Abstract

According to the classical view of imagery, images have analog properties and are like internalized perceptions. Mental representations of some spatial situations, such as the one investigated here, do not conform to that view. In the present research, people studied spatial scenes consisting of an array of objects either around themselves or around a doll. To direction probes, they reported the objects lying in different directions from the body either from observation or from memory. The patterns of retrieval times from observation and memory differed, and were not analog in memory, indicating that mental representations of these scenes are more like mental models than like images. Three models that account for the behavior in the different tasks are presented. In contrast to images, mental models reflect conceptions of space rather than perceptions of it. Mental models are more schematic or categorical than images and incorporate knowledge about the world that is not purely perceptual.

Retrieving Spatial Relations from Observation and Memory

Knowledge about our spatial surroundings is acquired from many different modalities, from sight, from sound, from touch. Knowledge about space may also be acquired from language. Retrieval time and memory accuracy data indicate that vivid spatial descriptions can induce mental representations that reflect relative directions (e. g., Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990; Mani & Johnson-Laird, 1982; Perrig & Kintsch, 1985; Taylor & Tversky, 1992) and relative distance (e. g., Denis & Cocude, 1989, 1992; Glenberg, Meyer, & Lindem, 1987; Morrow, Bower, & Greenspan, 1989; Wagener-Wender & Wender, 1990; but see also, Gray Wilson, Rinck, McNamara, Bower, & Morrow, 1993). Mental representations induced by language not only represent the spatial arrangement of characters, objects, and landmarks, but also allow changes in perspective (Franklin, Tversky, & Coon, 1992) and updating relative positions as new information becomes available (e. g., Bryant et al., 1992; Franklin & Tversky, 1990; Glenberg et al., 1987; Morrow et al., 1989).

The fact that complex spatial properties are spontaneously and coherently represented and updated in mental representations constructed from discourse suggests that these representations resemble those constructed from experience. In the present experiments, we directly compare mental representations of space that are constructed from descriptions with those that are constructed from actual experience. Several approaches - specifically imagery, spatial frameworks, and intrinsic computation - predict different patterns of results. We will describe each in turn.

Imagery

The similarities between spatial properties of mental representations formed from language and spatial properties of the world have led many to conclude that these mental representations are like mental images. The classic accounts of imagery have stressed the similarity of images to perceptions, and of transforming images to transforming perceptions (see Farah, 1988; Finke, 1980; Finke & Shepard, 1986; Kosslyn, 1980; Paivio, 1986; Shepard & Podgorny, 1978, among others). In support of this position, several studies have demonstrated that examining an image is like examining a percept of an object. For example, searching an image of an animal for a larger part takes less time than searching an image for a smaller part (Kosslyn, 1976). In visual search, too, larger things are located faster than smaller ones due to visual acuity. Other experiments have demonstrated that transforming an image is like perceiving a transformation. For example, the time it takes to mentally rotate an image is proportional to the angle of rotation (Shepard & Cooper, 1982). Similarly, mentally scanning a long distance takes longer than mentally scanning a short distance (Kosslyn, Ball & Reiser, 1978). Finally, images, like percepts, have a specific spatial perspective (Pinker, 1980).

Together, these findings have been used to support the view that imagery is like internalized perception (Kosslyn, 1980; Shepard & Podgorny, 1978). Images and transformations of images appear to be analog, as they are assumed to be in perception. They are bound to a specific perspective, like perception. Recent neuroimaging data indicate that imagery can activate the same brain structures as perception (Kosslyn, et al., 1993). This account of imagery has intuitive appeal as well. It provides an account for the origin of mental images and a mechanism for the transformation of them, through internalizing perception. It also provides a rationale for their existence, to support visual and spatial thinking and memory in the absence of perception. A strong prediction from this position is analog performance. A weaker prediction is identical performance in perception and imagery.

With rare exceptions, imagery experiments have not included perception conditions, so what happens in perception can only be surmised (but see Denis & Cocude, 1992; Denis, Goncalves & Memmi, 1995; Denis & Zimmer, 1992). Moreover, because using imagery is an optional and often burdensome strategy, participants in imagery experiments have frequently been instructed, and sometimes trained, in the use of mental images. Often, the subjective experience of having a mental image is equated with the processing that is revealed in reaction time tasks.

Mental Models

Appealing as the imagery-as-internalized-perception view is, there is growing evidence for mental representations of objects and space that do not have the properties ascribed to images. The inspirations for the classical view of imagery have been among those who have noted this (e.g., Kosslyn, 1987; Finke & Shepard, 1986). An alternative kind of mental representation has been proposed for such situations, namely, spatial mental models. The term mental model has been used in different ways, sometimes causing confusion. By some, the term has been used to characterize people's mental representations of dynamic in contrast to static systems, such as a doorbell, a calculator, or a steam plant (Gentner & Stevens, 1983; Halasz & Moran, 1983; Kieras & Bovair, 1984; Miller, 1979). Others have used the term used to characterize a component of deductive reasoning, in contrast to formal logic (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). Finally, the term mental model has been used to characterize the mental representations readers construct of situations described by discourse, in contrast to a mental representation of the language of discourse (e. g., Bransford, Barclay & Franks, 1972; Clark & Haviland, 1974; Glenberg, et al., 1987; Tversky, 1991; van Dijk & Kintsch, 1983). Despite these differences, for all, the term mental model is meant in two of its senses, in the sense of capturing some, but not all, aspects of the world and in the sense of being a system of postulates and inferences. A mental model consists of elements and the relations between them. Elements are typically objects in the world, or things that can be conceptualized as objects. Various properties may be ascribed to the objects, though for many purposes, their existence in space and perhaps time is all that is imputed. Relations between objects may be spatial, temporal, causal, or something else.

Mental models of spatial relations among objects are the case considered here. Spatial mental models are more abstract, more schematic, than images (Johnson-Laird, 1983; Taylor & Tversky, 1992). The spatial relations within or between objects need not be analog. Objects may be represented incompletely, even one dimensionally, as markers. An image, then, can be viewed as a special case of a mental model, where the spatial relations are analog and there is a single perspective. Although mental models represent states of the world, they are not necessarily equivalent to perceptions of the world.

Spatial Framework Model

Spatial mental models have received support from a paradigm examining the spatial situation that people are usually in - standing, sitting, or reclining in a setting surrounded by objects (Franklin & Tversky, 1990; Bryant et al., 1992). In the prototypic internal situation (Franklin & Tversky, 1990), participants read narratives describing themselves in a scene, such as a hotel lobby or workshed, with objects, such as a hammer, basket, or saw, located to their fronts, backs, left and right sides, above their heads, and below their feet. After studying a narrative, participants turned to a computer that presented further text orienting participants to face one of the objects, and then probing them with probe terms; *front, right, head*, etc. Participants responded with the object located in that direction. After all directions were probed, participants were reoriented and probed again. Participants made few errors, so the data of interest were response times to the six directions.

Franklin and Tversky (1990) considered three possible models to account for the pattern of reaction times. The data rejected an equiavailability model, according to which all directions should be equally fast, as would be expected in scanning a picture or looking at a scene. The data also rejected a mental transformation model, derived from theories of imagery. According to a mental transformation model, participants should imagine themselves in the scene, and then imagine themselves turning to mentally inspect the probed directions. If that model held, times to respond to probes to the *front* would be fastest, followed by probes 90° from *front*, namely, *left*, *right*, *head*, and *feet*. Probes to *back*, 180° from *front*, would be slowest. This pattern was not obtained. Instead, the data supported the spatial framework model.

According to the spatial framework model, participants construct a mental spatial framework consisting of extensions of the three body axes, head/feet, front/back, and left/right, and associate objects to the framework. The accessibility of an axis depends on characteristics of the body and the world, and the posture of the observer in the narrative (this model is partially based on work by Clark, 1973; Fillmore, 1976; Levelt, 1984; Miller & Johnson-Laird, 1976; and Shepard & Hurwitz, 1984). Table 1 shows the predictions of response times for the spatial framework model.

For an upright observer, the head/feet axis is most accessible because of properties of the world and properties of the body. For the canonically oriented upright person, the head/feet axis is correlated with the environmental axis of gravity, which is the only salient, fixed asymmetric axis of the world. As observers navigate the world, vertical relations among objects remain largely constant with respect to the viewer whereas relations in the horizontal plane change. The head/feet axis is also physically asymmetric, so that physically distinct body sides correspond to head and feet. Both the biological asymmetry and the correlation with gravity impart a special status to the head/feet axis, making it a salient indicator of location. This leads to easy and rapid access of objects beyond the head and feet. The front/back axis is next most salient. Its asymmetry separates the world that can be seen and manipulated from the world that cannot be easily perceived or manipulated. The left/right axis is least accessible as it has no salient asymmetries. Thus, for an upright observer, the spatial framework model predicts that participants will be fastest to identify objects to the *head* and *feet*, followed by *front* and *back*, followed by *left* and *right*. Franklin and Tversky (1990) obtained this pattern of data in four experiments, and parts of the pattern had been obtained by others (e. g., Farrell, 1979; Hintzman, O'Dell, and Arndt 1981; Maki & Braine, 1985; Sholl, 1987; Sholl & Egeth, 1981). In addition, because perceptual and behavioral asymmetries so strongly favor *front* over *back*, the spatial framework model predicts faster response times to *front* than *back* when the observer is surrounded by objects.

Insert Table 1 about here.

When the observer in the scene reclines, the situation changes. In this case, no axis of the body is correlated with gravity, so the relative salience of axes depends solely on the importance of their asymmetries. The perceptual and behavioral asymmetries of the front/back axis are stronger than those of the head/feet axis, so front/back should be faster than head/feet for the reclining case. The left/right axis has the weakest asymmetries, so it should be slowest. This pattern was obtained in two experiments by Franklin and Tversky (1990).[1]

The pattern of retrieval times is not a simple consequence of the verbal labels used. For one thing, the head/feet axis is fastest in the upright condition but the front/back axis is fastest in the reclining condition. Moreover, in several more complex situations with two characters, where participants were required to answer from both points of view, no differences in response times to the three axes were found (Franklin, et al., 1992; but see also Maki & Marek, 1997). Thus, the patterns of response times to retrieve objects located at different axes can be attributed to conceptual factors. These differences were accounted for by variants in the mental models induced by the situations, specifically in perspective in the situation and in interpretations of the axes (see Bryant et al., 1992; Bryant & Wright, 1999; Franklin et al., 1992).

Intrinsic Computation Model

The spatial framework model provides an account of the accessibility in memory of objects located beyond the intrinsic sides of a body. The model holds when participants take the point of view of the character in the scene. It also holds when the scene contains a central object surrounded by other objects rather than a central person, as long as participants adopt the object's point of view. However, a different model, the intrinsic computation model, may be appropriate when viewing a body or an object from an outside viewpoint. In that case, in order to identify the objects at specified directions from a body or object, participants may first determine the intrinsic sides of the body or object using general perceptual mechanisms, and then identify the objects located beyond each intrinsic side. Some intrinsic sides are more readily determined than others (see Table 1 for predictions of the intrinsic computation model). Several lines of research suggest that people first identify the top, and by contrast, the bottom of an object, followed by the *front* and *back*, and lastly, *left* and *right* (Braine, Plastow, & Greene, 1977; Corballis & Cullen, 1986; Jolicoeur, 1985; Jolicoeur, Ingleton, Bartram, & Booth, 1993; Maki, 1986; Rock, 1973). *Left* and *right* can be determined only after the top (*head*), bottom (*feet*), *front*, and *back* are known. Consequently, determining top and bottom should be faster than *front* and *back*, which in turn should be faster than *left* and *right*. Identifying the objects located beyond the intrinsic sides should be the same irrespective of side.

According to this model, the relative accessibility of the intrinsic sides does not depend on the posture of the body or the orientation of the object. Orientations other than upright may (and often do) yield longer times, but the increases should be the same for all sides. Logan (1995) adapted the situation of Franklin and Tversky (1990) and Bryant et al. (1992) to a diagram task. Participants saw diagrams of heads at varying orientations and identified what was located at specified directions from the heads. Logan found what we have termed the intrinsic computation pattern of data, namely, head/feet faster than front/back faster than left/right for all head orientations. Whether the retrieval times correspond to the spatial framework pattern or to the intrinsic computation pattern may be an consequence of the perspective taken on the scene.

The spatial framework and intrinsic computation analyses can easily be distinguished by their predictions for a character that is not upright. According to the spatial framework model, participants mentally take the perspective of the character in the scene. When the character is reclining, the head/feet axis is out of its canonical alignment with gravity and participants are faster for front/behind than head/feet relations. According to the intrinsic computation model, the participant identifies the sides of the character, beginning with the axis of the intrinsic top. As a consequence, participants should be faster to head/feet than front/back at all orientations of the person.

Present Experiments

The first question we posed is whether mental representations established from narrative are functionally equivalent to those established from observation. To address it, we compared participants' patterns of retrieval times for memory of scenes learned by observation to the patterns of retrieval times for memory of scenes learned from descriptions. The second question we posed is whether mental representations for these scenes are like internalized perceptions. To address it, we compared patterns of retrieval times for answering from memory to patterns of retrieval times for answering from observation. In the first experiment, the participant was internal to the scene, that is, surrounded by objects. In the second experiment, the participant was external to the scene, observing a doll surrounded by objects. Thus, we were able to determine the patterns of reaction times for actually observing scenes from two perspectives, one external, and one internal.

Experiment 1: Responding from Memory or Observation: Surrounding Spatial Array

In this experiment, participants learned a spatial array of objects that surrounded them by perceptual observation. Participants stood or reclined on a bench in an empty room. Large pictures of objects were hung on the walls, ceiling, and floor at the six directions from the participant's body. The objects were thematically related (e.g., kitchen objects), and were changed for each scene. Participants responded to direction probes either from memory, or while looking at the scene.

There are two issues of interest. The first is whether response times in the present memory condition will show the same pattern observed in previous research, where memory was established by narrative. The second is whether response times in the present memory condition will show the same pattern as the present perception condition. If the pattern of response times from memory of a perceived scene is the same as the pattern of response times from memory of a described scene, then the claim that spatial mental representations constructed from descriptions are in some way equivalent to those constructed from experience is supported. This inference relies on comparison of data across experiments, which does not allow for ready statistical analysis. The general spatial framework pattern, however, has been independently replicated in memory for narratives in a number of studies, each with several experiments (Bryant et al., 1992; Franklin & Tversky, 1990; Franklin et al., 1992). Although some variability exists among these studies, the critical differences between axes have appeared reliably in nearly a dozen experiments and seem to reflect a stable pattern for comparison.

For the case of perception, if participants look at the direction probed to find the object, then response times should conform to a physical transformation model. Specifically, times to *front* should be fastest, times to *back* should be slowest, and times to the other four directions, all 90 degrees from *front*, should be in between because the time to physically turn to a direction will determine response time. If memory is like internalized perception, as classical views of imagery maintain, the pattern of responding from perception should be like that of responding from memory.

In a pilot study for this experiment, we found that participants in the perception condition quickly learned the arrays and then stopped looking at them. Instead of using the perceptually available array to ascertain which object was at a probed direction, they relied on memory. Thus, in the perception condition of this experiment, we frequently changed the objects so that the array was difficult to learn.

Method

Participants.

Sixteen Stanford University undergraduates (eight male and eight female) participated in the memory condition. Twelve Northeastern University undergraduates (eight males and four females) participated in the perception condition. All participants received credit in an introductory psychology class.

Materials and Equipment.

A set of 42 black and white drawings of common objects was employed to represent seven scenes (one a practice scene) in both the memory and perception conditions. The objects and scenes are listed in Table 2. Each picture was roughly 15x22 cm and mounted on black poster board.

Insert Table 2 about here.

Direction probes were presented by a stereo tape recorder. One channel of the audiotape contained a sequence of direction probes for each scene. The probes were spoken by the experimenter and recorded on the tape. The second channel contained a series of tones that were coordinated with the auditory probes and controlled an electronic timing device. When a probe was presented, a tone on the second channel of the tape sent a signal to start an electronic timer. The timer in the memory condition measured response times accurately to 10 msec, whereas the timer in the perception condition measured response times accurately to 1 msec. Participants used a hand-held response button to stop the timer in the memory condition. In the perception condition, the timer was connected to a voice key. Participants spoke their response into a microphone connected to the voice key, which sent a signal to stop the timer.

The experimental situation is diagrammed in Figure 1. The experimental room for the memory condition was not perfectly square, measuring 9'5" by 11'7", with a ceiling 8'2" high. The experimental room for the perception condition measured 12’ by 12’ with a ceiling 8’1” high. In both conditions, a bench was placed such that one end was at the center of the room (i.e. at the midpoint horizontal and vertical distances of each wall), so that when a participant was standing on the bench he or she would be in the center, and when the participant reclined his or her head would be at the center. Pictures of objects were hung on hooks placed at the center of each wall and the ceiling. The object below the participant was simply placed face up on the floor immediately before the bench. Although distances between objects and the participant were not all equal, the relatively small deviations do not affect predictions of either the spatial framework model or the physical rotation models.

Insert Figure 1 about here.

Procedure.

Memory condition. For each scene, the participant stood on the bench, facing one of four directions. The experimenter then placed pictures of six objects on the walls, ceiling, and floor of the experimental room, so that the objects were above, below, to the front, back, left, and right of the participant. Participants were told the theme of the scene and the name of each object and instructed to study the scene so they could answer questions concerning the whereabouts of objects in the scene. They were allowed to view the scenes as long as they wished and to turn to inspect all directions. When participants indicated that they were ready, the objects were removed, and the questioning sequence begun.

For the questioning procedure, participants stepped off the bench and sat at a table in one corner of the room, facing directly into the corner so their orientation never matched any orientation they were to imagine. Once seated, participants were instructed to think of themselves standing on the bench in the center of the room. At the beginning of each block, the experimenter told the participant to imagine that he or she had turned to face a new object, either standing or reclining. Then the audiotape was started and six direction probes were presented, each six seconds apart. At each probe, participants were to simultaneously: a) say the name of the object located at that direction, and b) press the response button. Participants were instructed to respond as quickly as possible while maintaining accuracy. The experimenter recorded the participant's response time and accuracy. The relative looseness in coupling participants’ spoken and button responses could add variability to response times. The results of this experiment, however, have been independently replicated in similar designs which have collected responses by voicekey or other procedures that collect responses in one step (Experiment 2; Bryant & Tversky, 1999).

There were eight blocks of trials, four in the upright posture and four in the reclining. The order of postures was counterbalanced across scenes. For the upright posture, participants imagined themselves turning to face each of the four objects on the walls, rotating either clockwise or counterclockwise with each new block of probes. For the reclining posture, they imagined themselves turning around their head/feet axis, facing two objects on the walls and the objects on the ceiling and floor.

Participants began the questioning procedure in either the upright or reclining postures. Posture was alternated between scenes and the order of posture was counterbalanced across participants. The participant turned clockwise in half the scenes and counterclockwise in the other half. Direction probes within a given block were assigned to one of six counterbalanced orders, which assured that each probe appeared in each serial position an equal number of times across all scenes. The objects in a scene were placed in one of six counterbalanced sets of locations. Locations were fully counterbalanced across the first 12 participants, and four sets of locations were randomly selected for the last four participants. The first object faced during questioning was randomly selected and was never the object faced during initial learning.

Perception Condition. Participants were not allowed to study the scenes prior to the probing procedure. The experimenter placed each object in its location, naming it for the participant, then immediately started the audiotape containing the direction probes. The objects in the scene were left visible for the questioning procedure and the participant physically stood while responding to direction probes. The participant was told to face a particular object, answer six direction probes, turn to face another object, and so on. The participant turned to face each of the four objects on the walls, rotating clockwise or counterclockwise. Participants did not recline in this condition; participants responded to just four blocks of probes from the upright posture. Participants in the pilot study were found to have had limited movement while reclining, making physical rotation difficult. As in the memory condition, participants turned clockwise in half the scenes and counterclockwise in the other half. The order of probes and the placement of objects were counterbalanced as in the memory condition as well.

The procedure for responding to direction probes was the same as in the memory condition, except that participants spoke their responses into a microphone. Participants were not instructed to visually inspect probed directions, nor were they told not to do so. After finishing all eight blocks of probes in a scene, the experimental room was set up as the next scene, and the procedure repeated. The pilot experiment revealed that participants physically turn to respond to some probes but not others. This behavior was coded to allow comparison of response times for these two different strategies. For each probe, the experimenter coded whether the participant turned to look at least once in the probed direction (scanning strategy) or not (memory strategy).

Design.

The independent variables were direction (*front*, *back*, *head*, *feet*, *left*, or *right*), posture (upright or reclining), and response condition (memory or perception). Direction and posture were varied within participant and response condition varied between participants. The dependent variable was the time participants took to press the response button when probed.

Results

Data Treatment

In the memory condition, 3.1% of participants' responses were lost either because the participant made an inappropriate response (e.g., giving an answer without pressing the response button or pressing the button before saying the answer) or because the timing device malfunctioned. The remaining data were adjusted according to the following criteria. Participants made errors in response to 1.9% of the probes. Outliers, defined as response times greater than a participant's direction cell mean plus two standard deviations, accounted for 4.6% of the data. Errors and outliers were discarded from analysis. Response times were collapsed across blocks within each scene to form participant means.

In the perception condition, 5.9% of participants' responses were lost due to inappropriate responses or timing device malfunction. Participants made errors on an additional 0.6% of probes. Outliers accounted for 4.2% of responses. Errors and outliers were discarded from analysis.

Memory Condition

Mean direction by posture response times are presented in Table 3.

Insert Table 3 about here.

Effect of direction and posture. Participants responded faster overall when imagining themselves upright than reclining and the pattern of response times conformed to the spatial framework model for both postures. A two-factor analysis of variance (ANOVA) with repeated measures revealed significant effects of posture [$F(1,15) = 5.31$, $MSe = 0.39$, $p < .05$], direction [$F(5,75) = 18.50$, $MSe = 0.48$, $p < .05$], and their interaction [$F(5,75) = 2.84$, $MSe = 0.06$, $p < .05$]. Differences between subsets of levels of direction in this and subsequent experiments were tested by contrasts. For the upright posture, participants were faster to head/feet than front/back [$F(1,15) = 8.97$, $MSe = 0.16$, $p < .01$], which was faster than left/right [$F(1,15) = 46.43$, $MSe = 0.85$, $p < .01$]. For the reclining posture, participants were slightly faster to front/back than head/feet, but this difference was not significant [$F(1,15) = 2.30$, $MSe = 0.04$, n.s.], although responses to front/back were faster than those to head/feet for a significant majority of participants (see below). Participants were reliably faster to head/feet than left/right [$F(1,15) = 36.34$, $MSe = 0.66$, $p < .01$].

Constant and vertical dimensions. In the upright and reclining postures, the objects located to the head and feet did not change with reorientations, whereas objects located did change with reorientations. One might hypothesize that the advantage of the head/feet axis might be due to this constant association. If so, however, that advantage should occur for both upright and reclining postures. In the reclining posture, however, participants were faster to respond to front/back than head/feet. Thus, constancy of objects in and of itself did not make objects at the head and feet more accessible. In the reclining posture, all directions except head and feet were sometimes associated with the gravitational axis, depending on which side participants imagined themselves reclining. To examine the effect of this, a mean response time for the vertical axis was calculated by averaging response times for left/right when the participant reclined on his or her side with response times to front/back when the participant reclined on his or her front or back (i.e., the axes were aligned with the vertical gravitational axis). The mean response time for objects located on the vertical axis (1.40 s) was slower than response times to front/back (1.28 s) and head/feet (1.34 s). Being associated with gravity did not in itself convey fast access to *front*, *back*, *left*, or *right*.

Individual effects. Data of individual participants were generally consistent with the predictions of the spatial framework. To assess whether individual participants tended to display the predicted pattern, participants' response times were treated as the product of a random binomial process. There were six possible orders of response times to the three axes so that the spatial framework pattern had a 1/6 probability of occurring by chance. The binomial probability indicated is the probability that the given number of participants exhibited the predicted pattern by chance. Fifteen of 16 participants produced the general spatial framework pattern of response times for the upright posture (head/feet < front/back < left/right) (binomial probability < .001). The remaining participant did not exhibit the pattern predicted by the physical transformation model. Ten of the 16 participants were faster to *front* than *back* (binomial probability < .05). None of the remaining participants exhibited the pattern by the physical transformation model. Their response times displayed no consistent or readily interpretable pattern. For the reclining posture, 10 of 16 participants produced the predicted pattern (front/back < head/feet < left/right) (binomial probability < .01). There was no effect of participant gender [$F(1, 14) = 0.12$, $MSe = 0.05$, n.s], and this factor did not interact with any other.

Perception condition

Participants' data were divided into two conditions on the basis of the experimenter's observation of participant behavior on each trial. Response times for trials in which the participant gave no indication of physically turning to inspect the probed direction were assigned to the memory strategy condition. Response times for trials in which the participant physically turned to the probed direction (even if they did not fully turn) were assigned to the scanning strategy condition. All response times to *front* were assigned to the scanning strategy condition because participants could always see that object.

Participants used the scanning strategy on 71.1% of trials and the memory strategy on 28.9% of trials. Participants tended to use the memory strategy more for later blocks of probes. Otherwise, loss of data to the memory strategy was not differential across any counterbalancing factors. A repeated measures ANOVA revealed no significant effect of direction in the memory strategy data [$F(5,55) = 0.37$, $MSe = 0.01$, n.s.]. This is probably due to the fact that participants used the memory strategy on so few trials (average of 38.9 per participant) and provided too few data points for a reliable spatial framework effect.

Table 4 displays mean response times when participants employed the scanning strategy. Because participants scanned in 71.1% of trials, these means are based on an average of 96.34 responses per participant.

Insert Table 4 about here.

Effect of direction. A repeated measures ANOVA revealed a significant effect of direction [$F(5,55) = 30.12$, $MSe = 0.82$, $p < .01$]. Response times conformed entirely to predictions of the physical transformation model.

Individual patterns. All twelve participants displayed the general pattern predicted by the physical transformation model (*front* < *head*/*feet*/*left*/*right*/ < *back*) (binomial probability < .001). Participant gender did not affect response times [$F(1,10) = 1.79$, $MSe = 0.09$, n.s.], nor did this factor interact with direction [$F(5,50) = 2.10$, $MSe = 0.05$, n.s.].

Discussion

Participants learned an array of objects beyond their heads, feet, fronts, backs, and sides by observing a real scene. They were later tested either from memory or while the scene remained visible. In the memory condition, response times exhibited the same pattern as that for scenes acquired from narratives, namely the spatial framework pattern. When upright, participants were fastest to head/feet, followed by front/back, followed by left/right. When reclining, participants were fastest to front/back, followed by head/feet, followed by left/right.

In the perception condition, participants were induced to look at the array in order to respond to direction probes. The pattern of response times for trials in which participants physically scanned the probed direction did not conform to the spatial framework pattern. Instead, as expected, the pattern conformed to the physical transformation model. Participants were fastest to identify objects to the *front*, followed by directions offset by 90° (*head*, *feet*, *left*, and *right*), followed by the direction offset by 180° (*back*). Response times depended on how far the participant had to turn to the specified direction. Thus, the patterns of responses from perception and memory differ, indicating that spatial mental models are not like internalized perceptions.

The results do not imply that perception operates without reference to a mental model, only that that it operates with reference to a different mental model than that used in memory of these scenes. Thus, perception of the scenes was organized on the basis of different physical and spatial factors, or with a different emphasis on the salience of factors. The physical dynamics of turning were obviously crucial for the perception condition but not salient in memory where participants could inspect locations without mentally simulating physical transformations.

Experiment 2: Responding from Memory or Observation: External Spatial Array

The first experiment showed that when participants learn a surrounding array of objects by observation and identify objects in the directions around their bodies from memory, the pattern of response times is the same as when participants learn the array from narrative. The current experiment demonstrates the same phenomenon in a different situation. The array of objects is entirely in front of the participant, surrounding a doll, and the participants respond either from memory or from observation of the scene. This situation corresponds to the third person narratives employed by Bryant et al. (1992). Participants view the model from an external viewpoint, but the direction probes require participants to locate objects with respect to the doll's intrinsic body sides. This experiment explores whether people employ spatial frameworks or intrinsic computation for external arrays.

Based on the results of Experiment 1, we expected that participants who had learned the scene from observation and responded from memory would respond like participants in the studies of Bryant et al. (1992) who learned scenes from narratives and responded from memory. Specifically, we expected that they would construct mental spatial frameworks from the doll's point of view to keep track of the directions of objects relative to the doll and to update them as the doll is turned in the scene.

Participants in the perception condition responded to direction probes from the doll's perspective while observing the model scene. The doll was physically rotated and reclined in the model. Participants could perform this task in one of two ways. They could adopt the perspective of the doll, as they did in the third-person narratives. If so, the spatial framework pattern of data would be expected. However, this would entail mentally adopting a perspective that conflicts with the participant's own perspective on the scene. For the case of reading narratives or responding from memory, there is no such conflict. Alternatively, participants could keep their own perspective by determining the intrinsic sides of the doll and then searching for the object lying beyond the probed intrinsic side. This is the intrinsic computation model described earlier. It makes the same predictions as the spatial framework model for the upright case, but predicts faster responses to head/feet than front/back for the reclining case, unlike the spatial framework model.

Method

Participants.

Forty-eight Northeastern University students participated in the experiment for credit in an introductory psychology class or pay. Eleven men and 13 women served in the memory condition and 12 men and 12 women in the perception condition.

Materials and Equipment.

A physical model portrayed the scenes. In the center of each scene was a "Homer Simpson" doll (28 cm tall). The doll stood on a platform 14 cm high, and could be rotated to face four directions or reclined and rotated to face in four directions. Drawings of objects were hung from narrow wooden shafts to the front, back, head, feet, left, and right of the doll, such that they faced the participant at all times. A set of 42 object drawings was used to represent objects in seven scenes (one a practice trial). The objects and scenes (shown in Table 5) were meant to represent common situations. The same apparatus from Experiment 2 was used to present probes, and the voicekey used to collect participants' response times.

Insert Table 5 about here.

Procedure.

Two variants of the general procedure were employed. In both the memory and perception conditions, participants responded to either eight blocks of probes for a scene or to just one block per scene. Pilot studies indicated that participants might be able to predict or precompute object locations prior to receiving probes if the participants respond to the same objects in the same positions over several blocks of trials. Thus, we included one condition with eight blocks of probes per scene to encourage participants to treat scenes as stable and realistic, but a second condition with a single block per scene to prevent precomputation of responses. This factor, however, turned out not to affect participants' performance and is not reported in the analyses of data.

Memory Condition

Multiple Blocks per Scene Condition. For each scene, the experimenter placed six objects around the Homer figure. The participant sat about two feet from the model, which rested on a table. Participants were allowed to study the scene for as long as they wished. They were instructed to study the model until they were confident that they could answer questions concerning the whereabouts of the objects in the scenes. When participants indicated that they were ready, they were told to turn their seat so that they faced away from the model and could not see it.

Participants responded to eight blocks of probes per scene from memory. Rotations of the doll were described verbally to participants. For four blocks, the doll was said to be upright and for the other four it was reclining. The order of upright and reclining postures was counterbalanced across scenes and participants. Four blocks of trials were completed within a posture before the doll was changed to a new posture. A block began when the experimenter told the participant that the doll had turned to face a new object. For each block, six direction probes (*front*, *back*, *head*, *feet*, *left*, and *right*) were presented by audiotape, separated by six seconds of silence. Participants were instructed that the directions were to be interpreted with respect to the doll's body sides and current direction it was facing. They were also instructed to say aloud, upon hearing a probe, the name of the object that was located at the probed direction, as quickly as possible, without sacrificing accuracy. The experimenter recorded the participant's response time and accuracy. Participants completed six scenes in this fashion, after an initial practice scene.

The objects in a scene appeared in one of six counterbalanced sets of locations around the doll so that each object occupied each location an equal number of times across participants. The first object faced by the doll during questioning was randomly selected and was never the object faced during initial learning. In half of the scenes, the doll was rotated clockwise about its head/feet axis, and in the other half counterclockwise, for both postures. The direction of rotation was counter-balanced across scenes and participants.

Single Block per Scene Condition. The procedure was generally the same as the multiple block condition except that an entirely new scene was set up after each block of six direction probes. Participants completed 48 separate scenes, responding to six direction probes. The objects in each scene were randomly selected from the total set of 42 objects and randomly located in the model. The direction the Homer doll faced was randomly determined for each scene, with the provisions that it was upright and reclining in half the scenes and that it faced each of the four directions of rotation within a posture six times during the entire procedure.

Perception Condition

Multiple Blocks per Scene Condition. For each scene, the experimenter placed six objects around the Homer figure. As in the memory condition, the participant sat about two feet from the model. The model with Homer and objects was visible to the participant during the entire procedure. Participants were instructed to look at Homer between probes and were seated at a slight angle relative to the model so that they could see all the objects surrounding the doll. Participants were not given time to study the scene prior to the probing procedure.

The probing procedure was similar to that of the memory condition except that rotations of the doll were performed physically in the model and participants responded while viewing the scene. Participants responded to eight blocks of probes for each scene, counterbalanced for upright and reclining postures as in the memory condition. A block began when the experimenter rotated the doll to face a new object. Then six probes were presented by audiotape and participants named the object currently at the probed location with respect to the doll’s body sides. Within a block, participants received the six direction probes one after the other, separated by four seconds (participants generally needed less time to respond from perception than from memory). The experimenter recorded the participant's response time and accuracy.

Single Block per Scene Condition. The procedure was the same as the single block per trial memory condition except that participants responded to probes while viewing the scenes.

Design.

The independent variables were response condition (perception or memory), blocks per scene (multiple or single), posture (upright and reclining), and direction (*front*, *back*, *head*, *feet*, *left*, and *right*). Response condition and blocks per scene were varied between participants, and posture and direction were varied within participant. The dependent variable was the time participants took to say the name of the object located at a probed direction. Direction probes within a block were assigned one of six counterbalanced orders that assured that each probe appeared in each serial position an equal number of times.

Results

In the memory/ multiple blocks per scene condition, 6.8% of participants’ responses were lost, either because the participant made an inappropriate response or because of a timing device malfunction, 7.7% were errors, and 3.0% were outliers. In the memory/ single block per scene condition, nine participants failed to complete a total of 79 scenes (474 probes or 13.7% of possible responses of all participants) 1.8% of responses were lost, 4.4% were errors, and 3.5% were outliers. In the perception/ multiple block per scene condition, 0.7% of responses were lost, 1.4% were errors, and 2.8% were outliers. In the perception/ single block per scene condition, 1.7% of responses were lost, 1.4% were errors, and 2.4% were outliers. All errors and outliers were discarded from analysis. The remaining response times were collapsed to form participant direction by posture means for the four conditions.

An ANOVA with response condition and blocks per scene as between participants variables and direction and posture revealed no significant main effect of blocks per scene [$F(1,44) = 0.18$, $MSe = 0.350$, n.s.], nor did this factor interact with any other. Thus, it made no difference whether participants responded to only one or to eight blocks of trials for a scene. A significant main effect of test condition was observed [$F(2,44) = 59.96$, $MSe = 118.51$, $p < .01$], and this factor interacted with both posture [$F(2,44) = 8.69$, $MSe = 2.21$, $p < .01$] and direction [$F(10,220) = 11.12$, $MSe = 0.89$, $p < .01$]. In light of these results, data were analyzed separately for the memory and perception response conditions but combined across levels of blocks per scene within each response condition. Table 6 presents mean direction by posture response times for the memory and perception conditions.

Insert Table 6 about here.

Memory Condition

Effect of posture and direction. Participants' pattern of response times conformed to predictions of the spatial framework model for both postures, and participants responded faster when the doll was upright than reclining. A two-factor ANOVA with repeated measures revealed significant effects of posture [$F(1,23) = 33.54$, $MSe = 15.74$, $p < .01$], direction [$F(5,115) = 52.22$, $MSe = 6.96$, $p < .01$], and their interaction [$F(5,115) = 14.39$, $MSe = 0.99$, $p < .01$]. When the doll was upright, head/feet was faster than front/back [$F(1,23) = 36.63$, $MSe = 2.53$, $p < .01$], which was faster than left/right [$F(1,23) = 107.47$, $MSe = 7.42$, $p < .01$]. When the doll reclined, however, front/back was faster than head/feet [$F(1,23) = 31.03$, $MSe = 2.14$, $p < .01$], which was faster than left/right [$F(1,23) = 121.82$, $MSe = 8.41$, $p < .01$].

Constant and vertical dimensions. As in Experiment 1, participants were faster to front/back than head/feet in the reclining posture, even though the objects to head and feet were constant across rotations. In the reclining posture, the mean vertical response time, calculated as in Experiment 1, (2.74 s) was slower than that of front/back (2.29 s) and head/feet (2.59 s). Being associated with gravity did not itself convey fast access to *front*, *back*, *left*, and *right*.

Individual effects. Individual participants' data were generally consistent with the spatial framework model. In the upright posture, 22 of 24 participants produced the general spatial framework pattern (head/feet < front/back < left/right) (binomial probability < .001). Neither of the remaining participants exhibited the intrinsic computation pattern. Twenty of 24 participants were faster to *front* than *back* (binomial probability < .02). In the reclining posture, 20 of 24 participants displayed the general spatial framework pattern (front/back < head/feet < left/right) (binomial probability < .001). Only one of the remaining participants exhibited the intrinsic computation pattern; the rest showed no consistent pattern. Only 14 of 24 participants were faster to *front* than *back* (binomial probability > .05). There was no effect of participant gender on response times [$F(1,22) = 0.34$, $MSe = 1.128$, n.s.], and this factor did not interact with any other.

Perception Response Condition

Effect of posture and direction. Participants responded faster when the doll was upright than reclining, and the pattern of response times predicted by the internal spatial framework model was not observed for the reclining posture. Instead, participants were faster to head/feet than front/back in both postures. A two-factor ANOVA with repeated measures revealed significant effects of posture [$F(1,23) = 52.19$, $MSe = 0.986$, $p < .01$] and direction [$F(5,115) = 37.74$, $MSe = 0.689$, $p < .01$], but their interaction was not significant [$F(5,115) = 0.13$, $MSe = 0.0004$, n.s.]. When the doll was upright, head/feet was faster than front/back [$F(1,23) = 38.86$, $MSe = 0.14$, $p < .01$], which was faster than left/right [$F(1,23) = 225.39$, $MSe = 0.84$, $p < .01$]. Likewise, when the doll reclined, head/feet was faster than front/back [$F(1,23) = 42.76$, $MSe = 0.16$, $p < .01$], which was faster than left/right [$F(1,23) = 198.57$, $MSe = 0.74$, $p < .01$]. This pattern is consistent with the intrinsic computation hypothesis.

Constant and vertical dimensions. Unlike the memory condition, participants were fastest to head/feet in both postures. Consequently, we cannot rule out the possibility that the advantage of head/feet relative to other directions was due to the fact that the objects located to the head and feet were constant across rotations of the doll. Bryant and Tversky (1999), however, have observed that participants responding in a perception condition to a physical model or a diagram of a scene are fastest to head/feet in all orientations of a doll when the objects are not constant along the head/feet axis across rotations. Logan (1995) has also found head/feet to be faster than front/back for all orientations in a dot localization task. Thus, it seems unlikely that the advantage of head/feet in this particular experiment is due to the constancy of objects. In the reclining posture, all directions except head and feet were sometimes associated with the gravitational axis, depending on which side the doll was reclining. The mean vertical response time (1.30 s), however, was slower than that of front/back (1.20 s) and head/feet (1.12 s). Thus, neither the front/back nor left/right axes gained any advantage by being temporarily aligned with gravity.

Individual effects. In the upright posture, 20 of 24 participants produced the overall observed pattern (head/feet < front/back < left/right), which is consistent with both the spatial framework and intrinsic computation models (binomial probability <.001). The remaining participants exhibited no consistent pattern. However, only 14 of 24 participants were faster to *front* than *back* (binomial probability > .05), which is inconsistent with spatial framework model but consistent with the intrinsic computation model. In the reclining posture, 18 of 24 participants displayed the overall pattern (binomial probability < .001), which is consistent with the intrinsic computation model. Only 4 of 24 participants displayed the pattern predicted by the spatial framework model (front/back < head/feet < left/right) (binomial probability > .05) and the rest exhibited no consistent pattern. There was no effect of participant gender on response times [$F(1,22) = 0.023$, $MSe = 0.014$, n.s.], and this factor did not interact with any other.

Discussion

In this experiment, participants learned an array of objects located to all sides of a doll from observation, and responded to direction probes from memory or from observation. When participants responded from memory, their response times conformed to predictions of the spatial framework for both upright and reclining postures. This pattern has been obtained in experiments where participants learned arrays from narratives rather than observation (Bryant et al., 1992; Franklin & Tversky, 1990). The results suggest that, as for narratives, when participants observe a model of a doll in a scene, they mentally adopt the doll's perspective and construct a mental spatial framework centered on the doll. These data add to the evidence that for these types of spatial arrays, mental representations of spatial relations established from experience are functionally identical to those induced by description.

When participants viewed the same arrays but responded to direction probes while viewing the scene, their responses conformed to the intrinsic computation pattern. The pattern for the reclining posture was the same as for the upright posture, consistent with the intrinsic computation model but not the spatial framework model. These results indicate that participants do not employ spatial frameworks for responding during perception of a person in a scene.

General Discussion

Summary

In two experiments, participants learned a spatial array of objects from observing a scene. In the first two experiments, participants were embedded in an array of objects to all six sides of their bodies. In the second experiment, participants viewed an array consisting of a doll with objects to its head, feet, front, back, left, and right, and responded from the doll's internal perspective. For both situations, participants' knowledge of the spatial arrays was tested either from observation or from memory. In all cases, participants were given probe terms and asked to name the objects in those directions. The data of interest were the patterns of response times to those directions.

Two major findings emerged from these experiments. First, in both situations, when participants responded from memory, the pattern of response times was identical to the spatial framework pattern found in previous work where participants learned the scenes from descriptions rather than from direct experience (Bryant, et al., 1992; Franklin & Tversky, 1990). This implies that the spatial mental representations constructed from perceptual observation are functionally similar to those constructed from descriptions. Second, the patterns of response times obtained while participants observed the scene was considerably different from the patterns obtained when participants responded from memory. This was true when participants scanned scenes surrounding themselves (Experiment 1) and when they viewed a model scene (Experiment 2). For arrays of objects surrounding the participant, retrieval times corresponded to the physical transformation model, with longer retrieval times for longer distances from *front*. For perception of model scenes, retrieval times corresponded to the intrinsic computation model according to which participants determine the intrinsic sides of an object in an order from top and bottom to front and back to left and right. Thus, spatial mental representations are not like images or internalized perceptions.

Spatial Representations

We contrasted the physical transformation, spatial framework, and intrinsic computation models of spatial representation. Each posits a particular perspective, which is needed to establish a frame of reference. The frame of reference defines the spatial directions used to locate objects. The models differ with respect to frame of reference and/or means of accessing directions.

The physical transformation model posits an internal perspective. A person uses his or her own body axes as a frame of reference and mentally projects the body frame into a scene. Directions are accessed by imagining a rotation from one's front to a probed direction, which is an analog process. The spatial framework model also assumes an internal perspective; one's own body axes serve as a reference frame that can be projected into a scene. This model, however, posits categorical access to directions. Physical and perceptual asymmetries affect the salience of the body axes and determine how quickly one can access information from the mental model. The intrinsic computation model posits an external perspective. The frame of reference is the set of body axes of another person or a set of intrinsically defined object axes outside the self. Participants maintain their outside perspective and compute directions within the object-centered frame of the other person. Directions are accessed as they are in spatial frameworks but different physical and perceptual factors determine the salience of directions, leading to different patterns of responding.

Using Memory instead of Looking

Although participants in a pilot and the first experiment could inspect the scene around them to answer the direction probes, they rarely did so after learning the scene. This implies that while perceiving the scene, participants constructed mental spatial representations of the scene. Once the mental representations were constructed, information about the directions of objects in the scene was apparently more accessible from memory of the scene than from looking at the scene. Hence, it cannot be assumed that information is derived from perception simply because a scene is perceptually available to a viewer. Memory is part and parcel of perception, especially for maintaining awareness of elements of a scene that are not in direct line of sight. Information that is in the world may be more easily accessed by searching memory than by searching the world.

Mental Models and Images

The memory representations participants used in order to provide the objects at the probed directions are more like mental models than like images. Images are usually conceived of as internalized perceptions. Like perceptions, they have a point of view, and like perceptions, they have analog properties. However, performance in this task entailed taking many different perspectives. Moreover, performance in this task was not analog; in particular, times to retrieve objects displace 180° was less than times to retrieve objects to right and left, displaced only 90°. Furthermore, when reorientation times are measured, 180° reorientations are even shorter than 90° reorientations (Tversky, Kim & Cohen, in press). In contrast, studies of pointing to locations in arrays have shown what looks like analog processing (e.g., Rieser, 1989).

The present situation may have encouraged more categorical representations by using only categorical spatial relations and aligning the array with the observer's body. It is certainly possible that language and memory can capture and convey more refined spatial information than projections of the six body surfaces. The language people spontaneously use to describe spatial locations around the body is category plus hedge, for example, "slightly to the *right*," where "*front*" is presupposed. Memory errors are biased consistently with the descriptions (Franklin, Henkel, & Zangas, 1995). A model that incorporates both categorical and graded effects has been developed by Huttenlocher, Hedges, and Duncan (1991).

For two different situations, performance from memory was not like performance from perception so that the mental representations are not like internalized perceptions. Mental models are more schematic than images, capturing some elements and relations in a situation, but not all. They are not restricted to a specific point of view, and they may contain categorical rather than analog properties. Though the mental model account given here differs from an imaginal account, it also differs from a propositional account. For one thing, a propositional account cannot naturally incorporate the differences in accessibility of body axes or the interaction of that with body posture.

The theory of imagery as internalized perception is both elegant and simple. It accounts for the origins of images and provides insight into their nature. Although it explains behavior in many contexts, it does not provide an account of all spatial thinking. Mental representations of the situation under investigation, the situation that people find themselves in most of their lives, are not like images. Rather than deriving from our momentary perceptions of the spatial world, mental representations of this situation derive from our conceptions of the spatial world. Our conceptions of the spatial world are based on our extended interactions with the world (cf., Clark, 1973; Shepard, 1984). In those interactions, we are three-dimensional creatures, with a front/back axis that is asymmetric and orients both our perception and our behavior, a head/feet axis that is asymmetric and canonically upright, and a left/right axis that is more or less symmetric. The world we interact in is three-dimensional with a single asymmetric axis determined by gravity. These properties of the world and ourselves constrain our perception and our behavior and form the foundation for our conceptions of the world and the mental representations we create of it.

References

Braine, L. G., Plastow, E., & Greene, S. L. (1987). Judgments of shape orientation: A matter of contrasts. *Perception & Psychophysics, 41,* 335-344.

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

Bryant, D. J., & Tversky, B. (1992). Assessing spatial frameworks with object and direction probes. *Bulletin of the Psychonomic Society, 30,* 29-32.

Bryant, D. J., & Wright, W. G. (1999). How body asymmetries determine accessibility in spatial frameworks. *Quarterly Journal of Experimental Psychology, 52A,* 487-508.

Bryant, D. J., & Tversky, B. (1999). Mental representations of spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 137-156.

Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language, 31,* 74-98.

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

Clark, H. H. and Haviland, S. E. (1974). Psychological processes as linguistic explanation. In D. Cohen (Ed.), *Explaining linguistic phenomena* (pp. 91-124). Washington, D. C.: Hemisphere Publishing.

Corballis, M. C., & Cullen, S. (1986). Decisions about the axes of disoriented shapes. *Memory & Cognition*, *14*, 27-38.

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

Denis, M. and Cocude, M. (1992) Structural properties of visual images constructed from well- or poorly-structured verbal descriptions. *Memory & Cognition, 20,* 497-506.

Denis, D., Goncalves, M-R., & Memmi, D. (1995). Mental scanning of visual images generated from verbal descriptions: Toward a model of image accuracy. *Neuropsychologia*, *33*, 1511-1530.

Denis, M., & Zimmer, H. D. (1992). Analog properties of cognitive maps constructed from verbal descriptions. *Psychological Research, 54,* 286-298.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95*, 307-317.

Farrell, W. S. J. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 42-51.

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

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, *87,* 113-132.

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 performance: Volume II, cognitive processes and performance*. New York: John Wiley and Sons.

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

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models acquired from text. *Memory & Cognition, 20,* 507-518.

Franklin, N., Henkel, L. A., & Zangas, T. (1995). Parsing surrounding space into regions. *Memory & Cognition, 23*, 397-407.

Gentner, D., & Stevens, A. L. (Ed.). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Gray Wilson, S., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1993). Mental models and narrative comprehension: Some qualifications. *Journal of Memory and Language*, *32*, 141-154.

Halasz, F. G., & Moran, T. P. (1983). Mental models and problem solving in using a calculator. In A. Janda (Ed.), *CHI'83 Conference on Human Factors in Computing Systems* (pp. 212-216). New York: ACM.

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

Huttenlocher, J. Hedges, L. V., and Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*, 352-376.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, Massachusetts: Harvard University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13,* 289-303.

Jolicoeur, P., Ingleton, M., Bartram, L., & Booth, K. S. (1993). Top-bottom and front-behind decisions on rotated objects. *Canadian Journal of Experimental Psychology*, *47,* 657-677.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*, 255-273.

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. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review, 94*, 148-175.

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

Kosslyn, S. M., Alpert, N. M, Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E., Rauch, S. L. and Buonanno, F. S. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, *5,* 263-287.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 47-60.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van de Grind, & J. J. Koenderink (Eds.), *Limits in perception* (pp. 328-358). Utrecht: VNU Science Press.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, *28*, 103-174.

Maki, R. H. (1986). Naming and locating tops of rotated pictures. *Canadian Journal of Psychology*, *40*, 368-387.

Maki, R. H., & Braine, L. G. (1985). The role of verbal labels in the judgment of orientation and location. *Perception, 14,* 67-80.

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

Maki, R. H., & Marek, M. N. (1997). Egocentric spatial framework effects from single and multiple points of view. *Memory & Cognition*, *25*, 677-690.

Miller, G. A. (1979). Images and models, similes and metaphors. In A. Ortony (Eds.), *Metaphor and thought* (pp. 202-250). Cambridge, MA: Cambridge University Press.

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

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

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

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

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

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 1157-1165.

Rock, I. (1973). *Orientation and form*. New York: Academic Press.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*, 1-24.

Shepard, R. N. (1984). Ecological constraints on internal representations: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*, 417-447.

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

Shepard, R. N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, *18*, 161-193.

Shepard, R. N., & Podgorny, P. (1978). Cognitive processes that resemble perceptual processes. In W, K, Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 189-237). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Sholl, M. J., & Egeth, H. E. (1981). Right-left confusion in the adult: A verbal labeling effect. *Memory & Cognition*, *9*, 339-350.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, *31*, 261-292.

Tversky, B. (1991). Spatial mental models. In G. H. Bower (Ed.), *The psychology of learning and motivation, vol. 27* (pp. 109-145). San Diego: Academic Press.

Tversky, B., Kim, J., & Cohen, A. (in press). Mental models of spatial relations and transformations from language. In C. Habel and G. Rickheit (Eds.), *Mental models in discourse processing and reasoning*. Amsterdam: John Benjamins.

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

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

Table 1

Predicted Response Times (RT) for the Spatial Framework and Intrinsic Computation Models

	Posture	
Prediction	Upright	Reclining
Spatial Framework Model		
P1	RT(Head/Feet) < RT(Front/Back)	RT(Front/Back) < RT(Head/Feet)
P2	RT(Head/Feet) < RT(Left/Right)	RT(Head/Feet) < RT(Left/Right)
P3	RT(Front/Back) < RT(Left/Right)	RT(Front/Back) < RT(Left/Right)
Intrinsic Computation Model		
P1	RT(Head/Feet) < RT(Front/Back)	RT(Head/Feet) < RT(Front/Back)
P2	RT(Head/Feet) < RT(Left/Right)	RT(Head/Feet) < RT(Left/Right)
P3	RT(Front/Back) < RT(Left/Right)	RT(Front/Back) < RT(Left/Right)

“<” indicates “significantly faster than.”

Table 2

Scenes and Objects Used in Experiment 1

Scene	Objects
Kitchen	cake, calendar, hotdog, kettle, pear, tomato
Parent's Bedroom	boots, glasses, hat, pants, purse, shirt
Child's Bedroom	bed, chair, globe, microscope, radio, raincoat
Living Room	book, candle, flowers, lamp, telephone, television
Backyard	broom, flashlight, hose, lawnmower, pitchfork, rake
Laundry Room	clock, iron, sewing machine, table, towel, vacuum
Workshed	axe, crowbar, hammer, pliers, screwdriver, wrench

Table 3

Mean Response Times (in Seconds) in the Memory Condition (Experiment 1)

	Direction					
Orientation	Head	Feet	Front	Back	Left	Right
Upright	1.14	1.14	1.18	1.29	1.48	1.46
Mean	1.14		1.24		1.47	
Reclining	1.32	1.35	1.26	1.31	1.52	1.46
Mean	1.34		1.28		1.49	

Table 4

Mean Response Times (in Seconds) for Participants Who Used the Scanning Strategy in the Perception Condition (Experiment 1)

	Direction					
	Front 0°	Head 90°	Feet 90°	Left 90°	Right 90°	Back 180°
Scanning Strategy	0.89	1.21	1.15	1.28	1.24	1.70

Table 5

Scenes and Objects Used in Experiment 2

Scene	Objects
Kitchen	bread, fork, pie, plate, pot, spoon
Bedroom	cap, dress, pants, purse, shirt, sock
Living Room	chair, clock, lamp, painting, table, vase
Backyard	bird, cat, drum, flower, kite, (toy) truck
Zoo	bear, camel, elephant, lion, monkey, tiger
Workshed	axe, desk, ruler, saw, scissors, soap
Rec. Room	bell, (toy) boat, dice, glass, shoes, (toy) top

Table 6

Mean Response Times (in Seconds) for Memory and Perception of a Model Scene (Experiment 2)

	Direction					
Orientation	Head	Feet	Front	Back	Left	Right
MEMORY CONDITION						
Upright	1.80	1.84	2.02	2.27	2.67	2.73
Mean	1.82		2.15		2.70	
Reclining	2.58	2.60	2.28	2.31	3.13	3.24
Mean	2.59		2.29		3.18	
PERCEPTION CONDITION						
Upright	1.01	1.00	1.07	1.08	1.25	1.28
Mean	1.00		1.08		1.26	
Reclining	1.12	1.12	1.20	1.21	1.36	1.39
Mean	1.12		1.20		1.38	

Footnote

1. The reclining situation eliminates two alternative explanations for the primacy of head/feet in the upright situation. That primacy cannot be due to verticality or to the fact that objects at *head* and *feet* were constant. In the reclining case, response times to vertically arranged objects were not fastest. Also, in the reclining case, objects to *head* and *feet* were constant, but response times to the front/back axis were faster than those to the head/feet axis. Both upright and reclining patterns were replicated using objects to probe for directions rather than vice versa (Bryant & Tversky, 1992).

Figure Caption

Figure 1. Sketch of the experimental situation of Experiment 1. The dark rectangles surrounding the person represent the object pictures placed on the walls, ceiling, and floor. The drawing is not to scale.