Spatial Mental Representations

Barbara Tversky
Department of Psychology, Stanford University

Imagery as Internalized Perception. The classic work on imagery has emphasized the similarity of images to percepts and of operations on images to perceptual or physical operations. Elegant experiments have supported this view: smaller imagined features take more time to verify, larger imagined distances take more time to scan, and figures displaced by greater angles take longer to mentally rotate (see reviews by Finke & Shepard, 1986; Kosslyn, 1980). These demonstrations, however, have been done under special circumstances, such as instructing and training subjects to use imagery, and keeping pictures present. Do visual or spatial representations acquired and used under more natural circumstances have this character? Two classes of case, cognitive maps and spatial mental models, reveal other characteristics.

Cognitive Maps. Like many useful concepts, the term *cognitive map* has many senses, leading to inevitable misunderstandings. One common use is an image of a map or environment, a coherent whole that can be mentally inspected. Since that use is more specific than necessary and violated by data, the term is used here very broadly, to refer to whatever mental apparatus, representations or processes, underly the task behaviors.

First, an observation about the (non-psychological) task of locating landmarks in a town, cities in a state, continents in the world, stars in the sky, analogous tasks, and analogous, save scale, to locating objects in a room. Such a task cannot be performed absolutely, but only relatively, to other objects, landmarks, cities, continents, and stars, on the same level of analysis, and/or to a frame of reference, the room, highways, canonical coordinates, on a more general level of analysis. Like measurement, memory for spatial location is relative to other locations and to a frame of reference. This constructivist view can be contrasted with a snapshot theory of environmental memory, that we remember what we see. It would not be effective to remember snapshots of the world we explore because we are likely to have different viewpoints the next time. We need to know where things are relative to one another, not relative to our own particular viewpoint at a particular moment in time. Despite controversy on many other issues, researchers from widely different perspectives agree that more general representations of environments are constructed from particular views.

The strongest evidence for the constructivist view in general, and for the use of other landmarks and a frame of reference in particular, comes from systematic errors in memory for maps and environments (Tversky, 1981, 1991a). The use of standards, both landmarks and frames of reference, yield errors following a similar pattern. When an entity is represented with respect to some standard or anchor, it is distorted in the direction of that standard. The error that occurs as a consequence of representing one landmark in terms of another was termed *alignment.* To illustrate, college students were asked to select which of two maps of the Americas was correct, the correct one, or one altered by moving South America more directly south of North America. In fact, the east coast of North America barely overlaps the west coast of South America. A majority of students selected the incorrect, aligned map. Similarly, students preferred a map of the world in which Europe was more aligned with the United States and Africa with South America to the true world map.

The error that occurs as a consequence of encoding a land mass in terms of a frame of reference was termed *rotation*. The remembered orientation of South America serves to illustrate this. Students were given a cutout of South America and asked to orient it in a north-south east-west frame. Most of them put South America more upright than it actually is. In fact, South America appears quite tilted relative to the canonical directions. In general, elongated figures appear tilted when their primary axis is not the same as one of the canonical axes. Likewise, students' estimates of directions between Bay Area cities revealed that they remembered the San Francisco Bay Area as oriented more north-south and less east-west than it actually is.

Alignment and rotation appeared for a variety of stimuli and tasks: direction estimates as well as map recognition; artificial and real maps; environments learned by direct experience and those learned from maps. They are not the only systematic errors and biases that have been demonstrated in cognitive maps (Tversky, in press). People incorrectly believe that Reno is east of San Diego; direction estimates between cities are distorted toward the overall directions of their respective states (Stevens & Coupe, 1978). Those taking an Atlantic Coast perspective judge the distance from San Francisco to Salt Lake City to be smaller than those taking a Pacific coast perspective. This reverses for judgments of the distance between New York City and Pittsburgh (Holyoak & Mah, 1982). Finally, people judge the distance from an ordinary location to a salient landmark to be shorter than the (same) distance from the landmark to the ordinary location (Sadalla, Burroughs & Staplin, 1980).

Such errors not only constitute evidence for a constructivist position as opposed to a snapshot position, they also constitute evidence against a view of a cognitive map as image-like, as internalized perception. For one thing, systematic distortions are not consistent with internalized perceptions because they do not occur in perception, or occur only in a highly reduced fashion. Finally, such errors contradict the idea of a cognitive map as a coherent whole. Rather, our representations of maps and environments appear to be piecemeal, and put together from different bits and kinds of information, with no guarantee of consistency.

Spatial Mental Models. Cognitive maps are prototypically acquired from exploring an environment or from studying a map, that is, from visual or spatial experience. One of the major functions of language is to convey experience vicariously. From descriptions of simple environments--a zoo, a convention center, a town, a recreation area--readers drew maps where landmarks and the spatial relations among them were nearly as accurate as those drawn by subjects who studied maps instead of descriptions (Taylor & Tversky, 1992).

What is the nature of spatial mental models acquired from text (Tversky, 1991b)? First, there seem to be representations at several levels of abstraction. The descriptions in Taylor and Tversky's experiments took one of two perspectives, route or survey. The route descriptions took readers on a mental tour of the environment, describing landmarks in terms of right, left, front, and back, with respect to the reader. The survey descriptions took a bird's eye view, and described landmarks relative to one another in terms of north, south, east, and west. Subjects read either a survey or a route description of each environment, and then responded true or false to verbatim statements from the read text and the unread text, and to inference statements from both perspectives. In four experiments, subjects were faster and more accurate to verbatim statements, but, for inference statements, responded as quickly and accurately to statements from the read perspective as from the other perspective. This, along with other aspects of the data, led us

to conclude that from descriptions with either perspective, subjects construct mental representations that are more abstract than either perspective, representations similar to structural descriptions, that include all the parts and spatial relations among them in a perspective-free manner. Presumably, to verify a particular statement, readers use the more general mental model to take the specific perspective required to respond. Unlike images, such representations cannot be visualized as wholes, though particular perspectives on them can be.

Another project has examined spatial perspective-taking in more detail in a small-scale environment (Franklin & Tversky, 1990; Bryant, Tversky & Franklin, 1992; Bryant & Tversky, 1992; Franklin, Tversky & Coon, in press). Typically, narratives described a scene, in an opera house, say, or hotel lobby, where a character is surrounded by six objects, located at head, feet, front, back, left, and right. The narrative periodically reoriented the character in the environment, and subjects were timed to identify the objects currently in the six directions from the character. According to the *mental transformation* model, based on research and theory in imagery, subjects would perform this task by imagining themselves facing the given object and mentally turning to mentally inspect the probed direction. If subjects did that, they should be fastest to identify the object straight ahead, slowest to the object behind, and intermediate to the objects displaced by 90 degrees, head, feet, left, and right. In a dozen experiments, one explicitly directing subjects to adopt this strategy, a different pattern of data emerged.

Instead of reflecting the time it takes to mentally inspect the described environment, times to identify objects along the three body axes fit the pattern predicted by the *spatial framework* analysis of the relation of the body to the psychological world. The analysis is based in part on analyses of spatial language by Clark (1973), Fillmore (1975), Garnham (1989), Levelt (1984), Miller and Johnson-Laird (1976), and Shepard and Hurwitz (1984). In simple situations, subjects adopt the perspective of the character in the narrative. As before, for the discussion of cognitive maps, it is proposed that people use a reference frame to keep track of the objects. In this case, the reference frame is what we called a spatial framework, a mental scaffolding constructed from the axes of the body, head-feet, front-back, left-right, with the objects "attached" to the specified places. For an upright character, the head-feet axis should be most salient because it corresponds to the only natural axis of the world, the gravitational axis, and because the head-feet axis is asymmetric. Next most salient is the front-back axis, which is asymmetric and splits the psychological world into the half that can be easily perceived and manipulated and the half that cannot. The left-right axis neither corresponds to an axis of the world, nor has salient asymmetries. When the character reclines, no body axis corresponds to gravity, so the head-feet axis loses its primacy. The physical, perceptual, and behavioral asymmetries of the front-back axis dominate, so that axis is most salient, followed by head-feet, and then left-right. Retrieval times followed these patterns in a large number of studies.

Clearly such a pattern, while deriving from people's conceptions of the perceptual world, does not correspond to internalized perception. Moreover, it does not correspond to what subjects do "for real," when they are actually surrounded by objects and probed for them by directions (Bryant & Tversky, 1991). For the traditional imagery-as-internalized-perception view to hold, behavior in imagined situations must be the same as behavior in perceived situations, though this is rarely tested.

Mental Reference Frames. The classic view of imagery is as internalized perception, and it has been demonstrated in many clever experiments. Besides its empirical support, this position has theoretical appeal. Its predictions are straightforward and easily derived: mental representations will resemble physical ones, and mental processing physical processing. Despite its elegance and simplicity, it does not seem to generalize to other spatial and visual situations of interest. Rather than being based on internalized perceptions, many spatial mental representations appear to be based on our conceptions of the visual-spatial world. Our conceptions of the world are at once more complex, more situation-bound, and more interesting. One pervasive feature of spatial representations is that they appear to be constructed relative to mental reference frames. Both perceptual and conceptual factors influence selection of a mental reference frame (Tversky, 1981, 1991a; Tversky & Schiano, 1989; Schiano & Tversky, 1992): for nonsense blobs or lines, the sides of the page; for geographical entities, the north-south east-west coordinates; for lines perceived as graphs, the imaginary diagonal; for objects surrounding a person, the three body axes; for objects in front of a person, the three axes of the world. The selection of a mental reference frame has consequences for errors in memory and for information retrieval times in a wide variety of situations.

References

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. (1991). Locating objects from memory or from sight. Paper presented at the meetings of the Psychonomic Society, San Francisco.

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

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

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

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

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

Franklin, N., Tversky, B., & Coon, V. (In press). Switching points of view in spatial mental models acquired from text. *Memory and Cognition.*

Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition, 31,* 45-60.

Holyoak, K. J. & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic

magnitude. *Cognitive Psychology, 14,* 328-352.

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

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind, & J. J. Koenderink (Eds.), *Limits on perception* (pp. 323-358). Utrecht, The Netherlands: VNU Science Press.

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

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 516-528.

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

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

Stevens, A. & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 13,* 422-437.

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

Tversky, B. (In press). Distortions in cognitive maps. *Geoforum.*

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology, 13,* 407-433.

Tversky, B. (1991a). Distortions in memory for visual displays. In S. R. Ellis (Ed.) and M. K. Kaiser and A. Grunwald (Associate Editors), *Pictorial communication in virtual and real environments.* Pp. 61-75. London: Taylor and Francis.

Tversky, B. (1991b) Spatial mental models. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 27. Pp. 109-145. N. Y.: Academic Press.

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.