated behind the participant (as judged from the starting orientation) to be more difficult to remember in the presence of choppy optic flow.

Table 2. Euclidean distance between remembered and actual target locations for fluid and choppy optic-flow conditions. The first column gives the Cartesian coordinates of each target position.

Coordinates	Optic flow		t_{62}, p (two-tailed)
	fluid	choppy	
$(7, 4)$	2.57	2.69	0.165, ns
$(3, 8)$	3.88	2.89	-0.87, ns
$(-6, 9)$	3.40	2.97	-0.430, ns
$(-9, 4)$	2.37	4.27	1.756, $p < 0.1$
$(-8, -7)$	2.68	5.82	2.047, $p < 0.05$
$(-2, -9)$	3.33	5.90	1.896, $p < 0.1$
$(5, -8)$	3.61	6.21	1.789, $p < 0.1$
$(8, -3)$	3.73	5.12	1.180, ns

9 Discussion

Participants navigated to the remembered target position with greater accuracy in the fluid optic-flow condition than in the choppy optic-flow condition. They also required less time to locate the remembered target position with fluid optic flow. Why then, did participants use approximately equal numbers of keystrokes in both conditions?

Participants used two main strategies to navigate to the remembered target location. Some participants attempted to retrace one of the study routes (which were all composed of 28 keystrokes), while other participants took a more direct, L-shaped path to the remembered location. An L-path is composed of two segments (not necessarily of the same length) separated by a single turn. Navigating to a target location by using an L-path strategy required an average of 12.5 keystrokes. Figure 9 illustrates both an L-shaped path and a path that mimics an arrow route to the correct target position.[11]

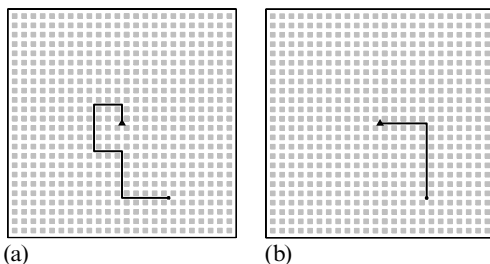(a)   (b)

Figure 9. Illustrations of two possible strategies a participant might have taken to get to the correct target location. Panel (a) illustrates a participant who retraced one of the studied paths. Panel (b) illustrates a participant who attempted to take the most direct, L-shaped, path from the start to the target position.

[11] All of the study routes for this target are diagrammed in figure 7.

Participants were remarkably consistent in the strategy they employed to locate the target position. Forty-eight of the sixty-four participants used the same strategy every time they navigated to a correct target location. Of these, thirty-two used an L-path strategy, whereas only sixteen retraced one of the study routes. Of the remaining sixteen participants, twelve never landed correctly on a target location, and four never used either of the two dominant strategies.

Figure 10 illustrates how performance varies with strategy. Overall, participants who used the L-path strategy were more accurate in remembering the target location than were participants who used the retracing strategy ($F_{1,46} = 12.79$, MSE = 10.88, $p < 0.01$). However, for both groups, participants landed closer to the target location with fluid optic flow than with choppy optic flow ($F_{1,46} = 11.49$, MSE = 5.23, $p < 0.01$). A marginally significant strategy × optic flow interaction effect ($F_{1,46} = 3.24$, MSE = 5.22, $p = 0.079$) was also observed.
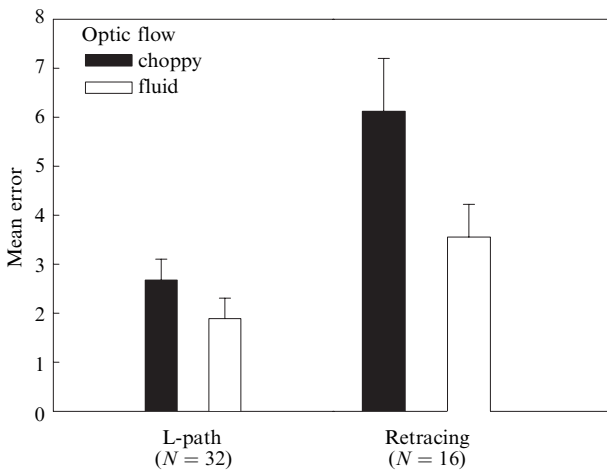

Figure 10. Mean error (Euclidean distance from target) as a function of strategy (L-path versus retracing one of the studied routes). Participants were included in a strategy group if they used that strategy every time they navigated directly to a target position. The number of participants in each group is indicated below the strategy labels. Fluid optic flow is represented by the white bars and choppy optic flow by the black bars. Error bars represent the standard error of the mean.

When successfully retracing one of the four study routes, participants followed the first or last route about twice as often as the intermediate routes. Owing to the small number of participants who adopted this strategy, this effect did not reach significance ($\chi^2 = 6.00$, $p = 0.11$). Participants' tendency to follow the first or last route more often than the middle routes is consistent with the general phenomena of serial position effects in memory. Even in very short lists, there is a reliable tendency for early and late list items to be remembered better than interior list items (see Murdock 1974, for a review).

In order to correctly navigate to the target position, a participant must put together (or perform path integration on) all pieces of the studied paths. This leads to an explanation for why participants used approximately the same number of keystrokes in both the fluid and choppy optic-flow conditions, yet navigated closer to the actual target position with fluid optic flow. The presence of choppy optic flow interrupts path integration, making it difficult to put the pieces of the path together. Participants knew the lengths of the different segments of the path and, therefore, knew the distance to the target, but were unable to determine the correct direction of the target (ie participants knew the lengths of the segments between the five turns, but could not deduce which direction to turn at those junctions). Approximately the same number of keystrokes was used in both optic-flow conditions irrespective of whether the participant landed directly on the target location. The approximate equivalence in number of keystrokes for the two optic-flow conditions also held individually for forward/backward movements ($F_{1,63} = 0.10$, MSE = 0.63, ns) and left/right turns ($F_{1,63} = 0.41$, MSE = 0.16, ns).

This suggests that participants knew how far away the target position was and how far they had to navigate to get there, but they had greater difficulty constructing the proper path when the salience of optic flow was diminished.

10 General discussion

Studies of wayfinding in natural environments sometimes describe people's behavioral strategies in terms that are starkly binary. For example, wayfinders are portrayed as depending either on landmark-related information, or upon information derived from dead reckoning or from some general sense of direction. Moreover, these strategies have been ascribed to variables such as gender (Astur et al 1998). Our experiments, which showed no gender-related differences in performance or information usage, show that people can and will exploit various sources of information to help them learn their way around an environment. The relative importance of particular sources of information varies with the properties of the environment, and with the availability of alternative sources of information. For example, in the mazes of our first experiment, the effects of optic flow suggest that participants are using some form of spatial information to orient themselves. However, the IRT data and participants' self-reports suggest that the maze is ultimately learned as a series of left and right turns. The temporal output pattern, characterized by bursts of three responses followed by a pause, is one that is usually associated with verbal sequence memory (Kahana and Jacobs 2000). The spatial dimension of the maze reasserts itself when participants make an error and need to know where they are in the sequence. In this case, either motion or landmarks may provide orienting information and prevent the participant from becoming lost in maze. When landmarks were present, they were not used to cue a particular direction of turn, say right or left. Instead, the landmarks served as an indicator to participants of whether they had made the correct decision at that junction. If a correct turn was made, a landmark would be visible in the next junction.

Our second experiment eliminated some constraints on participants' behavior. The results provided striking evidence that participants can integrate separate experiences with several different paths into a new, more efficient path. This behavior is reminiscent of the path integration behavior that Gallistel (1995) observed in desert ants and other creatures. However striking the parallel, it is important to remember that humans are not ants. In our experiment, for example, many people consistently chose another strategy, namely retracing one of the paths they had been led along during the training phase of the experiment. So, in our species, both strategies—path integration and retracing a previously taken path—coexist, and are available. One can imagine conditions that might modulate the relative frequencies with which the two are used. For example, if excess or unnecessary steps (keystrokes) carried a large penalty, one would expect a greater tendency to attempt to rely on path integration.

The two types of artificial environments we used far from exhaust the set of all virtual environments that could be valuable in research on wayfinding. The same can be said for the two types of tasks that we set for our participants. Because our results suggest that both environment and task influence wayfinding strategies and performance, it is important not to overgeneralize from the outcomes described above. For example, both our virtual environments were highly contrived, having relatively flat spatial organizations, with uniformity of spatial detail. Many natural or man-made environments have levels of detail that vary from one region of the environment to another, and this variation itself should influence strategies for learning and wayfinding. Consider, for example, geographical units that are laid out as a series of interconnected center-spoke units. The spokes connecting two units may be relatively undifferentiated (unpopulated by people or significant man-made landmarks), while each of the centers (such as town centers) are filled with both people and landmarks. The strategy needed

to navigate through the centers of such environments may well differ from the strategy needed to navigate along a spoke, from one center to another. Gillner and Mallot (1998) studied participants exploring such center-spoke towns, in a virtual environment, and showed that the overall layout of a complex environment can be learned by piecing together local information. Finally, as we mentioned before, the learning and navigational strategies that people choose are likely to be influenced by payoff structures that include costs and benefits defined by travel time, effort, and energy. These could be readily examined with specially constructed videogame tasks in which participants try to maximize an experimenter-defined reward, and incidentally learn the layout of complex spatial environments in the process.

11 Conclusions

Optic flow helps participants learn a series of left and right turns (experiment 1) and spatial locations (experiment 2) while navigating through synthetic environments. Although participants seemed to rely heavily on symbolic coding strategies in experiment 1, the absence of optic flow resulted in participants becoming disoriented and getting lost within our virtual mazes. The presence of visual landmarks at maze junctions attenuated the effect of optic flow on performance. Experiment 2 more directly tested the effects of optic flow on participants' ability to learn spatial locations within a large-scale synthetic environment. In this experiment, we show that memory of spatial locations is better when the environment is navigated with fluid optic flow than when optic flow is choppy.

When other cues are unavailable, optic flow can be a significant aid to wayfinding. In these experiments, we have shown that salient optic flow can facilitate the learning of specific locations in synthetic environments. Additionally, this optic flow aids in path integration and in forming mental representations of spatial environments.

References
Aguirre G K, Detre J A, Alsop D C, Esposito M D, 1996 "The parahippocampus subserves topographic learning in man" *Cerebral Cortex* **6** 823 – 829
Astur R S, Ortiz M L, Sutherland R J, 1998 "A characterization of performance by men and women in a virtual Morris water task: A large reliable sex difference" *Behavioral Brain Research* **93** 185 – 190
Berthoz A, 1997 "Parietal and hippocampal contribution to topokinetic and topographic memory" *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* **352** 1437 – 1448
Burgess N, Hitch G J, 1999 "Memory for serial order: A network model of the phonological loop and its timing" *Psychological Review* **106** 551 – 581
Chance F S, Kahana M J, 1997 "Testing the role of associative interference and compound cues in sequence memory", in *Computational Neuroscience, Trends in Research* Ed. J Bower (New York: Plenum Press) pp 599 – 603
Gallistel C R, 1995 "The replacement of general-purpose theories with adaptive specializations", in *The Cognitive Neurosciences*, volume 1, Ed. M S Gazzaniga (Cambridge, MA: MIT Press) pp 1255 – 1267
Gibson J J, 1958 "Visually controlled locomotion and visual orientation in animals" *British Journal of Psychology* **49** 182 – 194
Gillner S, Mallot H A, 1998 "Navigation and acquisition of spatial knowledge in a virtual maze" *Journal of Cognitive Neuroscience* **10** 445 – 463
Johnson N F, 1972 "Organization and the concept of a memory code", in *Coding Processes in Human Memory* Eds A W Melton, E Martin (Washington, DC: Winston) pp 125 – 159
Kahana M J, Jacobs J, 2000 "Inter-response times in serial recall: Effects of intraserial repetition" *Journal of Experimental Psychology: Learning, Memory and Cognition* **26** 1188 – 1197
Kahana M J, Sekuler R, Caplan J B, Kirschen M, Madsen J R, 1999 "Human theta oscillations exhibit task dependence during virtual maze navigation" *Nature (London)* **399** 781 – 784

Klatzky R L, Loomis J M, Beall A C, Chance S S, Golledge R G, 1998 "Spatial updating of selfposition and orientation during real, imagined, and virtual locomotion" *Psychological Science* **9** 293 – 298

Loftus G R, Masson M E J, 1994 "Using confidence intervals in within-subject designs" *Psychonomic Bulletin and Review* **1** 476 – 490

Maguire E A, Burgess N, Donnett J G, Frackowiak S J, Frith C D, O'Keefe J, 1998 "Knowing where and getting there: A human navigation network" *Science* **280** 921 – 924

Maguire E A, Frackowiak S J, Frith C D, 1996 "Learning to find your way: A role for the human hippocampal formation" *Proceedings of the Royal Society of London, Series B: Biological Sciences* **263** 1745 – 1750

Maguire E A, Frackowiak S J, Frith C D, 1997 "Recalling routes around London: activation of the right hippocampus in taxi drivers" *Journal of Neuroscience* **17** 7103 – 7110

Martin E, Noreen D L, 1974 "Serial learning: Identification of subjective subsequences" *Cognitive Psychology* **6** 421 – 435

Miles W R, 1928 "The high relief finger maze for human learning" *Journal of General Psychology* **1** 3 – 14

Murdock B B, 1974 *Human Memory: Theory and Data* (Potomac, MD: Lawrence Erlbaum Associates)

O'Keefe J, Nadel L, 1978 *The Hippocampus as a Cognitive Map* (New York: Oxford University Press)

Perrin F A C, 1914 "An experimental and introspective study of human learning process in the maze" *Psychological Monographs* **16** 1 – 97

Samsonovich A, McNaughton B L, 1997 "Path integration and cognitive mapping in a continuous attractor neural network model" *Journal of Neuroscience* **17** 5900 – 5920

Stone C P, Nyswander D B, 1927 "The reliability of rat learning scores from the multiple T-maze as determined by four different methods" *Journal of Genetic Psychology* **34** 497 – 524

Terrace H S, 2000 "The comparative psychology of serially organized behavior", in *Biomedical Implications of Model Systems of Complex Cognitive Capacities* Eds S Fountain, J H Danks, M K McBeth (New York: Sage)

Vincent S B, 1912 "The function of the vibrissae in the behavior of the white rat" *Behavioral Monographs* **5** (whole issue)

Warren W H J, 1998 "Visually controlled locomotion: 40 years later" *Ecological Psychology* **10** 177 – 220

Witmer B B, Bailey J H, Knerr B W, 1996 "Virtual spaces and real world places: transfer of route knowledge" *International Journal of Human – Computer Studies* **45** 413 – 428