Navigation and Acquisition of Spatial Knowledge in a Virtual Maze

Sabine Gillner and Hanspeter A. Mallot
Max-Planck-Institut für biologische Kybernetik

Abstract

■ Spatial behavior in humans and animals includes a wide variety of behavioral competences and makes use of a large number of sensory cues. Here we studied the ability of human subjects to search locations, to find shortcuts and novel paths, to estimate distances between remembered places, and to draw sketch maps of the explored environment; these competences are related to goal-independent memory of space, or cognitive maps. Information on spatial relations was restricted to two types: a visual motion sequence generated by simulated movements in a virtual maze and the subject's own movement decisions defining the path through the maze. Visual information was local (i.e., no global landmarks or compass information was provided). Other position and movement information (vestibular or proprioceptive) was excluded. The amount of visual information provided was varied over four experimental conditions. The results indicate that human subjects are able to learn a virtual maze from sequences of local views and movements. The information acquired is local, consisting of recognized positions and movement decisions associated to them. Although simple associations of this type can be shown to be present in some subjects, more complete configurational knowledge is acquired as well. The results are discussed in a view-based framework of navigation and the representation of spatial knowledge by means of a view graph. ■

INTRODUCTION

Spatial Memory and Cognitive Maps

All organisms capable of locomotion have to deal with space and spatial relations within their environment. Simple tasks like efficient grazing and foraging, path integration, or systematic search can be achieved without a mental representation of space, whereas more advanced competences require the recognition of places as well as knowledge of spatial relations, such as the distance and bearing of a goal, routes, or configurations of places. In this paper, we address the problem of exploration, path planning, and navigation in a virtual maze (i.e., in an environment composed of streets and junctions and with goals that are not generally visible from the starting position). The knowledge or mental representation required for this task is studied by behavioral experiments with human subjects navigating in a virtual environment simulated on a computer screen.

Mental representations of space are often called *cognitive maps*. More specifically, there seem to be at least three more or less independent ideas related to the concept of a cognitive map:

1. *Cognitive map as a spatial reasoning stage.* Tolman's original notion (Tolman, 1948) considers the ability to find (or infer) novel shortcuts as crucial for the presence of a cognitive map.
2. *Cognitive map as a cue integration stage.* Spatial behavior rests on a fair number of different information sources that are not easily combined. At the stage where the integration occurs, all information has to be present in a compatible way. This interaction stage may be called a cognitive map (see Gallistel, 1990).
3. *Cognitive map as goal-independent memory of space.* Information about spatial relations can be acquired in neutral (unrewarded) situations and can be used for goal-directed behaviors later (latent learning). In contrast, routes are always headed toward a goal. See O'Keefe and Nadel (1978) for a detailed discussion.

Clearly, the above definitions are not mutually exclusive but simply highlight different aspects of cognitive maps. In terms of the underlying mechanisms, the third notion seems to allow the most clear-cut distinctions: If spatial learning is achieved by a mere modification of the mechanism generating the behavior, it will be stereotyped, and we will not call this a cognitive map. If, however, a separate storage is involved that does not itself produce behavior but is "loaded" into flexible mechanisms or referred to during planning, the term appears to be appropriate. This distinction is akin to the procedural versus declarative memory dichotomy as discussed by Squire (1987).

Types of Spatial Memory

What types of spatial behavior can be achieved without a cognitive map, and which ones cannot? We will split the discussion of this question into three parts, related to three basic navigational mechanisms: (1) path integration, (2) approaching recognized views (e.g., "homing"), and (3) route and graph memory.

Path Integration

In insect navigation, it has been shown that many important tasks can be achieved by some kind of working memory such as a continuously updated "home vector" holding the egocentric coordinates of the starting position of the current excursion (Wehner & Menzel, 1990). The current position of some starting point in egocentric coordinates can easily be computed by triangulation (see Maurer & Séguinot, 1995, for review). Path integration has been studied in blind and blindfolded human subjects by Loomis et al. (1993) and in sighted subjects using virtual reality by May, Wartenberg, and Péruch (1997). The representation required for path integration is a simple buffer storing the two vector components (Mittelstaedt & Mittelstaedt, 1972/73; see also Touretzky, Redish, & Wan, 1993). Recently, McNaughton et al. (1996) have proposed an alternative mechanism based on hippocampal place cells. In all models, storage is achieved by neuronal activity (rather than synaptic plasticity), that is, by some kind of working or short-term memory.

The memory involved in repetitions of a previously traveled distance can be based on more elaborate mechanisms as well. Recent results by Berthoz, Israël, Georges-François, Grasso, and Tsuzuku (1995) indicate that, in humans, the repetition of short distances involves not just a continuously updated vector buffer but uses a stored velocity profile. It is not clear, however, how this result extends to longer routes.

An intriguing property of path integration is its close relation to metric information. Although it is sometimes assumed that the access to metric information requires highly sophisticated cognitive maps, it appears that metric is in fact one of the most basic properties of spatial short-term memory.

Approaching Recognized Views

Recognizing and approaching views (local landmarks) requires a long-term memory of the view or some of its features. A strictly associative mechanism for this task has been proposed by Barto and Sutton (1981). It actually stores the required approach direction for every position identified by its local position information. A more general mechanism for homing that computes the approach direction from the comparison of current and stored views has been proposed by Cartwright and Collett (1982). This scheme involves long-term memory of the approached view, but not of the required movements, which are computed. If only one view is to be approached (homing in a strict sense), memory can be realized in a procedural and stereotyped way (e.g., by some sort of matched filter for the home view). If, however, the same machinery is to be used for many different approach tasks, the appropriate target views would have to be "loaded" into a comparison stage as needed. In the meantime, they must be kept in some long-term memory. The same argument applies to the somewhat more powerful model by Benhamou, Bouvet, and Poucet (1995) describing homing behavior in mammals. (See also Franz, Schölkopf, Mallot & Bülthoff, 1998, for an alternative implementation of this approach mechanism.)

Routes and Configurations

As the basic element of route memory and configuration memory, we consider an association of the form

$$\text{(current view, (movement direction, expected next view))} \qquad (1)$$

which is illustrated in Figure 1d. Associations between views and movement decisions have been demonstrated, for example, in bees (Collett & Baron, 1995) and have been used in the associative schemes of Barto and Sutton (1981) and McNaughton and Morris (1987). When going from one view to the next, navigation can initially follow the movement direction associated with the present view. A scheme for robot navigation based on recognized landmarks and movement behaviors associated with them has been suggested by Kuipers and Byun (1991). The additional information on what view to expect next is required in order to switch to the appropriate approach behavior when arriving in the neighborhood ("catchment area") of that view. Alternatively, stereotyped approach behaviors for all known views could be active in parallel. In this case, they would need to produce a confidence measure allowing the selection of the correct one.

Chains of such association structures implement a route memory. If different routes are to be learned that share some common section, the decision at the crossroads requires more complicated memory. One way to think of this memory is to store all possible connections

$$\begin{array}{ll} \text{(current view,} & \text{(movement direction 1,} \\ & \quad \text{expected next view)} \\ & \quad \vdots \\ & \text{(movement direction } n\text{,} \\ & \quad \text{expected next view))} \end{array} \qquad (2)$$

and have a separate planning device select one of the possible movements. A neural network theory for storing the required information in the form of a labeled graph has been presented by Schölkopf and Mallot (1995). For

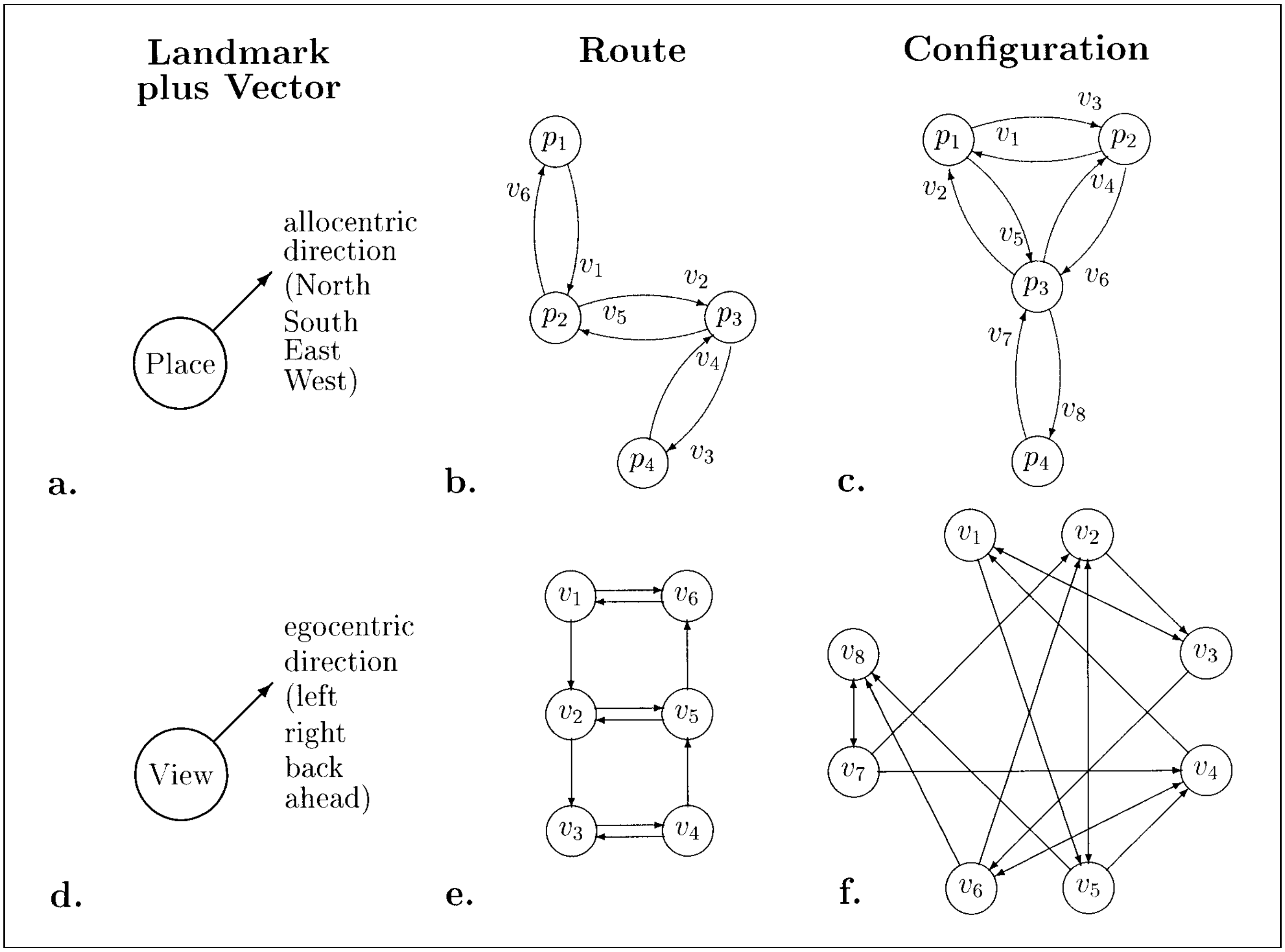


Figure 1. The graph approach to space representation. Top row (a–c): Place graphs. The nodes are places recognized irrespective of body orientation, the links (arrows) between them carry allocentric direction information. Bottom row (d–f): View graphs. The nodes are recognizable views or other positional information (i.e., depend on the observer's viewing direction), and the arrows carry directional information relative to gaze. Each view v_i in parts e and f corresponds to a directed connection in parts b and c. From left to right, increasingly more complicated spatial layouts are shown.

related approaches including hippocampal modeling, see Muller, Stead, and Pach (1996), Prescott (1996) and Touretzky and Redish (1996).

An Ecological View of Spatial Memory

In this paper, we will deal mostly with memory of routes and configurations (i.e., relational knowledge of position and the movements leading from one position to another). A further breakdown of this problem is given in Table 1. The behavioral competences have been arranged in order of increasing complexity. Cognitive maps may be unnecessary for the first two but become increasingly more relevant for the more complex tasks. The list of sources of information usable in navigation tasks is probably not complete. Again, there are trivial cases like pointers, which do not require any spatial knowledge or map, as well as more complicated cues that can only be interpreted correctly if map information is available. Note that we have included "path integration" as a source of information. Simple path integration does not require a cognitive map and can thus be considered a separate mechanism feeding into the map module. Possible representations acquired by spatial learning are listed in the third column of Table 1. As was discussed earlier, the home vector is a form of working memory. Associations and simple (nonbifurcating) chains of associations can be learned in a stereotyped or procedural way. If the same knowledge is to be used in the pursuit of different goals, a goal-independent, graphlike memory is required. Finally, a topographic map with coordinates and distances is the richest but rather unlikely representation.

The View-Based Approach to Navigation

The problem, then, is to find the minimum representation required to explain an animal's or human's

Behavioral Competences	Sources of Information	Possible Representations
• Repeat a previously travelled path • Find known target from new starting point • Find shortcuts • Bypass newly blocked connections • Latent learning: use previously acquired knowledge for new searches • Communicate about paths	• Pointers (guidances) and general rules • Global compasses (sun, chemical gradients, etc.) • Distance estimates (optical flow, effort, etc.) • Path integration • Sequences of views (landmarks) and movements	• Home vector • Associations between views and motor commands • Route memory • Relational information on places, connections, and views (e.g., graph structures) • Topographic maps representing metric relations

Table 1. An ecological view of cognitive maps. For explanation, see text.

behavioral competences in the presence of a certain type of environmental information. This idea of economic or parsimonious explanations of spatial behavior is especially well developed for insect navigation.[1] For the type of knowledge studied here (i.e., the expectation of the next snapshot generated from the current snapshot and the intended movement), the most simple element is shown in Figure 1a and d and Equation 1. In Figure 1a it is assumed that places are recognized irrespective of the observer's direction of gaze. Intended movements are then represented in a global coordinate system (i.e., in relation to an additional system such as global landmarks or path integration). These elements can be combined into chains (Figure 1b) or graphs (Figure 1c). In contrast to this "place-based" approach, the view-based approach (lower row of Figure 1) assumes that views, rather than places, are recognized and movements are represented in egocentric coordinates (i.e., without reference to an independent compass system). This approach can be extended to chains and graphs just like the place-based approach. For a mathematical analysis of the resulting view graphs, see Schölkopf and Mallot (1995).

Both the place- and the view-graph approaches are local in the sense that bits and pieces of spatial information can be accumulated without checking for global consistency. They focus on topological properties (connectivity); metric relations can be added as labels to the links. The main differences between the two approaches are (1) that metric labels of the place graph have to be allocentric (world-centered), whereas those of the view graph are egocentric (observer-centered) and (2) that the place graph is planar and symmetric (knowledge of a connection implies how to return), whereas the view graph is not.

It should be noted that the view-based approach to navigation is closely related to view-based mechanisms in direction-invariant object recognition (see Bülthoff, Edelman, & Tarr, 1995). Places and objects can be represented by their respective views in quite similar ways. The graph structure resulting for a maze with many places is generally not planar (cf. Figure 1f), whereas the view graphs for object recognition are.

Behavioral Experiments in Virtual Reality

In the work reported in this paper, we chose interactive computer graphics, or virtual reality (VR), as our experimental method. Previous studies using virtual reality have focussed on the transfer of knowledge between different media used for acquisition and testing. May, Péruch, and Savoyant (1995) and Tlauka and Wilson (1996), for example, have tested map-acquired knowledge in a pointing task performed in virtual reality. Tong, Marlin, and Frost (1995), using a VR bicycle, showed that active exploration leads to better spatial knowledge than passive stimulus presentation. Sketch maps produced after exploration of various virtual environments have been studied by Billinghurst and Weghorst (1995). Design principles for constructing easy-to-navigate virtual environments have been studied by Darken and Sibert (1996). In the present paper, we use virtual reality to isolate the various cues used for the build-up of spatial knowledge and to study the underlying mechanisms. The advantages of virtual reality for this application are (1) the high controllability of computer graphics stimuli and (2) the easy access to behavioral data, such as the subject's movement decisions.

Stimulus Control

When investigating the information sources used in navigation, it is advantageous to be aware of the exact movement trajectories of the subjects and the visual information available along these trajectories. This can easily be achieved with interactive computer graphic (see "Methods" section). The various parameters of the sensory input can be easily separated. For instance, in our experiments, we varied the number of buildings visible simultaneously in one view without changing the illumination, etc. In real-world experiments, such separate stimulus conditions are much harder to realize. Another interesting experimental paradigm is the modification or exchange of various features of the environment after learning. Aginsky, Harris, Rensink, and Beusmans (1996) exchanged landmarks after training in a route-learning task. The effects of landmark exchange on navigation have been addressed by Gillner and Mallot (1996).

The method also allows complete control over vestibular and proprioceptive feedback. In our experiments, both were completely absent, allowing the effects of visual input to be studied in isolation.

Measuring Behavior

Navigation performance can be accessed most directly by the paths or trajectories that the subjects take during the exploration. In virtual reality experiments, egomotion is very simple to record, because it is equivalent to the course of the view port used for rendering the computer graphics. In this paper, we present a number of novel techniques for data evaluation that are particularly suited for the virtual reality experiments described.

Plan of the Paper

The virtual reality setup and the procedure used in the experiments are described in the "Methods" section at the end of the paper. In the "Results" section, we present subjects' trajectories obtained during a search task, as well as two derived measures for transfer of knowledge between routes and for persistent associations of views to particular movements. In addition, distance estimates collected from the subjects after exploration are compared to theoretical distances from various candidate representations. Finally, some examples of subjects' sketch maps will be presented.

RESULTS

Exploration and Search

Performance

Figure 2 shows an example for the trajectory taken by a single subject when searching view number 5 from start view 15. In the first trial, the subject made a complete turn in the starting position and then started the exploration via view 17. At view 11, he performed a loop,

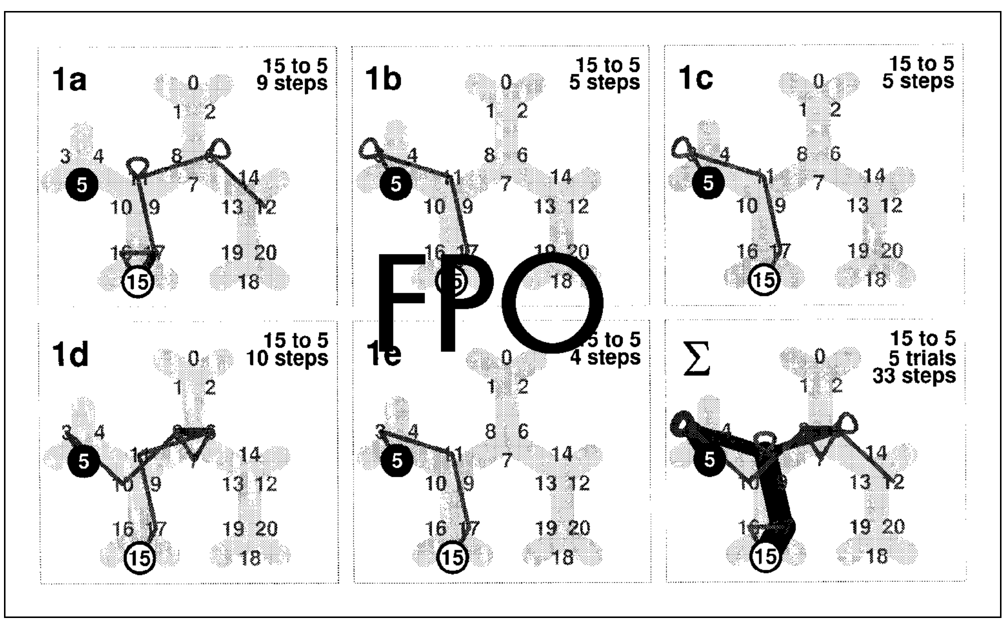


Figure 2. Sample trajectories for subject GPK (condition: dark) searching the way from view 15 (start) to view 5. 1a–1e: Search trials. In part 1e the shortest path is found for the first time completing the task 15 → 5. Σ: Accumulated trajectory from all five trials. This plot appears again in part 1 of Figure 3.

turning 60° to face a street and back again. He then proceeded to view 6, where he performed the same search behavior. At view 12, the trial was stopped because the subject deviated from the shortest path by more than one segment. In the second trial (part 1b in Figure 2), he finds the goal, though not the shortest possible path. Interestingly, the third trial is an exact replication of the second one. The first time he finds the shortest way is trial 5 (part 1e), which thus terminates the exploration of that route.

The cumulative trajectory shown in the lower right panel of Figure 2 appears again in the upper left panel of Figure 3. The other panels in this figure show the cumulative trajectories for the other routes performed subsequently in the sequence indicated by the number in the upper left of each panel. Paths 1 through 4 are excursions, 5, 7, 9, and 11 are returns, and 6, 8, 10, and 12 are novel routes. Overall, there is a tendency for lower error rates in the search tasks performed later. That is to say, there is a transfer of knowledge obtained in earlier searches to the later searches. The decrease of errors is not monotonic, though. Note, however, that the three last routes were found in just five trials. In some subjects, no such decrease of the error rate is found.

Error Rates

Errors were defined locally as decisions that do not reduce the distance of the goal. Each movement decision equals clicking the mouse buttons twice (cf. Figure 12 in the "Methods" section). Distance to the goal is measured as the minimum number of decisions needed to reach it ("decision distance"). Thus, if a subject enters a street leading away from the goal, the return from that street will be counted as a correct decision even though the current position is not part of the shortest path. In cases where the correct decision is a 60° turn left followed by a "go," the 120° turn left would leave the decision distance to the goal unchanged. This decision (and the mirror-symmetric case) is also counted as an error.

Average error rates for each path type are shown in Figure 4. For each viewing condition (1 through 4; see "Methods" section), the excursions, returns, and novel paths were lumped into groups of four. As mentioned above, the excursions were preformed first, and the novel and return paths were performed alternatingly, starting with a return in one group of subjects and starting with a novel route in a second group. The data

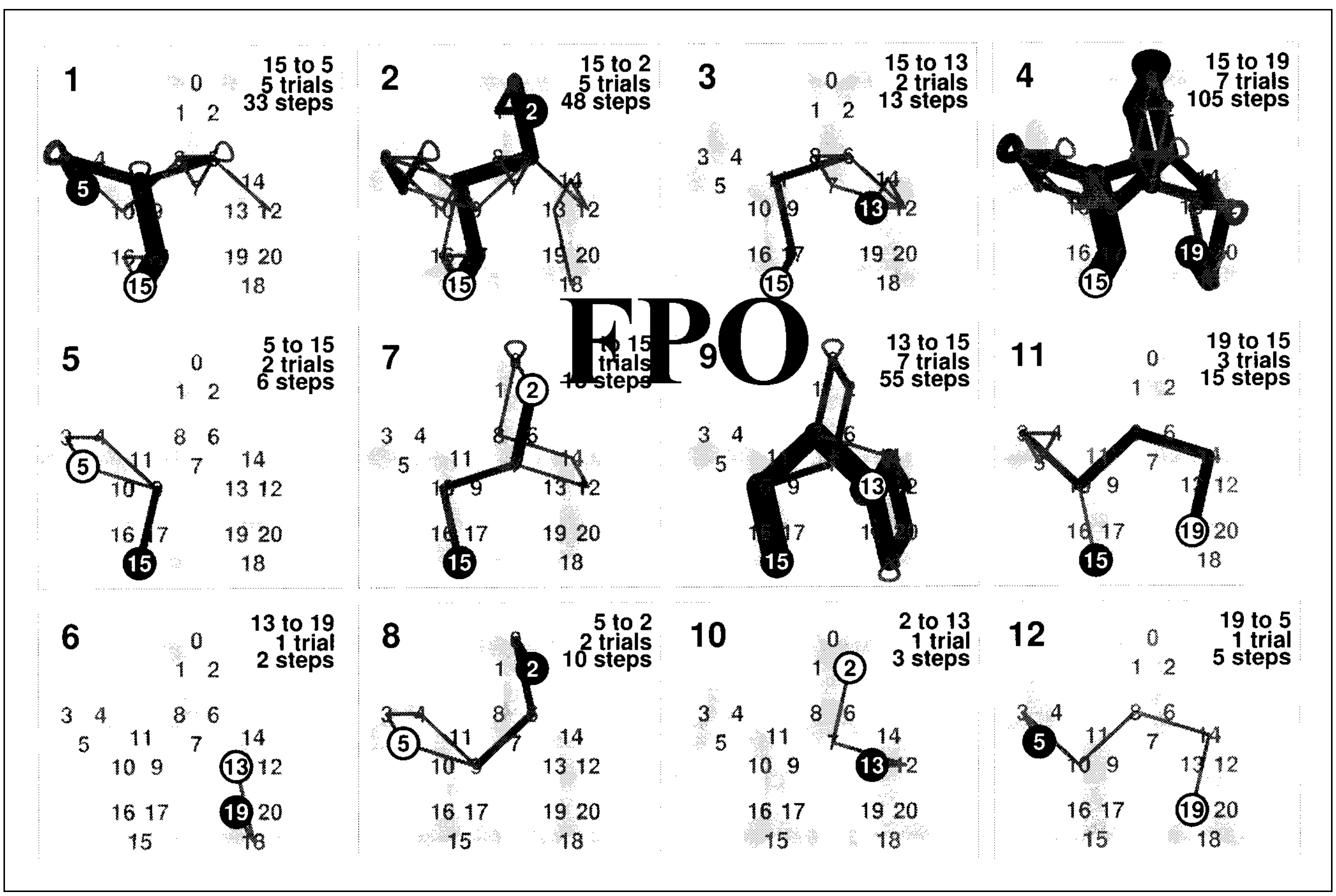


Figure 3. Traveling frequencies for each view transition for the 12 paths, subject GPK (viewing condition: 1; sequence condition: returns first). Top row: excursions, middle row: returns, bottom row: novel paths. The overall number of errors decreases at later stages of exploration.

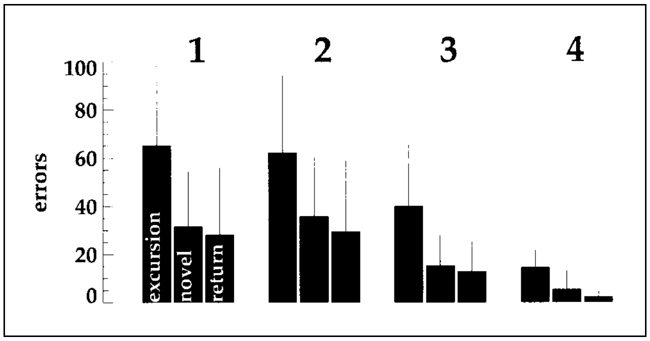


Figure 4. Total number of wrong movements performed in the different path types (excursion, novel, return). Numbers are averaged over 20 subjects, error bars are standard deviations. 1 through 4: viewing conditions.

show a learning effect in the sense that excursions take more errors than the later paths. They also show a clear effect of condition: Higher visibility results in lower error rates. This general relation does not hold for the comparison of conditions 1 and 2, however, which differ in the visibility of the neighboring places.

A three-way analysis of variance (ANOVA, 4 conditions × 3 path types × 2 genders) of error rate as the dependent variable reveals significant effects of condition ($F(3, 72) = 17.31$, $p < 10^{-4}$) and path type ($F(2, 144) = 60.65$, $p < 10^{-4}$) but not of gender ($F(1, 72) = 0.22$, *n.s.*). Additionally we found an interaction of condition and path type ($F(6, 144) = 2.66$, $p = 0.018$). The error rates of novel paths are slightly higher than those of the returns (see Figure 4). This effect is not significant, however.

Transfer

In our procedure, learning occurs on two time scales. During each of the 12 tasks, a route is learned as illustrated in the example in Figure 2. When switching from one route to the next, part of the knowledge acquired in the earlier routes might be transferred to the new ones. To test this, we define a transfer coefficient τ in the following way:

Let R and N denote two routes, for instance the first return and novel path, respectively. Our group of subjects is divided into two subgroups, one of which explores R first and N second, whereas the second group explores N and then R. As can be seen from Figure 15, four such pairs of routes have been tested. We accumulate the data from these four tested pairs of returns and novel paths:

$E_{R,1}$ Errors in returns in the returns-first condition
$E_{N,1}$ Errors in novel paths in the novel-first condition
$E_{R,2}$ Errors in returns in the novel-first condition
$E_{N,2}$ Errors in novel paths in the returns-first condition

Thus, $E_{R,1}$ and $E_{N,2}$ refer to the first group of subjects (returns-first condition) and $E_{N,1}$, $E_{R,2}$ to the second. If transfer occurs, the route explored first should have higher error rates in both cases. We define

$$\tau = \frac{E_{R,1} - E_{R,2} + E_{N,1} - E_{N,2}}{E_{R,1} + E_{N,1}} \tag{3}$$

If error rates do not depend on position, τ will be zero; if everything is learned already when exploring the first route, $E_{R,2}$ and $E_{N,2}$ will be zero and τ evaluates to 1. Statistical significance of transfer is tested by comparing the various error rates with the t test.

For this evaluation, the subjects from viewing conditions 1 and 2 were pooled because there were no significant differences between the respective error rates (three-way ANOVA 2 conditions × 4 routes × 2 gender, $F(1, 36) = 0.014$, $p = 0.9075$). If we take the average over all 40 subjects, no significant effect of transfer is found. If, however, only the 20 subjects with the lowest overall error rate are considered, a transfer effect with $\tau = 0.4$ is found (see Figure 5). In this case, 11 subjects were from the returns-first condition and 9 subjects from the novel-first condition. The result indicates that the good navigators show significant transfer learning even from one route to the next. Transfer across more steps of the exploration procedure is not visible in this evaluation, which does not mean that we exclude such a transfer.

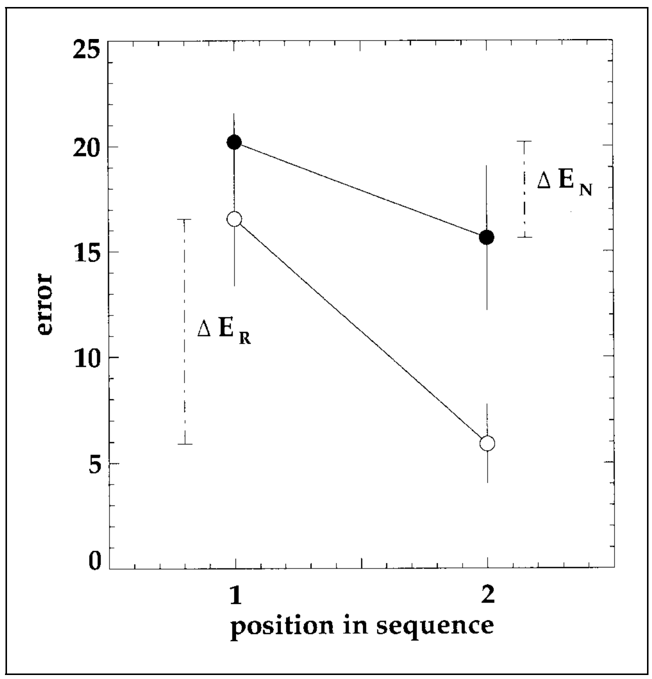


Figure 5. Average error rates for novel paths and returns in the novel-first and returns-first sequence conditions. All subjects from viewing conditions 1 and 2 with an overall error rate below the median were included in this plot. • novel routes; ○ returns. For both returns and novel routes, error rate drops when other routes are explored before. The transfer coefficient (Equation 3) is $\tau = 0.4$.

The transfer-coefficient τ averages the transfer in both directions. In order to look at direction-specific transfer effects, let us consider the two subject groups separately. In Figure 5, this would amount to connecting the left open dot with right filled dot, etc. It turns out that there is a much stronger improvement in the novel-first condition (t test: $t = 6.13$, $FG = 16$, $p < 0.001$) but no improvement in the returns-first condition ($t = 0.19$, $FG = 20$, *n.s.*).

Persistence

An inspection of Figure 2 shows that the subject repeated the first route that led to the goal (trial 2) exactly in the following trial. Similarly, it can be seen from Figure 3 that in almost all cases where the subject started from view 15, the first movement decision was LL even though RR would have been just as good. These and similar observations from other subjects lead to the conjecture that at least some movement decisions reflect simple, fixed associations between the current view and some motion that is performed whenever the view occurs. In order to test this in more detail, we analyzed the return statistics of the decision sequences.

Let $m_{\eta,\upsilon} \in \{LL, LG, LR, RL, RG, RR\}$ denote the movement decision taken at the ηth encounter of view υ (see Figure 12 for possible movement decisions). We are interested in cases where the movement chosen at the ηth encounter of view υ is the same as that taken at the $\eta - 1$th encounter. More generally, we count the cases where movement j is taken at encounter $\eta - 1$ and movement i at encounter η ($i,j \in \{LL, LG, LR, RL, RG, RR\}$):

$$F_{i,j} := \#\{(\eta, \upsilon) \mid m_{\eta,\upsilon} = i,\ m_{\eta-1,\upsilon} = j\}. \qquad (4)$$

It is important to note two points: First, the two encounters η and $\eta - 1$ do not occur in subsequent time steps (unless $m_{\eta-1,\upsilon} \in \{RL, LR\}$). Rather, long sequences of other views may occur in between. Second, the frequency $F_{i,j}$ is accumulated over all views. Thus we are looking for an average persistence rate rather than for a view-specific one.

In the experiments, each search task is repeated until the subject finds the shortest possible path. This procedure can in itself produce repetition rates above chance if parts of the path are created correctly several times. To exclude this type of error, we restrict our analysis to repetitions where both decisions were false in the sense that they did not lead to an approach to the goal (local definition of errors). Finally, we dropped the cases involving the decisions LR and RL because these are quite rare.

Example data from individual subjects are shown in Figure 6. The numbers on the diagonal correspond to cases where the same decision was chosen in two subsequent encounters even though the decision was false in both cases. From these matrices, we can estimate average movement transition probabilities $p_{ij} := P(m_{\eta,\cdot} = i \mid m_{\eta-1,\cdot} = j)$; averaging is performed with respect to the different views involved. A simple statistical model for these transition probabilities is

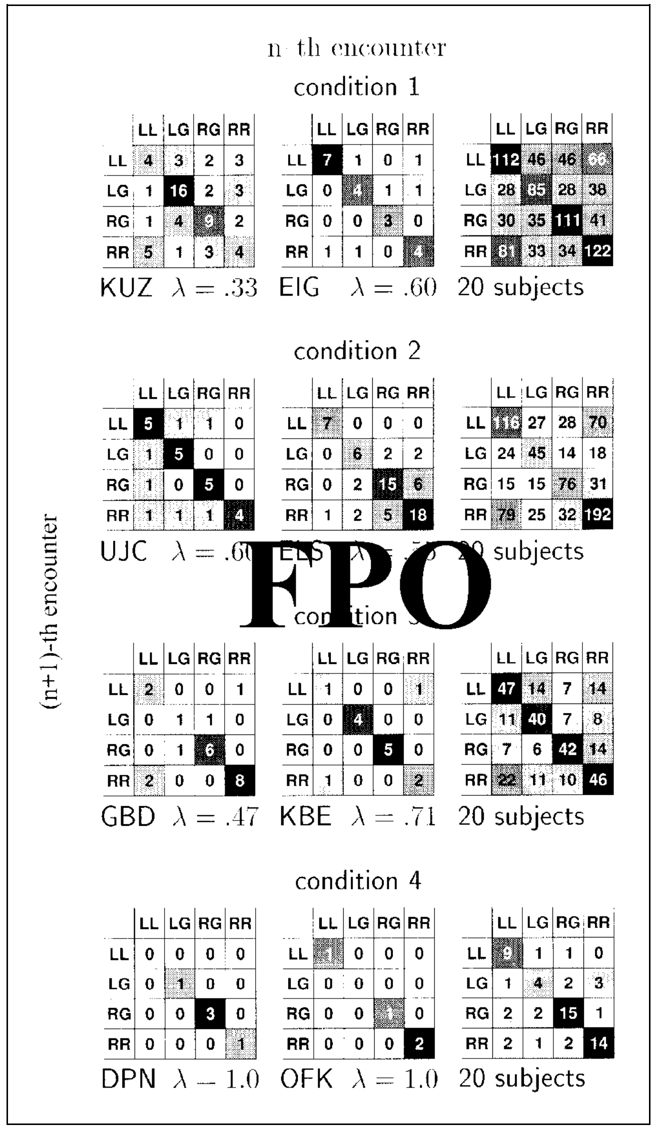


Figure 6. Examples of return statistics for selected subjects for the four viewing conditions. In the subjects shown for the first three conditions, hypothesis 1 could be rejected in all cases (i.e., persistence rate λ was significantly different from zero). For condition 4, where error rates were generally very low, hypothesis 1 could not be rejected for any subject.

$$p_{ij} = \begin{cases} \lambda + p_i & \text{if } \quad i = j \\ p_i & \text{if } \quad i \neq j \end{cases} \qquad (5)$$

where $\lambda, 0 \leq \lambda \leq 1$ and $\lambda + \Sigma_{i=1}^{4}\ p_i = 1$. It states that there is a bias λ for the repetition of the movement chosen at the previous encounter. Other than that, the decisions at subsequent encounters are independent. If $\lambda = 0$, true independence is obtained.

This model was fitted to the data by a maximum likelihood procedure, that is, by minimizing

$$\chi^2 := \sum_{i=1}^{4} \sum_{j=1}^{4} \frac{(F_{ij} - p_{ij}F_{i\cdot})^2}{p_{ij}F_{i\cdot}} \tag{6}$$

where $F_{i\cdot}$ denotes the marginal frequencies $\Sigma_{j=1}^{4} F_{ij}$. If $F_{i\cdot} = 0$ (empty columns in the matrices of Figure 6), the corresponding terms were deleted from the above sum.

The analysis could be applied to data from 67 out of 80 subjects. For the remaining 13 subjects, the number of total errors was too low to fit the model. Ten of these had been tested in viewing condition 4, where the error rates were lowest. Goodness of fit was tested with the χ^2 test; choosing a significance level of 5%, the best-fitting model could not be rejected in any of the 67 subjects. Figure 7 shows the histogram for the best-fitting persisting rates.

In order to get an impression of the confidence intervals for λ, we repeated the analysis with fixed $\lambda = 0$ in Equation 5. By testing goodness of fit for this model with the χ^2 test, 18 cases could be rejected on the 10% level, 9 of which could be rejected on the 1% level as well. The 18 cases are highlighted in Figure 7. Here, persistence rate is significantly different from zero.

Average persistence rate over all subjects was 0.33, indicating that about one-third of the decisions were based on persistence. A regression analysis of persistence rate λ with the overall number of errors for each subject did not reveal a significant correlation.

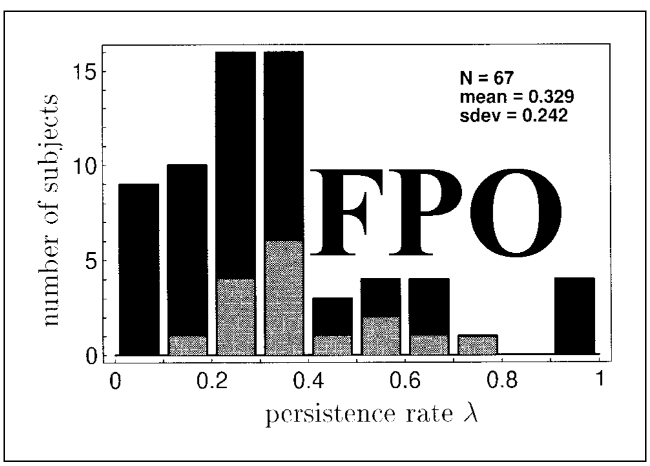


Figure 7. Histogram of best-fitting persistence values λ. Dark columns: Data from $n = 67$ subjects where the analysis could be applied (mean = 0.329, $\sigma = 0.242$). Light columns: Data from $n = 18$ subjects where λ was significantly different from zero. Data are cumulated from viewing conditions 1 through 4.

Judged Distances

Analysis of Variance

Following the exploration phase, subjects were asked to rate the distances between 20 pairs of views (see Table 6) on an ordinal scale from 0 to 4. A $20 \times 4 \times 2$ ANOVA on ranks as dependent variable and view pair, viewing condition, and instruction as independent variables reveals a significant effect of view pair ($F(19, 1368) = 38.7$, $p < 10^{-4}$) but no effect of viewing condition ($F(3, 72) = 1.63, p = 0.19$) or instruction ($F(1, 72) = 1.39, p = 0.24$). In addition, a significant interaction of view pair and viewing condition was found ($F(57, 1368) = 1.74, p = 0.0007$). Thus, the instruction ("distance" versus "airline distance") did not influence the result.

In the following, we discuss the two significant effects separately.

Judged and True Distance

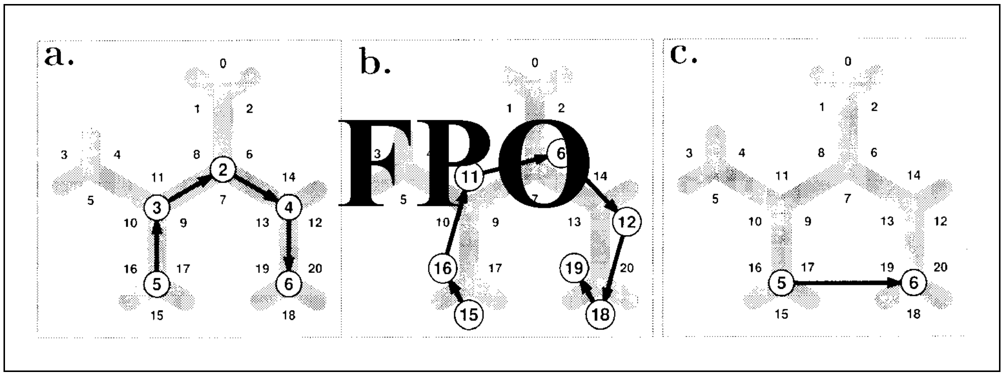
Depending on the type of representation acquired, different distance estimates could be expected (Figure 8):

1. *Walking distance.* This is the length of the minimum path connecting the two views. Because all segments have the same length in our model, it is equivalent to the number of streets traveled or the number of "Go"-decisions taken, that is, to the graph distance in the place graphs (top row of Figure 1).
2. *Decision distance.* This is the minimum number of decisions (mouse clicks) required to travel from one view to the other. It is also the number of views encountered and thus the length of a chain of view-movement associations or the graph distance in the view graph. Using the conventions of Figure 12, we take the unit to be a view-to-view transition (i.e., two subsequent mouse clicks).
3. *Metric or euclidian distance.* Metric distance can be measured in meters in the three-dimensional model underlying the simulation.

Figure 9 shows the average distance ratings from all 80 subjects as functions of each of the three possible distance measures. Judged distance increases with true distance, indicating that subjects have in fact learned some of the distance relations. This dependence of the ratings on the actual distance between the views of the pair accounts at least in part for the effect of view pair found in the ANOVA. However, it is not obvious from the data presented in Figure 9 which of the three theoretical measures is closer to the subject's sense of distance. Correlation with the data is highest for decision distance, whereas standard deviations are smallest for walking distance. One reason for this poor discrimination lies in the fact that all three theoretical measures are closely correlated to each other.

Clearer distinctions between the three theoretical dis-

Figure 8. Possible distance measures. a. Walking distance, b. Decision distance, c. Metric distance. For further explanation see text.



tance measures can be achieved by selecting discriminating view pairs. For example, the decision distance of view pairs $a \rightarrow b$ and $b \rightarrow a$ is often different, whereas walking and metric distances of both directions are of course the same. Table 2 shows the results for the four such cases tested in our experiments. The ratings do not depend on direction (i.e., they do not reflect the decision distance). This result was also obtained when separately comparing the ratings from the different viewing conditions.

For the comparison of metric and walking distance, we pooled the ratings from both directions of each view pair, which were shown to be equal in Table 2. The results from discriminating cases are shown in Table 3. The first two rows compare view pairs with equal metric distance and different walking distance. Here, the ratings differ significantly and are in agreement with walking distance. The next two rows of Table 3 show the reverse case (i.e., equal walking distance but different metric distance). Here again, a significant difference is found, which, however, does not agree with metric distance: pair $5 \leftrightarrow 19$ is rated closer than pair $15 \leftrightarrow 19$, in contradiction to the metric distances. Thus, the difference between these two ratings does not indicate an influence of metric distance. In the last row of the table, a possible alternative explanation is illustrated. Here, a significant difference between two pairs with equal walking and metric distance is found. The pair involving view 15 ("home") is rated further apart. This might indicate a perceptive expansion of the area around view 15, which would also explain the ratings found in rows 3 and 4 of Table 3. However, further experiments are needed to clarify this point.

The effects illustrated in Table 3 do not depend on viewing condition, even though the significances are weaker when analyzing the four groups separately.

Interaction of View Pair and Condition

In the ANOVA including all ratings (Analysis of Variance section), no effect of viewing condition was found, indicating that average ratings were the same in all condi-

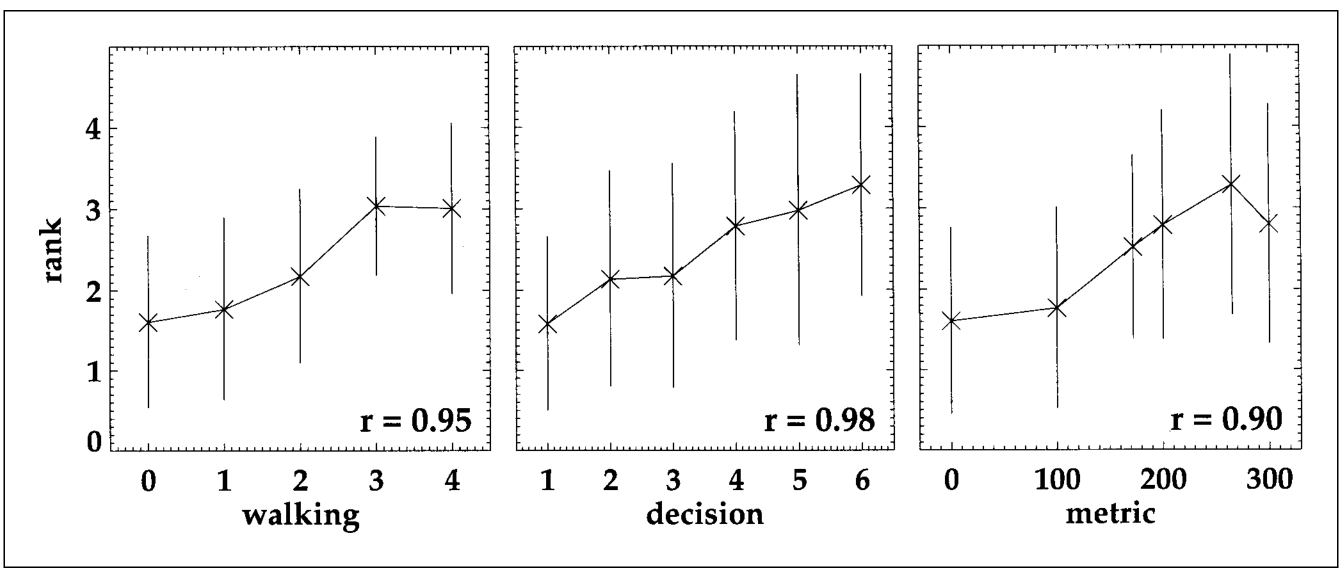


Figure 9. Average distance ratings from all subjects and all viewing conditions (n = 80) plotted as a function of a. walking distance, b. decision distance, and c. metric distance. r: Pearson correlation (ranks treated as numbers). Error bars are standard deviations.

Table 2. Distance rating in view pairs with different decision distance and equal walking and metric distance (different directions on the same path). d_w: walking distance (number of segments). d_d: decision distance (number of double mouse clicks required). p significance of difference as obtained from the Mann-Whitney U-test (n = 80). Errors are standard deviations. No significant differences are found.

Pair	d_w	d_d	*Rating*	p
7→10	1	1	1.48±1.36	0.39
10→7	1	3	1.51±1.09	
5→15	2	2	2.33±1.01	0.31
15→5	2	4	2.43±0.98	
2→15	3	3	3.24±0.72	0.22
15→2	3	5	3.33±0.73	
19→15	4	4	3.14±0.96	0.22
15→19	4	6	3.29±0.83	

tions. More interestingly, however, a significant interaction between viewing condition and view pair could be demonstrated. One possible explanation of this interaction is that in one viewing condition, ratings correlate more strongly with true distance than in another viewing condition. To test this possibility, we calculated Pearson (product moment) correlations individually for each viewing condition (Table 4). The correlation is smallest

Table 3. Distance rating in view pairs discriminating for walking and metric distance. d_w: walking distance (number of segments). d_m: metric distance (meters). p significance of difference as obtained from the Mann-Whitney U-test (n = 160). Errors are standard deviations. For explanation see text.

Pair	d_w	$d_m(m)$	*Rating*	p
2 ↔ 13	2	173	1.96±1.12	$<10^{-5}$
15 ↔ 19	4	173	3.21±0.90	
5 ↔ 15	2	173	2.38±1.00	$<10^{-5}$
15 ↔ 19	4	173	3.21±0.90	
2 ↔ 5	3	200	2.74±0.91	$<10^{-5}$
2 ↔ 15	3	265	3.28±0.72	
15 ↔ 19	4	173	3.21±0.90	0.001
5 ↔ 19	4	300	2.80±1.15	
2 ↔ 13	2	173	1.96±1.12	0.0005
5 ↔ 15	2	173	2.38±1.00	

Table 4. Pearson correlation r of distance ratings with walking distance in the four viewing conditions.

Condition	*1*	*2*	*3*	*4*
Correlation, r	0.99	0.92	0.94	0.89

in condition 4 (open environment) and highest in condition 1 (dark). Additionally, the interaction might be due to condition-dependent rating differences of view pairs with equal true distances; such dependencies have not been found, however (see Judged and True Distance section).

Sketch Maps

As a final part of the experiments, subjects were asked to draw a map of the explored maze. Three subjects refused to draw a map (i.e., 77 maps have been collected). Each row of Figure 10 shows examples from one viewing condition, a good navigator (few errors in exploration phase) on the left side and a poor navigator (many errors in exploration phase) on the right side. In each viewing condition, subjects were ranked according to the number of errors that occurred during the exploration phase. The best navigator is ranked 1, the poorest is ranked 20. The position of view 15 (often labeled "MP," "Institut," or "Schild" by the subjects) is marked with a circle. It has been chosen as the start of the drawing by 74 out of 77 subjects. In Figure 10 all maps have been oriented in roughly the same way.

Good navigators often produce sketch maps that are topologically or even metrically correct and contain identifiable objects. Subject CBK, for example, drew a perfectly correct map except for four missing objects whose locations are included. An equally good map had been drawn by two other subjects, whose maps are not included in the figure.

A frequent deficit of maps are omissions or additions of places. For example, subject LIS drew a good map with the rightmost place missing. Subject GBC, on the other hand, included two nonexisting places in a map with otherwise correct connectivity: one in the lower left and one in the spiral part on the right side. This map (GBC) also shows another interesting feature: The regular Y junctions (120°) are represented as T junctions. This locally feasible assumption leads to global problems such as nonexistent intersections. In her drawing, GBC solved this problem by rolling the right branch to a spiral. T junctions were found in 11 out of the 77 sketch maps. In most maps (43 out of 77) three streets meet at each place. Examples of four- and five-way junctions appear in the map of VOJ.

The number of structurally correct places N_p (i.e., identifiable three-way junctions meeting nonorthogonally) has been determined for each sketch map. The

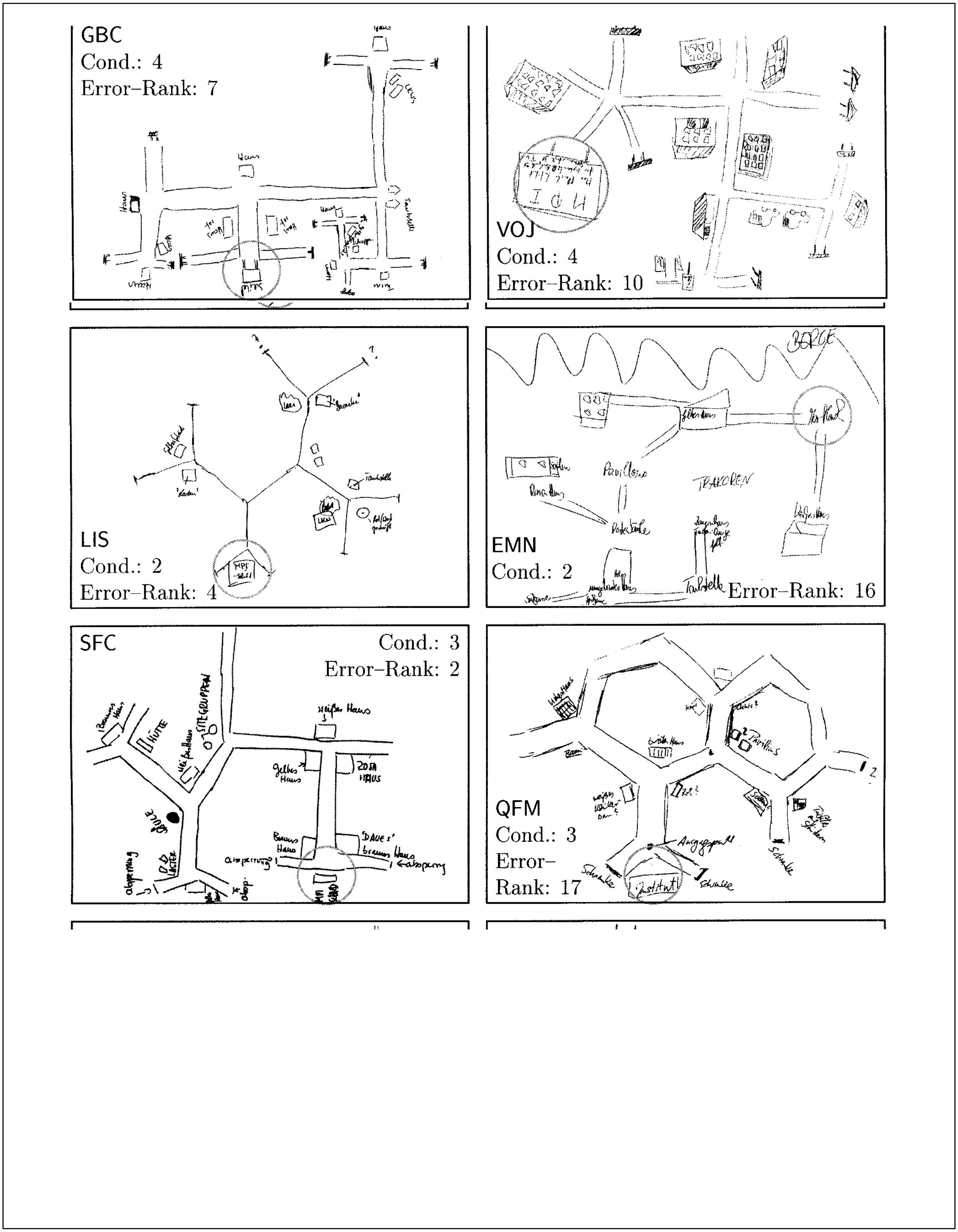


Figure 10. Sample sketch maps from eight subjects. Cond: viewing condition. Error-rank: subject's rank in terms of navigation errors during the exploration phase. Rank 1 indicates lowest number of errors in the respective condition group, and rank 20 indicates highest error number. For further explanation see text.

"place error," $|N_p - 7|$, is taken as a measure of map quality. It correlates moderately with the navigation errors in the exploration phase (Table 5), indicating that good navigators tend to include the correct number of places in their sketch maps. An ANOVA over the viewing conditions revealed a significant dependence of N_p on condition ($F(3, 76) = 3.18$, $p = 0.029$). As can be seen from Table 5, this relation is not monotonic. Place error is high in conditions 2 and 4 and low in conditions 1 and 3.

Figure 10 also illustrates two kinds of global errors. Subject SFC drew a map with largely correct connectivity that is basically a mirror image of the correct map. The lower left place (with two trucks, identified by the word "Laster"), however, has been replaced from its correct position at the upper end of the drawing without any changes to its local structure. Altogether, three mirror-inverted maps were drawn. Another example of global errors is the map of subject QFM, who invented closed hexagonal loops.

Most maps distinguish places and objects (67 out of 77). In the remaining cases, each junction or place is identified by just one object, resulting in a structure resembling a view graph (EMN). Subject TER has a reduced number of connections, so that the map consists mainly of isolated objects.

DISCUSSION

Navigation in Virtual Environments

The results presented in this paper clearly show that spatial relations can be learned from exploration in a virtual environment even under rather restricted viewing conditions. Here, we briefly summarize the most important findings:

Effect of Viewing Condition

The four viewing conditions differ in the amount of information available to the subjects. Not surprisingly, the number of errors during the search phase decreases as more information is provided. This is in spite of the fact that the field of view was the same in all four conditions. Bringing the objects closer to the places in condition 3 and removing the occluding hedges in condition 4 are reminiscent of zooming out the whole scene with a wide-field lens. May et al. (1997) showed that this zooming does not improve path-integration performance in a triangle-completion task. This discrepancy may characterize a difference between path integration and landmark navigation. Alternatively, it may be due to the marked errors in perspective associated with zooming.

Table 5. Average number of structurally correct places N_p in the sketch maps and the correlation of $|N_p - 7|$ with the number of errors made in the exploration phase. r coefficient of correlation. $n = 20$.

Condition	$N_p \pm \sigma$	r
1	6.05±3.70	0.61
2	3.10±3.24	0.48
3	5.75±3.21	0.47
4	3.95±3.45	0.35

The comparison between conditions 1 and 2 (night and day) does not show an improvement in error rates. This is surprising because more information is available in condition 2 (objects at the far end of the streets become visible). This finding may be related to the fact that the local structure of the maze becomes more complicated in condition 2, where six objects are visible from each place.

The correlation of distance estimates with walking distance in the maze decreases from viewing condition 1 to viewing condition 4 (Interaction of View Pair and Condition section). We take this as evidence that less information is stored in the open environment where navigation need not rely on memory as strongly as in the other conditions. This interpretation is also in line with the observation that sketch maps from condition 4 are not better than those from condition 1. In conclusion, it appears that the amount of knowledge acquired is determined not by its availability but by the different needs in the four conditions.

Irrespective of this difference of correlation between the viewing conditions, the analysis of discriminating view pairs shows that walking distance is the theoretical distance measure closest to the subject's ratings. This may indicate that the structure of the representation acquired from all four conditions is the same.

Transfer and Latent Learning

The overall number of errors was smaller for the later search tasks. For the 50% best subjects, this effect was already clearly visible for the comparison of one search task with the next (Figure 5). If subjects simply learned a set of independent routes (e.g., by reinforcement learning), each search would be a new task and no such transfer would be expected. The knowledge being transferred from one route to the next is not just a route memory but involves the recombination of route segments; this is to say, it is of the configuration type. Its acquisition is akin to latent learning, because knowledge obtained during one search can be employed later in other, unrelated search tasks.

As can be seen from Figure 5, transfer was strong from the novel to the return paths but not the other way around. One possible explanation of this finding is that the novel paths are more difficult than the returns. When considering the shortest possible paths, the novel paths involve 14 different views, 8 of which also occur in the returns. The returns involve only 9 different views (i.e., almost all of their views are already known from the novel paths). The only view not occurring in the novel paths is the final goal of the returns, view 15. The transfer asymmetry may thus be due to the fact that the novel routes contained more information about the returns than vice versa.

The occurrence of transfer from one route to another is also evidence for the presence of goal-independent memory of space (i.e., a cognitive map).

Persistence

Along with these arguments for configuration knowledge, evidence for simpler types of spatial learning was also found. The persistence rates presented in Figure 7 indicate that at least some of the subjects based a considerable part of their movement decisions on simple associations of views with movements. This strategy is efficient for learning nonintersecting routes but will lead to errors for views at a crossroads where the correct motion decision depends on the current goal. We speculate that the persistence rate will decrease if longer training sequences are used.

Distance Measures

Judged distance agrees best with the walking distance in the maze, rather than with metric or decision distance. This result is based on the rather small number of "discriminating view pairs" where the theoretical measures disagree. It holds irrespective of the instruction ("distance" versus "airline distance"). The analysis of distance ratings presented here rests on the assumption that distances in different parts of the maze can be compared. As was discussed with respect to Table 3, this need not be the case. Rather, the perceived distances in the vicinity of well-known places (home) might be increased.

Subject Differences

Subjects differed strongly in terms of the number of errors made when searching a goal as well as in the quality of their distance estimates. However, no clustering in different groups can be obtained from our data. In particular, no significant gender differences were found.

View-Based Navigation

In viewing conditions 1 and 2, subjects had to rely on local views as their only position information. Their performance and the transfer learning is therefore view-based in an obvious sense. However, this result does not exclude the possibility that some more complicated representation of space is constructed from the local view information. Here we summarize the evidence against such a representation (i.e., evidence for a view-based mechanism of navigation).

Returns Aren't Easy

After having learned the four excursions, the returns to the starting point along the very same paths are almost as difficult as novel paths (Figure 4). The advantage on the order of just one error per search task is not significant ($F(1, 76) = 2.860$, $p = 0.095$). If the subjects acquired a place-based representation of space, it would be the same for excursions and returns, because the corresponding place graphs are symmetric (see Figure 1). In this case, we would therefore expect that returns should be much easier and more reliable than novel paths. The weak difference between the number of errors occurring in returns and novel paths seems to indicate that this is not the case. It is rather more in line with a view-based mechanism. because the views occurring along the return path are as different from the original views as are any other views in the maze.

Recognition and Action

The average persistence rate of 32.9% (Figure 7) indicates that direct associations of views to movement decisions can be learned. As was pointed out in the introduction, the association pair of view and motion decision is the basic element of a view-based memory of space.

Local Information Combined to a Graph?

If the representation is in fact view-based, a graph structure is the only representation we can think of that would account for the transfer and planning behavior observed. Independent evidence for a graphlike representation comes mainly from the sketch maps: as was point out in the Sketch Maps section, maps are often locally correct but globally inconsistent. Also, places with correct local connectivity have been translocated to erroneous positions. Connectivity can be correct even though metric properties of the sketch maps, such as angles and lengths, are grossly mistaken.

The distance estimates do not reflect the decision distance, which is the graph distance of the view graph, but correlate better with walking distance (i.e., the graph distance of the place graph). It therefore appears that we cannot decide between the view- and place-graph representations at this point. In ongoing work (Mallot & Gillner, 1997), this question is addressed with additional experiments.

CONCLUSION

In our view, the most important result of this study is the fact that configuration knowledge can be acquired in virtual environments. This is in spite of the fact that the subjects did not actually move but were interacting with a computer graphics simulation. With respect to the high controllability of visual input, this result may well make virtual reality a valuable addition to more realistic field studies, where stimulus control is often a problem. As a novel experimental opportunity, we are presently using object transpositions after learning, which could be done in real environment only with great difficulties (Gillner & Mallot, 1996).

With respect to our starting point (i.e., view-based navigation), we think that three conclusions can be drawn:

1. *Views suffice.* Map learning is possible if only local (i.e., view information) is provided. In this sense, navigation can be view-based.
2. *Graph versus view from above.* The representation contains local elements (i.e., a place or view with one or several movement decisions and the respective outcome associated with it). These local elements need not be globally consistent, and they need not combine into a metric survey map. Rather, a graphlike representation is sufficient to account for our results.
3. *Places versus views.* It is not clear from our data whether the nodes of this graph are places or views. We have not found evidence that the local views are combined into a representation of space independent of the orientation of the viewer. So, a view-based representation seems more likely at this point.

METHODS

The Virtual Maze (Hexatown)

The virtual town was constructed using Medit software and animated with a frame rate of 36 Hz on a SGI Onyx RealityEngine^2 using IRIX Performer software. A schematic map of the town appears in Figure 11. It is built on a hexagonal raster with a distance between two places of 100 m. At each junction, one object, normally a building, was located in each of the 120° angles between the streets; so each place consisted of three objects. In the places with less than three incoming streets, dead ends were added instead, ending with a barrier at about 50 m. The hexagonal layout was chosen to make all junctions look alike. In contrast, in Cartesian grids (city-block raster), long corridors are visible at all times and the possible decisions at a junction are highly unequal: going straight to a visible target or turning to something not presently visible. The whole town was surrounded by a distant circular mountain ridge that showed no salient features. The mountains were constructed from a small model that was repeated periodically every 20°.

Subjects could move about the town using a computer mouse. In order to have controlled visual input and not to distract subject's attention too much, movements were restricted in the following way. Subjects could move along the street on an invisible rail right in the middle of each street. This movement was initiated by hitting the middle mouse button and was then carried out with a predefined velocity profile without further possibilities for the subject to interact. The translation took 8.4 sec with a fast acceleration to the maximum speed of 17 m/sec and a slow deceleration. The movement ended at the next junction, in front of the object facing the incoming street. Sixty-degree turns could be performed similarly by pressing the left or right mouse button. Again, the simulated movement was "ballistic" (i.e., following a predefined velocity profile). Turns took 1.7 sec with a maximum speed of 70°/sec and symmetric acceleration and deceleration.

Figure 12 shows the movement decisions that subjects could choose from. Each transition between two views is mediated by two movement decisions. When facing an object (e.g., the one marked "a" in Figure 12), 60° turns left or right (marked "L", "R") can be performed; they will lead to a view down a street. If this is not a dead end, three decisions are possible: the middle mouse button triggers a translation down the street (marked "G" for go), whereas the left and right buttons lead to 60° turns. If the street is a dead end, turns are the only possible decision. In any case, the second movement will end in front of another object.

An aerial view of Hexatown is shown in Figure 13. It gives an impression of the objects used. The spacing and position of the trees correspond to viewing conditions 1 and 2 (see below).

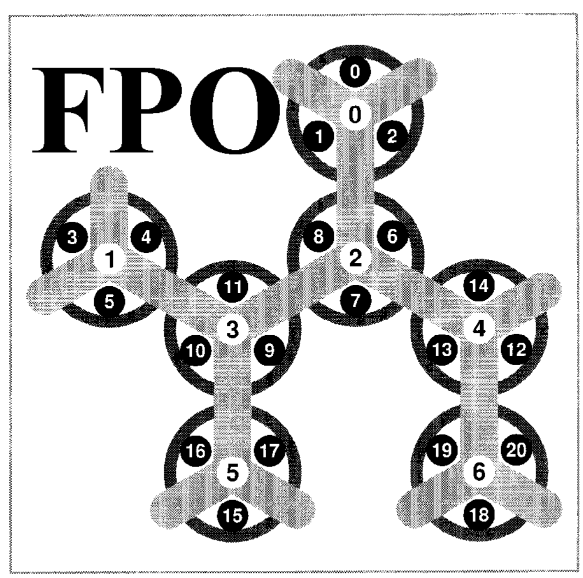


Figure 11. Street map of the virtual maze with 7 places numbered 0 through 6 and 21 views numbered 0 through 20. The ring around each place indicates the hedges used in viewing conditions 1 through 3 to occlude distant places.

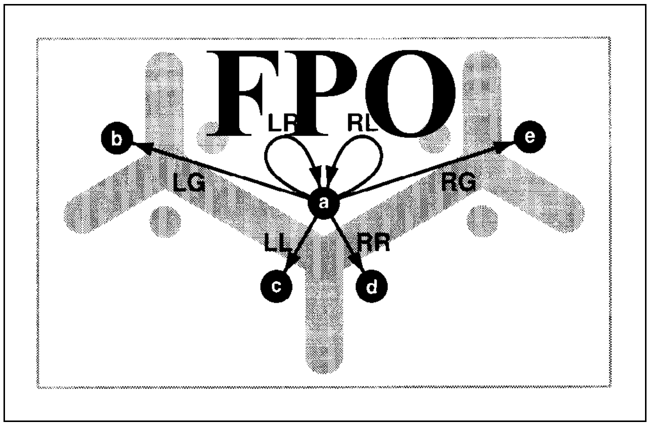


Figure 12. Possible movement decisions when facing the view marked a. L: turn left 60°. R: turn right 60°. G: go ahead to next place.

We used four stimulus conditions with varying degrees of visibility of the environment. In conditions 1, 2, and 3, a circular hedge or row of trees was placed around each junction with an opening for each of the three streets (or dead ends) connected to that junction. This hedge looked the same for all junctions and prevented subjects from seeing the objects at more distant junctions. In conditions 1, 2, and 4, the objects were placed 22 m away from the center of the junction. The arrangement was such that when entering the circular hedge in conditions 1 and 2, the buildings to the left and right were already outside the observer's field of view (60°). Thus, the three buildings at one junction could never be seen together in these conditions. In condition 3, the buildings were at a distance of only 15 m from the junction. In this case, all three buildings were seen at once when entering the place. The town was illuminated from the bright sky, except in condition 1, where an exploration at night was simulated. Here, illumination was as with a torch or the headlights of a car and reached about 60 m. Thus, the building at the far end of a street was not visible in this condition. A summary of the viewing conditions and variants of Hexatown is given in Figure 14.

Procedure

Experiments were performed using a standard 19-in SGI monitor. Subjects were seated comfortably in front of the screen and no chin rest was used. They moved their heads in a range of about 40 to 60 cm in front of the screen, which results in a viewing angle of about 35 to 50°.

The experiment was performed in three phases. In the *exploration phase,* subjects found themselves facing some view v_1. They were then presented with a target

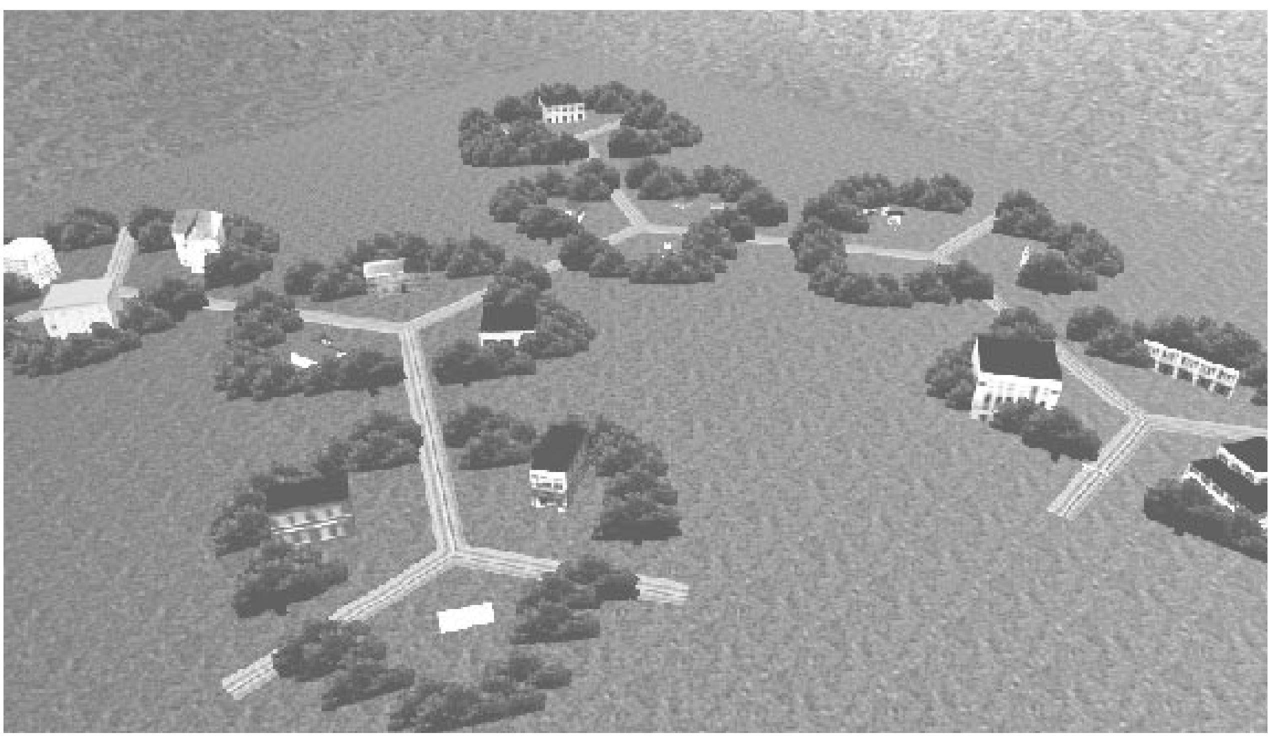


Figure 13. Aerial view of Hexatown. Orientation as in Figure 11. The white rectangle in the left foreground is view 15, used as "home" position in our experiments. The aerial view was not available to the subjects. Object models are courtesy of Silicon Graphics, Inc., and Prof. F. Leberl, Graz.

view v_2 printed out on a sheet of paper and asked to find this view in the virtual town (task $v_1 \rightarrow v_2$). When they found the view, feedback was given in the form of a little sign appearing on the screen. If they got lost in the maze (i.e., if they deviated from the shortest possible path by more than one segment), the trial was stopped and another sign announced the failure. This feature of the procedure was included in order not to discourage subjects by long unsuccessful searches. In pilot experiments with free exploration (no goals specified), learning was much harder. The number of terminations varied from 9.9 times per subject in viewing condition 1 to 1.5 in viewing condition 4; terminations occurred almost exclusively during the excursion phase (see below). If a trial was terminated, or if the way found was not the shortest possible way, the subject was relocated to the starting point and a new trial for the same goal started. The sequence was terminated when the shortest possible way, that is, the way involving the minimal number of decisions (mouse clicks), was found. The whole exploration phase contained 12 such search tasks, or ways to be found. The first four ways were excursions from view 15, which served as a "home" position. View 15 showed a poster wall saying "Max-Planck-Institut für biologische Kybernetik." The following eight searches were either returns to home or novel paths not touching on view 15. The return and novel path tasks were presented alternatingly in two sequence conditions: in the *returns-first condition,* the first task was a return, whereas in the *novel-first condition,* the sequence started with a novel path. In both conditions, the four excursions were performed prior to both returns and novel paths (Figure 15).

In the *distance estimation* phase, following immediately after the exploration, subjects were presented with pairs of views on the screen. The first view was shown without time limit; after hitting the spacebar, the second view was presented again without time limit. Subjects were asked to rate the distance between the two views on an integer scale ranging from 0 to 4, where 0 meant "very close" and 4 "very far apart." One of two instructions was used: In the first, subjects were asked to rate according to distance ("Abstand"); in the second, they were told to estimate the airline (metric) distance ("Luftlinie"). After hitting the appropriate number button, the next view pair was presented. Ranking data from a total of 20 view pairs was collected from each subject (Table 6).

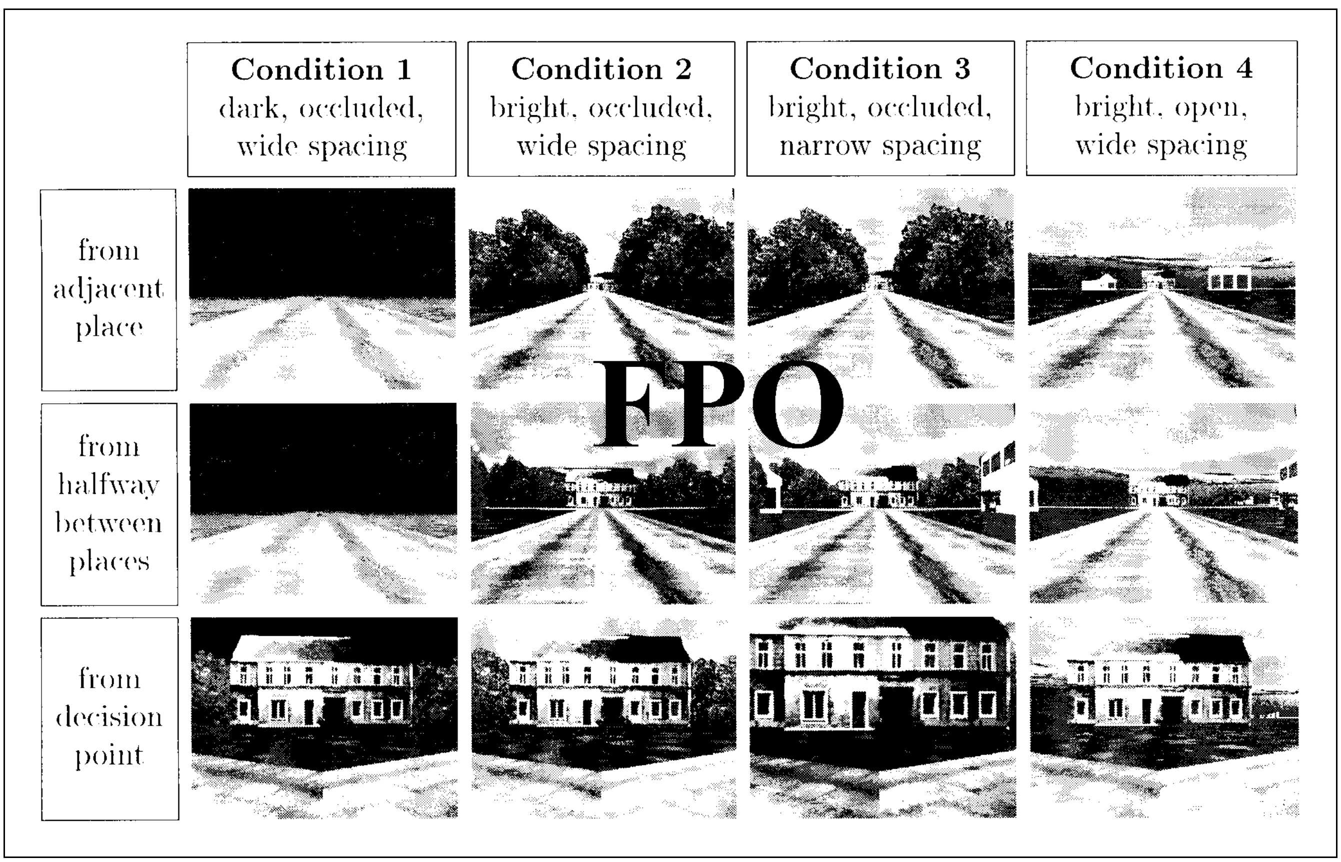


Figure 14. Viewing conditions used in the experiments. For each condition, three views ar shown. The views in the top row occur when looking from a place (no. 5) into a street. The views in the middle row can be seen during the motion along a street, in this example from place 5 to place 3. The views in the bottom row show an object (view 11) as seen from the corresponding junction (place 3). Object models are courtesy of Silicon Graphics, Inc., and Prof. F. Leberl, Graz.

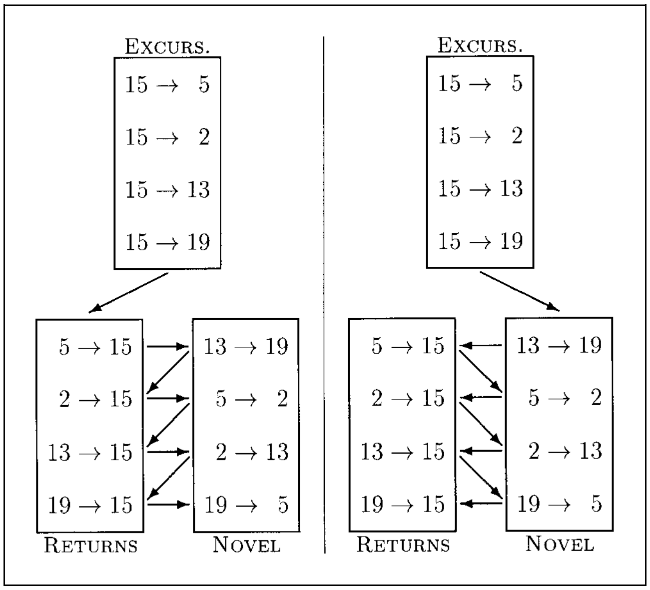


Figure 15. Sequence of search tasks identified by the numbers of start and goal view. Left: Returns-first condition. Right: Novel-first condition.

Table 6. View pairs tested in the distance estimation phase. d_w, d_d, and d_m are theoretical distance measures; see Judged and True Distance section.

Pair	d_w	d_d	d_m
6 → 8	0	2	0
8 → 6	0	2	0
10 → 11	0	2	0
11 → 10	0	2	0
7 → 10	1	2	100
10 → 7	1	6	100
13 → 19	1	4	100
19 → 13	1	4	100
2 → 13	2	6	173
13 → 2	2	6	173
5 → 15	2	4	173
15 → 5	2	8	173
2 → 5	3	8	200
5 → 2	3	8	200
2 → 15	3	6	265
15 → 2	3	10	265
15 → 19	4	12	173
19 → 15	4	8	173
5 → 19	4	10	300
19 → 5	4	10	300

In the *sketching* phase, subjects were asked to draw a sketch map of the virtual town on a sheet of paper.

The experiment was run on 80 paid volunteers, 40 male and 40 female, aged 15 to 38. Twenty subjects (10 male, 10 female) took part in each of the four viewing conditions (Figure 14). Within each viewing condition, the group of subjects was split equally (10 to 10) between the two sequence conditions (returns first and novel first, Figure 15), as well as the two instructions for the distance estimation ("distance" and "airline distance").

Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft, Schwerpunktprogramm Raumkognition. We are grateful to Silicon Graphics Inc. and Prof. F. Leberl (University of Graz) for providing VR-models used in the experiments. We are also grateful to Tordis Philipps for help with the experiments and to Heinrich H. Bülthoff, Bernhard Schölkopf, and Hendrik-Jan van Veen for comments on prior versions of the manuscript. The writing of the final text was supported by a fellowship from the Institute of Advanced Study Berlin to HAM.

Reprint requests should be sent to H. A. Mallot, Max-Planck-Institut für biologische Kybernetik, Spemannstr. 38, D-72076 Tübingen, Germany, or via e-mail: hanspeter.mallot@tuebingen.mpg.de.

Notes

1. Bennett (1996) has driven the minimalistic view to the point of denying the very existence of declarative memory of spatial relations. Although this may be true for insects (Wehner & Menzel, 1990), we argue that such memory does exist in humans. The view-graph model is a simple model of declarative memory of spatial configurations that builds on features of insect spatial memory.
2. Additional material on Hexatown is available from http://www.kyb.tuebingen.mpg.de/links/hexatown.html.

REFERENCES

Aginsky, V., Harris, C., Rensink, R., & Beusmans, J. (1996). Two strategies for learning a route in a driving simulator. Technical Report CBR TR 96-6, Cambridge Basic Research, 4 Cambridge Center, Cambridge, Massachusetts 02142 U.S.A.

Barto, A. G., & Sutton, R. S. (1981). Landmark learning: An illustration of associative search. *Biological Cybernetics, 42,* 1–8.

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

Bennett, A. T. D. (1996). Do animals have cognitive maps? *The Journal of Experimental Biology, 199,* 219–224.

Berthoz, A., Israël, I., Georges-François, P., Grasso, R., & Tsuzuku, T. (1995). Spatial memory of body linear displacement: What is being stored? *Science, 269,* 95–98.

Billinghurst, M., & Weghorst, S. J. (1995). The use of sketch maps to measure cognitive maps of virtual environments. *Proc. IEEE 1995 Virtual Reality Annual International Symposium,* 40–47.

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are

three-dimensional objects represented in the brain? *Cerebral Cortex, 5,* 247–260.
Cartwright, B. A., & Collett, T. S. (1982). How honey bees use landmarks to guide their return to a food source. *Nature, 295,* 560–564.
Collett, T. S., & Baron, J. (1995). Learnt sensori-motor mappings in honeybees: Interpolation and its possible relevance to navigation. *Journal of Comparative Physiology A, 177,* 287–298.
Darken, R. P., & Sibert, J. L. (1996). Navigating large virtual spaces. *International Journal of Human-Computer Interaction, 8,* 49–71.
Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H. (1998). Learning view graphs for robot navigation. *Autonomous Robots, 5,* 111–125.
Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.
Gillner, S., & Mallot, H. A. (1996). Place-based versus view-based navigation: Experiments in changing virtual environments. *Perception, 25,* 93.
Kuipers, B. J., & Byun, Y.-T. (1991). A robust exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems, 8,* 47–63.
Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122,* 73–91.
Mallot, H. A., & Gillner, S. (1997). Psychophysical support for a view-based strategy in navigation. *Investigative Ophthalmology and Visual Science, 38,* S1007.
Maurer, R., & Séguinot, V. (1995). What is modelling for? A critical review of the models of path integration. *Journal of Theoretical Biology, 175,* 457–475.
May, M., Péruch, P., & Savoyant, A. (1995). Navigating in a virtual environment with map-acquired knowledge: Encoding and alignment effects. *Ecological Psychology, 7,* 21–36.
May, M., Wartenberg, F., & Péruch, P. (1997). Raumorientierung in virtuellen Umgebungen. [Spatial orientation in virtual environments]. In R. H. Kluwe, (Ed.), *Kognitionswissenschaft: Strukturen und prozesse intelligenter systeme.* Wiesbaden: Deutscher Universitätsverlag.
McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., Kudrimoti, H., Qin, Y., Skaggs, W. E., Suster, M., & Weaver, K. L. (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *The Journal of Experimental Biology, 199,* 173–185.
McNaughton, B. L., & Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences, 10,* 408–415.
Mittelstaedt, H., & Mittelstaedt, M.-L. (1972/73). Mechanismen der orientierung ohne richtende außenreize [Mechanisms of orientation without orienting external stimuli]. *Fortschritte der Zoologie, 21, 46–58.*
Muller, R. U., Stead, M., & Pach, J. (1996). The hippocampus as a cognitive graph. *Journal of General Physiology, 107,* 663–694.
O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon.
Prescott, T. (1996). Spatial representation for navigation in animals. *Adaptive Behavior, 4,* 85–123.
Schölkopf, B., & Mallot, H. A. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior, 3,* 311–348.
Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.
Tlauka, M., & Wilson, P. N. (1996). Orientation-free representations from navigation through a computer-simulated environment. *Environment and Behavior, 28,* 647–664.
Tolman, E. C. (1948). Cognitive maps in rats and man. *Psychological Review, 55,* 189–208.
Tong, F. H., Marlin, S. G., & Frost, B. J. (1995). Visual-motor integration and spatial representation in a visual virtual environment. *Investigative Ophthalmology and Visual Science, 36,* S1679.
Touretzky, D. S., & Redish, A. D. (1996). Theory of rodent navigation based on interacting representations of space. *Hippocampus, 6,* 247–270.
Touretzky, D. S., Redish, A. D., & Wan, H. S. (1993). Neural representation of space using sinusoidal arrays. *Neural Computation, 5,* 869–884.
Wehner, R., & Menzel, R. (1990). Do insects have cognitive maps? *Annual Review of Neuroscience, 13,* 403–414.