Journal of Experimental Psychology:
Learning, Memory, and Cognition
1999, Vol. 25, No. 1, 137–156

Copyright 1999 by the American Psychological Association, Inc.
0278-7393/99/$3.00

Mental Representations of Perspective and Spatial Relations From Diagrams and Models

David J. Bryant
Humansystems

Barbara Tversky
Stanford University

In previous research (D. J. Bryant, B. Tversky, & N. Franklin, 1992; N. Franklin & B. Tversky, 1990), the authors showed that spatial knowledge conveyed by descriptions and direct experience induces participants to take the perspective of a character surrounded by objects. In this study, the authors used models and diagrams to convey the same information. With models, as with descriptions and experience, participants adopted the character's perspective (the *spatial framework* analysis). With diagrams, participants took an outside perspective (the *intrinsic computation* analysis). Even when informationally equivalent, different depictions made salient different aspects of the world. When instructed, however, participants were able to take either the inside or the outside perspective in memory for both diagrams and models. Depth cues in depictions also govern participants' perspective. When diagrams contained rich pictorial depth cues, participants used the spatial framework analysis, and when models were viewed without access to depth cues, participants relied on the intrinsic computation analysis.

People's knowledge of the world comes not only directly, from experiencing the world, but also indirectly, from descriptions and depictions of the world. Perhaps because of its significance, spatial knowledge has been conveyed by external representations since prehistory. Maps, whether from stone, clay, wood, bark, or paper, have been invented by many cultures (e.g., Brown, 1949; Wilford, 1981). Spatial language alone can act like a map, effectively conveying spatial relations and relative distances (e.g., Bryant, Tversky, & Franklin, 1992; Denis & Cocude, 1989; Franklin & Tversky, 1990; Glenberg, Meyer, & Lindem, 1987; Mani & Johnson-Laird, 1982; Morrow, Greenspan, & Bower, 1987; Taylor & Tversky, 1992) and allowing updating of relative positions and perspectives as new information becomes available (e.g., Bryant et al., 1992; Franklin & Tversky, 1990; Franklin, Tversky, & Coon, 1992; Glenberg et al., 1987; Morrow et al., 1987). Language describing space is so fundamental that it is used to express other, nonspatial concepts, such as time, mood, and power (e.g., Clark, 1973; Lakoff & Johnson, 1980).

David J. Bryant, Humansystems, Inc., Guelph, Ontario, Canada; Barbara Tversky, Department of Psychology, Stanford University.

This research was conducted at Northeastern University and sponsored by a Research and Scholarship Development Fund grant from Northeastern University and by the Air Force Office of Scientific Research, Air Force Systems Command, U.S. Air Force, under grant or cooperative number AFOSR 94-0220. We are grateful to Sophia DeRosa, Carol Gapsky, Marnie Nilsson, and Virginia Ricker for their assistance in conducting this research and to Maggie Lanca, Gordon Logan, Ruth Maki, three anonymous reviewers, and the Language and Cognition Group at Northeastern University for their helpful comments.

Correspondence concerning this article should be addressed to David J. Bryant, Humansystems, Inc., 111 Farquhar Street, Second Floor, Guelph, Ontario, Canada NIH 3N4. Electronic mail may be sent to dbryant@humansys.com.

External Representations: Diagrams and Models

Diagrams and models are external graphic representations or depictions that consist of elements and the spatial relations among them (Tversky, 1995a). As such, they are external stimuli with their own spatial properties. In particular, relations in the represented world are mapped onto spatial relations in the graphic representation (Tversky, 1993, 1995a; Tversky, Kugelmass, & Winter, 1991). For example, in corporate organization diagrams, the vertical spatial relations represent power. In the case of the depictions studied here, spatial relations in the represented world are mapped onto spatial relations in the diagrams and models.

Depictions, however, schematize the situations they represent and require interpretation. What is schematized, and how it is schematized, can affect how a depiction is interpreted and used. The third dimension, for example, is an important factor that must be schematized in depictions of spatial situations. As we use the terms, *models* convey all three spatial dimensions directly; *diagrams,* by contrast, may depict three-dimensional (3D) relations but are themselves two dimensional (2D). They may use a number of conventions for conveying depth, including relative size, occlusion, height in the picture plane, and converging lines, but they necessarily lack binocular cues. Also, diagrams more often than models use verbal and symbolic information to convey spatial information. Thus, diagrams convey the 3D structure of an environment indirectly, whereas 3D models convey that information directly.

One goal of the present research is to explore how the differences between diagrams and models lead people to create different kinds of mental models of depicted environments. A second goal is to determine whether diagrams and models necessarily induce one kind of mental representation or whether individuals can alter their representations on the basis of instructions. A third goal is to explore what features of diagrams and models, especially those related to conveying depth, lead to differences in mental representation.

Mental Representations of a Paradigmatic Spatial Situation

Paradigmatic Scene

One spatial situation people carry around with them at all times is their own bodies surrounded by the objects in their immediate environment. This is a useful situation for studying people's spatial mental models, because people appear to keep track of objects effortlessly as they move about the world. In our studies, we investigated participants' understanding of depictions of a scene in which a character is situated in a natural setting, such as a kitchen or living room, with objects, such as a pot or spoon, located to the six body sides (head, feet, front, back, left, and right). The situation is illustrated in Figure 1, which is also an example of one of the diagrams used in our experiments. After learning a scene, participants were informed that the character had moved to face a new object, and participants were then probed for the objects lying in the six directions from the character's body.

Because certain body axes have a favored status in people's interactions with the world, they are more salient to thinking about spatial relations, and this leads to differences in retrieval times for spatial relations. Which axes are most salient depends on the mental perspective one adopts. There are at least two perspectives that one could effectively use for our paradigmatic scene, the *inside* and *outside* perspectives. Which perspective a participant adopts will determine his or her mental representation and how he or she accesses spatial directions, which will in turn affect retrieval of spatial information.

Spatial Framework Analysis

To represent the scene, participants could adopt the inside perspective of the character. This perspective leads to a particular organization of objects in participants' mental models of scenes. Franklin and Tversky (1990) developed the *spatial framework analysis* (based in part on previous analyses of spatial language and cognition by Clark, 1973; Fillmore, 1976; Levelt, 1984; Miller & Johnson-Laird, 1976; and Shepard & Hurwitz, 1984; among others) to explain patterns of retrieval times of spatial information from narrative descriptions of scenes. According to this analysis, people mentally place themselves in the place of the character and use their own head/feet, front/back, and left/right axes to code locations of objects. In other words, they apply an egocentric frame of reference.

The accessibility of objects in memory in this analysis

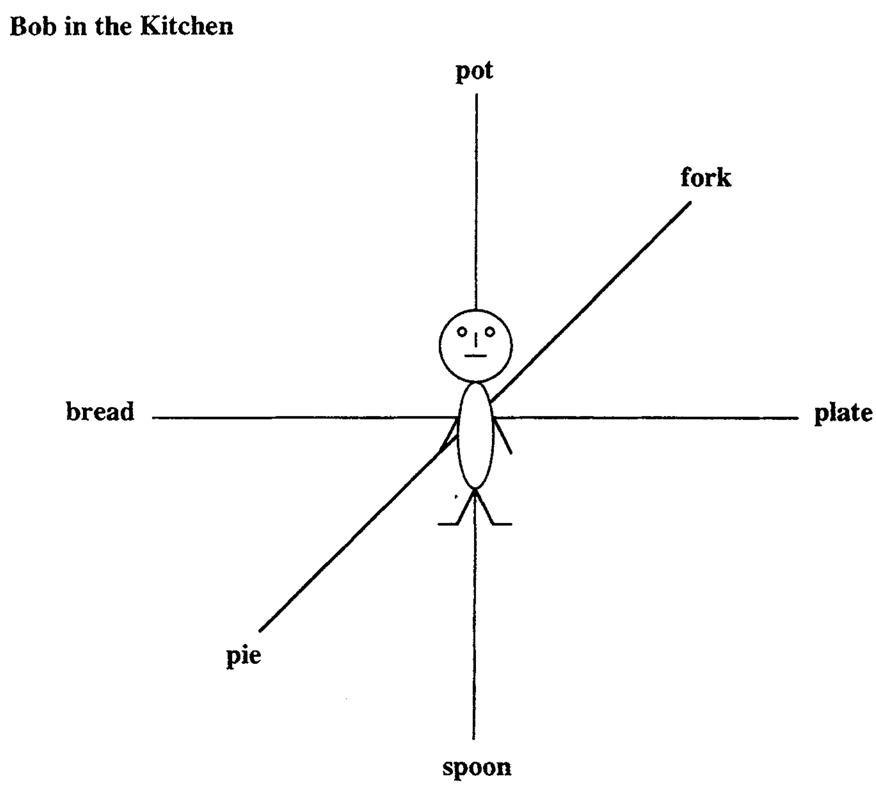

Figure 1. "Bob in the Kitchen": Paradigmatic scene and example of a diagram used in Experiments 1–3. Depth is indicated by the diagonal line; the pie to the lower left projects out of the page and the fork to the upper right projects into the page.

depends on the characteristics of the participant's own body and perception of space, rather than those of the character. For the body, the front/back and head/feet axes are more salient than the left/right axis because of their biological asymmetries. Because the perceptual and motor apparati are oriented forward, the front/back axis separates the world that can be easily perceived and manipulated from the world that cannot. This gives that axis a slight advantage over the head/feet axis, but one that can be countered by the asymmetries of the physical world. The world has one salient asymmetric axis: the vertical axis of gravity. Thus, when the character in the scene is described as upright with the body's natural axis of rotation aligned with gravity, the asymmetries of the body and world combine to make the head/feet axis more accessible than the front/back axis. Thus, participants should be faster to identify objects to the head/feet than front/back, and faster to identify objects to the front/back than left/right.

When the character in the scene is described as reclining, however, and turning from back to side to front, the axis of gravity no longer corresponds to a body axis, so the front/back body axis is most accessible. In this case, the biological asymmetries of front/back, which are more extreme than those of head/feet, predominate. Thus, participants should be faster to respond to objects to front/back than head/feet, and slowest to objects to left/right. These predictions have been upheld in more than a dozen experiments (Bryant, Lanca, & Tversky, 1995; Bryant & Tversky, 1992; Bryant et al., 1992; Bryant, Tversky, & Lanca, 1998; Franklin & Tversky, 1990; Franklin et al., 1992).[1]

Participants construct similar mental models from experiencing a situation as from reading about a situation. When participants in one experiment learned an environment by standing in it themselves and responded to direction probes from memory, their patterns of reaction times (RTs) to retrieve objects were the same as when learning was from description (Bryant et al., 1998).

Intrinsic Computation Analysis

A second way that participants could represent scenes is by taking the outside perspective and imagining themselves looking at the entire scene from some external position. This perspective affords a view of the character and objects. Given the outside vantage point, participants do not rely on their own body axes to code space but instead use a frame of reference centered on an external object or person. In our paradigm, participants locate objects by determining where the character's body sides are and naming the objects associated with each side. Accessibility of objects depends on the cognitive and perceptual mechanisms used to identify the sides of the character.

We call this procedure the *intrinsic computation analysis,* and it can be invoked for any object that has acknowledged intrinsic sides, front, back, top, bottom, and consequently left and right, such as people and cars, but not balls or trees (e.g., Fillmore, 1975; Levelt, 1984; Levinson, 1996; Miller & Johnson-Laird, 1976).[2] Some intrinsic sides are more readily determined than others. Research indicates that people first determine the top/bottom axis of an object, which is the head/feet axis of a person (Braine, Plastow, & Greene, 1987; Jolicoeur, 1985; Maki, 1986; Rock, 1973). In particular, Rock showed that to identify what an object is, one needs to know how it is oriented (i.e., where its top is). The front/back axis of an object must be determined prior to the left/right axis as the left/right axis can only be defined with respect to the top/bottom and front/back axes. Corroborating this analysis, people are faster to identify the tops and bottoms of objects than the fronts and behinds of objects at all orientations (Jolicoeur, Ingleton, Bartram, & Booth, 1993). They are also faster at identifying asterisks at the top and bottom than at the left and right for all orientations (Corballis & Cullen, 1986). Identification of sides in the intrinsic frame does not depend on the orientation of the object or viewer. Thus, participants using the intrinsic computation analysis should always be fastest to identify objects at the head/feet, then the front/back, and finally the left/right of the person in the scene, irrespective of orientation.

The intrinsic computation pattern has been observed when participants respond to direction probes while viewing 2D diagrams. Logan (1995) presented diagrams of 2D slices of our 3D situation, consisting of schematic heads in front, profile, or top views, with colored dots located in the appropriate directions. The heads were presented upright or rotated 90°, 180°, or 270°. The participant's task was to make judgments about the directions of colored dots from the heads. Logan's data fit the intrinsic computation pattern of response times, head/feet fastest followed by front/back followed by left/right, for all orientations of the head. His data did not fit the spatial framework pattern. Although he referred to the spatial framework analysis of Franklin and Tversky (1990) to explain his data, Logan did not note the inconsistency of his data with theirs for the cases that were not upright. Because Logan did not recognize the inconsistency, his explanation of his data is inadequate and does not

[1] The reclining situation, in which gravity and verticality no longer correlate with any body axis, eliminates two alternative explanations for the primacy of head/feet in the upright situation (see Bryant et al., 1992; Franklin & Tversky, 1990). The primacy cannot be due to the fact that the head/feet axis happened to be the vertical axis for the upright orientation. In the reclining case, left/right corresponded to the vertical axis for half of the trials, but reaction times (RTs) were slower than those to head/feet and front/back in the horizontal plane. The primacy also cannot be attributed to the fact that objects to the head and feet were constant with rotations of the character. In the reclining case, objects to head and feet were still constant with rotations, but RTs to head/feet were slower than those to front/back where objects changed with each rotation.

[2] The intrinsic computation analysis is not the same as the external spatial framework described in previous research (Bryant et al., 1992). Both assume a mental perspective outside or external to the scene. In the external spatial framework, however, participants locate objects with respect to their own body sides. Participants using intrinsic computation locate objects in relation to the intrinsic sides of the character in the scene.

account for the inconsistency of his results with those of Franklin and Tversky. The intrinsic computation analysis, however, does provide an adequate explanation for Logan's data and indicates that his participants used an outside rather than inside perspective.

The spatial framework and intrinsic computation analyses can easily be distinguished by their predictions for a character who is reclining. According to the spatial framework analysis, participants mentally place themselves in the position 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 analysis, the participant identifies the sides of the character from an outside perspective, 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 Research

Selecting a physical model of the paradigm situation for the present research was not difficult, as a model is inherently 3D. The model consisted of a doll with schematic pictures of objects suspended in the appropriate positions. Selecting a diagram was more difficult. A diagram is necessarily 2D, but in this case, it needed to convey a 3D situation. In Western art, the third dimension has been expressed by convergent perspective since the Renaissance. However, this convention has not been universally adopted (e.g., Asian art) because it can distort shape and size. Rather than using a biased, complex, and cluttered converging perspective drawing to convey the scene, we schematized the depth dimension by adapting another common convention for conveying depth—a diagonal line—and eliminated irrelevant detail. This convention is used spontaneously by children (Braine, Schauble, Kugelmass, & Winter, 1993) and is standard in Chinese art (Willats, 1990).

In the first experiment, we examined mental representations spontaneously established from diagrams and models. In the second experiment, we examined the effects of instructions to interpret the diagrams and models on the mental representations established. In the final two experiments, we examined the characteristics of diagrams and models that induce the use of intrinsic computation or spatial framework analyses. Because models convey 3D information directly, but diagrams do not, we expected models to induce taking the perspective of the character, yielding the spatial framework pattern of RTs, and diagrams to induce taking the outside perspective, yielding the intrinsic computation pattern.

The current experiments consider the upside-down orientation, which has not been examined in previous studies (e.g., Bryant et al., 1992; Franklin & Tversky, 1990). Predictions of the intrinsic computation analysis are the same as for all orientations, and predictions of the spatial framework analysis match the upright orientation. The head/feet axis is again aligned with gravity, but in a noncanonical orientation. The asymmetries of front/back could render this axis most salient, as it does for the reclining posture. People, however, naturally rotate around the head/feet axis as they navigate the world, as when a person turns to walk in another direction. In the current paradigm, too, the character rotates around the head/feet axis. Having the natural axis of rotation aligned with gravity, even in the opposite direction, should render it more salient than front/back. Thus, for an upside-down character, participants should show fastest access to head/feet, followed by front/back, followed by left/right. Participants, however, should be slower overall because of the character's noncanonical orientation.

General Method

All experiments followed the same general method. Changes in materials and procedures specific to each experimental manipulation are discussed separately for each experiment. The materials and manipulations are summarized in Table 1, with predictions for each condition.

Participants

Participants in all experiments were Northeastern University undergraduates with normal or corrected-to-normal vision who participated for credit in an introductory psychology class. The numbers of participants in each experiment are listed in Table 1.

Scenes

Diagram Condition

Participants learned four critical scenes, indicated in Table 2, and one for practice. Each scene depicted a setting with a character surrounded by six objects (see Figure 1). The name of the character and the type of setting were printed at the top of the diagram. In half the scenes, the name given to the character was female and in the other half it was male. The settings and the objects were selected to be familiar and common and to form a coherent scene. The locations of objects were selected randomly. In all diagrams, the character was shown facing forward.

The character was 3.9 cm (1.5 in.) long. The vertical and horizontal axes were 11.7 cm (4.5 in.) long, and the diagonal axis was slightly shorter (approximately 11.05 cm, or 4.25 in.). In all diagrams, the diagonal was drawn from the lower left to the upper right of the page. Braine et al. (1993) observed an early tendency in children to interpret objects to the left and lower in pictures as being closer than objects to the right and higher. This suggests a bias to interpret the left end-point of a diagonal as nearer than the right end-point. The name of an object was printed at the end of each axis.

Model Condition

Participants learned four critical scenes plus one for practice. Four of the scenes (bedroom, construction site, kitchen, and living room) were the same as those of the diagram condition. The other (backyard) was adapted from materials used by Bryant et al. (1998). A "Homer Simpson" doll (28 cm tall) was placed in the center of the model. The doll stood on a platform 14 cm high and

Table 1
Summary of Experimental Conditions

Experiment and n[a]	Study material	Experimental manipulation[b]	Predicted mental representation
Experiment 1[c]			
32[d] (13/19)	Diagram		Intrinsic computation
24 (12/12)	Model		Spatial framework
Experiment 2[c]			
20[e] (10/10)	Diagram	Inside perspective instructions	Spatial framework
20 (11/9)	Model	Outside perspective instructions	Intrinsic computation
Experiment 3[f]			
16 (8/8)	Standard diagram	No depth cues	Intrinsic computation
16 (8/8)	Intermediate diagram	Converging lines	Spatial framework
16 (8/8)	Perspective diagram	Convering lines, relative size, texture gradient	Spatial framework
Experiment 4[f]			
16 (8/8)	Standard model	Normal depth cues	Spatial framework
16 (8/8)	Impoverished model	Minimal depth cues	Intrinsic computation

[a]Numbers of male and female participants are indicated, respectively, in parentheses. [b]The orientation of the character and the objects' directions relative to the character were manipulated within-participant in all experiments. This column lists additional manipulations. [c]Between-participants design. [d]Data of 1 participant were discarded due to high error rate. [e]Data of 3 participants were discarded due to high error rate. [f]Within-participant design.

could be rotated to face four directions or reclined to face in four directions. Drawings of the objects in each scene 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.

Procedure

Diagram Condition

Participants received detailed instructions about the procedure before beginning. Participants were instructed that the diagonal axis represented an axis in depth and were asked to think of the lower end as projecting out of the page toward them. Participants were told that the diagrams conveyed 3D environments and that they should attempt to think of what that setting would be like. They were encouraged to think of the character as standing or reclining on some kind of platform in the center of the scene and to elaborate the setting to help them remember the scene. They were instructed to study each diagrammed scene, taking care to learn the names of the objects and where they were located. Participants were allowed to study the diagram for as long as they wished, then returned it to the experimenter. They then proceeded to the direction probes, which were presented by computer.

A block of direction probes always began with a sentence presented on the computer screen telling the participant that the character had turned to face another object and/or changed orientation (from upright to reclining, reclining to upside down, etc.). The sentence specified the direction in which the character turned and stated explicitly the character's resultant orientation and which object the character now faced. When participants understood this sentence, they pressed the space bar of the computer keyboard to receive a series of six direction probes. Probes consisted of the names of the six body directions in relation to the character (front, back, head, feet, left, and right). Participants were specifically instructed not to interpret the probes in relation to themselves. In response to a probe, participants pressed the space bar as soon as they knew which object was located at that direction, without sacrificing accuracy. The time participants took to do this was the critical RT. After participants pressed the space bar, the names of the six objects appeared in a line on the screen in random order, numbered 1 to 6. Participants pressed a numbered key corresponding to the correct object as quickly and accurately as possible. This served as an accuracy check. Direction probes were separated by 500 ms of blank screen, and a series continued until all six directions had been probed.

The character changed orientation four times during the probing procedure: upright, reclining with the head pointed left (in the picture plane), upside down, and reclining with the head pointed right. Changes in orientation proceeded counterclockwise in the picture plane in all scenes. After participants completed the first block of six probes, the character was rotated around the head/feet axis in that same orientation in three subsequent blocks of trials. The character rotated counterclockwise in all scenes. After four blocks in one orientation, the participant was told that the character had changed orientation and completed four rotations in that

Table 2
Scenes and Objects Used in Experiments 1 and 2

Scene	Character	Objects
Construction site	Harry	axe, bucket, jackhammer, ladder, shovel, wheelbarrow
Barn	Nancy	brush, hay, lantern, pail, saddle, shears
Bedroom	Steve	dress, hat, pants, purse, shirt, sock
Kitchen	Bob	bread, fork, pie, plate, pot, spoon
Living room	Sally	chair, clock, lamp, painting, table, vase

posture, and so on. Participants completed 16 blocks of probes for each scene.

Model Condition

The procedure was the same as that of the diagram condition, except that participants studied a physical model of scenes, rather than a diagram, and the character in the scene was always Homer Simpson.

Design

In all experiments, within-participant independent variables included direction (front, back, head, feet, left, and right) and orientation (upright, upside down, reclining to the left, and reclining to the right). Manipulations specific to each experiment are listed in Table 1. The critical dependent variable in all experiments was RT. Orders of presentation of scenes, of orientation within scenes, of rotation within block of probes, and of direction probes were counterbalanced in the same way for all experimental conditions. Equal numbers of participants were assigned to eight random orders of presentation of the four scenes. Four versions of the block sequence were constructed for each scene. In one version, the character began upright and rotated counterclockwise across blocks of probes. In other versions, the character began reclining to the left, upside down, and reclining to the right. Version was counterbalanced across participants such that each participant received one scene in which the character began the probing procedure in each orientation. Within a scene, the character rotated around its head/feet axis in a clockwise fashion, and in the other half counterclockwise. 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.

Experiment 1: Locating Objects From Memory of Diagrams and Models

This experiment documents differences in representations of diagrams and models. Because of the strong cues to depth, participants who learn scenes from a model should adopt the inside perspective of the character in the scene and use spatial frameworks. Depth cues convey a detailed 3D environment, making it easy for participants to mentally place themselves in the scenes. Because of the weak depth cues and small size of diagrams, participants who learn scenes by diagram should adopt the outside perspective and use intrinsic computation.

Results

In this and all subsequent experiments, a probability criterion of .05 was assumed for statistical tests, unless otherwise stated.

Data Treatment

Diagram condition. Participants made errors on 7.4% of probes, and these data points were discarded from analysis. Outliers, defined as RTs greater than a participant's depiction by orientation by direction cell mean plus two standard deviations, accounted for 5.1% of the data and were also discarded. In addition, all RTs from a total of three scenes from 3 participants were discarded because participants made more than 16 errors in these scenes. The data of 1 participant were discarded in its entirety because the participant averaged more than 16 errors per scene. Outliers in this and all conditions of all experiments were generally equally distributed across direction and orientation conditions. In particular, there were no more outliers to the relatively difficult left and right probes than other directions.

Men and women generally displayed the same patterns of RTs. There was no overall effect of participant gender, $F(1, 30) = 2.56$, $MSE = 51.70$, but the three-way interaction of gender, orientation, and direction was significant, $F(15, 450) = 2.25$, $MSE = 1.08$. This finding seems to reflect that men were slightly faster overall for the upright orientation, but not for others, and that women tended to show extreme RTs to right and left probes for reclining orientations. The effect of gender in the diagram condition was considered in the analysis of response patterns for each orientation.

Model condition. Participants made errors on 5.1% of probes, and 5.1% of the data were outliers. Men and women displayed the same patterns of RTs. There was no overall effect of participant gender, $F(1, 21) = 0.03$, $MSE = 0.58$, and this factor did not interact with any other.

Effect of Type of Depiction

RT data were subjected to an analysis of variance (ANOVA) with type of depiction as a between-participants variable and orientation and direction as within-participant variables (see Table 3). RTs did not differ overall between the diagram and model conditions, but different patterns of RTs were observed in the two conditions. For this reason, the effects of orientation and direction were examined in separate ANOVAs for the diagram and model conditions. The results of these ANOVAs are also shown in Table 3.

Diagram Condition

Mean RTs are presented in Table 4. The pattern of RTs for reclining to the right did not differ from the pattern for reclining to the left, and these two conditions were collapsed to a single reclining condition. Similarly, in all subsequent experiments, data were collapsed to form a single reclining condition because there were no significant differences between reclining to the right and reclining to the left.

Because the diagrams contained few depth cues, participants should have adopted an outside perspective and used intrinsic computation to locate objects. Thus, for all orientations, participants should have been faster to identify objects to the head/feet than front/back, and slowest to identify objects to left/right. Participants should also be slower overall for nonupright orientations because it is more difficult to identify the sides of the character when it is reclining or upside down.

Direction significantly affected RTs in the diagram condition (see Table 3). When the character was upright, participants responded faster to head/feet than front/back,

Table 3
Analyses of Variance for Experiments 1 and 2

Variable	Experiment 1 *df*	*F*	*MSE*	Experiment 2 *df*	*F*	*MSE*
Overall						
Type of depiction (TD)	1, 47	0.11	1.58	1, 35	0.34	2.04
TD × Orientation (O)	2, 94	1.67	1.29	2, 70	0.15	0.05
TD × Direction (D)	5, 24	0.11	1.94	5, 175	2.16	0.53
TD × O × D	10, 470	3.23*	1.74	10, 350	2.10*	0.33
Diagram condition						
O	3, 93	24.21*	18.36	3, 48	16.06*	5.47
D	5, 155	28.51*	16.22	5, 80	23.72*	6.53
O × D	15, 465	8.87*	4.45	15, 240	6.25*	0.95
Model condition						
O	3, 69	27.01*	18.96	3, 57	23.10*	6.37
D	5, 115	35.95*	25.68	5, 95	63.71*	13.38
O × D	15, 345	10.19*	4.40	15, 285	3.08*	0.41

*$p < .05$.

$F(1, 31) = 14.80$, $MSE = 6.56$, and faster to front/back than left/right, $F(1, 31) = 42.86$, $MSE = 19.00$.[3] For the reclining character, RTs conformed to the intrinsic computation pattern. Participants responded faster to head/feet than front/back, $F(1, 31) = 5.60$, $MSE = 2.09$, and faster to front/back than left/right, $F(1, 31) = 39.64$, $MSE = 11.05$. For the upside-down orientation, participants responded faster to head/feet than front/back, although not reliably so, $F(1, 31) = 2.73$, $MSE = 5.25$. Participants did respond significantly faster to front/back than left/right, $F(1, 31) = 63.18$, $MSE = 121.36$.

Data of individual participants in this and all subsequent experiments were subjected to a binomial test to determine whether participants tended to display the intrinsic computation or spatial framework patterns. Participants' RTs were treated as the product of a random binomial process. There were six possible orders of RTs to the three axes so that the intrinsic computation or spatial framework pattern each have a one in six probability of occurring by chance. A significant majority of participants exhibited the predicted pattern in all conditions of all experiments. In no condition of any experiment did the remaining participants exhibit a systematic pattern of RTs.

To distinguish the spatial framework and intrinsic computation analyses, one must perform a crucial comparison between the relative RTs to head/feet and front/back for upright and reclining orientations. The spatial framework analysis predicts a crossover interaction with RTs to head/feet faster than RTs to front/back for the upright orientation, but RTs to front/back faster than RTs to head/feet for the reclining orientation. The intrinsic computation analysis predicts no such interaction. We tested this in the current data using an interaction contrast (Maxwell & Delaney, 1990, p. 268). It revealed no significant interaction between RTs to head/feet and front/back with orientation, $F(1, 465) = 0.86$, $MSE = 0.43$, which is consistent with the intrinsic computation analysis.

Participants were fastest overall when the character was upright (3.96 s), next fastest when reclined with the head to the right (4.30 s), followed by reclining with the head to the left (4.34 s), and slowest when upside down (4.72 s). Participants responded significantly faster for upright characters than for those reclining to the right, $F(1, 31) = 14.65$, $MSE = 11.11$, and all other orientations. RTs for both reclining orientations were significantly faster than those for upside-down orientations: reclining to right, $F(1, 31) = 21.91$, $MSE = 16.61$; reclining to left, $F(1, 31) = 18.54$, $MSE = 14.06$. The difference between the upside-down and two reclining orientations is due primarily to the especially long RTs to left and right for the upside-down character.

There was no effect of participant gender for the upright, $F(1, 30) = 2.13$, $MSE = 7.94$, or reclining, $F(1, 30) = 1.80$, $MSE = 9.72$, orientations, nor did this variable interact with direction or orientation. For the upside-down orientation, there was no main effect of participant gender, $F(1, 30) = 3.89$, $MSE = 28.38$, but gender did interact with direction, $F(5, 150) = 3.36$, $MSE = 6.00$. The mean pattern was observed for both men and women, but the patterns of individual directions differed between the two. Men were faster to left than right, whereas women were faster to right than left. Also, men were faster to head than feet, but women to feet than head.

Model Condition

Because models provide direct 3D information, participants should have been able to adopt the inside perspective and use spatial frameworks to locate objects. The data do, indeed, conform to predictions of the spatial framework analysis (see Table 4; again, RTs were collapsed to form a single reclining condition). When the character was upright, participants responded faster to head/feet than front/back, $F(1, 23) = 19.16$, $MSE = 5.72$, and faster to front/back than left/right, $F(1, 23) = 16.30$, $MSE = 4.86$. In contrast, when the character reclined, participants responded faster to

[3] Differences between subsets of levels of direction in this and all subsequent experiments were tested by contrasts.

Table 4
Mean Reaction Times (in Seconds) for Memory of Diagrams and Models (Experiment 1)

	Direction					
Orientation	Head	Feet	Front	Back	Left	Right
Diagram						
Upright	3.44	3.37	3.78	3.94	4.69	4.57
M	3.40		3.86		4.63	
Reclining	3.96	4.05	4.10	4.36	4.80	4.73
M	4.01		4.23		4.77	
Upside down	3.80	3.80	4.06	4.35	6.25	6.05
M	3.80		4.20		6.15	
Model						
Upright	3.47	3.44	3.85	4.04	4.16	4.63
M	3.46		3.94		4.39	
Reclining	4.24	4.32	3.80	3.75	4.80	4.77
M	4.28		3.78		4.79	
Upside down	3.84	3.96	4.54	4.49	6.42	5.56
M	3.90		4.52		5.99	

front/back than head/feet, $F(1, 23) = 33.97$, $MSE = 6.02$, and faster to head/feet than left/right, $F(1, 23) = 35.08$, $MSE = 6.21$. When the character was upside down, participants responded faster to head/feet than front/back, $F(1, 23) = 7.50$, $MSE = 8.73$, and faster to front/back than left/right, $F(1, 23) = 43.17$, $MSE = 50.26$. An interaction contrast revealed a significant interaction effect, $F(1, 345) = 13.561$, $MSE = 5.84$, which is consistent with predictions of the spatial framework analysis.

Participants were fastest overall when the character was upright (3.92 s), next fastest when the character reclined to the right (4.21 s), followed by reclining to the left (4.33 s), and slowest when upside down (4.80 s). Participants responded faster for upright characters than those reclining to the left, $F(1, 23) = 17.03$, $MSE = 11.95$, and all other orientations. RTs to both reclining orientations were significantly faster than those to upside down: reclining to left, $F(1, 23) = 22.23$, $MSE = 15.60$; reclining to right, $F(1, 23) = 34.51$, $MSE = 24.22$. The difference between reclining and upside-down orientations appears to be attributable to responses to left and right.

Discussion

The major prediction was upheld. Retrieval times of participants learning the scenes from models displayed the spatial framework pattern, suggesting that they adopted the perspective of the character in the scene. In contrast, retrieval times of participants learning scenes from diagrams displayed the intrinsic computation pattern, suggesting that they treated the entire scene as an object to be viewed from outside. Thus, models and diagrams of the same spatial situation are interpreted differently. The strong depth cues of the model apparently induced participants to adopt a perspective embedded in the scene, whereas the flatness and integrated nature of the diagrams induced participants to regard the scene as an external whole.

Experiment 2: Effect of Inside Perspective and Outside Perspective Instructions

The difference between memory for models and diagrams reflects the adoption of two different perspectives: the inside and the outside. If the adopted perspective, rather than the kind of depiction per se, determines the nature of one's mental representation, people should be able to alter how they represent scenes in our paradigm. The purpose of this experiment was to determine whether interpretations of models and diagrams can be manipulated. In this experiment, one group of participants viewed diagrams and were explicitly told to create a mental model of themselves in the scene. Likewise, another group was instructed to adopt the outside perspective from a model. Depictions are often accompanied by instructions on how to interpret them. Are such instructions effective for this case in particular?

Method

Diagrams With Inside Perspective Instructions

To promote the perspective of the character, the names of characters were removed from the diagrams, and participants were told that the character depicted in the diagram "was" the participant. During the presentation of direction probes, descriptions of reorientations and rotations were in the second character, referring to "you" in the scene. In addition, participants received special instructions about how to study the diagrams. Participants were explicitly instructed to think of the diagrams as depicting scenes around themselves and to build a mental model with themselves immersed in the scene. To do this, participants read the following paragraph:

> When you study the diagram, I want you to think of yourself being in that scene. You should imagine yourself being in that place and create a model in your mind of what the place is like. Try to think of what the objects look like, where they are around you, and what it would be like to be there. To help yourself remember the scene, you should think of yourself standing on some kind of platform so that you are directly in the center of the six objects. All this will allow you to

remember what the scene was like and where the objects are located. During the questions, you will be told that you have turned to face different objects, so you must be able to update your mental model of the scene by imagining yourself turning to face different objects.

Models With Outside Perspective Instructions

Participants were explicitly instructed to treat the model as an integrated object and to encode the model as whole with respect to themselves. To do this, participants read the following paragraph:

> When you study the model, we would like you to create a visual image or mental picture of what it looks like *from where you sit.* Look at the model, then see if you can picture it in your mind. Try to make your visual image as vivid and detailed as possible—picturing where the objects are located, what they are, and what they look like. This will allow you to remember what the model looked like and where the objects were located. During the questions, you will be told that Homer has turned to face different objects, so you must be able to update your mental image by imaging Homer turning to face different objects. Try to learn the model well enough that you can form a mental picture of it with Homer in any position.

Results

Data Treatment

Diagrams with inside perspective instructions. Participants made errors on 3.3% of probes, and outliers accounted for 4.9% of the data. Data of 1 male and 2 female participants were discarded because they averaged more than 45 errors per scene (roughly one in six probes). There was no overall effect of participant gender, and this variable had no effect in either the diagram or model condition. Gender did not interact with any variable in any condition.

Model with outside perspective instructions. Participants made errors on 2.9% of probes, and outliers accounted for 4.7% of the data. Men and women displayed the same patterns of RTs. There was no overall effect of participant gender, and this variable had no effect in either the diagram or model condition. Gender did not interact with any variable in any condition.

Effect of Type of Depiction

RT data were analyzed in the same way as in Experiment 1 (see Table 3). Overall RTs did not differ between the diagram and model conditions, but different patterns of RTs were observed in the diagram and model conditions. Consequently, RTs were analyzed separately in the diagram and model conditions.

Diagrams with inside perspective instructions. We predicted that inside perspective instructions would induce participants to use spatial frameworks to locate objects. Thus, participants should respond faster to head/feet than front/back for upright and upside-down orientations but faster to front/back than head/feet for the reclining orientations.

Mean RTs are presented in Table 5, and the data are consistent with predictions of the spatial framework analysis. When the character was upright, participants responded faster to head/feet than front/back, $F(1, 16) = 12.43$, $MSE = 1.60$, and faster to front/back than left/right, $F(1, 16) = 9.01$, $MSE = 1.16$. When the character reclined, participants responded faster to front/back than head/feet, $F(1, 16) = 12.50$, $MSE = 0.90$, and faster to head/feet than left/right, $F(1, 16) = 25.42$, $MSE = 1.83$. For upside-down characters, participants responded faster to head/feet than front/back, $F(1, 16) = 7.22$, $MSE = 2.79$, and faster to front/back than left/right, $F(1, 16) = 33.19$, $MSE = 12.82$. An interaction contrast revealed a significant interaction effect, $F(1, 240) = 8.026$, $MSE = 1.22$, which is consistent with predictions of the spatial framework analysis.

Participants were fastest overall when the character was

Table 5
Mean Reaction Time (in Seconds) for Memory of Diagrams With Inside Perspective Instructions and Models With Outside Perspective Instructions (Experiment 2)

Orientation	Head	Feet	Front	Back	Left	Right
	Direction					
Diagram with inside perspective instructions						
Upright	3.58	3.52	3.74	3.97	4.06	4.17
M	3.55		3.85		4.12	
Reclining	4.14	4.10	3.90	3.88	4.46	4.44
M	4.12		3.89		4.45	
Upside down	3.87	3.76	4.13	4.31	5.05	5.12
M	3.81		4.22		5.09	
Model with outside perspective instructions						
Upright	3.44	3.35	3.64	3.81	4.22	4.03
M	3.39		3.73		4.12	
Reclining	3.67	3.65	3.90	3.93	4.51	4.37
M	3.66		3.92		4.44	
Upside down	3.74	3.64	4.08	4.22	4.94	5.05
M	3.69		4.15		4.99	

upright (3.84 s), next fastest when reclining to the right (4.06 s), followed by reclining to the left (4.24 s), and slowest when upside down (4.37 s). Participants responded significantly faster for upright characters than those reclining to the right, $F(1, 16) = 17.03$, $MSE = 2.59$, and all other orientations. RTs for both reclining orientations were significantly faster than those for upside down: reclining to right, $F(1, 16) = 31.82$, $MSE = 4.84$; reclining to left, $F(1, 16) = 5.50$, $MSE = 0.84$.

Models with outside perspective instructions. In this condition, we predicted that outside perspective instructions would induce participants to use intrinsic computation to identify objects. Thus, for all orientations, participants should respond faster to head/feet than front/back and slowest to left/right.

The data are consistent with predictions of the intrinsic computation analysis (see Table 5). When the character was upright, participants responded faster to head/feet than front/back, $F(1, 19) = 14.00$, $MSE = 2.26$, and faster to front/back than left/right, $F(1, 19) = 19.44$, $MSE = 3.14$. Similarly, when the character was reclining, participants responded faster to head/feet than front/back, $F(1, 19) = 19.48$, $MSE = 1.34$, and faster to front/back than left/right, $F(1, 19) = 78.54$, $MSE = 5.41$, which indicates the intrinsic computation pattern. For the upside-down character, participants responded faster to head/feet than front/back, $F(1, 19) = 17.56$, $MSE = 4.24$, and faster to front/back than left/right, $F(1, 19) = 58.86$, $MSE = 14.21$. An interaction contrast revealed no significant interaction effect, $F(1, 285) = 0.86$, $MSE = 0.029$, which is consistent with predictions of the intrinsic computation analysis.

Participants were fastest overall when the character was upright (3.75 s), next fastest when reclining to the right (3.90 s), followed by reclining to the left (4.10 s), and slowest when upside down (4.28 s). Participants responded significantly faster for upright characters than those reclining to the right, $F(1, 19) = 5.37$, $MSE = 1.48$, and all other orientations. RTs to both reclining orientations were significantly faster than those to upside down: reclining to right, $F(1, 19) = 29.91$, $MSE = 8.25$; reclining to left, $F(1, 19) = 6.39$, $MSE = 1.76$.

Discussion

When instructed to interpret the diagrams by placing themselves in the scenes, participants adopted the inside point of view rather than the outside point of view adopted without instructions in Experiment 1. Consequently, the pattern of RTs corresponded to the spatial framework analysis rather than the intrinsic computation analysis, even though the diagrams used here were the same as those in Experiment 1. Thus, the diagrams alone do not determine the mental representations of participants. Rather, the diagrams as interpreted by the viewer determine the mental representations, and the interpretation can be altered by instruction. The instructions to create a 3D inside perspective were sufficient to overcome the lack of perceived 3D structure in the diagrams themselves, presumably because people possess a very good understanding of 3D space.

As for diagrams, instructions on how to interpret the models countered participants' spontaneous interpretations of the models. In the present experiment, instructions to interpret the models as objects viewed from an outside perspective led participants to adopt that perspective, and their RTs corresponded to the intrinsic computation analysis. This finding contrasts with the spatial framework pattern observed spontaneously for models in Experiment 1. Models, like diagrams, need to be interpreted.

Experiment 3: Degree of Pictorial Depth Cues in Diagrams

The most salient difference between the diagrams and models used in the current experiments was the strength of cues to depth. The diagrams contained a simple diagonal line that stood for the third dimension. The line was a symbolic cue and did not give the perception of the 3D structure of scenes. The models were actual 3D scenes with binocular and some monocular cues to depth.

If depth cues are responsible for the different mental representations induced by diagrams and models, then adding depth cues to the diagrams should lead participants to respond to diagrams as they do to models. Specifically, depth cues in diagrams should encourage participants to adopt an inside perspective on the scene. In this experiment, participants studied diagrams that varied in the number of monocular depth cues. Standard diagrams from earlier experiments contained no depth cues. Intermediate diagrams used converging lines and, to a lesser extent, relative size and occlusion to convey depth. These diagrams also contained symbolic information such as labels for objects. Perspective diagrams used converging lines, texture gradients, and, to some extent, relative size, to convey depth. Objects were indicated by drawings rather than labels to enhance realism. Note that the perspective diagrams did not necessarily contain better depth cues than the intermediate diagrams, only different cues. The terminology was adopted solely for convenience.

Method

Scenes

Scenes and objects are listed in Table 6.

Standard diagrams. Scenes were conveyed by the same 2D drawings used in previous experiments.

Intermediate diagrams. The standard schematic figure indicated the position of the character in scenes (see Figure 2). The character stood in a room frame with walls that provided convergent line cues to depth. The side and back walls were colored in gray tone to make them appear solid. The character was shown on a bench to indicate the need for support from gravity. No axes were drawn in the diagram. Instead, object labels were located along virtual axes from the person's torso. The objects were indicated by names, but the labels varied in size to indicate depth. The closest

Table 6
Scenes and Objects Used in Experiment 3

Scene	Objects
Standard diagrams	
Lagoon	bottle, frisbee, paddle, shell, snorkel, towel
Halloween party	bowl, ghost, mask, pumpkin, skeleton, stereo
Space exhibit	map, meteorite, portrait, rocket, satellite, spacesuit
Navy ship	anchor, antenna, cannon, flag, lifeboat, siren
Intermediate diagrams	
Barn	brush, hay, lantern, pail, saddle, shears
Bedroom	dress, hat, pants, purse, shirt, sock
Kitchen	bread, fork, pie, plate, pot, spoon
Living room	bookcase, clock, lamp, painting, table, vase
Perspective diagram	
Backyard	(toy) car, cat, kite, drum, flower, bird
Opera	bouquet, curtain, plaque, violin, sculpture, stereo
Workshed	saw, axe, desk, ruler, soap, scissors
Child's bedroom	bed, chair, globe, microscope, radio, raincoat

object was printed in 24-point font, the farthest in 10-point font, and the others in 14-point font. The front object label slightly occluded the person to further indicate depth.

Perspective diagrams. The diagrams showed a more realistic human silhouette in a scene frame (see Figure 3). The person was shown on a bench to indicate the need for support from gravity. Drawings of objects were placed along virtual axes from the character. The diagrams used converging lines, texture gradient, and relative size to convey depth.

Design

Type of depiction (standard, intermediate, and perspective) was varied within participant. Because of time constraints, participants

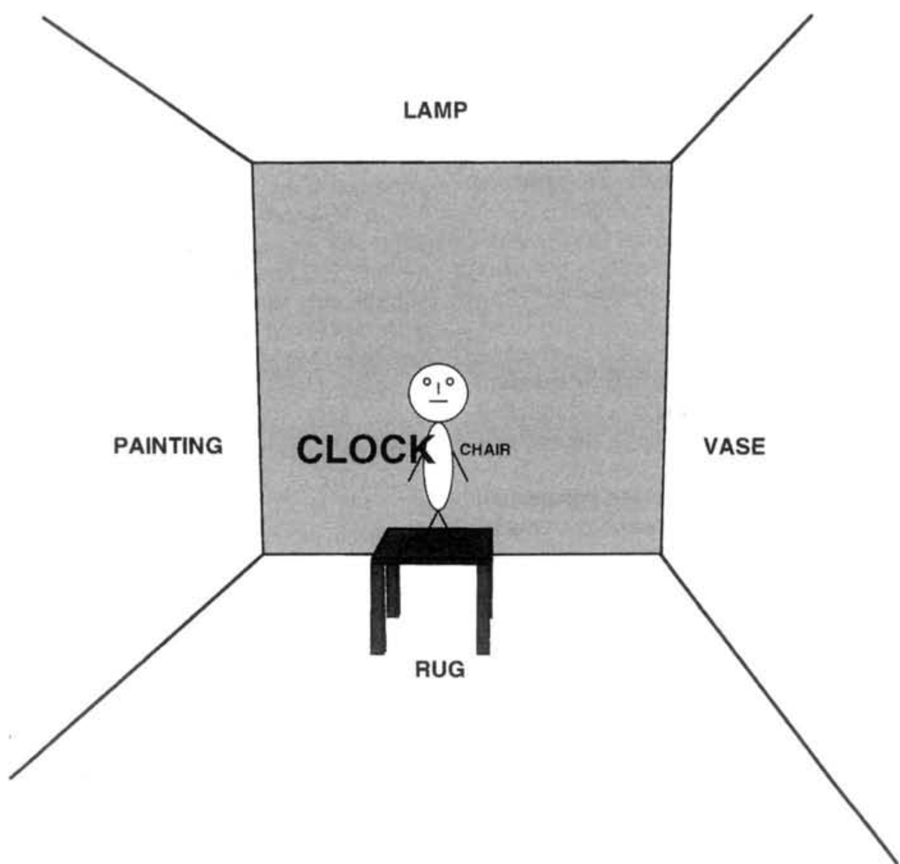

Figure 2. Example of an intermediate diagram used in Experiment 3. Depth is conveyed primarily by converging lines, but relative size of labels and occlusion provide weak cues as well.

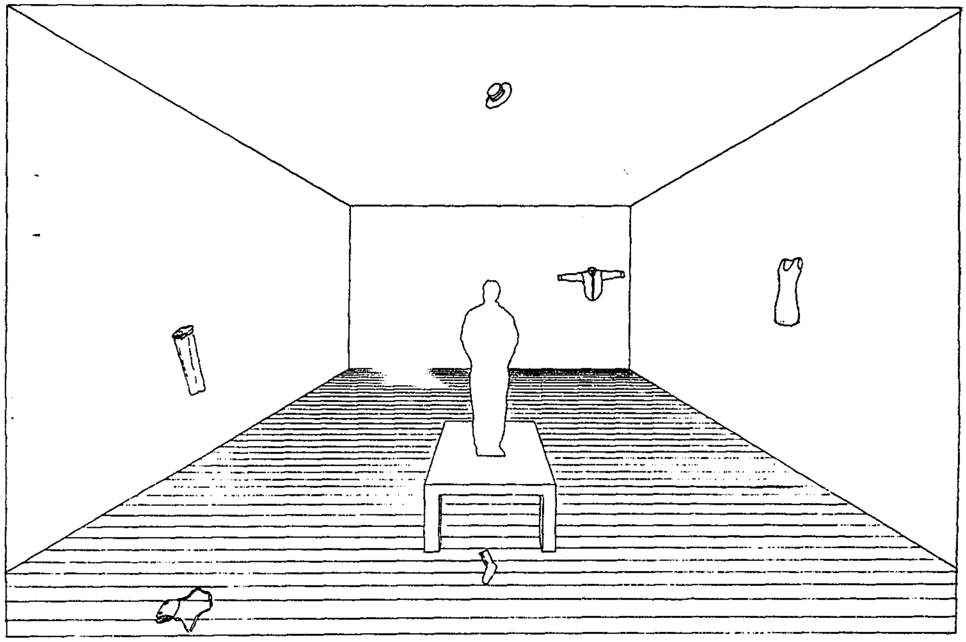

Figure 3. Example of a perspective diagram used in Experiment 3. Depth is conveyed by converging lines, texture gradient, and relative size. This figure is precisely what participants saw, including any apparent flaws in the drawing.

completed the experiment in two sessions. During the first session, participants completed two diagram conditions, then returned within 3 days to complete the third. Order of diagram condition was completely counterbalanced across the first 12 participants. Four different orders were randomly chosen for the remaining participants.

Results

Data Treatment

Participants made errors on 4.2% of probes in the standard condition, 5.1% of probes in the intermediate condition, and 4.4% of probes in the perspective condition. Of the remaining data, 4.9% were outliers in the standard condition, 5.0% in the intermediate condition, and 5.1% in the perspective condition. Men and women displayed the same patterns of RTs. There was no overall effect of participant gender, and this variable had no effect within any diagram condition. Gender did not interact with any other variable in any condition.

Effect of Diagram Condition

RT data were subjected to an ANOVA with type of depiction, orientation, and direction as within-participant variables (see Table 7). Because the ANOVA revealed an interaction of orientation and direction with type of depiction, these variables were examined separately for the standard, intermediate, and perspective diagram conditions. Results of these analyses are also shown in Table 7.

Standard Diagrams

This condition replicated the diagram condition of Experiment 1, and participants were expected to adopt the outside perspective and use intrinsic computation to identify objects. Mean RTs are presented in Table 8, and the data are generally consistent with the intrinsic computation pattern. When the character was upright, participants did not respond significantly faster to head/feet than front/back, $F(1, 15) = 3.19$, $MSE = 1.49$, but did respond faster to head/feet than left/right, $F(1, 15) = 64.99$, $MSE = 30.31$. Participants also responded faster to front/back than left/right, $F(1, 15) = 39.40$, $MSE = 18.37$. For the reclining character, participants respond faster to head/feet than front/back, $F(1, 15) = 6.25$, $MSE = 2.91$, and faster to front/back than left/right, $F(1, 15) = 33.55$, $MSE = 15.64$. Similarly, for the upside-down character, participants responded faster to head/feet than front/back, $F(1, 15) = 3.52$, $MSE = 1.64$, and faster to front/back than left/right, $F(1, 15) = 36.93$, $MSE = 63.85$. An interaction contrast revealed no significant interaction effect, $F(1, 225) < 1$, $MSE = 0.08$, consistent with predictions of the intrinsic computation analysis.

Participants were fastest overall when the character was upright (3.93 s), next fastest when the character reclined with its head to the left (4.12 s), followed by reclining to the right (4.19 s), and slowest when upside down (4.57 s). Participants did not respond significantly faster for upright characters than for reclining to the left, $F(1, 15) = 2.35$, $MSE = 1.89$, but were significantly faster for upright than all other orientations. RTs for both reclining orientations were significantly faster than those for upside down: reclining to

Table 7
Analyses of Variance for Experiments 3 and 4

Variable	Experiment 3			Experiment 4		
	df	*F*	*MSE*	*df*	*F*	*MSE*
Overall						
Type of depiction (TD)	2, 30	0.02	0.14	1, 15	1.68	19.25
TD × Orientation (O)	6, 90	1.82	1.15	3, 45	1.90	1.08
TD × Direction (D)	10, 150	3.94*	2.54	5, 75	3.04*	1.41
TD × O × D	30, 450	1.82*	0.80	15, 225	4.86*	0.99
Standard diagram						
O	3, 45	8.72*	7.02			
D	5, 75	22.41*	29.13			
O × D	15, 225	3.64*	1.70			
Intermediate diagram						
O	3, 45	14.26*	4.93			
D	5, 75	23.74*	6.73			
O × D	15, 225	5.82*	0.93			
Perspective diagram						
O	3, 45	18.15*	13.27			
D	5, 75	31.06*	17.30			
O × D	15, 225	8.36*	2.82			
Standard model						
O				3, 45	26.36*	20.67
D				5, 75	25.10*	15.40
O × D				15, 225	14.14*	4.67
Impoverished model						
O				3, 45	23.01*	6.01
D				5, 75	51.51*	9.59
O × D				15, 225	4.01*	0.52

*$p < .05$.

Table 8
Mean Reaction Times (in Seconds) for Memory of Standard, Intermediate, and Perspective Diagrams (Experiment 3)

Orientation	Direction					
	Head	Feet	Front	Back	Left	Right
Standard diagrams						
Upright	3.39	3.35	3.55	3.79	4.67	4.82
M	3.37		3.67		4.74	
Reclining	3.71	3.73	3.90	4.14	4.81	4.64
M	3.72		4.02		4.72	
Upside down	3.65	3.74	3.89	4.13	5.99	6.04
M	3.69		4.01		6.01	
Intermediate diagrams						
Upright	3.61	3.54	3.75	4.01	4.12	4.23
M	3.58		3.88		4.18	
Reclining	4.19	4.14	3.93	3.91	4.51	4.49
M	4.16		3.92		4.50	
Upside down	3.87	3.78	4.13	4.30	5.10	5.16
M	3.83		4.21		5.13	
Perspective diagrams						
Upright	3.43	3.34	3.72	3.86	4.11	4.19
M	3.39		3.79		4.15	
Reclining	4.08	4.08	3.70	3.67	4.73	4.53
M	4.08		3.68		4.63	
Upside down	3.77	3.88	4.31	4.29	6.52	5.31
M	3.81		4.30		5.92	

the left, $F(1, 15) = 11.94$, $MSE = 8.60$; reclining to the right, $F(1, 15) = 8.94$, $MSE = 7.20$.

Intermediate Diagrams

These diagrams contained some depth cues but also symbolic elements, so it was unclear whether they would allow an inside perspective. RTs, however, did correspond to the spatial framework pattern (see Table 8). For the upright character, participants responded significantly faster to head/feet than front/back, $F(1, 15) = 9.40$, $MSE = 1.50$, and faster to front/back than left/right, $F(1, 15) = 8.62$, $MSE = 1.38$. For the reclining character, however, participants responded faster to front/back than head/feet, $F(1, 15) = 12.06$, $MSE = 1.93$, and faster to head/feet than left/right, $F(1, 15) = 22.44$, $MSE = 3.59$. When the character was upside down, participants responded faster to head/feet than front/back, $F(1, 15) = 14.72$, $MSE = 2.35$, and faster to front/back than left/right, $F(1, 15) = 84.06$, $MSE = 12.43$. An interaction contrast revealed a significant interaction effect, $F(1, 225) = 7.59$, $MSE = 1.21$, indicating different patterns of RTs in the upright and reclining conditions. This result is consistent with the spatial framework analysis.

Participants were fastest overall when the character was upright (3.88 s), next fastest when the character reclined to the right (4.10 s), followed by reclining to the left (4.29 s), and slowest when the character was upside down (4.39 s). Participants responded significantly faster for upright characters than for reclining to the right, $F(1, 15) = 6.69$, $MSE = 2.31$, as well as for all other orientations. RTs for reclining to the right were significantly faster than those for upside down, $F(1, 15) = 11.91$, $MSE = 4.11$; RTs for reclining to the left were not, $F(1, 15) = 1.33$, $MSE = 0.46$.

Perspective Diagrams

These diagrams presented at least three cues to depth and minimized the use of symbolic cues. Participants in this condition were expected to adopt the inside perspective and use spatial frameworks to identify objects. RTs in this condition were, in fact, consistent with the spatial framework pattern (see Table 8). When the character was upright, participants responded significantly faster to head/feet than front/back, $F(1, 15) = 7.70$, $MSE = 2.60$, and faster to front/back than left/right, $F(1, 15) = 27.66$, $MSE = 9.33$. In contrast, when the character reclined, participants responded faster to front/back than head/feet, $F(1, 15) = 14.83$, $MSE = 5.00$, and faster to head/feet than left/right, $F(1, 15) = 28.76$, $MSE = 9.70$. For the upside-down character, participants responded faster to head/feet than front/back, $F(1, 15) = 11.25$, $MSE = 3.79$, and faster to front/back than left/right, $F(1, 15) = 209.80$, $MSE = 70.74$. An interaction contrast revealed a significant interaction effect, $F(1, 225) = 7.56$, $MSE = 2.55$, indicating different patterns of RTs in the upright and reclining orientations, which is consistent with the spatial framework analysis.

Participants were fastest overall when the character was upright (3.78 s), next fastest when the character reclined to the right (4.12 s), followed by reclining to the left (4.14 s), and slowest when upside down (4.68 s). Participants responded significantly faster for upright characters than for reclining to the right, $F(1, 15) = 7.67$, $MSE = 5.60$, as well as for all other orientations. RTs for both reclining orientations were significantly faster than those for upside down: reclining to the right, $F(1, 15) = 20.49$, $MSE = 14.98$; reclining to the left, $F(1, 15) = 18.61$, $MSE = 13.61$.

Discussion

Participants in this experiment and Experiment 1 who viewed standard diagrams took the outside perspective and used intrinsic computation to identify objects. In contrast, the results of the intermediate and perspective diagram conditions show that diagrams with greater depth cues led participants to take an inside perspective on the scenes so that RTs conformed to the spatial framework pattern rather than the intrinsic computation pattern. Depth cues conveyed the 3D structure of scenes and encouraged participants to mentally place themselves in the scenes. The depth cues in the intermediate diagrams were relatively weak but sufficient to allow participants to adopt the inside perspective. A possible explanation is that people are well trained at interpreting 2D diagrams as depictions of 3D spaces. Thus, people probably do not need very many depth cues to engage in 3D visualization for familiar situations such as these.

Experiment 4: Binocular Versus Monocular Viewing of Models

Adding depth cues to diagrams altered the perspective people took on the scenes. Can subtracting depth cues from models do the same thing? In particular, will eliminating depth cues from models lead participants to take an outside perspective on the scene? In this experiment, one group of participants viewed models under standard conditions, with full access to binocular and monocular depth cues. In the impoverished condition, participants wore an eye patch over one eye to eliminate binocular cues. Furthermore, the model itself was enclosed in a black field and lit from directly overhead to reduce shadows and other monocular cues. Participants viewed the model through an aperture so that the model appeared without context. The impoverished condition should render the model like a diagram and promote an outside perspective, leading to intrinsic computation.

Method

Scenes

Scenes and objects are listed in Table 9.

Procedure

The general procedure was followed in both viewing conditions; the only difference was how participants viewed models. In the normal viewing condition, participants sat about 0.6 m (or 2 ft) from the model with their chair adjusted so that Homer was at eye

Table 9
Scenes and Objects Used in Experiment 4

Scene	Objects
Standard model	
Backyard	(toy) car, cat, kite, drum, flower, bird
Bedroom	dress, hat, pants, purse, shirt, sock
Kitchen	bread, fork, pie, plate, pot, spoon
Living room	bookcase, chair, lamp, painting, table, vase
Impoverished model	
Zoo	lion, monkey, elephant, camel, bear, giraffe
Opera	stereo, sculpture, violin, plaque, bouquet, curtain
Barn	lantern, pail, rake, saddle, shears, hay
Laundry room	clock, iron, sewing machine, table, towel, vacuum

level. The model was presented in normal room light. In the impoverished condition, participants wore an eye patch over their nonpreferred eye to eliminate binocular depth cues. Participants sat about 0.6 m (or 2 ft) from the model, with Homer at eye level. The model was placed in a black hemispherical cardboard enclosure, and participants viewed the model through a circular hole. Thus, the model was not seen in any context. The model was illuminated by a single light from directly above.

Results

Participants made errors on 4.4% of probes in the standard condition and 4.7% of probes in the impoverished condition. Of the remaining data, 4.8% were outliers in the standard condition and 4.9% in the impoverished condition. Men and women displayed the same patterns of RTs. There was no overall effect of participant gender, and this variable had no effect within any diagram condition. Gender did not interact with any other variable in any condition.

Effect of Type of Depiction

RT data were analyzed in the same way as Experiment 3 (see Table 7). Because differing patterns of RTs emerged in the standard and impoverished conditions, the effects of orientation and direction were examined separately for standard and impoverished models.

Standard Models

This condition replicates the model condition of Experiment 1, and participants were expected to adopt the inside perspective and use spatial frameworks to identify objects. Mean RTs are presented in Table 10 and are consistent with the spatial framework pattern. When the character was upright, participants responded significantly faster to head/feet than front/back, $F(1, 15) = 6.77$, $MSE = 2.23$, and faster to front/back than left/right, $F(1, 15) = 4.56$, $MSE = 1.50$. In contrast, when the character reclined, participants respond faster to front/back than head/feet, $F(1, 15) = 5.43$, $MSE = 1.79$, and faster to head/feet than left/right, $F(1, 15) = 13.78$, $MSE = 4.55$. For the upside-down character, participants responded faster to head/feet than front/back, $F(1, 15) = 241.52$, $MSE = 79.70$, and faster to front/back than left/right, $F(1, 15) = 158.62$, $MSE = 52.35$. An interaction contrast revealed a significant interaction effect, $F(1, 225) = 7.31$, $MSE = 0.28$, indicating that RTs conformed to predictions of the spatial framework analysis.

Participants were fastest overall when the character was

Table 10
Mean Reaction Times (in Seconds) for Memory of Models Under Standard and Impoverished Viewing Conditions (Experiment 4)

	Direction					
Orientation	Head	Feet	Front	Back	Left	Right
Standard viewing						
Upright	3.44	3.38	3.70	3.86	3.90	4.28
M	3.41		3.78		4.09	
Reclining	4.09	4.10	3.79	3.72	4.66	4.59
M	4.09		3.76		4.63	
Upside down	3.74	3.86	4.24	4.21	6.66	5.41
M	3.80		4.22		6.03	
Impoverished viewing						
Upright	3.31	3.22	3.50	3.71	3.97	3.85
M	3.27		3.61		3.91	
Reclining	3.59	3.54	3.77	3.79	4.27	4.22
M	3.56		3.78		4.24	
Upside down	3.61	3.49	3.98	4.08	4.90	5.00
M	3.55		4.03		4.95	

upright (3.76 s), next fastest when the character reclined to the right (4.13 s), followed by reclining to the left (4.19 s), and slowest when upside down (4.68 s). Participants responded significantly faster for upright characters than for reclining to the right, $F(1, 15) = 8.55$, $MSE = 6.44$, as well as for all other orientations. RTs for both reclining orientations were significantly faster than those for upside down: reclining to the right, $F(1, 15) = 19.90$, $MSE = 14.99$; reclining to the left, $F(1, 15) = 15.45$, $MSE = 11.64$.

Impoverished Models

In this condition, the viewing situation deprived participants of depth cues. We expected that this would induce the outside perspective and lead participants to use intrinsic computation to identify objects. The pattern of RTs is consistent with this prediction (see Table 10). For the upright character, participants responded significantly faster to head/feet than front/back, $F(1, 15) = 14.14$, $MSE = 1.84$, and faster to front/back than left/right, $F(1, 15) = 11.26$, $MSE = 1.47$. Similarly, when the character reclined, participants respond faster to head/feet than front/back, $F(1, 15) = 5.58$, $MSE = 0.73$, and faster to front/back than left/right, $F(1, 15) = 26.81$, $MSE = 3.50$. For the upside-down character, participants responded faster to head/feet than front/back, $F(1, 15) = 28.65$, $MSE = 3.74$, and faster to front/back than left/right, $F(1, 15) = 103.40$, $MSE = 13.49$. An interaction contrast revealed no significant interaction effect, $F(1, 225) = 3.94$, $MSE = 0.13$, which is consistent with the intrinsic computation analysis.

Participants were fastest overall when the character was upright (3.60 s), next fastest when the character reclined to the right (3.77 s), followed by reclining to the left (3.96 s), and slowest when upside down (4.18 s). Participants responded significantly faster for upright characters than for reclining to the right, $F(1, 15) = 5.40$, $MSE = 1.41$, as well as for all other orientations. RTs for both reclining orientations were significantly faster than those for upside down: reclining to the right, $F(1, 15) = 31.06$, $MSE = 8.13$; reclining to the left, $F(1, 15) = 9.14$, $MSE = 2.39$.

Discussion

In conjunction with the results of Experiment 3, the present results provide further evidence that cues to depth determine the preferred mental perspective and representation of scenes. Special viewing conditions that reduced depth cues in models led participants to adopt the outside perspective and use intrinsic computation as they spontaneously do for diagrams. When depth cues are available (standard models), participants spontaneously use the inside perspective and spatial framework analysis. Without depth cues, models become like the standard diagrams used in previous experiments. In particular, the third dimension must be inferred. The array of objects was also presented without context, which presumably also made it easier for participants to view it as an object in space.

General Discussion

Mental Representations of Scenes

Spatial Framework and Intrinsic Computation Analyses

In previous research conveying the spatial information by description or by experience, participants' retrieval times corresponded to the spatial framework pattern (e.g., Bryant et al., 1992; Franklin & Tversky, 1990). This pattern was seen as the result of taking the inside perspective of the character in the scene and constructing a mental spatial framework from extensions of the body axes. Recent work suggested that perception of some diagrams (Logan, 1995) and models (Bryant et al., 1998) is based on a different perspective on the scene and, concomitantly, a different method of locating objects around the central character. We termed this procedure the intrinsic computation analysis, according to which participants adopt a perspective outside of the scene and use an intrinsic or object-centered reference frame to retrieve objects.

Both the spatial framework and intrinsic computation analyses seem to be used in real life to determine spatial relations of objects with respect to another character (or object), for example, in describing where things are located or how to get somewhere. The present research explored the conditions under which each kind of analysis is invoked. We contrasted presentation by diagram or model and examined the effect of instructions to adopt a specified reference frame and the presence of depth cues in diagrams and models.

The patterns of results are summarized in Table 11, along with results of previous research, to give an overview of manipulations that affect mental representation of scenes. We found that learning from diagrams (Table 11, A: Experiment 1) spontaneously induced the intrinsic computation pattern consistent with taking an outside viewpoint on the scene and regarding the central character as an object. In contrast, learning from models (Table 11, B: Experiment 1) spontaneously induced the spatial framework pattern consistent with taking the central character's inside viewpoint. Instructions on how to interpret the diagram or model reversed the effects. Under instructions to take the inside perspective of the central character in a diagram (Table 11, C: Experiment 2), participants' RTs corresponded to the spatial framework pattern. Similarly, under instructions to take an outside stance on models and regard the doll and surrounding objects as a whole (Table 11, D: Experiment 2), participants' RTs fit the intrinsic computation pattern. Features of diagrams and models suggest a particular perspective to participants. Notably, depth cues that convey the 3D layout of scenes favor the inside perspective and spatial frameworks (Table 11, E: Experiment 3), whereas the absence of depth cues favors the outside perspective and intrinsic computation (Table 11, F: Experiment 4).

Both mental representations—the spatial framework and the intrinsic computation—led to successful performance, generally equally fast and accurate. What is more, both mental representations can be constructed from both kinds

Table 11
Summary of Effects From the Four Experiments and Previous Studies

Orientation	Pattern of RTs	Mental frame indicated by pattern	Individual data[a]
A. Experiments 1 and 3: Memory of standard 2D diagrams			
Upright	H/F < Fr/B < L/R[b]	Intrinsic computation	24/36**
Reclining	H/F < Fr/B < L/R		21/36**
Upside down	H/F < Fr/B < L/R[c]		19/36**
B. Experiment 1: Memory of models			
Upright	H/F < Fr/B < L/R	Spatial framework	22/36**
Reclining	Fr/B < H/F < L/R		22/36**
Upside down	H/F < Fr/B < L/R		18/36**
C. Experiment 2: Memory of diagrams with inside perspective instructions			
Upright	H/F < Fr/B < L/R	Spatial framework	12/17**
Reclining	Fr/B < H/F < L/R		12/17**
Upside down	H/F < Fr/B < L/R		11/17**
D. Experiment 2: Memory of models with outside perspective instructions			
Upright	H/F < Fr/B < L/R	Intrinsic computation	16/20**
Reclining	H/F < Fr/B < L/R		16/20**
Upside down	H/F < Fr/B < L/R		16/20**
E. Experiment 3: Memory of diagrams with intermediate and perspective depth cues			
Upright	H/F < Fr/B < L/R	Spatial framework	25/32**
Reclining	Fr/B < H/F < L/R		24/32**
Upside down	H/F < Fr/B < L/R		19/32**
F. Experiment 4: Memory of models viewed without depth cues			
Upright	H/F < Fr/B < L/R	Intrinsic computation	13/16**
Reclining	H/F < Fr/B < L/R		12/16**
Upside down	H/F < Fr/B < L/R		15/16**
G. Franklin and Tversky (1990; Experiment 5); Bryant et al. (1992; Experiment 2): Memory of narrative descriptions			
Upright	H/F < Fr/B < L/R	Spatial framework	18/35**
Reclining	Fr/B < H/F < L/R		18/35**
H. Bryant et al. (1998; Experiment 1): Memory of direct experience			
Upright	H/F < Fr/B < L/R	Spatial framework	15/16**
Reclining	Fr/B < H/F < L/R		10/16**
I. Logan (1995); Bryant (in press): Perception of 2D diagrams			
Upright	H/F < Fr/B < L/R	Intrinsic computation	N/A
Reclining	H/F < Fr/B < L/R		N/A
Upside down	H/F < Fr/B < L/R		N/A
J. Bryant et al. (1998; Experiment 2): Perception of 3D models			
Upright	H/F < Fr/B < L/R	Intrinsic computation	20/24**
Reclining	H/F < Fr/B < L/R		18/24**

Note. H = head; F = feet; Fr = front; B = back; L = left; R = right. "Less than" sign indicates significantly faster reaction times (RTs) at the .05 level. "Equals" sign indicates that RTs did not differ significantly.
[a]Proportion of participants exhibiting the predicted pattern, which by chance would occur in 1 out of 6 participants. [b]H/F not significantly faster than Fr/B in Experiment 1. [c]H/F not significantly faster than Fr/B in Experiment 3.
$^{**}p < .001$ (binomial).

of graphic depiction. Why then did the models encourage taking the inside viewpoint of the character but the diagrams encourage taking an outside viewpoint? It cannot be the viewpoint on the physical depictions, as the actual viewpoints of participants on both models and diagrams were, in fact, outside. It cannot be the information represented in the external representations, as both models and diagrams represented the essential information, namely, the spatial relations of the objects to the central character. It cannot be the fidelity of the graphic (Ferguson & Hegarty, 1995; Gibson, 1966; Schwartz, 1995). The least "faithful" and most abstract way of conveying the spatial information is through verbal description. The most faithful method is through models, with diagrams somewhere in between. Yet the mental representations adopted from descriptions and models are the same (Bryant et al., 1998) and differ from the mental representations adopted from diagrams.

When do people use spatial frameworks and when do they use intrinsic computation? The broad spectrum of findings summarized in Table 11 suggests an answer. People use spatial frameworks when they can readily create an understanding of a 3D layout and imagine themselves in it. Thus, the inside perspective is crucial. Well-crafted narratives can impart a 3D understanding with an inside perspective (Table 11, G: Bryant et al., 1992, Experiment 2; Franklin & Tversky, 1990, Experiment 5). Although narratives present no perceptual information, they invoke extensive knowledge of environmental space. People can draw upon this knowledge to build an inside world. Experiencing a scene or viewing a physical model or depth-enriched diagram of one can also convey the complex 3D relations among the character and objects (Table 11, H: Bryant et al., 1998, Experiment 1).

In contrast, people use intrinsic computation when they can only rely on a representation of a scene from a particular vantage point without good cues to depth. In other words, the outside perspective leads participants to use intrinsic computation in this paradigm. This is especially true when people are actually observing a display (Table 11; I, J: Bryant, in press; Bryant et al., 1998, Experiment 2; Logan, 1995); then it is difficult to ignore one's own perspective and mentally take another. Yet even when working from memory, when the diagram or model has poor cues to depth, it seems to be easier to use intrinsic computation. When a depiction does not facilitate a 3D perception, as in the flat diagrams, people use an outside perspective. Then they treat the diagram as an object to be mentally examined as a whole.

Responses to displays are not immutable. People use spatial frameworks for flat diagrams when instructed to take an insider's perspective, and they use intrinsic computation from 3D models when instructed to take an outsider's perspective. Thus, the adopted perspective determines the times to retrieve spatial information, and both the display and the instructions affect the adopted perspective.

Perspective

Ascertaining the directions among elements of a scene requires several component processes: Observers need to (a) interpret the scene, (b) distinguish target and reference elements, (c) take a viewpoint, and (d) determine the origin for the description of the spatial relations. Each of these components has been called *perspective,* separately and together. Note that the viewpoint and reference elements are aspects of the spatial array of objects, which are, of course, open to influence by linguistic and other factors. The description of spatial relations, however, is purely linguistic. Levinson's (1996) analysis, which updates Levelt's (1984) and Talmy's (1983) analyses, among others, includes most of these components but in a different configuration that does not conveniently include all the cases that have been investigated. The spatial framework and intrinsic computation analyses have included all of these processes. In many cases, some of the aspects of perspective—the viewpoint, the reference object, and the origin of the spatial relations—coincide. This is the case for the simplest version of the spatial framework analysis, the internal spatial framework analysis developed by Franklin and Tversky (1990) and Bryant et al. (1992). For this case, the viewpoint, reference object, and origin coincide and are embodied in the central person, surrounded by target objects. This simplest case is also the one investigated here.

It is not necessary, however, for the viewpoint, reference object, and origin to coincide. Recent work has considered situations in which they do not. Bryant et al. (1992) explored the so-called external spatial framework analysis, in which only the viewpoint and origin coincide. The viewpoint and origin were embodied in an observer who was regarding a reference element (a person or object) surrounded by target objects. The task of the participant was to report the directions of the targets to the reference object from the viewpoint or origin of the observer (i.e., using the observer's body sides, not those of the person in the array of objects). This case led to slight variations in the pattern of response times to retrieve target objects in specified directions. For the internal case, in which the viewpoint, reference, or origin was surrounded by target objects, RTs to front were faster than RTs to back. For the external case, where the reference element was surrounded by targets but the viewpoint or origin was not, RTs to front and back did not differ.

The intrinsic computation analysis encompasses yet another way of interpreting a scene and imposing a viewpoint, reference element, and origin on it. In this case, the reference element and the origin coincide in the character at the center of the array of objects, but the viewpoint is external. The origin and the viewpoint are no longer embodied, as in previous cases. Consequently, the method of determining directions from reference to the targets differs. In the case of the spatial framework analysis, the origin and viewpoint coincide, so directions are determined in relation to a mental framework based on the body experienced from inside. RTs depend on the relative accessibility of the body axes. In the case of the intrinsic computation analysis, the viewpoint and origin differ, so directions are determined in relation to the sides of a character experienced from outside. RTs depend on the order of determining the object's sides. Thus, the spatial framework analysis applies when viewpoint and origin coincide, whereas the intrinsic computation analysis applies when reference object and origin coincide.

The cases that have been explored are rooted in natural situations, such as keeping track of locations of objects surrounding one's self (spatial framework analysis) or keeping track of the positions of objects in immediate space (intrinsic computation analysis). Other combinations of viewpoint, reference object, and origin of direction terms remain to be explored. Some of these combinations seem less tractable, perhaps because they are less rooted in natural situations. Consider, for example, the case in which viewpoint and reference object coincide but the origin of the spatial direction terms does not coincide with either. Describing such a case is awkward, if at all possible. Imagine the following scene, constructed to fit those requirements. A chair and a house are lined up in tandem so that the chair is in front of the house from the point of view of the house, and the back of the chair is closest to the house. The challenge is to describe the location of the house with respect to the chair using the house as the origin of the spatial expressions. "The house is in back of the chair," though true, does not fit the case under consideration because it treats the chair as both reference object and origin. "The chair is in front of the house," though true, also does not fit because it treats the house as both reference object and origin. "The house's front faces the back of the chair" is arguably an example of using the house as origin and the chair as referent, but the sentence does not have a clear viewpoint. The difficulty in formulating different combinations of viewpoint, reference object, and origin suggests that there is a limited number of perspectives people use to describe space.

External Representations

External representations, whether graphic or linguistic, serve many purposes (e.g., Larkin & Simon, 1987; Mayer & Gallini, 1990; Stenning & Oberlander, 1995; Suwa & Tversky, 1996; Tversky, 1995a, 1995b; Winn, 1989). They can facilitate memory. They can organize information and focus attention on particular elements of information. Of special concern here, they can facilitate inference and insight.

External representations, like diagrams, models, and language, are schematic. Schematizations by nature simplify the world they represent; consequently, they may omit, add, and distort information about the world that they represent. From external representations, people form mental representations, which further schematize the information about the world. As the present research confirms, however, even when external representations are informationally equivalent (see Larkin & Simon, 1987), they can encourage quantitatively (e.g., Dwyer, 1978; Ferguson & Hegarty, 1995) and qualitatively (e.g., Bauer & Johnson-Laird, 1993; Gattis & Holyoak, 1996; Schiano & Tversky, 1989; Tversky & Schiano, 1992) different mental representations.

In the present experiments, we studied the effects of different depictions on conveying a paradigmatic 3D scene of a character surrounded by objects to six sides of the body. Although 3D models and depth-enriched diagrams promoted a 3D mental representation but flat diagrams did not, instructions to adopt an inside perspective enabled 3D mental representation from flat diagrams. This stands in contrast to the studies of Shah and Carpenter (1995) in which instruction to visualize the third dimension from flat graphs were unsuccessful. Visualization in three dimensions from a flat diagram, then, is not a general cognitive ability. Rather it depends on experience with the 3D domain depicted. Despite some naturalness of mapping space to space in diagrams, diagrams are schematic and may require tutelage to be fully exploited.

References

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 6,* 372–378.

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

Braine, L. G., Schauble, L., Kugelmass, S., & Winter, A. (1993). Representation of depth by children: Spatial strategies and lateral biases. *Developmental Psychology, 29,* 466–479.

Brown, L. A. (1949). *The story of maps.* New York: Dover.

Bryant, D. J. (in press). Mental frames for interpreting direction terms. In P. Olivier & W. Maass (Eds.), *Vision and language.* New York: Springer-Verlag.

Bryant, D. J., Lanca, M., & Tversky, B. (1995). Spatial concepts and perception of physical and diagrammed scenes. *Perceptual and Motor Skills, 81,* 531–546.

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

Bryant, D. J., Tversky, B., & Lanca, M. (1998). *Retrieving spatial relations from observation and memory.* Manuscript submitted for publication.

Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 26–63). New York: Academic Press.

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.

Dwyer, F. M. (1978). *Strategies for improving visual learning.* State College, PA: Learning Services.

Ferguson, E. L., & Hegarty, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems. *Cognition and Instruction, 13,* 129–160.

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

Fillmore, C. J. (1976). Frame semantics and the nature of language. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Annals of the New York Academy of Sciences: Vol. 280 Origins and evolution of language and speech* (pp. 20–32). New York: New York Academy of Sciences.

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.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1–9.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

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

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.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–100.

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, the Netherlands: VNU Science Press.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT 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 the tops of rotated pictures. *Canadian Journal of Psychology, 40,* 368–387.

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

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective.* Belmont, CA: Wadsworth.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82,* 715–726.

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

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

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

Schiano, D., & Tversky, B. (1992). Structure and strategy in viewing simple graphs. *Memory & Cognition, 20,* 12–20.

Schwartz, D. L. (1995). Reasoning about the referent of a picture versus reasoning about the picture as the referent: An effect of visual realism. *Memory & Cognition, 23,* 709–722.

Shah, P., & Carpenter, P. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124,* 43–61.

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

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19,* 97–140.

Suwa, M., & Tversky, B. (1996). What architects see in their sketches: Implications for design tools. *Proceedings of the 13th conference on human factors in computing systems* (pp. 191–192). New York: Association for Computing Machinery.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225–282). New York: Plenum Press.

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

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 14–24). Berlin, Germany: Springer-Verlag.

Tversky, B. (1995a). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), *Understanding images* (pp. 29–53). New York: Springer-Verlag.

Tversky, B. (1995b). Perception and cognition of 2D and 3D graphics. *Proceedings of the 13th conference on human factors in computing systems* (p. 175). New York: Association for Computing Machinery.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23,* 515–557.

Tversky, B., & Schiano, D. (1989). Perceptual and conceptual factors in distortions in memory for maps and graphs. *Journal of Experimental Psychology: General, 118,* 387–398.

Wilford, J. N. (1981). *The mapmakers.* New York: Alfred A. Knopf.

Willats, J. (1990). The draughtsmans's contract: How an artist creates an image. In H. Barlow, C. Blakemore, & M. Weston-Smith (Eds.), *Images and understanding* (pp. 235–254). Cambridge, England: Cambridge University Press.

Winn, W. (1989). The design and use of instructional graphics. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 125–143). Amsterdam: Elsevier.

Received October 25, 1996
Revision received June 9, 1998
Accepted June 9, 1998 ■