SPATIAL MENTAL MODELS

Barbara Tversky

I. Overview

There are many simple, everyday tasks, such as following road directions, using instructions to assemble a bicycle, reading a novel, or helping to solve your child's geometry homework, that seem to entail constructing a spatial mental model from a description. In order to comprehend *Go straight till the first light, then turn left, go down about three blocks to Oak, and make a right,* it is useful to have a spatial representation. Of course, the gist of the message could be remembered instead, but incorporating the instructions into a mental model helps, especially when things don't quite turn out as expected, such as encountering a "No Left Turn" sign at the light. Indeed, there is evidence that people do construct such spatial models. The nature of such models is the topic of this article.

Ample research in memory and comprehension of text supports the assertion that listeners or readers form not only representations of the language of the text—of sound or graphemic properties, of actual words or sentences, of gist—but also of the situation described by the text (Bransford, Barclay, & Franks, 1972; Garnham, 1981; Johnson-Laird, 1983; van Dijk & Kintsch, 1983; among others). Because they are familiar, universal, and objective, we have chosen to investigate descriptions of spatial environments. People have considerable experience converting spoken or written communications about environments into mental representations, and then acting on them. People then get feedback—they either get lost or find their way—and can correct their models. In addition, there is a large body of data on how people learn and remember environments from experience or from maps that can be compared to acquiring environments from descriptions. Just as for maps, learning environments from narratives can be assessed by measuring speed and accuracy to make judgments of spatial relations, distance, and direction, as well as by style and accuracy to make productions, such as maps.

We have developed two separate but related experimental paradigms to investigate spatial mental models constructed from text. In the first paradigm, we vary characteristics of the descriptions and observe the consequent mental models. This work has been done with Holly Taylor. In the second paradigm, we examine in great detail the spatial characteristics of a particular but very common situation, the one people are in most of the time, of having objects at different places around them. Much of this work has been done with Nancy Franklin and, more recently, David Bryant.

This research program has several goals. The first is to demonstrate that the mental models constructed from text with neither visual displays nor special instructions to image nevertheless reflect spatial properties described in the text. Many of the early and elegant demonstrations of imagery and spatial thinking per force used contrived situations. Now that a body of techniques for exploring spatial thinking has been developed, such techniques can be applied to more natural situations, and especially to cases where neither visual information nor instructions to image are given. Another aim is to discover which spatial properties are preserved, and how they are organized and accessed, and to investigate the effects of discourse organization and spatial organization on that. Studies by Denis and Denhiere (1990), Foos (1980), Mani and Johnson-Laird (1982), Ehrlich and Johnson-Laird (1982), and Perrig and Kintsch (1985) have shown that when descriptions are complete and coherent, readers' mental models preserve information about the spatial relations among the objects in a described scene. Studies by Denis and Cocude (1989), Franklin (1991), Glenberg, Meyer, and Lindem (1987), Morrow, Bower, and Greenspan (1989), Morrow, Greenspan, and Bower (1987), and Wagener-Wender and Wender (1990) indicate that some distance information described in text is preserved in mental models. The first set of studies addresses the issue of the generality and perspective of spatial mental models constructed from different text perspectives. Specifically, are they like structural descriptions (e.g., Marr, 1982; Minsky, 1975; Palmer, 1977; Pinker, 1984; Ullman, 1989), i.e., perspective-free representations of the spatial relations of parts of a scene that allow viewers to take different perspectives on them? Or are they like images (e.g., Kosslyn, 1980; Shepard & Podgorny, 1978), i.e., internalized perceptions, representing a scene from a particular viewpoint, namely, the one described in the text? The second set of studies investigates representation and access of particular spatial relations from particular perspectives.

II. Survey and Route Descriptions

When tourists visit a new place, they often buy guidebooks to let them know what is worth seeing and doing, and how to get there. An informal review of guidebooks reveals that they tend to adopt one of two perspectives on the place described. Some take the reader on a mental tour or *route* through the environment. A route description of the Smithsonian in Washington, D.C. might proceed:

> As you leave the Capitol going along the Mall, the first building you pass on your right is the East Wing of the National Gallery. Continuing on, you come to the main building of the National Gallery. On your left, across the Mall, you can see the Air and Space Museum . . . until you reach the Washington Monument.

Another perspective commonly adopted is to give the reader a bird's eye view or *survey* of the place. A survey description of the same scene might proceed:

> At the east end of the Mall stands the Capitol and at the west end, the Washington Monument. Along the north side of the Mall, the eastern-most building is the East Wing of the National Gallery. Just west of it is the National Gallery. . . . On the south side of the Mall, the eastern-most building is the Air and Space Museum, directly south across the Mall from the National Gallery.

Survey descriptions take a perspective from above and describe the locations of landmarks relative to one another in canonical direction terms: north, south, east, and west. In addition, survey descriptions are often hierarchical, beginning with an overview of boundaries of large-scale regions, and becoming more specific. Route descriptions take the perspective of a moving observer in the environment, typically addressed as *you*, and describe the locations of landmarks relative to your (the observer's) changing position in terms of left, right, in front, and behind. Route descriptions are typically at a single level of analysis whose sequence is determined by the particular path. Thus, the description perspectives differ in spatial terminology, and whether locations of landmarks are described with respect to other landmarks or with respect to the location of an observer.

The initial question Taylor and I (in press) asked is: Do route and survey descriptions lead to different mental representations? That is, do the representations generated by each perspective preserve that perspective, or are they perspective-free? The question of perspectives of narratives and of mental representations is of more generality than just spatial models, as route-like and survey-like descriptions are appropriate for other topics as well, e.g., descriptions of time. Here we focus on spatial descriptions only.

Previous research on narrative comprehension and on learning actual environments suggests that different perspectives yield different representations. Readers remember details relevant to their own perspective better than those relevant to an alternate perspective for both physical (Abelson, 1975; Perrig & Kintsch, 1985) and character perspective (Anderson & Pichert, 1978; Bower, 1978). Some information about actual environments is better acquired by studying maps, such as Euclidean distance and direction, whereas other information is better acquired from actual navigation, such as traversal distance (Evans & Pezdek, 1980; Sholl, 1987; Streeter, Vitello, & Wonsiewicz, 1985; Thorndyke, 1981; Thorndyke & Hayes-Roth, 1982). Narratives, however, cannot easily present the continuous information available from maps and navigation. Narratives can easily convey categorical information: north, south, east, west, and right, left, front, back. Considerable research has shown that spatial information acquired from both maps and actual traversal is distorted toward these and other major spatial categories, though, of course, some more detailed information is retained and used (e.g., R. W. Byrne, 1979; Chase, 1983; Hirtle & Jonides, 1985; Maki, 1981; McNamara, 1986; Moar & Bower, 1983; Stevens & Coupe, 1978; Tversky, 1981; Wilton, 1979). In this research on narratives, we can only assess the global, categorical spatial relations easily conveyed by language.

A. Experiment 1: Route vs. Survey Descriptions

1. *Task*

Taylor and I (in press) developed four fictitious environments: two large-scale—one county-sized and the other a small town—and two small-scale—a zoo and a convention center—containing from 11 to 15 landmarks each. Depictions of these environments are in Figs. 1–4, but subjects in the initial experiments did not see these maps.

We wrote a survey and a route description of each environment. The

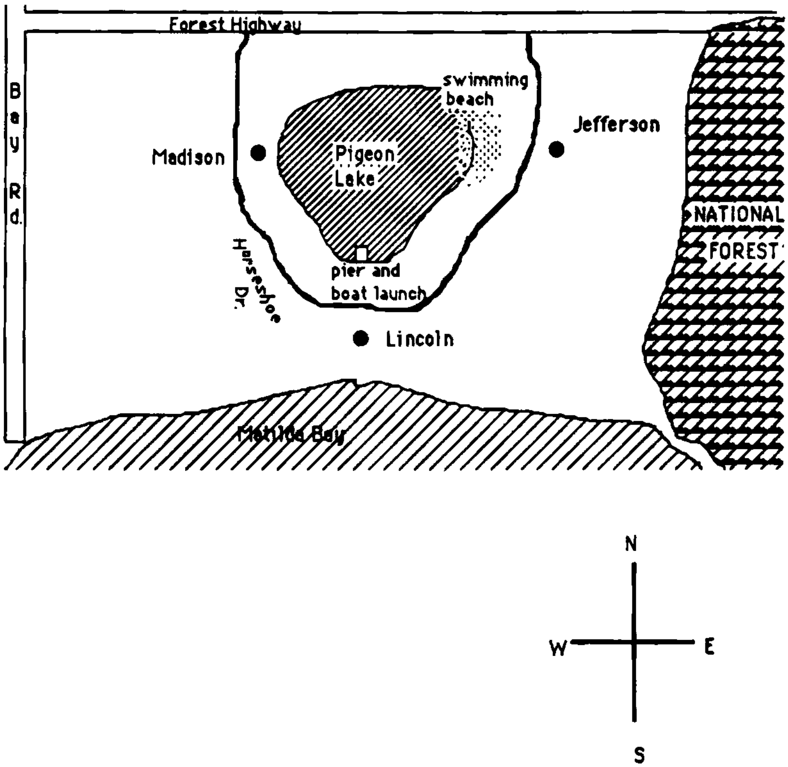

Fig. 1. Map of resort area. From Taylor and Tversky (in press). Reprinted by permission.

survey descriptions took a perspective from above, used a hierarchical organization, and adopted canonical direction terms to describe landmarks relative to each other in terms of north, south, east, and west. The route descriptions took a perspective from within the environment, used a sequential organization, and adopted egocentric direction terms to describe landmarks in relation to a moving ego, in terms of left, right, and front.

While we wished to make the alternative descriptions equally coherent, there is no widely applicable measure of discourse coherence. Coreference, i.e., linking sentences in sequence by referring to the same thing, has sometimes been suggested (Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Coreference may be appropriate for route or sequential organizations, but not for hierarchical descriptions, where a new descriptive part will refer back to the overview but not to the previous sentence. Lacking an objective measure, we asked a group of pilot subjects to evaluate the coherence of the texts, and they reported that the two types of

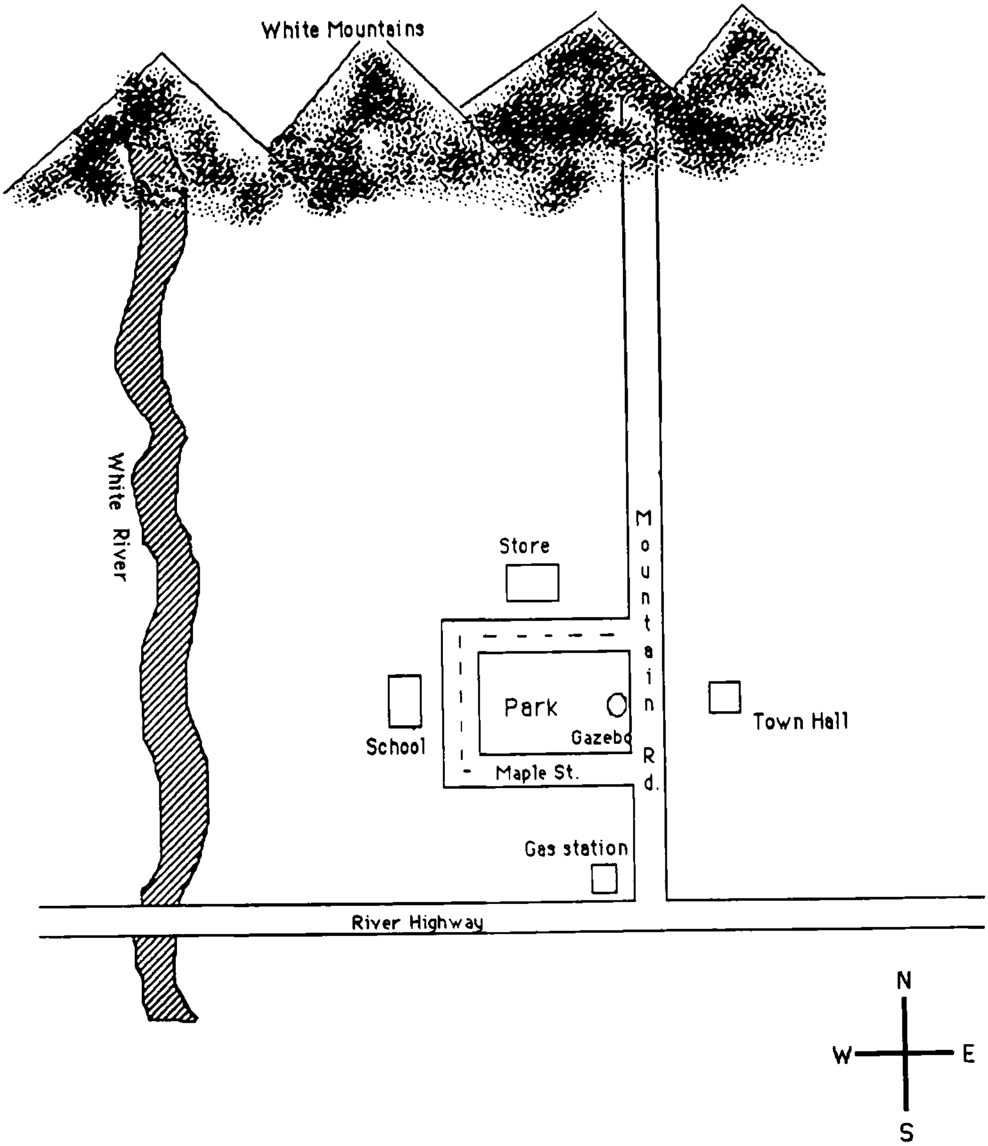

Fig. 2. Map of town. From Taylor and Tversky (in press). Reprinted by permission.

descriptions were equally coherent. We also pretested the descriptions to make sure that readers could correctly place all landmarks in sketches, i.e., that the information was in fact complete and determinate. In addition to the locative information, each description contained nonlocative information, e.g., relating activities that could be performed in different parts of

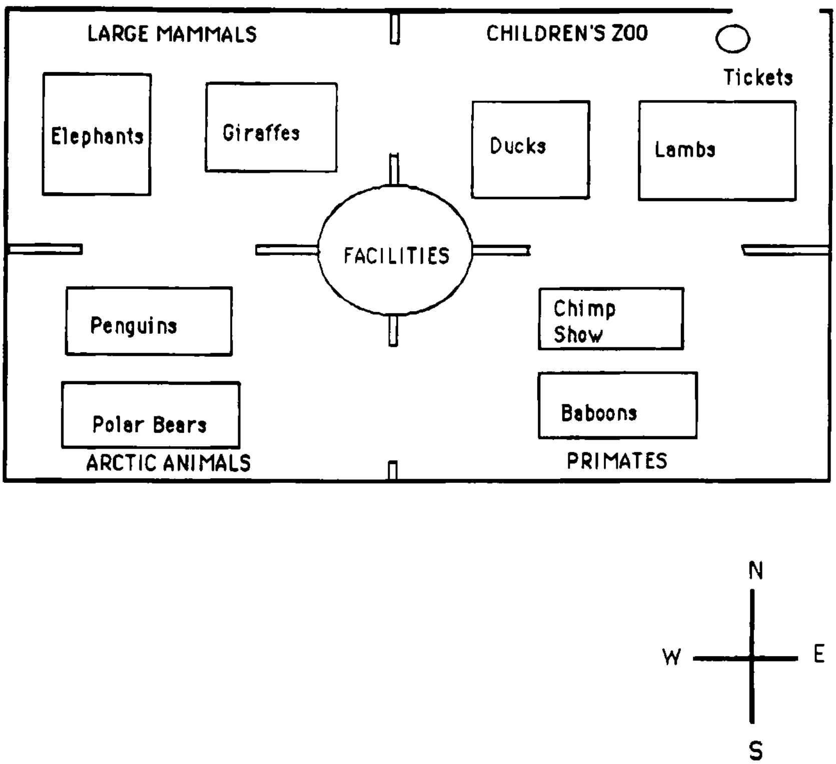

Fig. 3. Map of zoo. From Taylor and Tversky (in press). Reprinted by permission.

the environment, or giving elaborative details about landmarks. This information was identical for route and survey descriptions. As examples, the route and survey text for the resort area are presented as follows:

Survey Description of Resort Area

The Pigeon Lake resort area is well situated for people who are interested in a variety of outdoor activities. The resort area is bordered by four major landmarks: the National Forest, Matilda Bay, Bay Rd., and the Forest Highway. The eastern border is made up of the National Forest. The National Forest has facilities for camping, hiking, and rock climbing. The southern border is made up of Matilda Bay. Two major roads, Bay Road and the Forest Highway, form the other two borders of the region. Bay Rd., runs north–south along the western border of this region. Bay

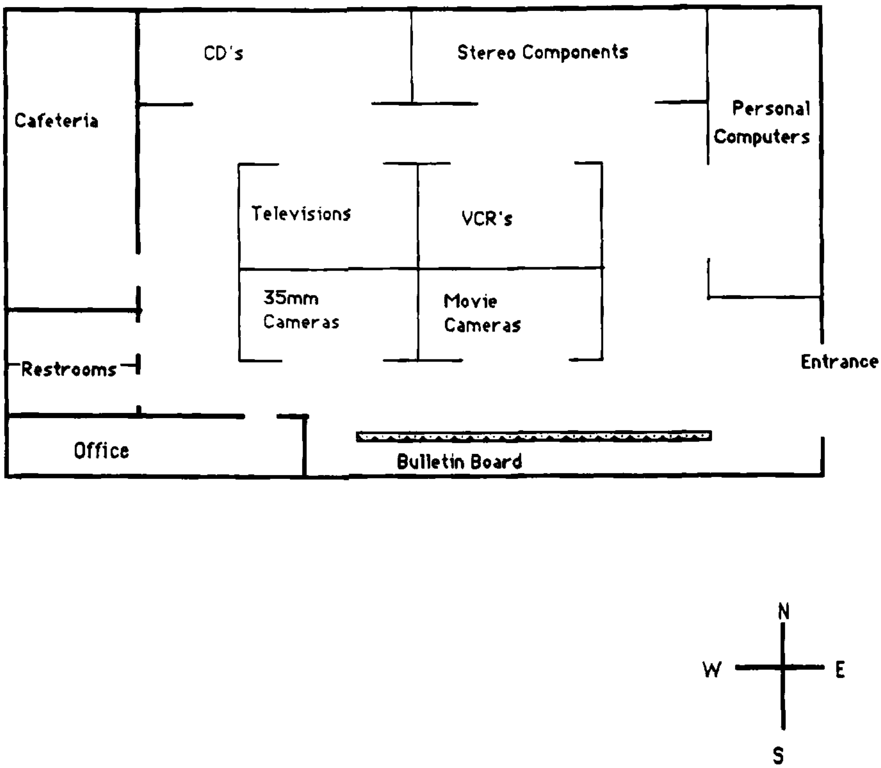

Fig. 4. Map of convention center. From Taylor and Tversky (in press). Reprinted by permission.

Rd. is the main access to the many recreational areas on Matilda Bay. Bay Rd. is also the main route in and out of this region. The Forest Highway forms the northern border and provides the 120-mile link between Bay Rd. and the National Forest. Pigeon Lake is a large recreational lake in the center of the region. There are many activities that center around Pigeon Lake. People enjoy boating, water skiing, and swimming on the lake. There is a fishing pier and boat launch at the southernmost point of the lake. Since this is the only place to launch boats, there is usually quite a bit of traffic near the launch site. On the east shore of the lake there is a swimming beach. In the busy summer tourist season, there are lifeguards on the beach. Horseshoe Drive follows the rounded outline of the lake and is connected at both ends to the Forest Highway. Horseshoe Drive begins about 40 miles east of Bay Rd. and ends about 40 miles west of the national forest. There are three small towns within the Pigeon Lake region that all lie along Horseshoe Drive. Madison lies on the west shore between the lake and Horseshoe Drive. Madison is directly across the lake from the swimming beach. Madison is the site of the annual seafood festival where the main event is the fishing contest. Jefferson lies on the east side of the lake on the National Forest side of Horseshoe Drive. Jefferson is the main center for hiking and cycling. Lincoln lies on the south side of the lake midway between Horseshoe Dr. and the Bay. Lincoln is considered by tourists to have the best location in the region because of its close proximity to the bay.

Route Description of Resort Area

The Pigeon Lake resort area is well situated for people who are interested in a variety of outdoor activities. To reach the Pigeon Lake region, drive south along Bay Rd. until you reach, on your left, the point where the Forest Highway dead-ends into Bay Rd. From this intersection, you can see in the distance that Bay Rd. continues to Matilda Bay and its many recreational areas. You turn left onto the Forest Highway and travel about 40 miles until, on your right, you reach Horseshoe Drive. Horseshoe Dr. is the only road that you can take to get into the Pigeon Lake region. Turning right onto Horseshoe Drive, from the Forest Highway, you see, on your left, Pigeon Lake. Pigeon Lake is a large recreational lake in the center of this region. There are many activities that center around Pigeon Lake. On the lake, people enjoy boating, water skiing, and swimming. After you drive for ten miles along Horseshoe Drive, you see, on your left, the small town of Madison. Madison is the site of the annual seafood festival where the main event is the fishing contest. As you continue along Horseshoe Drive, you notice that the road follows the rounded outline of the lake. Twenty miles after you leave Madison, you see, off Horseshoe Dr. on your right, the little town of Lincoln. From your position, only a short distance beyond Lincoln you can see Matilda Bay. Because of its close proximity to the bay, Lincoln is considered, by tourists, to have the best location in the region. From your position with Lincoln on your right, you see, on your left, the fishing pier and boat launch for Pigeon Lake. Since there is only one boat launch for Pigeon Lake, there is usually quite a bit of traffic near the launch site. Continuing around the shore of the lake on Horseshoe Dr., you drive about twenty more miles until you come to the swimming beach and the town of Jefferson. On your left is the swimming beach. In the busy summer tourist season, there are lifeguards on the beach. From your position with the swimming beach on your left, you see, on your right, the town of Jefferson. Jefferson is the main center for hiking and cycling for the area. You drive for another ten miles on Horseshoe Dr. until you return to the Forest Highway. To your right, about forty miles away, you can see the National Forest. The National Forest has facilities for camping, hiking, and rock climbing. Turning left onto the Forest Highway, you travel about 40 miles and again see, on your left, the beginning of Horseshoe Dr. Continuing along the highway, you return to Bay Rd., which leads you out of the region.

We modeled the design and memory tasks on those of Perrig and Kintsch (1985), who tested a similar hypothesis. Their results were inconclusive, partly because their descriptions were too difficult, hence poorly learned, and partly because their survey description's organization was derived from that of the route description and was consequently awkward as well as indeterminate, i.e., the locations of some of the landmarks could not be determined from the description. Our subjects read two route and two survey descriptions, one large-scale and one small-scale environment for each description type. Across subjects, each environment was presented equally often as a route and as a survey description. Subjects could read each description up to four times. Reading time was self-paced, and total times were recorded.

After reading each description, subjects were presented with statements to verify as true or false; reaction time and errors were recorded. Some statements tested the nonlocative information. Perspective should make no difference on performance on these questions. Other statements tested the locative information. The verbatim locative statements were taken directly from the texts. The inference locative statements were from the same perspective of the texts and contained information that could be inferred from the text but was not directly given in the texts. Half of both the verbatim and inference locative statements were from a route perspective and half from a survey perspective. Of the inference statements, half were true, half false. A true route inference statement from the convention center was: *Walking from the Personal Computers to the Televisions, you pass, on your right, the Stereo Components.* A false route inference statement from the resort area was: *Driving from Jefferson to Lincoln, Pigeon Lake is on your left.* A true survey inference statement from the town was: *The Gas Station is east of the river and south of Maple St.* A false survey inference statement from the zoo was: *The Giraffes' Cage is west of the Polar Bears' Cage and south of the Baboon Colony.* Readers answered all questions regardless of perspective read. Thus, a verbatim statement from a different perspective was in effect an inference statement for that reader. Following the questions, readers drew a map of each environment. This served to check that readers were able to form integrated and correct spatial models from the text, and to check if one type of description (or environment) had an advantage.

2. *Predictions*

Previous research indicates that readers form multiple representations of text and may verify statements against any or all of those representations. If readers use representations of the language of the text to answer the questions, verbatim questions should be faster and more accurate than inference questions. When verification statements are verbal, comparison to linguistic or propositional information is faster than to images or mental models (e.g., Kosslyn, 1976). Inference statements, on the other hand, cannot be verified directly by comparison to a representation of the language of the text. They can be verified either by comparison to a representation of the text plus rules of spatial inference, or by comparison to a mental model of the situation described by the text. Using descriptions of spatial arrays similar to but simpler than the present ones, R. M. D. Byrne and Johnson-Laird (1989) showed that readers verify by comparison to mental models rather than by applying spatial inference rules to representations of text. If the situation models readers construct depend on the particular perspective of the narrative, then readers should respond faster and more accurately to inference statements from the perspective read than to inference and verbatim statements from the other perspective. If, however, readers construct the same spatial mental models irrespective of the perspective of the text, then there should be no differences in speed or accuracy on the inference questions that depend on perspective read.

3. *Results*

Route maps took slightly but significantly longer to read. Subjects made more map errors on route descriptions (1.31) than on survey descriptions (0.68), but there were very few errors made on maps altogether, indicating that readers formed highly accurate situation models from the texts. The data of primary interest are the reaction times and error rates to the different types of questions, presented in Fig. 5. As in the case of the maps, overall performance was excellent. First, there were fewer errors and faster reaction times to verify the nonlocative statements than the locative statements. As expected, perspective had no effect on performance on nonlocative statements. We would not like to claim that nonlocative infor-

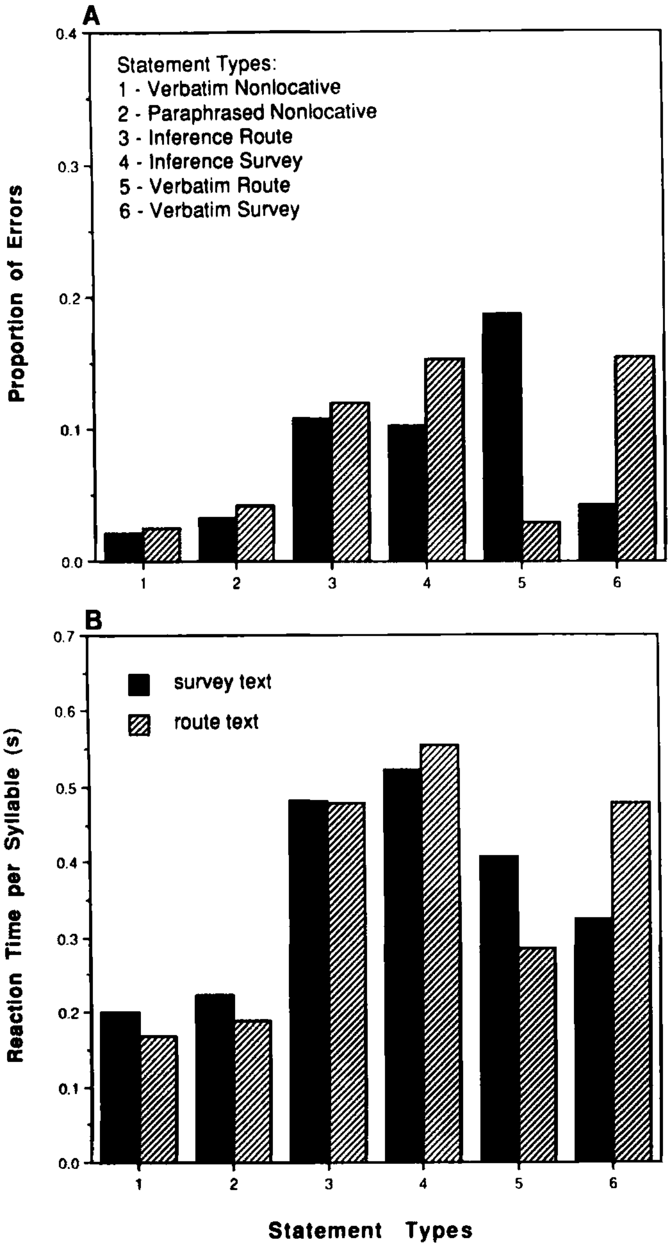

Fig. 5. Experiment 1. A, errors to question types by description type. B, reaction times to question types by description type. Adapted from Taylor and Tversky (in press). Reprinted by permission.

mation is generally easier than locative; surely one could write nonlocative statements that would be very difficult to remember. One possibility is that nonlocative statements can be verified by reference to a representation of the language of the text, which is faster than verification by reference to a mental model.

As for locative statements, the only differences to be found are in the verbatim statements. Subjects were faster and more accurate verifying statements that they had actually read than statements about inferences from information presented in the descriptions. Verbatim statements can also be verified more efficiently by reference to a representation of the language of the text than by reference to a mental model. For inference statements, however, perspective read made no difference. In other words, subjects were as fast and accurate on inference statements from the perspective read than from the other perspective, for both perspectives.

4. *Discussion*

Consistent with previous research, the present results support the establishment of multiple mental representations from text. The rapid and accurate performance on nonlocative and verbatim statements suggests that they were verified by comparison to a representation of the language of the text. How abstract that representation (or representations) is we cannot determine from these results. In contrast to verbatim and nonlocative statements, inference statements were verified more slowly and less accurately, suggesting that these are verified against a mental model of the situation described by the text.

The lack of any differences in verification time or accuracy of survey and route inference statements as a consequence of perspective of description read suggests that the situation model constructed does not depend on the perspective of the text. Because readers are just as good taking a new perspective as taking a previous perspective, their mental models must be general enough to allow the taking of different perspectives with equal ease. Readers of route and survey descriptions appear to have formed the same mental models of the spatial relations of landmarks regardless of perspective of text. Because this finding is on the surface contrary to previous work and because it is a null finding, we replicated it in three more experiments that also allowed exploration of the phenomenon.

B. Experiment 2: Verbatim vs. Paraphrased Statements

In the first experiment, readers were faster and more accurate verifying statements previously read than inference statements. Does the advantage to verbatim statements depend on the exact wording of the sentences or the gist of the information conveyed by them? The second experiment addressed this question by adding paraphrased statements to the set of statements readers were asked to verify. The paraphrased route and survey statements were exactly that, reversals of order of clauses. This was the only possible paraphrasing because there are no adequate synonyms for either the direction terms or the names of places. There was one other change in this experiment, the reason for which will become clear later; the descriptions were changed so that the orders of mentioning landmarks in survey and route versions were quite different. These new narratives were used in all subsequent experiments.

1. *Results*

All of the previous findings were replicated, as is evident in Fig. 6. Readers took longer to study route texts and made more errors on maps drawn from them. Performance was excellent, both in map drawing and in statement verification. Nonlocative statements were verified more quickly and accurately than locative statements. Subjects were equally fast and accurate with both types of inference locative statements regardless of perspective read. However, subjects were faster and more accurate with verbatim and paraphrased statements than with inference statements from either perspective; furthermore, there were no differences between verbatim and paraphrased sentences.

2. *Discussion*

Verbatim statements appear to be verified by comparison to a representation of the text, in contrast to inference statements, which took longer and appear to be verified by comparison to a representation of the situation described by the text, or a mental model. Like verbatim statements, paraphrased statements are verified more quickly and accurately than inference statements, and thus appear to be verified against a representation of the language of the text. Because only changes in word order and minor changes in wording could be used as paraphrases, no broad conclusions can be drawn about the nature of the representation of the language of the text beyond concluding that representation is not sensitive to large changes in word order and minor changes in wording.

C. Experiment 3: Texts vs. Maps

Is the mental representation of spatial relations induced by the two types of descriptions similar to that induced by studying a map? If so, then subjects

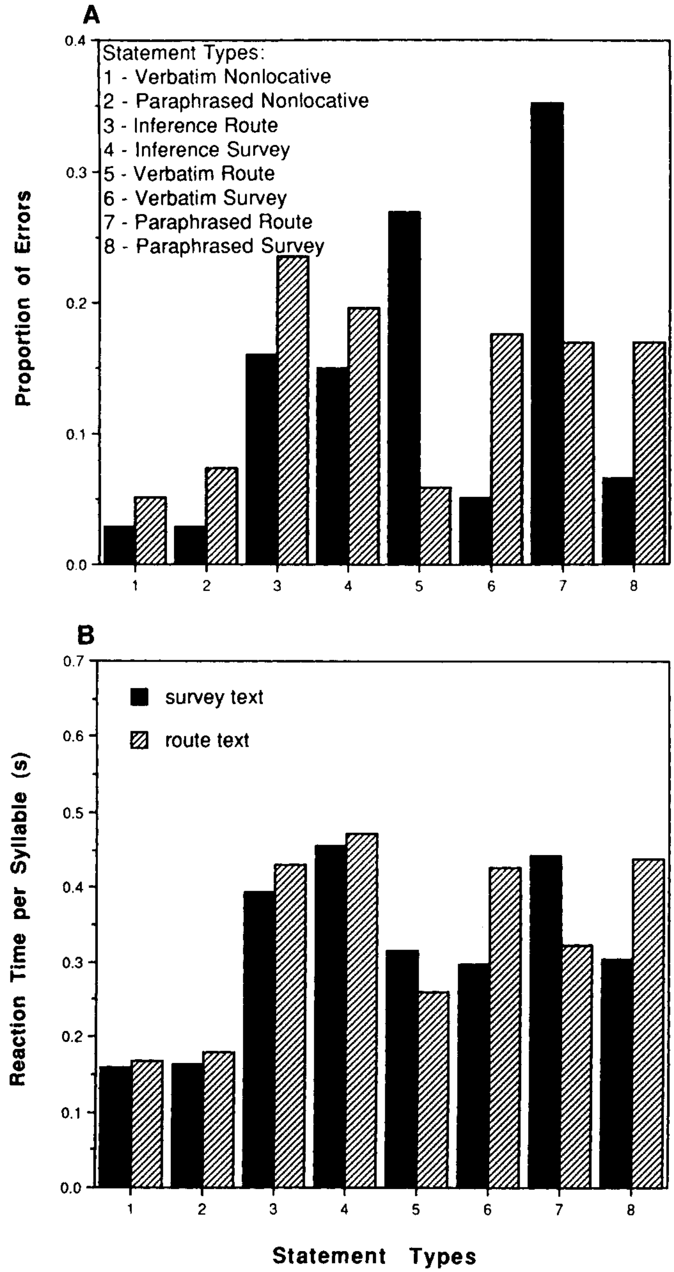

Fig. 6. Experiment 2. A, errors to question types by description type (paraphrased questions). B, reaction times to question types by description type (paraphrased question). Adapted from Taylor and Tversky (in press). Reprinted by permission.

who study a map should, like subjects who studied texts, do equally well on survey and route questions. On the other hand, a map, even more than a survey description, presents a survey perspective. Thus, if perspective is preserved in mental models of spatial relations, then subjects who study maps should perform better on survey than on route statements. This experiment was a replication of the first experiment with the addition of a second group of subjects who studied the maps presented in Figs. 1–4 in lieu of reading descriptions, for up to 10 min per map.

When the results of this experiment were analyzed, the main results of Experiment 1 were replicated a third time, as can be seen in Fig. 7. Locative statements were slower and less accurate than nonlocative statements. Route texts took longer to read, and subjects' maps of route texts were slightly less accurate than those of survey texts. The maps drawn by subjects who studied maps were the most accurate of all. For text subjects, verbatim sentences were faster and more accurate than inference statements, and there were no effects of perspective read on verification of inference statements.

As for map subjects, their pattern of reaction times to the statement types was comparable to that of text subjects. They responded equally quickly to both route and survey statements, indicating that their mental models were not biased toward either perspective. The pattern of errors for map subjects were slightly more complex. For route questions, accuracy of map subjects was at the level of survey text subjects and at that of inference questions for route text subjects, again supporting the claim that map subjects' mental models of spatial relations were comparable to those of text subjects. However, on survey questions, map subjects were more accurate than route subjects (though, again, note that the overall error rate is low). However, we are reluctant to take that as evidence that a survey perspective is inherent in the mental representations of map subjects for several reasons. First, there was no comparable effect for verification time, and second, map subjects' performance was highly similar to text subjects' performance. An explanation we prefer for the especially high accuracy of map subjects on survey statements is that the information required to verify these statements was given directly by the maps, whereas the information required to verify route statements was not. Thus, for map subjects, survey statements are analogous to verbatim statements, and route statements are analogous to inference statements. Such a stance seems to imply that just as there are multiple representations for text, for example, representations of the language of the text and representations of the situation described by the text, so there may be multiple representations of depictions, some closer to the actual visual

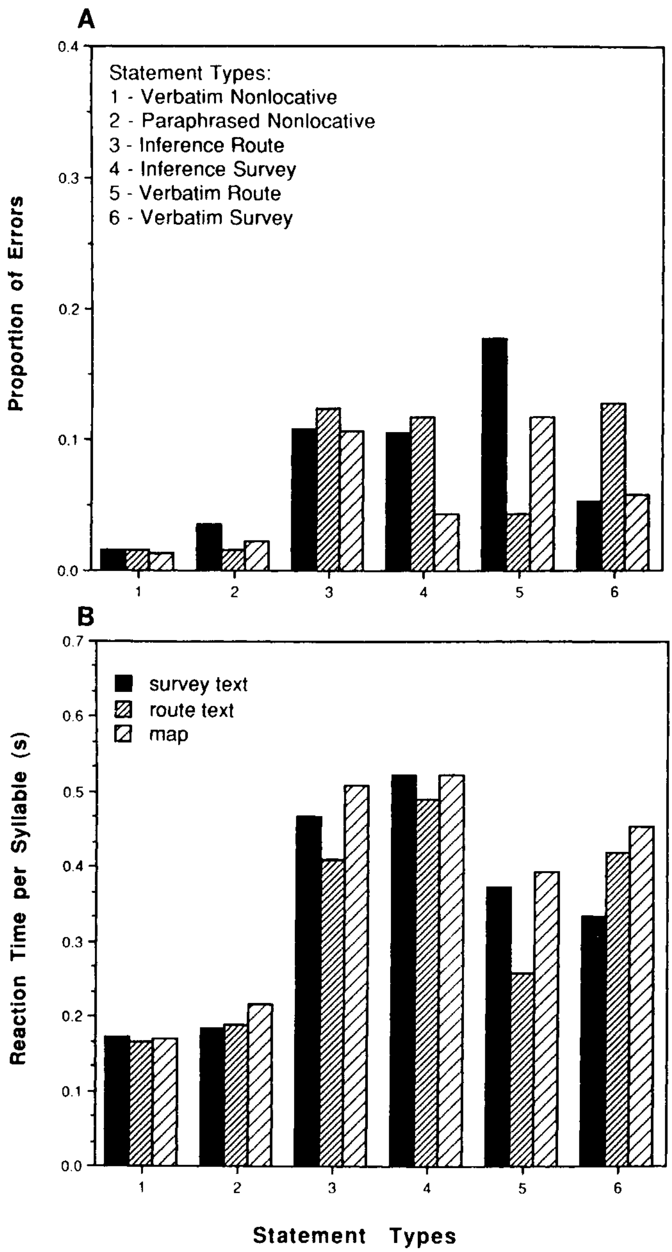

Fig. 7. Experiment 3. A, errors to question types by description type (text and map conditions). B, Reaction times to question types by description type (text and map conditions). Adapted from Taylor and Tversky (in press). Reprinted by permission.

display, and some more abstract representations constructed from integrating over the display.[1]

D. Experiment 4: Single Trial

In the previous experiments, subjects studied four texts (or maps) and knew that they would be asked to draw a map after each one. Furthermore, after the first trial, they knew that they would be asked to verify statements from both perspectives. Perhaps these expectations led them to construct mental models more abstract than they otherwise would. For this reason, another experiment was run in which subjects studied only a single text, and were told to study the text so that they could answer questions about the information presented in it. They were not told ahead of time about the map task, though they were asked to draw maps after the verification task.

The results of this experiment showed no differences in study time or in map accuracy due to text perspective. Otherwise the main results were replicated a fourth time, and are displayed in Fig. 8. Verification of nonlocative statements was faster and more accurate than that of locative statements. Performance was very high in both statement verification and map drawing. Verbatim statements were faster and more accurate than inference statements. There were no effects of perspective on inference statements despite no expectations of map drawing or of questions from a different perspective.

E. Spatial Mental Models

In four experiments, subjects read a route or a survey description of an environment. Route descriptions took readers on a mental tour and described environments in terms of left, right, front, and back, relative to a moving observer, addressed as "you." Survey descriptions described the environments from above, relating locations of landmarks to one another in terms of north, south, east, and west. After each description, subjects verified verbatim and inference statements from both perspectives for each of up to four environments. Following the verification task, subjects drew sketch maps of the environments. The maps subjects drew contained very few errors, indicating that readers formed accurate mental representations of the envionments from text alone. Readers were faster and more accurate to verify verbatim statements than inference statements, indicat-

[1] Data on eye movements indicate that a pictorial display is scanned part by part (e.g., Noton & Stark, 1971). This in turn suggests that complex visual displays such as maps and pictures are not encoded wholistically as snapshops but rather are encoded piecemeal and integrated.

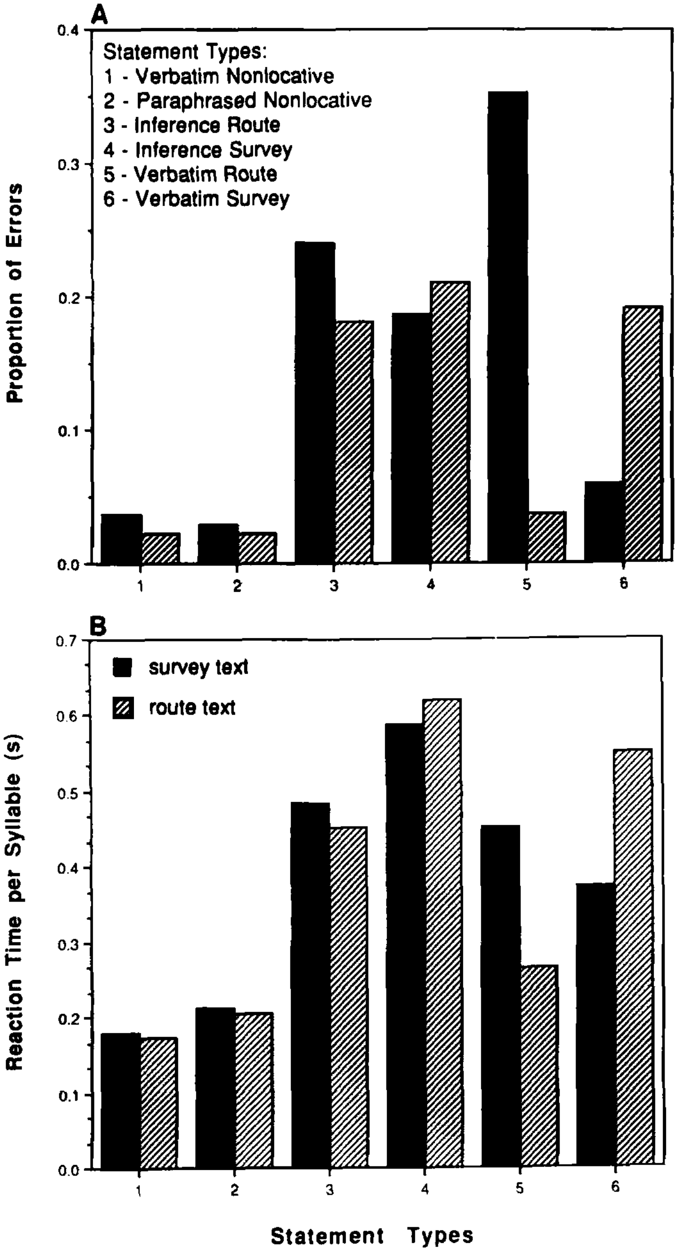

Fig. 8. Experiment 4. A, Errors to question types by description type (single description). B, Reaction times to question types by description type (single description). Adapted from Taylor and Tversky (in press). Reprinted by permission.

ing that verbatim statements were verified by comparison to a representation of the language of the description. For inference statements there was no effect of perspective read. That is, readers were as fast and as accurate with statements in the other perspective as in the read perspective. This finding was obtained four times under varying conditions. Inference statement verification and map drawing seem to have been done from mental models of the situation described in the texts, of the spatial relations between landmarks.

The simplest account of these findings (an account quite similar to that of Johnson-Laird, 1983) is that for this type of text and environment, spatial mental models induced by the two text perspectives are functionally the same, and perspective-free. Perspective is represented, but in representations of the language of the text, not in representations of the situation described by the text. What might such a spatial mental model look like? Taylor and I (in press) speculated that it might not look like anything that can be visualized. Rather, it might be something analogous to an architect's 3-D model of a town that can be viewed or visualized from many different perspectives, but cannot be viewed or visualized as a whole. In fact, answering the locative questions required taking a particular perspective, either from above or from within. Moreover, the questions required subjects to call up only a small part of the representation, typically about three landmarks, rather than the entire scene. To verify a route statement seemed to require imagining being in a particular location in the environment, facing a specific landmark and surrounded by others, and checking that the landmarks were in the proper directions—front, back, left, right—relative to the imagined viewpoint. To verify a survey statement seemed to require focusing on the locations of a particular set of landmarks as if from above to ascertain if they were in the proper relations—north, south, east, west. Subjects' mental models were abstract enough to allow either type of judgment with equal ease. These speculations correspond to many peoples' introspections about spatial environments they know very well: their homes, workplaces, neighborhoods. That is, they do not have a single mental representation of them but many, and they can adopt many different points of view on such well-known environments.

As such, the spatial mental models constructed by readers of route and survey texts are like structural descriptions of objects. Structural descriptions specify the spatial relations among parts of an object. Rather than being viewer-centered, they are object-centered (e.g., Marr & Nishihara, 1978), that is, perspective-free. That construct, structural description, was developed to account for our ability to recognize objects from many different perspectives. Both mental models and structural descriptions differ from the representations proposed in the classic work on imagery (e.g., Finke & Shepard, 1986; Kosslyn, 1980; Pinker, 1984; Shepard & Podgorny, 1978), which are perception-like and from a particular point of view. As Johnson-Laird (1983, p. 157) put it, "images correspond to views of models."

Of course, these speculations about spatial mental models apply only for the simple categorical spatial relations among landmarks that can be easily specified in descriptions and were tested in these experiments, and not for the more continuous information that can be acquired from maps and navigation. Distances were not specified, nor were directions specified in terms of degrees, so we could not expect these spatial mental models to contain that information. Similarly, we do not mean to imply that all of what has been termed "cognitive maps," whether formed from descriptions, maps, or actual navigation, have this same abstract character, but of course it is possible that some of them do.

F. New Directions: Drawing Order

From the first experiment a serendipitous finding emerged that Taylor and I have begun to chase. Taylor noticed that in producing maps, subjects seemed to draw the landmarks in the order of their mention in the descriptions. Halfway through the study, she began recording the drawing order, and analysis of those data confirmed her observation. To make absolutely sure there was no bias, two new experimenters and a new set of subjects were recruited. The experimenters were told to record the order of drawing the landmarks, but they did not know which type of description subjects read each trial or the general hypothesis. Some of the descriptions were rewritten to make the route and survey orders as different as possible. The results of this study (Experiment 3) confirmed those of the first. The correlation between order of mentioning landmarks in descriptions and order of drawing them in maps was very high (r = .72), significantly higher than the correlation between the other description order and map drawing order (r = .22). Approximately the same correlations were found in Experiment 2, where drawing orders were also recorded by experimenters blind to the hypothesis.

At first, finding correspondences between description order and map-drawing order may seem contrary to the findings for speed and accuracy in verifying locative statements from the two types of text, where we found no differences save faster times for paraphrased statements. Now, for maps, we find a large difference in drawing order due to description in a task that seems to depend on drawing a mental image or a cognitive map. It is that latter assumption that we question. We proposed earlier that the mental representations subjects construct from these types of narratives do not seem like images but rather like structural descriptions of the spatial relations between the landmarks of the scene. As such, unlike a mental image, they cannot be visualized as a whole. They can be imagined from particular perspectives. Moreover, because they contain a relatively large number of landmarks and spatial relations, they may not be imagined in whole, but rather in parts. So, we conjecture, when subjects are asked to draw a map of the described environments, they do so by reconstructing their mental models part by part, and they reconstruct the mental model of the environment in the same order as they originally constructed it, i.e., in the order of the description they read. Thus, this finding is indirect evidence in support of the contention that readers' mental models are not image- or map-like. If they were, there should be no differences in drawing order depending on description perspective; rather, drawing order should depend on characteristics of the image or map alone (e.g., Novick & Tversky, 1987).

G. New Directions: Descriptions and Depictions

Order of output has often been used as a clue to mental organization in unconstrained tasks (see, e.g., Tulving, 1962). With this in mind, we examined the order of drawing landmarks in the subjects in Experiment 3 who learned the environments from maps rather than descriptions. Their drawing orders contrasted the drawing orders of those who had read descriptions, whose orders corresponded to the order of mention in the descriptions. Subjects who had studied maps tended to draw maps in a hierarchical fashion, beginning with borders or large entities and working inward or toward smaller objects and parts. This pattern could be due to memory organization, i.e., spatial memory may be hierarchically organized (e.g., McNamara, 1986; Stevens & Coupe, 1978), or it could be due to demands of the drawing task, where borders and larger elements set the scale for internal and smaller elements, or both. It is also possible that in describing environments, subjects are implicitly aware that they are constructing mental models in the minds of their readers or listeners, and that many of the constraints of model construction are similar to those of actual construction (Novick & Tversky, 1987).

In another study of organization of descriptions and depictions of environments (Taylor & Tversky, 1990), subjects were asked to study one of three maps—a large-scale (town), a medium scale (a new map of an amusement park), or a small-scale (convention center)—in anticipation either of reproducing the map or of writing a description from memory. In fact, subjects were asked to do both, in counterbalanced order. Neither expectations nor order had effects on either descriptions or depictions. In both cases, almost all subjects recalled all or almost all of the landmarks.

Interestingly, there was a high correspondence between order of drawing landmarks and order of mentioning landmarks both within and across subjects, irrespective of the perspective of the description. Linde and Labov (1975) studied people's descriptions of their apartments, and Levelt (1982) studied people's descriptions of street map-like networks. Both studies found that subjects' descriptions took consistent perspectives on the environments. These descriptions for the most part took listeners on mental tours of the environments, i.e., subjects gave route descriptions. Linde and Labov, and Levelt, attributed this to the linear characteristic of language as opposed to pictures and environments. Language must serialize a two- or three-dimensional array, and the most natural way to serialize is to take a path through an environment.

Linde and Labov's and Levelt's findings and contentions notwithstanding, the descriptions our subjects produced used both route and survey perspectives (and no other perspective), both purely and in mixtures. About half of the subjects used a route perspective for the convention center; of the other half, a small proportion used a survey perspective, and the rest used mixed perspectives. The mirror image of this pattern emerged for the town: about half used a survey perspective, and of the other half, a small portion used a route perspective, and the rest used both. The amusement park, the medium scale environment, elicited a pattern in between that of the convention center and that of the town. Although all environments yielded route, survey, and mixed descriptions, the large-scale environment tended to elicit more survey descriptions and the small-scale environment more route descriptions. Several factors may contribute to this. The larger scale environment contained both large- and small-scale features, encouraging a hierarchical description. The large-scale environment had more than one route through it, and the small scale only a single route. In the real world, people are more likely to interact with a large-scale environment via many routes, and more likely to interact with a small-scale environment by a single path. It may be possible to disentangle these factors—features at more than one scale, single vs. multiple paths, and typical mode of interaction—using specially designed maps. Overall, these findings suggest that the spatial organization is primary for descriptions, and the linguistic devices, such as those used to establish perspective, are secondary.

III. Spatial Frameworks

In the research reviewed in Section II of this article, Taylor and I studied the spatial mental models constructed from route or survey descriptions of environments. We found evidence that such models captured the spatial relations of the parts of an environment in a perspective-free manner. However, we asserted that verifying statements required taking specific perspectives, and other research has shown that language induces specific spatial perspectives (e.g., Bly, 1989; Bower, 1978). In experiments to be reviewed now, Franklin, Bryant, and I have begun to examine specific spatial relations and specific perspectives in detail (Bryant, Tversky, & Franklin, in press; Franklin & Tversky, 1990a, 1990b).

The basic situation we chose to study is the one we humans find ourselves in for most of our lives, of being surrounded by objects, and keeping track of the directions of those objects from our bodies as we change position. Not only is this situation familiar, it also serves as the basis for techniques as old and revered as the method of loci for memorization, and as contemporary and popular as computer adventure games. Unlike most imagery tasks, it also has the interesting property of being three-dimensional, not in the sense of depth of field in front of the observer (as in the tasks of Pinker, 1980), but in the sense of surrounding the observer, where not all of the field can be "viewed" from any given position.

A. Task

As before, the scenes were described rather than actually viewed. Also as before, readers were not given any instructions to image (except in one study directed at that), or any diagrams or special training. The mental models constructed seem to be a natural consequence of instructions to comprehend and learn the narratives. In the first set of experiments, Franklin and I (1990b) developed 10 different narratives written in the second person; each first described the scene, including the locations of five critical objects in front of you, in back of you, to your left or your right, beyond your head, and beyond your feet. Next, the narrative oriented you the reader toward one of the three horizontal objects and queried you about the objects located at the five possible locations. Then, the narrative oriented you toward another of the three objects and repeated the questions in random order, and so on. Reaction times to identify the objects located at head, feet, front, back, right, and left were the dependent variable of interest.[2]

The hotel scene will serve as an example. The critical objects, those whose locations will be queried, are here in italics, but they were not in the versions subjects read.

[2] These experiments used only five objects to keep memory load at a minimum; later experiments using six objects obtained the same results.

> You are at the Jefferson Plaza Hotel, where you have just taken the escalator from the first to the second floor. You will be meeting someone for dinner in a few minutes. You now stand next to the top of the escalator, where you have a view of the first floor as well as the second floor. You first look directly to your left, where you see a shimmering indoor *fountain* about 10 yards beyond a carpeted walkway. Though you cannot see beyond the low stone wall that surrounds it, you suppose that its bottom is littered with nickels and pennies that hotel guests have tossed in. The view down onto the first floor allows you to see that directly below you is a darkened, candle-lit *tavern*. It looks very plush, and every table you see seems to be filled with well-dressed patrons. Looking directly behind you, you see through the window of the hotel's *barbershop*. You can see an older gentleman, whose chest is covered by a white sheet, being shaved by a much younger man. You next look straight ahead of you, where you see a quaint little *giftshop* just on the other side of the escalator. You're a sucker for little ceramic statues, and you squint your eyes to try to read the hours of operation posted on the store's entrance. Hanging from the high ceiling directly above you, you see a giant *banner* welcoming the Elks convention to the hotel. It is made from white lettering sewn onto a blue background, and it looks to you to be about 25 feet long.

Thus, you might first be oriented toward the barber shop, then the fountain, and then the gift shop (orders were counterbalanced), and at each point queried about what was to your head, feet, front, and so on (also counterbalanced).

B. Equiavailability and Mental Transformation Models

Three classes of models to account for the response times were considered. According to the *equiavailability* model, all locations are equally available to the observer, as they would be in a picture or viewed scene; that is, no direction has priority over any other (Levine, Jankovic, & Palij, 1982; Sholl, 1987). But the scene described is a three-dimensional one, with the observer embedded in it. The equiavailability model makes more sense for a situation where the observer is outside the scene, looking on. According to the *mental transformation* model, the reader imagines him- or herself facing the designated object, and then mentally turning to the cued direction to verify the object. The classical models of imagery (e.g., Finke & Shepard, 1986; Kosslyn, 1980) based on internalized perception and mental transformations suggest imposing a mental transformation on an internalized equiavailable scene. It is as if the reader were viewing the scene and turning to inspect the cued direction to see what object is there. In this case, reaction times should increase with the objects' angular disparity from straightahead. This "mental rotation" is similar to that studied by Shepard and Cooper (1982) in that the observer imagines her- or himself perceiving an environment, but different in that the observer imagines him- or herself turning rather than imagining an object rotating. In imagery tasks where subjects were asked to make left/right (or same/mirror image) judgments on pictures of hands (Cooper & Shepard, 1975; Parsons, 1987b; Sekiyama, 1982) or bodies with outstretched arms (Parsons, 1987a), reaction times indicated that subjects mentally moved their own bodies or parts of their bodies in an analog fashion to the depicted orientation to make the judgment.

Forming and transforming mental images (to transform a phrase of Shepard's) are effortful processes and may not be used when a simpler method of verification is available. Franklin and I (1990b) suggested that what readers in this task do is construct what we termed a *spatial framework,* or mental scaffolding, for keeping track of objects located in the directions of the three axes defined from our bodies. Our conceptions of space, unlike our perceptions of space, may give precedence to certain directions over others, rendering them more accessible. This is suggested by several analyses of spatial language, which in turn are based on asymmetries of the way human observers typically interact with the world (e.g., Clark, 1973; Levelt, 1984; Shepard & Hurwitz, 1984). These analyses served as a basis for our own.

C. Spatial Framework: Upright Case

According to the spatial framework model, the canonical position of the observer is upright, and the canonical world of the observer can be described by one vertical and two horizontal dimensions. The vertical dimension is correlated with the long axis of the body, an asymmetric axis. It is also correlated with gravity, which renders shapes of objects and movement in the world asymmetric along that axis. Canonical movement is horizontal, under which vertical spatial relations generally remain constant with respect to the observer, but horizontal spatial relations change. Whereas there are environmental reference points for the vertical dimension—the sky and the ground, for example—the reference points for the two horizontal dimensions are more arbitrary and changing, often defined only by the prominent dimensions of the observer's body. Thus, for the upright observer, the vertical dimension is predominant. Of the two horizontal dimensions, the front/back dimension predominates over the left/right. The former is asymmetric perceptually and functionally: the observer can more readily see, attend to, and move toward the front than the back. The left/right dimension, in contrast, is derived from the front/back, and has no salient asymmetries. Thus, for the upright observer, this model predicts that the vertical dimension should be fastest, followed by front/back, followed by left/right. In addition, it predicts that front should be faster than back. Data consistent with this analysis were obtained by Hintzman, O'Dell, and Arndt (1981), who asked subjects to point to real or imagined objects arrayed in a horizontal circle around the subject. They found correlations between degree of rotation and reaction time for a real scene (akin to mental rotation), but not for an imagined scene, where response times were fastest to front, followed by back, and slowest but equal to left and right.

D. Experiments 5–7

We ran three separate upright-only experiments. In the first, the direction terms were those most popular in pilot testing: above, below, ahead, behind, left, and right. In the second experiment, we switched to direction terms derived from body parts: head, feet, front, back, left, and right. This meant that the terms were homogeneous, i.e., all referred to surfaces of the body and would allow later comparison to cases where the observer is not upright. In the third experiment, subjects were given explicit instructions to imagine themselves in the scene and to imagine themselves mentally turning to inspect the cued direction for the object. Despite those differences, the pattern of data obtained was the same, corresponding to that predicted by the spatial framework model: fastest reaction times to the vertical dimension, head/feet, next fastest to front/back, with front faster than back, and slowest to left/right (see Table I). The equiavailability model was rejected by any systematic effects of direction on reaction time, and the mental transformation model was rejected both because reaction times to the smallest angular displacement (front) were not the fastest and because reaction times to the largest angular displacement (back) were faster than those to smaller angular displacements (left and right).

E. Spatial Framework: Reclining Case

There is potentially a problem with the previous conclusions. Times were fastest to the vertical axis, but the objects located at those axes were constant; unlike the other directions, the objects did not change as the reader/observer was reoriented. Moreover, the fastest times were to vertical, which may have a privileged status independent of asymmetries of the

TABLE I

MEAN RESPONSE TIMES (SEC) FOR EACH DIMENSION FOR ALL EXPERIMENTS[a]

Dimension	Experiment 1	2	3	4	5
Upright					
Head/feet	1.57	1.36	1.59	—	1.50
Front/back	1.84	1.58	1.81	—	1.72
Left/right	2.21	2.02	2.26	—	2.07
Reclining					
Head/feet	—	—	—	2.42	2.14
Front/back	—	—	—	2.26	1.82
Left/right	—	—	—	3.25	2.59

[a] From Franklin and Tversky (1990b). Copyright © 1990 by the American Psychological Association. Reprinted by permission of the publisher.

body. The solution to this problem is to use narratives in which the observer is horizontal, and reorients by rolling from side to front to back at random.

Although the predictions from the equiavailability and mental transformation models are the same for the reclining case, the predictions for the spatial framework are not. When the observer is reclining, the vertical axis of the world no longer corresponds to any axis of the body. For this reason, for reclining, according to the spatial framework model, only the relative salience of the body axes determines the speed of accessibility, not the relations of the body to the world. Clearly, the left/right axis is least salient, having no asymmetries and being dependent for definition on the front/back axis. Both the front/back and the head/feet axes have asymmetries; however, the front/back axis seems to dominate the head/feet axis, especially given that pedal locomotion is not possible reclining. The front/back axis still separates the world that can be perceived and manipulated from the world that cannot be perceived and manipulated. Thus, the spatial framework model predicts that for the reclining observer, accessing objects along the front/back axis should be fastest, followed by the head/feet, and the left/right last.

F. EXPERIMENTS 8 AND 9

Two experiments investigated the reclining case by adapting the previous narratives and using the previous procedures. In the first of these, all narratives used only the reclining position; in the second, all narratives used both reclining and upright positions, counterbalancing order. Within a narrative, all three reorientations for upright or reclining were blocked.

Adopting a reclining perspective and accessing information from it appear to be more difficult than adopting an upright posture. It took much longer to answer all questions when the observer was described as reclining, in both the pure and mixed experiments. As before, the pattern of responding conformed to the spatial framework predictions and not to the predictions of equiavailability or mental transformation (see the last two columns of Table I). For the reclining case, the objects located at head and feet were still constant but the head/feet axis was not the fastest, indicating that constancy was not responsible for the rapid reaction times for head/feet in the upright posture. To test whether there was a special advantage to verticality, we grouped the front, back, left, and right reaction times that were on the vertical axis. The responses to vertically oriented objects were slower than those to front and back, so there is no special status to vertical when it is not reliably aligned with a body axis. Finally, because of the interaction of posture and direction, the differences in reaction times cannot be attributed solely to the direction terms. Consistent with the spatial framework, then, when the observer is upright, times are fastest to access objects at head/feet, then front/back, and then left/right. When the observer is reclining, times are fastest for front/back, followed by head/feet, and then left/right.

G. Extensions

Thus, the spatial framework accounts for the pattern of responding for both upright and reclining observers. The narratives were written in the second person to induce the reader to put her- or himself inside the scene described by the narrative. Novelists and journalists induce readers to identify with their characters without using this device. In subsequent experiments, we (Bryant et al., in press) found that readers adopted the perspective of observers described in the third person, rather than taking an outsider's perspective, and could take the perspective of same-sex and opposite-sex observers with equal ease. We also found that readers spontaneously adopted the perspective of a central inanimate object. Inanimate objects necessitated some change of terminology, namely, "head" to "top" and "feet" to "bottom," and that change of terminology slowed the responses to those terms in the reclining case. This was probably due to the conflicting meanings of "top" and "bottom" both as certain sides of objects and as upward-pointing or downward-pointing sides (see Clark, 1973). In yet other extensions, we probed with objects for directions, rather than vice versa as in the previous studies, and obtained the same pattern of results.

H. New Directions: Other Perspectives

1. External vs. Internal Arrays

In the narratives we have studied so far, the central character was surrounded by an array of objects, and the questions put to subjects were about the spatial relations of the objects to that character. Readers adopted the perspective of the central character when the character was described as *you,* when the character was described as a third person, and even when the "character" was an inanimate object. Other perspectives and arrays are possible. In one study, we (Bryant et al., in press) described a cubic array of objects from the point of view of an upright outsider looking into the array, and questions were from the same point of view. As in the original experiments (Franklin & Tversky, 1990b), the use of spatial reference terms was deictic (Fillmore, 1975; Levelt, 1984; Miller & Johnson-Laird, 1976), but here the point of view was external to the array. Thus, *the pumpkin is left of the ghost* meant to the left from the observer's point of view (not from the pumpkin's), and *the pumpkin is in front of the ghost* meant that the pumpkin was closer to the observer along the same line of sight as the ghost (rather than the pumpkin was in front of the ghost's front).

Spatial framework reasoning can be adapted to this situation, where the array is external to the observer. Because the array is in front of rather than surrounding the observer, the reader can keep in mind a two-dimensional projection of a three-dimensional scene, rather than a three-dimensional scene surrounding the observer. The conceptual field of view in this case is smaller and more compact. This should be easier to keep in mind.

Again, because the array is in front of rather than surrounding the observer, the spatial framework analysis depends more on the axes of the world and the field of view of the observer than on the asymmetries of the observer's body axes. Many of the same predictions of the spatial framework for upright posture with surrounding array hold, but for different reasons. For the external array, above/below is determined in large part by the vertical axis of the world because the objects are not directly above or below the observer, but rather above or below each other. As before, gravity is aligned with the vertical axis of the world, and confers asymmetry on it. Of course, the head/feet axis of the observer is also aligned with the vertical. For an outsider, objects are directly in front of the body, but not directly behind; rather, objects described as behind objects in front are also in front, but farther from the observer than the objects described as in front. This axis still has an asymmetry, i.e., front objects are closer, and relatively larger in size than behind objects. This is a weaker asymmetry than in the internal case, where objects were in front and in back of the observer. In that case, objects to the front could be seen and objects to the back could not; here both front and back objects can be seen, but front objects are closer. As before, there is no asymmetry along the left/right axis; objects are equally close. Thus, the overall ordering of dimensions is predicted to be the same for the upright external perspective case as for the upright internal case: above/below should be fastest, followed by front/back, and then left/right. The advantage of front over back, however, should diminish or disappear. These predictions were obtained, and replicated in an experiment describing an outside observer examining an array of objects surrounding another character.

2. *Two-Person Situations*

Franklin and I have begun studying more complex cases with two observers (Franklin & Tversky, 1990a). Thus far, we have investigated narratives that described a set of objects around each of the characters, and readers were queried about the locations of the objects around each character relative to that character. Thus, one way readers could perform the task is by taking the perspectives of each of the characters in turn. If so, the upright internal spatial framework pattern of responses should appear for each character. Another strategy readers could take is to adopt a single survey perspective on both characters at once. In the former case, in order to answer questions, readers construct two smaller spatial mental models, one for each character, and switch back and forth. In the latter, readers construct a single large spatial mental model and switch focus within.

It appeared that readers adopted both of these strategies, depending on the situation described. In a "neutral" situation, where both characters were described as near each other but surrounded by different sets of objects, readers seemed to adopt a single survey perspective. Thus, they appear to prefer to use a larger, integrative mental model and constant perspective to shifting perspective between two smaller mental models. In a second study, both characters were described as being in such different scenes that it was difficult to construct a single unifying perspective, e.g., one person in a lagoon and the other in a museum. Then readers adopted the perspective of each of the characters in turn, yielding an upright spatial framework pattern around each character. In a third experiment, we provided readers with explicit bird's eye perspectives (e.g., from a helicopter; from a museum with a glass roof) on these scenes, and readers again chose a single integrative survey perspective. In this case, the data fit the equiavailability pattern, i.e., all directions were equally quick. This makes sense under analysis. If the reader's point of view is above the scene, then the axes of the observer and of the observer's world are not aligned with the axes of the characters and the axes of their world, unlike the previous external perspective. Given that the observer's body axes and world are misaligned with the characters' body axes and world, and given that questions are from the point of view of the character, there is no reason for any particular axis to predominate any other. In other words, in this situation the axes are treated arbitrarily and equally.

These extensions, to new arrays and to new perspectives, have led to modifications of the spatial framework analysis. That analysis is based on considerations of the body axes and the perceptual world from different perspectives. Thus, the spatial framework is more properly regarded as a family of related variants, deriving from the same set of general principles (similar to Lakoff's, 1987, "image schemas").

I. Spatial Frameworks

In these experiments, readers read narratives describing arrays of objects around observers, other characters, or other objects, and were later probed for objects by directions (or vice versa). The pattern of reaction times to access information from the spatial mental models did not show the analog, perceptual characteristics typical of imagery tasks. The pattern did correspond to the spatial framework model, according to which readers construct a mental scaffolding to keep track of the directions of objects from their bodies and each other, which can be updated as the situation changes.

The spatial framework derived from an analysis of our canonical interaction with the perceptual world, the asymmetries of that world and our bodies, posture, and perspective. The world as we view it has one vertical and two horizontal axes. The vertical axis is correlated with gravity, which exerts a considerable asymmetric force on the world, constraining how the world looks and how we maneuver in it. Moreover, the vertical has natural anchors in the environment: the ground and the sky for outside environments, floors and ceilings for indoors. In contrast, the two horizontal dimensions are not correlated with environmental forces or anchored to features in the environment. In many situations, then, two natural axes of our own bodies—the front/back and left/right axes—serve as reference points for horizontal axes of the world. Although the left/right axis is essentially symmetric, the front/back axis is not; both perception of the environment and manipulation of it are natural frontward, but difficult, if at all possible, backward. The third axis of the body, the head/feet axis, not only has asymmetry but also correlates with the vertical axis of the world, and with gravity, i.e., when we are in canonical upright orientation. For the upright observer surrounded by an array of objects, then, both body and environmental factors contribute to the predominance of the vertical head/feet axis. Body factors lead to the predominance of front/back over left/right.

The environmental and body factors change as the perspective of the observer changes. For a reclining observer, there is no body axis correlated with the distinguished environmental axis, gravity, so the predictions derive only from consideration of the body. Because the head/feet axis is not correlated with vertical when the body reclines, the front/back asymmetry looms larger than the head/feet, and the left/right remains least distinguished. The spatial framework analysis was confirmed for these two cases, upright and reclining observers, surrounded by arrays of object. Preliminary work has begun extending the spatial framework to other perspectives and arrays, yielding a family of spatial frameworks, i.e., situation-specific variants based on the same general principles. Thus, systematic exploration of people's responses to imaginary environments has revealed some of the ways we conceive of the visual world.

IV. Summary

Readers of spatial descriptions spontaneously construct spatial mental models of the described scenes as a natural consequence of reading for comprehension and memory, with no special training, instructions, or prior visual displays. Of courses, readers do not necessarily construct spatial models from all text; the text must be spatial, coherent, well integrated, and more or less determinate, among other characteristics (e.g., Denis & Denhiere, 1990; Ehrlich & Johnson-Laird, 1982; Mani & Johnson-Laird, 1982; Perrig & Kintsch, 1985). The spatial mental models constructed reveal people's conceptions of space, which, though built on their perceptions of space, are more abstract and general.

In the first set of experiments (Taylor & Tversky, in press), subjects read route or survey descriptions of four environments, and verified verbatim and inference statements about those environments from both the same and the other perspective. Subjects were equally fast and accurate in verifying inference statements from the read perspective and the other perspective. This led us to the conclusion that subjects' mental models capture the categorical spatial relations described in the text, but not from any particular perspective. Like structural descriptions, spatial mental models contain information about the parts of a scene and the relations between the parts. Unlike images, which have been likened to internalized perceptions, spatial mental models are perspective-free and allow the taking of many perspectives, required in order to verify the test statements.

The second set of studies examined perspective taking and information retrieval in a particular (imaginary) environment, one that is simple and common, that of an observer surrounded by objects (Franklin & Tversky, 1990b). We found that times to report what objects lie at six canonical directions from the observer (at head or feet, to the left or right, in front or back) differed reliably and systematically depending on the direction of the object and the posture of the observer. The times could not be accounted for by a model that assumed that readers imagined themselves in the place of the observer, and imagined themselves rotating in place to ascertain what objects are at what directions. Rather, the reaction times were accounted for by an analysis of how space is conceived in relation to the body, yielding a family of what we termed spatial frameworks.

We opened with the problem of understanding directions, instructions, and narratives, and observed that constructing a mental model of the situation described in the directions, instructions, or narrative not only seemed useful but also seemed to be what readers and listeners do when the conditions are right. The experiments reported here have added to that body of research, uncovering many features of spatial mental models in the process. Consider the ladder in the following passage from F. Scott Fitzgerald (1922/1950). "Fifth and Sixth Avenues, it seemed to Anthony, were the uprights of a gigantic ladder stretching from Washington Square to Central Park. Coming up-town on top of a bus toward Fifty-second Street invariably gave him the sensation of hoisting himself hand by hand on a series of treacherous rungs, and when the bus jolted to a stop at his own rung, he found something akin to relief as he descended the reckless metal steps to the sidewalk" (p. 10). On the one hand, the ladder describes the appearance of that part of the city, two broad avenues, anchored in one park and reaching toward another, with many narrow cross-streets. Yet the ladder is also used to convey the effort and precariousness of coming uptown, fighting against gravity on an unwieldy apparatus. Spatial mental models do more than capture a physical setting; instilled by a gifted writer, they are replete with meaning.

ACKNOWLEDGMENTS

I have been fortunate to have Nancy Franklin, Holly Taylor, and David Bryant as collaborators, and I am grateful for their insightful comments on a previous draft of this manuscript. The research reported here was supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under Grant #AFOSR 89-0076 to Stanford University.

REFERENCES

Abelson R, (1975). Does a story understander need a point of view? In R. Schank & B. L. Nash-Webber (Eds.), *Theoretical issues in natural language processing.* Washington, DC: Association for Computational Linguistics.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17,* 1–12.

Bly, B. (1989). *Perspective in mental models of text.* Unpublished manuscript, Stanford University, Stanford, CA.

Bower, G. H. (1978). Experiments on story comprehension and recall. *Discourse Processes, 1,* 211–231.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3,* 193–209.

Bryant, D., Tversky, B., & Franklin, N. (in press). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language.*

Byrne, R. M. J., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language, 28,* 564–575.

Byrne, R. W. (1979). Memory for urban geography. *Quarterly Journal of Experimental Psychology, 31,* 147–154.

Chase, W. G. (1983). Spatial representations of taxi drivers. In D. R. Rogers & J. A. Sloboda (Eds.), *Acquisition of symbolic skills.,* New York, Plenum.

Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language,* (pp. 27–63). New York: Academic Press.

Cooper, L., & Shepard, R. N. (1975). Mental transformations in the identification of left and right hands. *Journal of Experimental Psychology: Human Perception and Performance, 104,* 48–56.

Denis, M., & Cocude, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology, 1,,* 293–307.

Denis, M., & Denhiere, G. (1990). Comprehension and recall of spatial description. *European Bulletin of Cognitive Psychology, 10,* 115–143.

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior, 21,* 296–306.

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

Fillmore, C. J. (1975). *Santa Cruz lectures on deixis.* Bloomington: Indiana University Linguistics Club.

Finke, R. A., & Shepard, R. N. (1986). Visual functions of mental imagery. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance.* New York, Wiley.

Fitzgerald, F. Scott, (1950). *The beautiful and the damned.* New York: Scribner's (Original work published 1922).

Foos, P. W. (1980). Constructing cognitive maps from sentences. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 25–38.

Franklin, N. (1991). *Representation of spatial information in described routes: Distance, turns, and objects.* Manuscript, State University of New York-Stony Brook.

Franklin, N., & Tversky, B. (1990a). *Mental spatial frameworks for different perspectives.* Poster presented at the Psychonomic Society meetings, New Orleans, LA.

Franklin, N., & Tversky, B. (1990b). Searching imagined environments. *Journal of Experimental Psychology: General, 119,* 63–76.

Garnham, A. (1981). Mental models as representations of text. *Memory & Cognition, 9,* 560–565.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26,* 69–83.

Hintzman, D. L., O'Dell, C. S., & Arndt, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology, 13,* 149–206.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13,* 208–217.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? Evidence from studies of information retrieval time. *Memory & Cognition, 4,* 291–297.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL: University of Chicago Press.

Levelt, W. J. M. (1982). Linearization in describing spatial networks. In S. Peters & E. Saarinen (Eds.), *Processes, beliefs, and questions.* (pp. 199–220). Dordrecht, The Netherlands: Reidel.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. de Grind, & J. J. Koenderink (Eds.), *Limits in perception.* Utrecht, The Netherlands: VNU Science Press.

Levine, M., Jankovic, I., Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General, 111,* 157–175.

Linde, C., & Labov, W. (1975). Spatial structures as a site for the study of language and thought. *Language, 51,* 924–939.

Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 15–32.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10,* 181–187.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* New York: Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B, 200,* 269–291.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18,* 87–121.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision.*, New York: McGraw-Hill.

Moar, I., & Bower, G. H. (1983). Inconsistency in spatial knowledge. *Memory & Cognition, 11,* 107–113.

Morrow, D. G., Bower, G. H., & Greenspan, S. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28,* 292–312.

Morrow, D. G., Greenspan, S., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26,* 165–187.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American, 224,* 34–43.

Novick, L. R., & Tversky, B. (1987). Cognitive constraints on ordering operations: The case of geometric analogies. *Journal of Experimental Psychology: General, 116*, 50–67.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9*, 441–474.

Parsons, L. (1987a). Imagined spatial transformation of one's body. *Journal of Experimental Psychology: General, 116*, 172–191.

Parsons, L. (1987b). Imagined spatial transformation of one's hands and feet. *Cognitive Psychology, 19*, 178–241.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24*, 503–518.

Pinker, S. (1980). Mental imagery and the third dimension. *Journal of Experimental Psychology: General, 109*, 354–371.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18*, 1–63.

Sekiyama, K. (1982). Kinesthetic aspects of mental representations in the identification of left and right hands. *Perception & Psychophysics, 32*, 89–95.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press/Bradford Books.

Shepard, R. N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition, 18*, 161–194.

Shepard R. N., & Podgomy, P. (1978). Cognitive processes that resemble perceptual processes. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. Hillsdale, NJ: Erlbaum.

Sholl, M. J. (1987). Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 615–628.

Stevens, A. L., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422–437.

Streeter, L. A., Vitello, D., & Wonsiewicz, S. A. (1985). How to tell people where to go: Comparing navigational aids. *International Journal of Man-Machine Studies, 22*, 549–562.

Taylor, H. A., & Tversky, B. (1990). *Spatial descriptions and depictions*. Paper presented at the Psychonomic Society meetings, New Orleans, LA.

Taylor, H. A., & Tversky, B. (in press). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*.

Thorndyke, P. W. (1981). Spatial cognition and reasoning. In J. Harvey (Ed.), *Cognition, social behavior, and the environment*. Hillsdale, NJ: Erlbaum.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14*,, 560–589.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69*, 344–354.

Tversky, B. (1981). Distortions in cognitive maps. *Cognitive Psychology, 13*, 407–433.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193–254.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wagener-Wender, M., & Wender, K. F. (1990). Expectations, mental representations, and spatial inferences. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension*, (pp. 137–157). San Diego, CA: Academic Press.

Wilton, R. N. (1979). Knowledge of spatial relations: The specification of information used in making inferences. *Quarterly Journal of Experimental Psychology, 31*, 133–146.