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

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION, VOL. 27

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

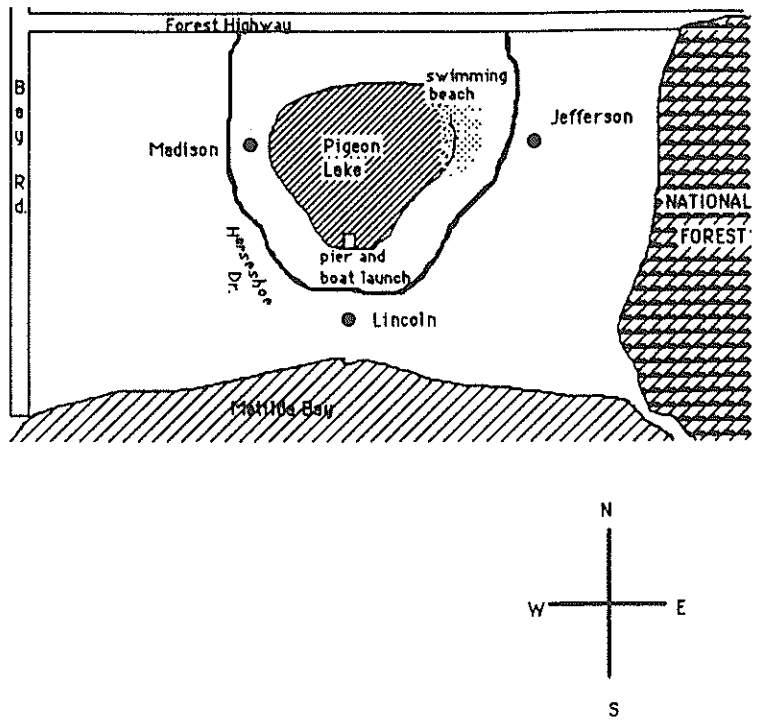

Fig. 1. Map of resort area. From Taylor and Tversky (in press). Reprinted by permission.

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

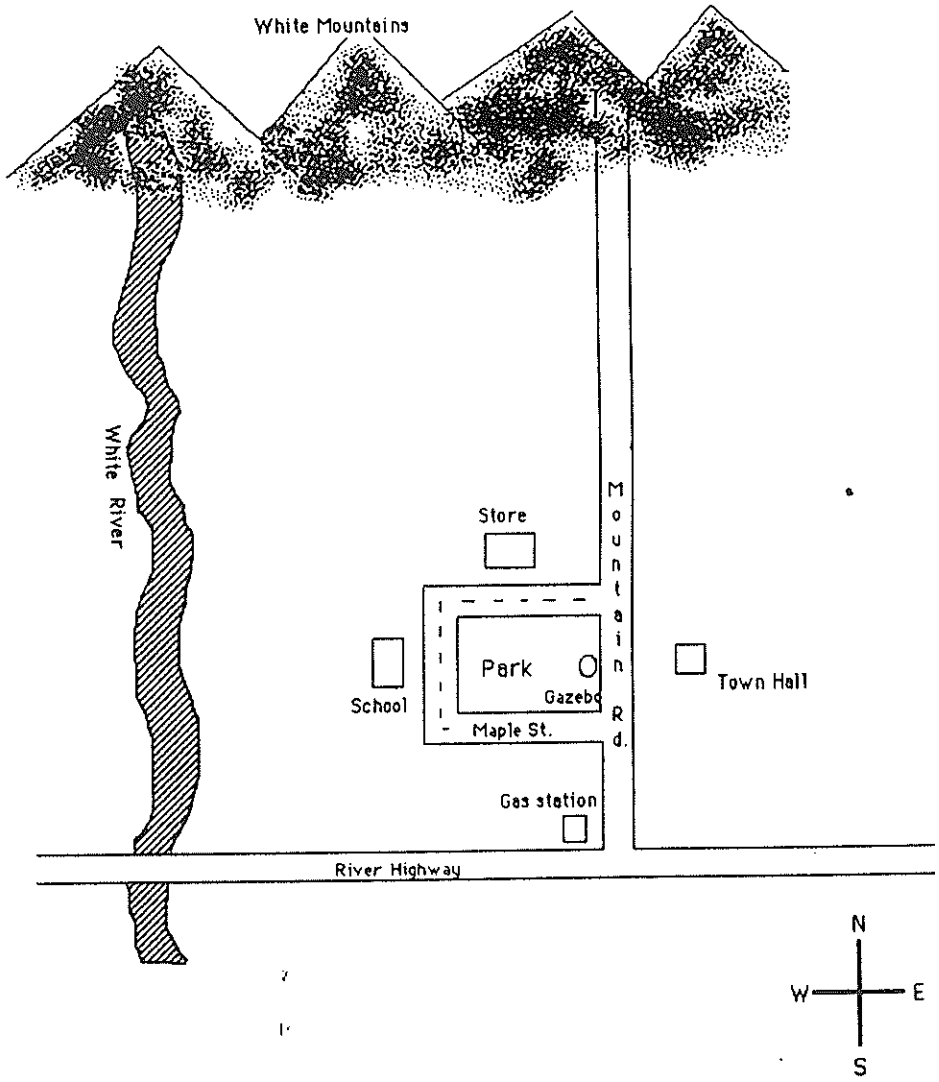

Fig. 2. Map of town. From Taylor and Tversky (in press). Reprinted by permission.

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

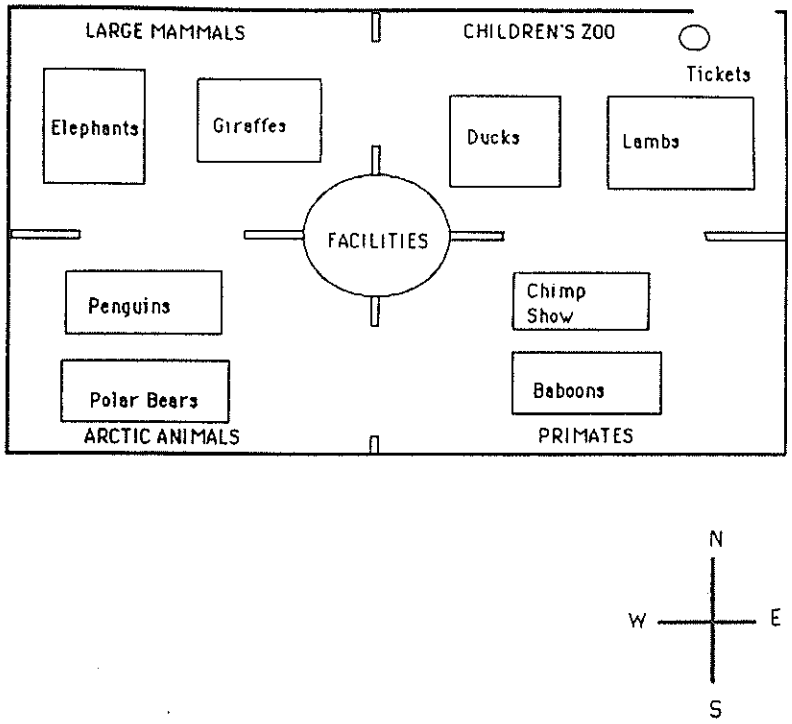

Fig. 3. Map of zoo. From Taylor and Tversky (in press). Reprinted by permission.

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

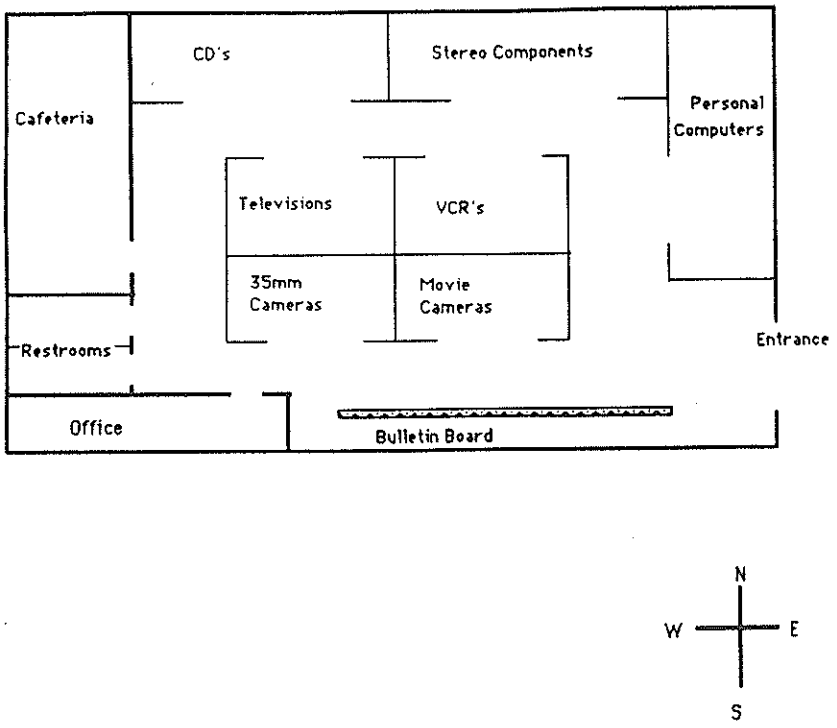

Fig. 4. Map of convention center. From Taylor and Tversky (in press). Reprinted by permission.

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-

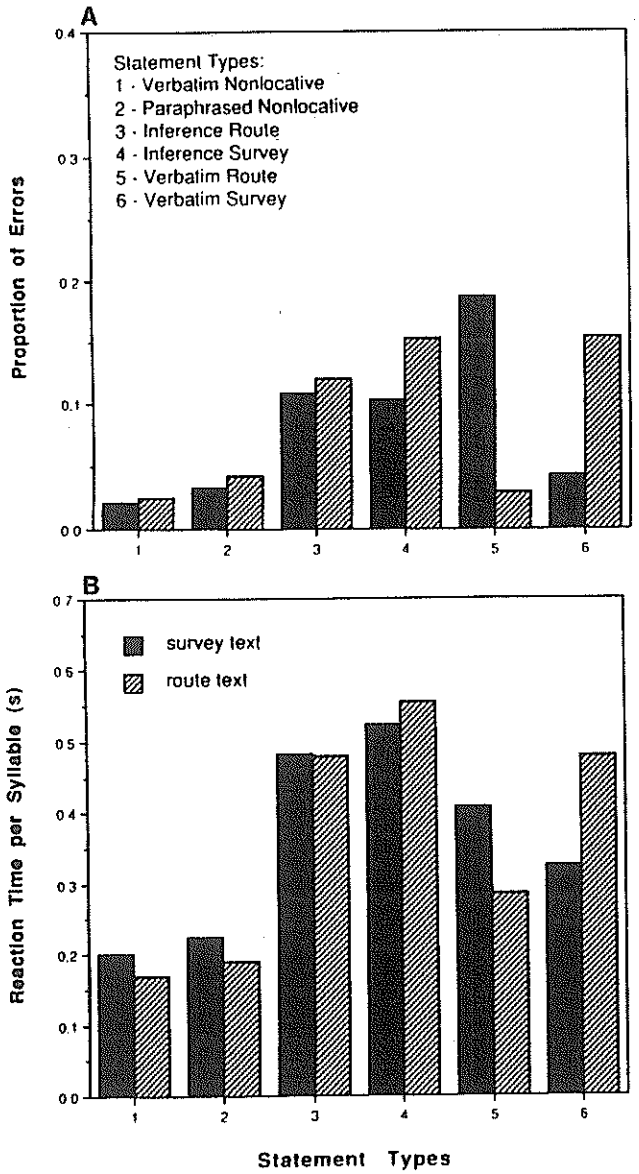

Fig. 5. Experiment 1. A, errors to question types by description type. B, reaction times to question types by description type. Adapted from Taylor and Tversky (in press). Reprinted by permission.

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

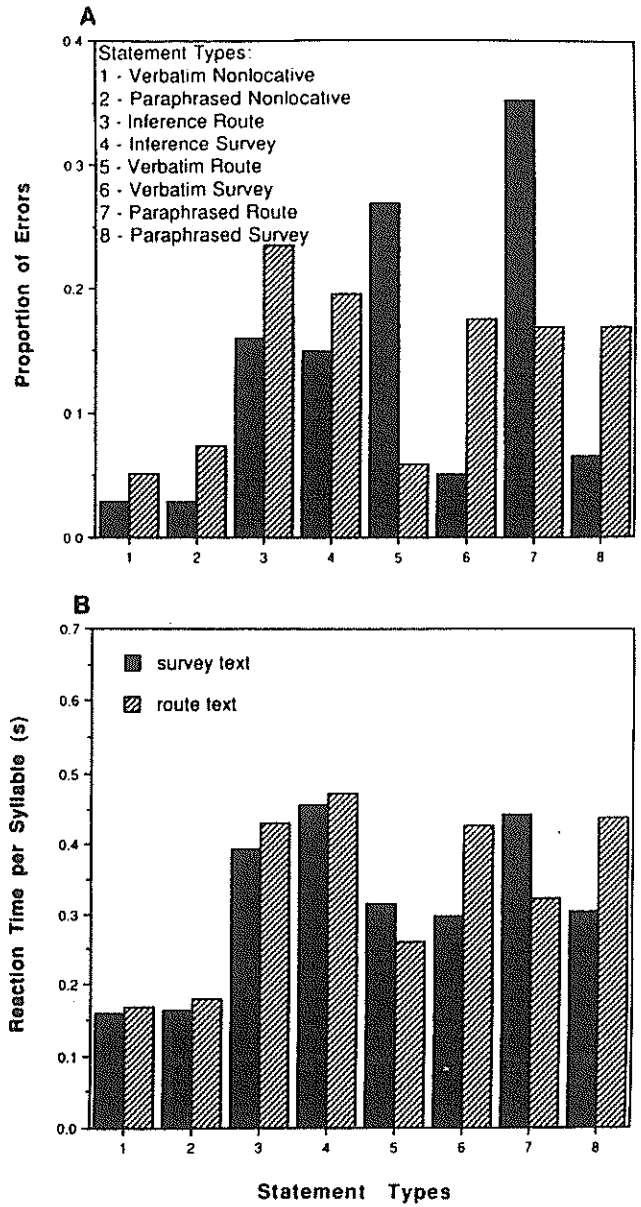

Fig. 6. Experiment 2. A, errors to question types by description type (paraphrased questions). B, reaction times to question types by description type (paraphrased question). Adapted from Taylor and Tversky (in press). Reprinted by permission.