

Multiple views of spatial memory

AMY L. SHELTON and TIMOTHY P. McNAMARA
Vanderbilt University, Nashville, Tennessee

Recent evidence indicates that mental representations of large (i.e., navigable) spaces are viewpoint dependent when observers are restricted to a single view. The purpose of the present study was to determine whether two views of a space would produce a single viewpoint-independent representation or two viewpoint-dependent representations. Participants learned the locations of objects in a room from two viewpoints and then made judgments of relative direction from imagined headings either aligned or misaligned with the studied views. The results indicated that mental representations of large spaces were viewpoint dependent, and that two views of a spatial layout appeared to produce two viewpoint-dependent representations in memory. Imagined headings aligned with the study views were more accessible than were novel headings in terms of both speed and accuracy of pointing judgments.

When people give directions or decide what route to take from one place to another, they often rely on memories of the locations of objects in the environment. To understand how these and other spatial problems are solved, we must understand how information about location is mentally represented. More specifically, we need to understand what frames of reference are used to encode location. Just as parallels of latitude and meridians of longitude provide a language for specifying locations on the earth, we need to discern what language is used in memory to specify where objects are located in the environment.

Memory for location could be encoded with respect to frames of reference that depend on the observer's point of view (e.g., retinal coordinates) or frames of reference that do not depend on the observer's point of view (e.g., scene-centered coordinates). In a viewpoint-dependent representation, familiar views should be more accessible than novel views, whereas in a viewpoint-independent representation, familiar and novel views should be equally accessible (Tarr, 1995).

An effective method of measuring the accessibility of spatial memories is to have observers make judgments of relative direction on the basis of their memories of a spatial layout (see, e.g., Levine, Jankovic, & Palij, 1982). For example, examine Figure 1 (with "viewing position a" at the bottom), and then make the following judgments of relative direction without looking at the figure: (1) Point to the shoe as if standing at the jar, facing the clock. (2) Point to the clock as if standing at the jar, facing the

book. The imagined heading in the first judgment is aligned with the viewing perspective, and hence is "familiar," whereas the imagined heading in the second judgment is not aligned with the viewing perspective, and hence it is "novel."¹ Judgments of the first kind are typically easier than judgments of the second kind (e.g., Levine et al., 1982; Presson & Hazelrigg, 1984). The difference in difficulty indicates that the underlying mental representation is viewpoint dependent. If the mental representation of Figure 1 were viewpoint independent, both of these judgments would be equally easy; the familiar perspective (1) and the novel perspective (2) would be equally accessible.

The viewpoint dependence of spatial memories has been well established for small spaces such as maps (e.g., Evans & Pezdek, 1980; Levine et al., 1982; Presson & Hazelrigg, 1984; Thorndyke & Hayes-Roth, 1982), and more recently for large spaces when observers are limited to a single view of the space (e.g., Presson & Montello, 1994; Rieser, 1989; Roskos-Ewoldsen, McNamara, Shelton, & Carr, in press). However, there are situations in which memories of large spaces seem to be viewpoint independent (e.g., Evans & Pezdek, 1980; Presson, DeLange, & Hazelrigg, 1989; Thorndyke & Hayes-Roth, 1982). The findings of such studies raise questions about when spatial memories appear to be viewpoint dependent and when they appear to be viewpoint independent.

Similar issues have arisen in investigations of visual object recognition. Recent studies indicate that mental representations of objects may consist of a collection of viewpoint-dependent representations (e.g., Edelman & Bülthoff, 1992; Tarr, 1995; Tarr & Pinker, 1989; but see Biederman & Gerhardstein, 1995). Each representation corresponds to a familiar view of the object. Judgments such as recognition and handedness are accomplished by matching the visual stimulus with the representations in memory. If there exists a viewpoint-dependent representation that corresponds to the orientation of the stimulus, recognition is fast and accurate. If no such representation exists, a normalization procedure ensues to match the

The research reported in this paper was supported by National Science Foundation Grant SBR-9222002. We are grateful to Ted Payne for writing the experimental program, to Jeffery Franks for the use of his laboratory space, and to Kyle Cave and Jeffrey Schall for their comments on the research. We also thank Nancy Franklin, Arthur Glenberg, and an anonymous reviewer for their comments and suggestions on an earlier draft. Correspondence should be addressed to T. P. McNamara, Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, TN 37240 (e-mail: mcnamara@ctrvax.vanderbilt.edu).

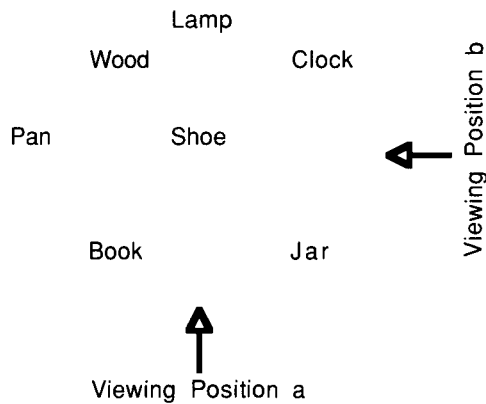


Figure 1. One of the spatial layouts used in the experiment.

stimulus to the representation that is most similar in orientation; this normalization procedure leads to longer recognition times and more errors. The normalization process has been characterized as mental rotation of the unfamiliar view to the familiar view (e.g., Tarr, 1995; Tarr & Pinker, 1989). Highly familiar objects can be recognized equally well from nearly any perspective because they have been experienced from nearly every perspective.

We hypothesized that similar processes may occur in spatial memory. Small spaces or maps are generally learned from a small number of views, usually just one. Hence, they would be mentally represented in a single viewpoint-dependent representation. Large (i.e., navigable) spaces, on the other hand, are generally learned by exploration and from many different perspectives. Memories of large spaces may appear to be viewpoint independent because of the presence in memory of multiple views (e.g., Evans & Pezdek, 1980). Retrieving and using information about spatial relations relative to a novel view may require minimal transformation of an existing view in memory.

The present study is a first step in the attempt to unravel the role of multiple views in spatial memory. Observers learned the locations of objects in a room from two views and subsequently made judgments of relative direction from several imagined headings. We hypothesized that two views would produce two viewpoint-dependent representations in memory. Imagined headings corresponding to the learned views should therefore show an advantage over all novel imagined headings.

Although several previous investigations have allowed the participants to experience a space from more than one orientation (e.g., Hintzman, O'Dell, & Arndt, 1981; Presson, DeLange, & Hazelrigg, 1987), various features of these studies have precluded a clear answer to the question posed in this paper. To our knowledge, the present experiment is the first in which observers learned a large (i.e., navigable) spatial layout by sight, but were limited to a small number of well-specified views.

METHOD

Subjects

Twenty-four students (16 undergraduate, 8 graduate) consented to participate in return for extra credit and/or research experience. Half the subjects were male and half were female.

Materials and Design

Two configurations of seven objects each were constructed. The objects were common, visually distinct, had monosyllabic names, and shared no primary semantic associations (e.g., a book, a lamp, etc.). Objects were placed on a 3×3 m clear plastic sheet in one half of a large room. Two viewing positions differing by 90° were selected and kept constant across the two configurations. Orthogonal views were chosen to maximize the unique information afforded by each view.

Each test trial was constructed from three objects in the display and required subjects to point to an object as if standing in a particular position within the display; for example, "Imagine you are standing at the jar and facing the shoe. Point to the clock." Two of the objects established the imagined station point and heading (e.g., jar and shoe), and the third object was the target (e.g., clock). Although these judgments may involve imagined translations (e.g., from the actual viewing position to the imagined station point), the effects of these translations would be very small compared with the effects of imagined rotation (e.g., Easton & Sholl, 1995; Rieser, 1989) and were not examined.

The principal independent variable was imagined heading. Eight equally spaced headings were used. To facilitate exposition, headings were arbitrarily labeled counterclockwise from 0° to 315° in 45° steps beginning with viewing position a (Figure 1). For example, 0° corresponds to all headings aligned with viewing position a (e.g., book-wood; jar-clock); 90° corresponds to all headings aligned with viewing position b (e.g., jar-book; clock-wood); and 270° corresponds to all headings contra-aligned with viewing position b (e.g., book-jar; wood-clock). Each configuration yielded 24 object-pair headings, three instances of each of the eight imagined headings. Each object-pair heading was used with three different targets for 72 trials, 9 at each imagined heading.

Pointing direction (the direction of the target object relative to the heading) was varied systematically by dividing the space into six areas, front-right (0° - 60°), right (60° - 120°), back-right (120° - 180°), and so on. Subjects experienced approximately equal instances of each pointing direction with at least one instance of each direction at every imagined heading.

Procedure

All subjects learned and were subsequently tested on one of the two configurations. The factorial combination of configuration and order of viewing (a-b vs. b-a) produced four groups of subjects. Subjects were randomly assigned to groups with the constraint that each group contain an equal number of males and females.

Learning phase. Each subject was instructed to learn the objects and their locations for a spatial memory test, and then was blindfolded and led to a viewing position (viewing position a or b in Figure 1). The subject viewed the display for 30 sec, and then was required to close his or her eyes and to point to and name each object. This study-test sequence was repeated until subjects could point to and name all the objects correctly two times in a row. Subjects were then blindfolded and walked directly to the second viewing position. The study-test sequence was repeated at the second viewing position. Before moving from a particular viewing position, subjects were offered additional viewing time, and most chose to view the display for 5-10 additional seconds at each position. All subjects were highly familiar with the locations of the objects at the end of the learning phase.

Test phase. After learning the spatial layout, subjects were taken to another room to be tested. The test trials were presented on a Macintosh computer. Subjects first received instructions on using the program and four practice trials involving locations on the campus. Each test trial was initiated by the subject and proceeded as follows: The heading was printed at the top of the screen (e.g., "Imagine you are standing at the

jar and facing the shoe.”). The subject clicked the mouse to indicate when he/she had adopted the heading. The name of the target object was then displayed (e.g., “Point to the clock”), along with a circle and a movable line. The subject used the mouse to position the line on the circle to represent the direction the target would be if the subject were in the imagined position. This circle-and-line display simulated the hand-held dials used in previous studies (e.g., Rieser, 1989).

RESULTS

Although each trial provided three measures, only two will be reported: pointing accuracy (the absolute angular difference between the actual and the judged direction of the target) and pointing latency (the time from the appearance of the target to the pointing response).² These dependent variables were analyzed separately using split-plot factorial analyses of variance (ANOVAs) with terms for group, sex, imagined heading, and pointing direction. Because of the large number of main effects and interactions, an alpha level of .01 was used. Only statistically reliable effects are discussed. Error bars in all figures are ± 1 SEM as estimated from the appropriate ANOVA.

The effect of imagined heading was significant for both pointing accuracy [$F(7,112) = 15.87$], and pointing latency [$F(7,112) = 7.91$]. As Figure 2 indicates, imagined headings aligned with the two familiar views (0° and 90°) show significantly lower angular errors [$F(1,112) = 98.78$] and shorter latencies [$F(1,112) = 36.26$] than all

other imagined headings. Mean angular errors for familiar and novel headings were 18.8° and 33.3° , respectively, and mean response latencies for familiar and novel headings were 16.8 and 22.7 sec, respectively. Pairwise comparisons indicated that the two familiar views did not differ ($z < 1$). Pointing judgments from the imagined heading of 45° (which is between the two study views) were faster than judgments from 315° or 135° [$F(1,112) = 8.87$], but not more accurate [$F(1,112) = 1.87$].

The main effect of pointing direction was also significant for both accuracy [$F(5,80) = 12.34$] and latency [$F(5,80) = 4.25$]. For accuracy (Figure 3a), the mean angular errors increased as pointing direction varied from in front of the subject to behind the subject. Contrasts revealed that front (FR and FL) was more accurate than back (BR and BL) and sides (R and L), and that sides were more accurate than back. For latencies (Figure 3b), the trend was much less pronounced. Front directions were faster than back and marginally faster than sides, but sides and back were not significantly different. The effects of pointing direction on accuracy and latency are broadly consistent with other investigations of how observers parse surrounding space into regions (e.g., Franklin, Henkel, & Zangas, 1995).

In accuracy, but not latency, pointing direction interacted with imagined heading [$F(35,560) = 3.36$]. This interaction did not compromise any of the conclusions reached from the omnibus effect of imagined heading (Figure 2a). In particular, angular error for all pointing directions was lower for imagined headings of 0° and 90° (the studied views) than for other headings. The interaction revealed that pointing judgments to the left or to the right were easier when they were made toward or away from a study view than when they were made in other directions. This effect accounts for most of the savings in angular error in Figure 2a for headings of 180° and 270° .

DISCUSSION

These results suggest that mental representations of large (i.e., navigable) spaces are viewpoint dependent, and that two views of a spatial layout produce two viewpoint-dependent representations in memory. Imagined headings aligned with the study views were easier to retrieve (as assessed by accuracy and latency) than were novel headings.

It is instructive to compare the results of the present study with the results of an experiment in which subjects received only one view of the space (Roskos-Ewoldsen et al., in press, Experiment 2). These two experiments used the same spatial layouts and very similar procedures. The difference in mean angular error between familiar and novel headings was 14.5° in the present study and 11.8° in the single-view study. Similarly, the difference in mean response latency was 5.9 sec in the present study and 5.5 sec in the single-view study. If multiple views of a space had resulted in generalization across novel headings, we would have expected the differences to decrease when we added a second study view. On the contrary, the difference between familiar and novel headings when subjects saw two views of the layout was as large as, if not larger than, the difference when subjects learned a single view. These data indicate that an additional view did not lead to the formation of a viewpoint-independent representation in memory.

To retrieve information about a novel heading, subjects must have transformed an existing representation in memory. An obvious candidate for this transformation is mental rotation (Shepard & Cooper,

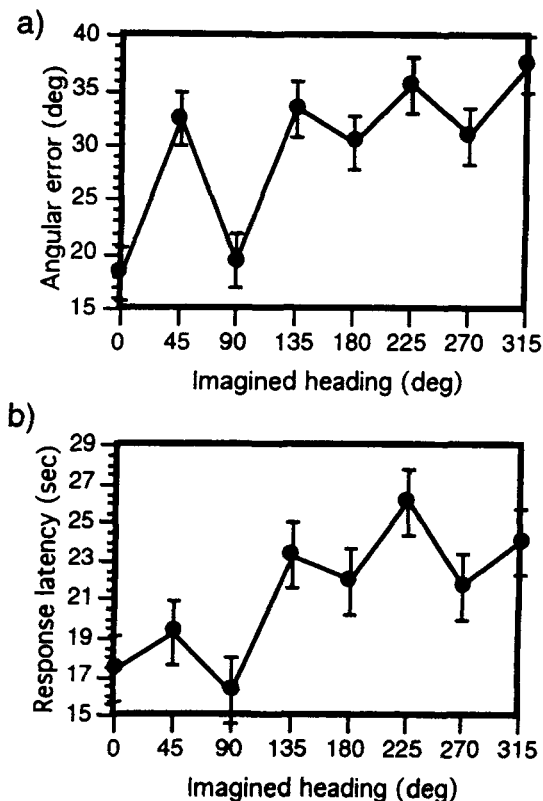


Figure 2. (a) Mean angular error and (b) latency of pointing judgments as a function of imagined heading (± 1 SEM).

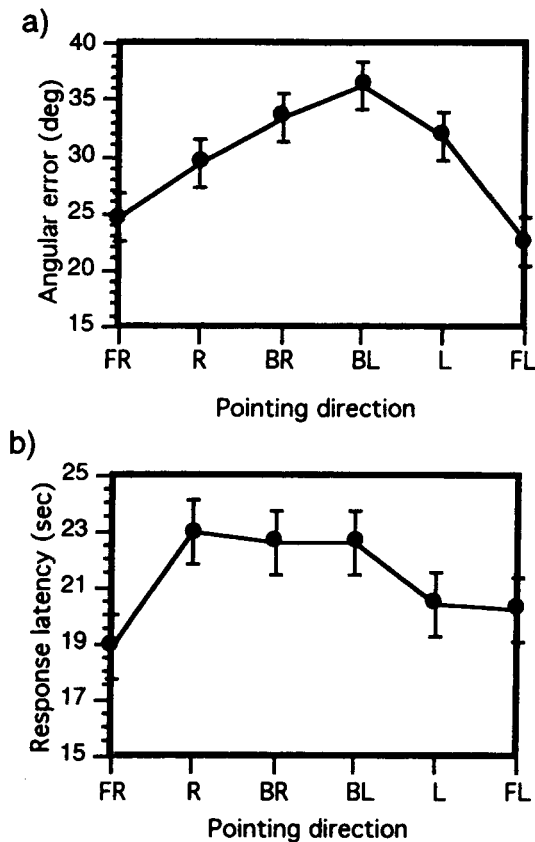


Figure 3. (a) Mean angular error and (b) latency of pointing judgments (\pm SEM) as a function of pointing direction: FR, front-right; R, right; BR, back-right; BL, back-left; L, left; and FL, front-left.

1982). In fact, most subjects reported after the experiment that they tried to imagine either rotating the configuration of objects or rotating themselves in the configuration. When the data from imagined headings were collapsed according to the angular distance to the nearest study view, the rate of rotation as indexed by the slope of the best fitting linear function was about 14°/sec. This rate is well within the limits proposed in previous work on mental rotation (e.g., Shepard & Hurwitz, 1984; Shepard & Metzler, 1971) and indeed is on the slow side.

The speeded performance at 45° compared with that at all other novel headings suggests that this heading was more accessible (Figure 2b). There are at least two plausible explanations of this savings in performance. First, when people were moved between the two viewpoints during the learning phase, they might have updated the locations of the objects relative to themselves. Given that the 45° heading was aligned with a perspective that would be available during movement, there might have been improved speed as a result of having locomoted through that region of the space (we emphasize, however, that subjects were blindfolded while they were moved and never experienced the space between 0° and 90° visually). This explanation is consistent with data on real versus imagined locomotion. Rieser (1989; see also, Presson & Montello, 1994) showed that judgments of relative direction were easy to make from novel imagined headings if observers were allowed to rotate their bodies before making the judgment.

A second explanation of the savings in response latency at headings of 45° follows from recent research on visual object recognition. Bühlhoff and Edelman (1992) and Tarr (1995), for example, have found that a novel view of an object “between” two closely spaced study views (referred to as “interpolation”) may be recognized as easily as the study views, whereas the difficulty of recognizing a novel view that is “be-

yond” two study views (referred to as “extrapolation”) is an increasing function of the angular distance to the nearest study view. In our experiment, an imagined heading of 45° corresponds to interpolation because it is between the study views of 0° and 90°. The mechanisms that give rise to these effects in object recognition are not well understood. However, these effects are predicted by viewpoint-dependent models of object recognition (e.g., Poggio & Girosi, 1990), and hence do not indicate that viewpoint-independent representations were being formed.

The present experiment cannot distinguish between these two alternative explanations of the speeded performance on imagined headings of 45°. This finding, however, does not compromise the conclusion that two viewpoint-dependent representations were formed in memory.

If multiple views of a spatial layout are represented in memory, then with a sufficient number of views, memory for a space might appear to be viewpoint independent (see, e.g., Evans & Pezdek, 1980; Thorndyke & Hayes-Roth, 1982). It is natural to ask, however, whether a small number of additional views, beyond just two, would produce a viewpoint-independent representation in memory. Several lines of evidence suggest that the answer to this question is “no.” The existence of binocular stereo demonstrates that the visual system can reconstruct much of the spatial structure of the world from two views (viz., the two eyes), and Koenderink and van Doorn (1991) have proven that the affine structure of a spatial array can be reconstructed from two views alone. In general, additional views in excess of two do not facilitate the perception of geometric structure. More direct evidence can be found in an experiment reported by Diwadkar and McNamara (in press). Observers saw a collection of objects from a single view and then learned to recognize the scene from this study view and three additional “training” views (presented as digitized images). Recognition performance indicated that all four familiar views were represented in memory, and that novel views were recognized by normalization to the nearest study or training view. Finally, comparison of the results of the single-view and the two-view experiments (discussed above) indicates that an additional view of a spatial layout does not facilitate the retrieval of information about novel views. Collectively, these observations suggest that a small number of additional views of a space would not produce a viewpoint-independent representation of the space.

The parallels between our results and those obtained in recent investigations of visual object recognition are striking (e.g., Edelman & Bühlhoff, 1992; Tarr, 1995). In both domains, spatial relations are encoded with respect to reference frames that depend on point of view. Multiple views of a scene or an object appear to produce multiple viewpoint-dependent representations in memory, not a single viewpoint-independent representation. Apparently, interobject and intraobject spatial relations are represented and processed in similar ways by the human brain.

REFERENCES

BIEDERMAN, I., & GERHARDSTEIN, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: Reply to Tarr and Bühlhoff (1995). *Journal of Experimental Psychology: Human Perception & Performance*, *21*, 1506-1514.

BÜLTHOFF, H. H., & EDELMAN, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, *89*, 60-64.

DIWADKAR, V. A., & MCNAMARA, T. P. (in press). Viewpoint dependence in scene recognition. *Psychological Science*.

EASTON, R. D., & SHOLL, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 483-500.

EDELMAN, S., & BÜLTHOFF, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*, 2385-2400.

EVANS, G. W., & PEZDEK, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 13-24.

FRANKLIN, N., HENKEL, L. A., & ZANGAS, T. (1995). Parsing surrounding space into regions. *Memory & Cognition*, *23*, 397-407.

HINTZMAN, D. L., O'DELL, C. S., & ARNDT, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology*, *13*, 149-206.

KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from

- motion. *Journal of the Optical Society of America A*, **8**, 377-385.
- LEVINE, M., JANKOVIC, I. N., & PALIJ, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, **111**, 157-175.
- POGGIO, T., & GIROSI, F. (1990). Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, **247**, 978-982.
- PRESSON, C. C., DELANGE, N., & HAZELRIGG, M. D. (1987). Orientation-specificity in kinesthetic spatial learning: The role of multiple orientations. *Memory & Cognition*, **15**, 225-229.
- PRESSON, C. C., DELANGE, N., & HAZELRIGG, M. D. (1989). Orientation-specificity in spatial memory: What makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 887-897.
- PRESSON, C. C., & HAZELRIGG, M. D. (1984). Building spatial representations through primary and secondary learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 716-722.
- PRESSON, C. C., & MONTELLO, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, **23**, 1447-1455.
- RIESER, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1157-1165.
- ROSKOS-EWOLDSSEN, B., McNAMARA, T. P., SHELTON, A. L., & CARR, W. S. (in press). Mental representations of large and small spatial layouts are viewpoint dependent. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- SHEPARD, R. N., & HURWITZ, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, **18**, 161-193.
- SHEPARD, R. N., & METZLER, J. (1971). Mental rotation of three-dimensional objects. *Science*, **171**, 701-703.
- TARR, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint-dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review*, **2**, 55-82.
- TARR, M. J., & PINKER, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, **21**, 233-282.
- THORNDYKE, P. W., & HAYES-ROTH, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, **14**, 560-589.

NOTES

1. These judgments also require that an equivalence be established between "up-down" and "forward-backward." Apparently this transformation is natural and effortless, because people have no difficulty with it.

2. We also measured the time from trial onset until the subject clicked the mouse indicating the adoption of a heading. An analysis of variance revealed no significant effects or interactions for this measure. Apparently, subjects did not adopt a heading until the target object was also present.

(Manuscript received February 26, 1996;
revision accepted for publication June 18, 1996.)