

STRUCTURES OF MENTAL SPACES

How People Think About Space

BARBARA TVERSKY is a professor of psychology at Stanford University. She earned her Ph.D. in cognitive psychology from the University of Michigan. Her research interests include spatial thinking and language, memory, event perception and cognition, categorization, and diagrammatic reasoning.

ABSTRACT: Human activity takes place in space. To act effectively, people need mental representations of space. People's mental representations of space differ from space as conceived of by physicists, geometers, and cartographers. Mental representations of space are constructions based on elements, the things in space, and the spatial relations among them relative to a reference frame. People act in different spaces depending on the task at hand. The spaces considered here are the space of the body, the space around the body, the space of navigation, and the space of graphics. Different elements and spatial relations are central for functioning in the different spaces, yielding different mental representations.

Humans act in space. Sometimes, interactions in space are explicit, as we grasp the things around us or find our ways inside and out. Other interactions are implicit, an awareness of where we are and what things surround us. Still other spatial activities are in imagination, when we estimate the distance from San Francisco to Los Angeles, or give directions from the airport to the hotel, or describe a hiking trip, or read a vivid novel. To act effectively in space, people need mental representations of space. The knowledge underlying mental representations of space and the things in it comes from many sources: from looking, from hearing, from touching, from imagining, and from language. The knowledge obtained from each source is different, sometimes integrated and coherent, other times not. Nevertheless, each source can support spatial activities, and for many actions, the sources are interchangeable. We can find the hotel from the airport using a map or using directions.

The mental representations that people form of space from these real and imagined interactions differ from the external representations of space of

geometry or of physics or of maps. For geometry, physics, and cartography, space is the foundation; it is typically metric, uniform, and unitary. Things can then be located in those measured spaces. In human conceptions of space, the things in space are fundamental, and the qualitative spatial relations among them with respect to a reference frame form a scaffolding for mental spaces. Which elements or things are selected and which spatial relations are chosen as relevant depend on the space and the functions it serves us. We interact with many spaces, the space of the body in yoga or dance, the space around the body in basketball or housecleaning, the space of navigation in wayfinding or estimating distances, and the space of graphics in reading maps or understanding diagrams. Each of these spaces is represented schematically in terms of the things and spatial relations that are important for functioning within it. Here we examine four spaces, the space of the body, the space around the body, the space of navigation, and the space of graphics (see Tversky, 2000b, for more detail). For each, we will consider the functions it serves, the elements and spatial relations important to those functions, and the mental representations consequentially engendered

THE SPACE OF THE BODY

One space critical for human interaction is the space of the body. The body is naturally divided by its joints into parts. We move the various parts independently depending on the activity we are engaged in. We keep track of where the parts of our body are as we move in space. The different parts of the body interact in different ways with the surroundings. Feet walk and kick, hands grasp and manipulate, heads sense through eyes and ears and communicate through mouth and expression. The different parts also vary in size, perceptual prominence, and functional significance. Which of these factors underlies mental representations of the body, that is, which factor determines the accessibility of the parts? According to the classical account of mental imagery, because large parts are detected faster in imagery, they should also be verified faster (e.g., Kosslyn, 1976). Theories of object recognition based on parts would predict that salient parts, those with greater contour discontinuity, should be verified more quickly (Biederman, 1987; Hoffman & Richards, 1984). Finally, theories of object understanding should predict that more significant parts should be verified faster (Tversky & Hemenway, 1984). More significant parts are those that enjoy both perceptual salience and functional significance. On the whole, parts that are more salient, that is, those with greater contour discontinuity, are also those that have greater significance;

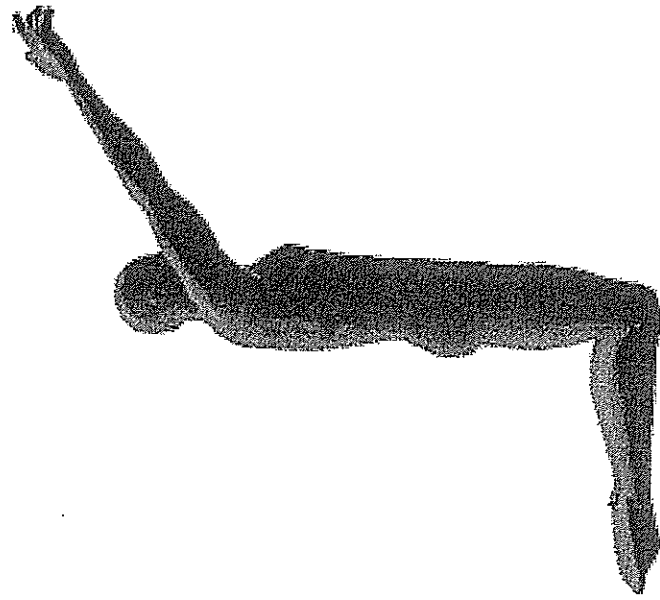


Figure 1: Examples of the Body Stimulus Used by Morrison and Tversky (1997)

think of the legs and back of a chair or the trunk of a tree. Thus, distinguishing accounts based on part salience from accounts based on part significance relies on subtle measures.

To determine which factor, part size, part salience, or part significance underlies mental representations of bodies, Morrison and I (1997; see also Tversky, Morrison, & Zacks, 2002) examined times to verify body parts in name-body and body-body comparisons. We chose parts that are named frequently across languages—head, arm, hand, chest, back, leg, and foot (C. H. Brown, 1976). In the *name* stimuli, the name of a body part appeared on the screen. In the *body* stimuli, a realistic rendering of a body in one of many poses and orientations appeared on the screen with a part highlighted by a white dot (see Figure 1 for examples of some of the body stimuli). Twenty participants responded to more than 300 pairs of name stimuli followed by body stimuli; 19 participants responded to more than 300 pairs of body stimuli. In each case, the task of participants was to say whether the parts referred to by name-body pairs or body-body pairs were the same or different. On half the trials, the correct answer was "same," on half, "different."

The data of interest are the relative speeds of verifying the different parts. If large parts are the essential elements of the way we think about bodies, then parts such as leg and back should be verified relatively faster. If part salience is the critical variable underlying mental representations of the body, then parts such as head and hand should be relatively faster to verify. Part significance, however, is correlated with part salience so that theory also predicts that head and hand should be relatively rapid. In fact, times to verify body parts were faster for the salient and significant parts, such as head and hand, and slower for the large parts, such as back and leg. For bodies as for some other things, the correlation between contour discontinuity and functional significance is not perfect; for the body parts considered here, "chest" is relatively low on contour distinctiveness, but it is relatively high on functional significance because it includes important internal body parts, such as lungs and heart, and is the front of the body. Interestingly, perceptual salience accounted for verification times better than part significance for the body-body comparisons. These can be done without accessing the meaning of the visual stimuli, purely on perceptual grounds. In fact, the impression of participants in the body-body comparison task was that the stimuli were perceived as shapes with dots, not interpreted as human bodies. For name-body comparisons, functional significance accounted for the verification times better than part salience. Names evoke function in this and in many other tasks. Names are abstract and required the participants to think of the stimuli as bodies in order to find the named part. The fact that perceptual features of objects and bodies correlate with more abstract, functional features is of fundamental importance, especially as it occurs in other domains. Because of this, humans can (and do) use perceptual salience as a sign of functional significance and use perceptual features of things to infer their functions. In sum, mental representations of bodies are organized around significant body parts.

THE SPACE AROUND THE BODY

Another space with functional significance for human activity is the space around the body, the space of things that can be seen and often reached from the current position. Franklin, Bryant, and I have investigated this space in more than a dozen experiments. We initially used narratives that described three-dimensional environments to examine the nature of spatial mental models created by language rather than perceptual experience.

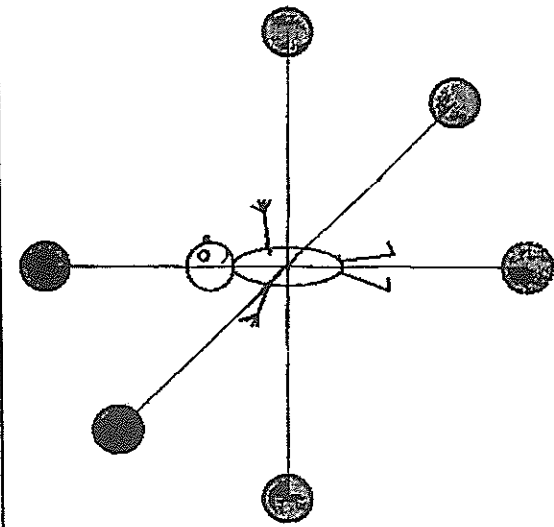


Figure 2: Schematic of the Spatial Framework Situation Used by Franklin and Tversky (1990) and Bryant, Tversky, and Franklin (1992)

We later used real scenes, diagrams, and models of scenes to instill the environments. The narratives described "you" in an environment like a museum or a barn surrounded by objects in front, back, left, right, above, and below you (Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990; Franklin, Tversky, & Coon, 1972; Tversky, 1991) (see Figure 2 for a schematic of the situation). The objects were typical of the scenes, but their locations were selected at random; that is, the chandelier in the opera was not necessarily above the observer. Participants, approximately 20 per experiment, first studied the descriptions of the environments. Then they turned to a computer for the experimental trials. On each trial, the computer reoriented them to "face" another object in the scene. Then the computer probed them for the objects currently located in the six directions from the body by naming the direction. Participants chose the correct object from the entire list as quickly and accurately as possible. After responding to probes from all the directions from the body, participants were turned to face another object and probed again. After all positions had been probed, participants were given a new environment to learn. In a typical experiment, participants studied 6 to 10 narratives.

Accuracy levels to provide the objects in the directions from the body were very high, showing that people are able to learn environments such as

these easily from descriptions and to easily update their positions in the environments. The data of interest, then, were the relative times to retrieve the names of the objects in the various directions from the body. Three theories of the retrieval times were considered. According to the equiavailability theory, all directions should be equally fast. This is because no region of space is privileged. According to the imagery theory, participants should imagine themselves in the environments facing toward the specified object. To retrieve objects in each direction, they should imagine themselves turning to face that direction. If so, retrieval times should increase the farther the turn so that front should be fastest, followed by head, feet, left, and right. Back should be slowest. The retrieval times differed systematically, rejecting the equiavailability theory. However, they did not support the imagery theory either as responses to left and right were slower than those to back.

The pattern of retrieval times indicated that the space around the body is conceived of three-dimensionally from a reference frame based on extensions of the three major body axes, head/feet, front/back, and left/right. The relative times to respond to each axis depended on asymmetries of the body and of the world. Times to report the objects to head and feet were fastest, presumably because the head/feet axis is asymmetric and is correlated with the only asymmetric axis in the world, the axis of gravity. Front/back was next as it is an asymmetric axis of the body but not correlated with an asymmetric axis of the world. Left/right is slowest as it lacks salient asymmetries and does not correlate with an asymmetric axis of the world. The situation changes slightly when "you," the observer in the situation, are described as reclining in the environment and turning from front to back to sides. In that case, no body axis is correlated with a salient axis of the world, so retrieval times depend only on the asymmetries of the body. Front/back is faster than head/feet for the reclining observer, presumably because the front/back axis is more significant than the head/feet axis; the front/back axis is not only asymmetric, it also separates the world that can be easily seen and manipulated from the world that cannot be easily seen and manipulated. Thus, the asymmetries of the front/back and head/feet axes of the body are both perceptual and functional. This account of the pattern of reaction times has been termed the *spatial framework theory*. It reflects people's enduring conceptions of the spatial world that they inhabit rather than momentary internalized imagery of the current scene.

Variations of this situation have been tried, with sensible variations in the patterns of retrieval times. The spatial framework pattern of retrieval times holds for arrays of objects in front of the observer as well as those surrounding the observer. The spatial framework pattern emerged when scenes were acquired from diagrams (such as that in Figure 2), models, or real life instead

of from narratives as long as the act of responding is conducted from memory (e.g., Bryant & Tversky, 1999; Bryant, Tversky, & Lanca, 2001). When the narratives described the environments as moving rather than the observer, the spatial framework pattern appeared as soon as new viewpoints were adjusted to. However, adjusting to new viewpoints took twice as long when the environment was described as moving as when the observer was described as moving (Tversky, Kim, & Cohen, 1999). This finding also illustrates the influence of enduring conceptions of the perceptual world on mental representations of the spatial world. In the world as experienced by people, it is people who move, not environments, save unusual circumstances such as earthquakes.

In short, mental representations of the space around the body appear to be three-dimensional, with the dimensions defined by extensions of the axes of the body. Times to retrieve objects in directions from the body can be accounted for by perceptual and functional asymmetries of the body axes and the axes of the world.

THE SPACE OF NAVIGATION

The space of navigation is the space we explore, the space we inhabit as we move from place to place, typically a space too large to be seen at once. One remarkable feat of the human mind is to conceive of some large spaces as integrated wholes rather than piecemeal as they are experienced. Similar to the space around the body, the space of navigation is a mental construction that is schematized. Certain information, such as exact metric information, is systematically simplified and even distorted. The critical elements of the space of navigation are landmarks and paths, links and nodes. Similar to the space around the body, spatial relations in the space of navigation are relative to a reference frame. In the case of the space of navigation, several reference frames are possible, primarily based on viewer, object, or environment (e.g., Taylor & Tversky, 1996). Directions and axes are not represented analogically or metrically in exact degrees or meters but rather somewhat categorically. It is this schematization into elements and paths relative to reference frames that allows integration of fragments into a whole.

Some evidence for the nature of the schematization of the space of navigation comes from an analysis of route maps and directions. Both sketches and descriptions schematize environments in the same way (Tversky & Lee, 1998, 1999). Denis (1997) analyzed a large corpus of spontaneous route directions, finding that they consist of sequences of segments of reorientations, actions,

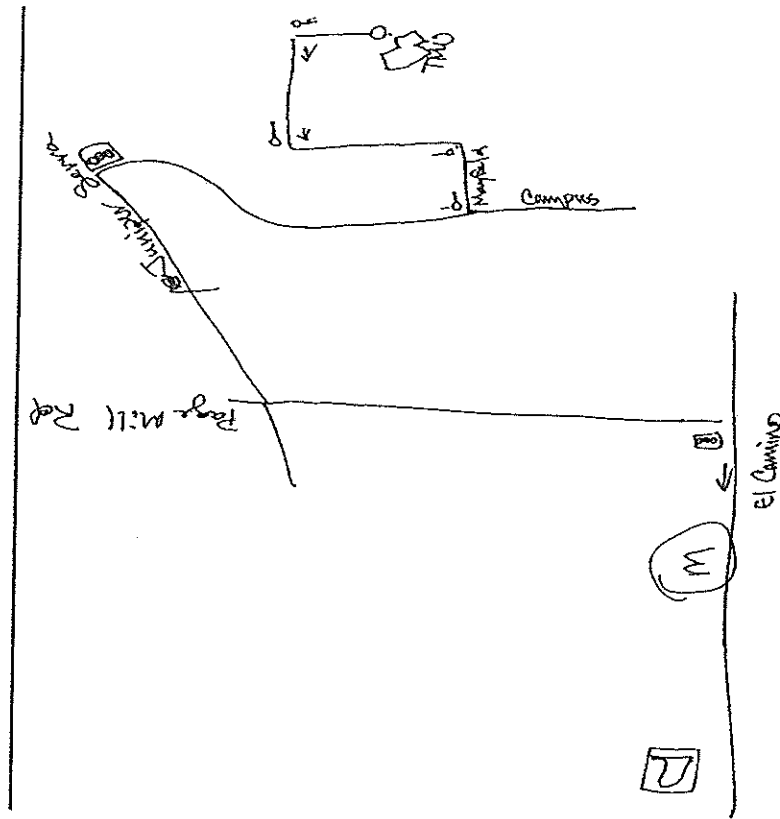


Figure 3: Example of Route Map From Tversky and Lee (1998)

Santa Teresa to Campus Dr, take a right, follow campus around until you get to Galvez, take a left; follow Galvez past stadium and take right on El Camino (stop light, busy st.). Follow El Camino for a while (1-1.5 miles?). Taco Bell will be on your right by Ernie's Liquor.

Figure 4: Example of Route Directions From Tversky and Lee (1998)

paths, and landmarks. Tversky and Lee (1998, 1999) asked more than 20 students outside a dorm if they knew how to get to a nearby fast-food place. If they did, they were asked to either write directions or sketch a map (see Figures 3 and 4 for an example of each). Both route directions and route maps consisted of the segments Denis found, suggesting that the same underlying mental representation generates both. In both descriptions and depictions, the angles of reorientations and distances of paths are represented

approximately as near right angles in maps and as words such as *turn* or *take a* or *make a* in directions. Unlike the space around the body, the space of navigation is generally conceived of in two dimensions rather than three. Another impressive feat of the mind is that it can convert a space experienced vertically as three surrounding dimensions to a space that is two-dimensional as if viewed from above.

Further evidence that the space of navigation is schematized comes from systematic errors in judgments on remembered spaces. Sketch maps of the local environment produced by several dozen students showed that large environmental features, such as roads or states or countries, are not remembered at their correct angles relative to an environmental reference frame or as randomly different from the environmental frame but rather as more north-south or east-west than they actually are (Tversky, 1981; see also Tversky, 2000a, for an overview of systematic errors). In sketch maps, people draw roads running at odd angles as more perpendicular and parallel to the dominant road structure. When asked to place South America in a north-south east-west frame, a significant majority of people upright South America. When viewed as it is on a map, South America appears tilted with respect to the cardinal axes. This error, in which large environmental features are remembered as closer to the axes of the overall reference frame, has been termed *rotation*. It occurs in a variety of judgments on real and artificial environments and on meaningless blobs as well.

Large environmental features are remembered relative to each other as well as relative to an encompassing frame of reference. When asked to choose which world map is correct, the correct map or one in which the Americas are moved northward so that the United States is more aligned with Europe and South America with Africa, a significant majority of people select the incorrect, more aligned map (see Figure 5 for the maps people judged). Similarly, when South America is moved westward to be more directly aligned with North America, more people select that as the correct map than the correct map (Tversky, 1981). This error, in which large environmental features are remembered as more aligned with each other, has been called *alignment*. It also appears in a variety of judgments on real and artificial environments and on meaningless blobs.

These are not the only systematic errors in memory for large spaces. Environments are organized around salient landmarks, such as the Eiffel Tower or Times Square. This leads to violations of metric assumptions in judgments; that is, people estimate the distance from a landmark to an ordinary building to be less than the distance from an ordinary building to a landmark (Sadalla, Burroughs, & Staplin, 1980). Imagined perspective also distorts judgments,

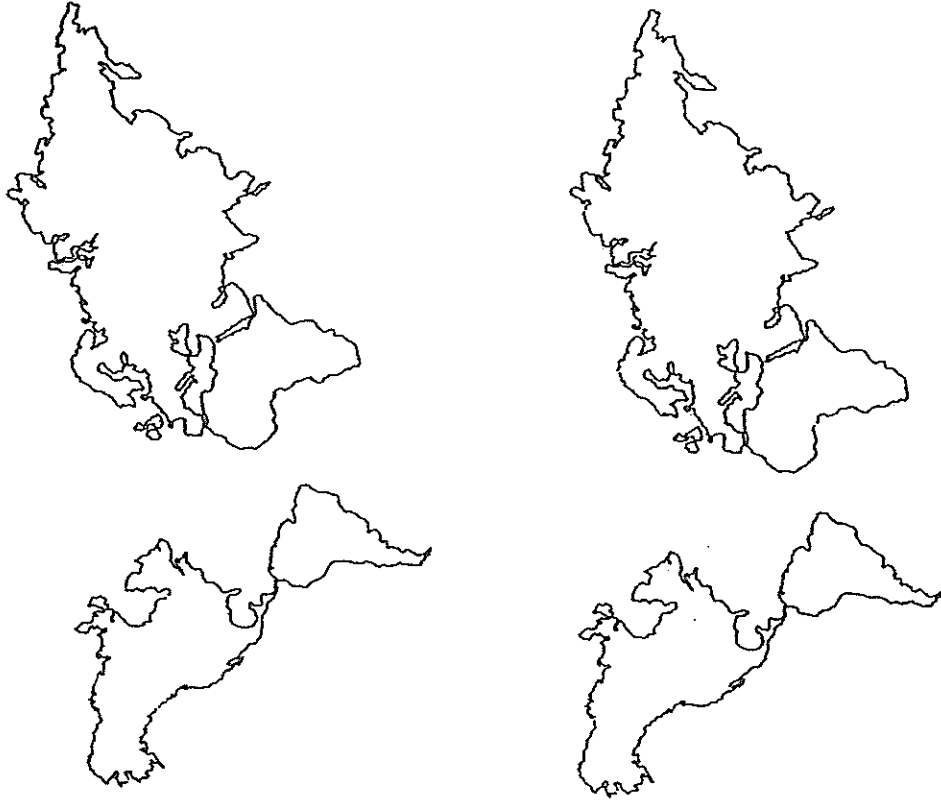


Figure 5: Correct and Aligned World Maps From Tversky (1981)
NOTE: Used with permission.

much along the way as Steinberg cartoons in the *New Yorker* did. Nearby distances loom larger than faraway distances (Holyoak & Mah, 1982).

In all, mental representations of the large environments that are pieced together from different views or different modalities are constructed from elements relative to a reference frame and a perspective. The constructions may use information from remembered experiences in the environments, memories of maps that have been inspected, of verbal descriptions, and more.

There is no guarantee that these mental constructions of large spaces are coherent. Given the multimodal nature of the information people have of large spaces and given the wide variety of systematic errors, there may be no way for people to integrate the information and reconcile the discrepancies. Rather than *cognitive map*, a more apt metaphor for people's mental representations of large spaces is *cognitive collage* (Tversky, 1993).

THE SPACE OF GRAPHICS

Humans also create external spaces, such as maps, architectural drawings, charts, diagrams, and graphs, as tools to augment cognition. Some of these graphics, for example, maps, use space to represent space. Such uses are ancient and pervasive across the world (e.g., L. Brown, 1977). Others, such as graphs that represent visually things that are not inherently visual, are modern inventions. Effective graphics schematize information in ways similar to the ways that mental representations schematize information. Schematization entails excluding information irrelevant to the task and simplifying, even distorting, the information important to the task. Consider maps. An aerial photograph, despite or rather because of being realistic, does not make a good map for most purposes; it is cluttered with unnecessary detail such as foliage and building tops that mask the detail that is necessary, typically the structure of the roads and large natural features. It shows the tops of buildings and trees when it is the frontal views that are typically experienced and recognized. Schematic maps are more useful for most tasks and have been found widely dispersed in time and space, whether incised in stone, carved in wood, or sketched in sand. Architectural plans are also ancient and widespread (Coulton, 1977). Graphics representing spatially things that are not inherently spatial began to appear in Europe in the late 18th century, typically representing economic data, such as balance of payments over time (Beniger & Robyn, 1978).

Similar to the natural spaces of the body, around the body, and of navigation, the invented space of graphics consists of elements and the spatial relations among them. Both can convey meaning more or less literally to more or less metaphorically. Elements may be icons, such as those used in ideographic scripts or in computer applications, or more abstract, as in points that represent cities on maps or data on graphs. Spatial relations, such as distance and direction, may represent literal spaces—on a smaller scale—or metaphorical spaces, such as time or value. That these are natural correspondences is supported by a study examining spontaneous use of space to

represent temporal, quantity, and preference relations by hundreds of people, from 4-year-olds to adults, in three language communities (Tversky, Kugelmass, & Winter, 1981). For temporal relations, for example, people were asked to place stickers or dots on a square piece of paper to indicate the times of breakfast, lunch, and dinner. Most of even the youngest children ordered the stickers or dots on an imaginary line, thus using increases in spatial distance to convey increases in temporal distance. The same held for quantitative and preference relations. That thinking about abstract dimensions spatially is cognitively compelling is evident in language as well. We say that someone is at the top of the heap, that the economy has fallen into a depression, that a new field of inquiry is wide open. These correspondences make graphical displays of abstract information easy to understand. In addition, graphics take advantage of human capacity to reason about space, to estimate distances and direction, to mentally transform spatial arrays, and to infer function from structure (Heiser & Tversky, 2002; Tversky, 1995, 2001).

Graphic displays serve a number of different functions. They preserve memory, acting as historical records. They reduce load on working memory by externalizing memory and cognitive operations, for example, in counting or doing arithmetic. They convey information in a compact way, as in maps or organization charts or diagrams. External representations facilitate inference and problem solving. They also serve as a platform for generating new ideas as in architecture or design (e.g., Suwa, Tversky, Gero, & Purcell, 2001).

MULTIPLE SPACES

Human activity occurs in a multitude of spaces. The space of the body, the space immediately around the body, the space of navigation, and the space of graphics are a few of them. Each of these spaces is conceptualized differently, depending on the functions it serves, the activities invoked, and the entities involved. The space of the body is important for movements and sensations of the body. It is thought of not in terms of size but in terms of the body parts that are perceptually salient and functionally significant. The space surrounding the body, the space that can be readily perceived and acted on, is organized in three dimensions, defined by extensions of the three axes of the body. Accessibility of objects in the space around the body depends on the perceptual and functional asymmetries of the body axes and on their relation to the only asymmetric axis of the world, that formed by gravity. The space of navigation is larger than that which can be seen at a glance. It is the space we navigate in and also the space we consider to make sense of distances in the world,

geological and meteorological phenomena, and political and economic theories. It is pieced together, often from several modalities, from navigation, from maps, from descriptions. Although large and of course three-dimensional, it is thought of as flat and compressed but not necessarily coherent. To make judgments about the space of navigation, people seem to extract the relevant information on the fly and integrate it using relations between elements, such as landmarks and cities, relative to a reference frame. These processes produce systematic errors in the judgments. The final space considered, the space of graphics, is external, created by humans to augment cognition, for example, to offload memory and processing, to communicate, to promote inference and discovery. It maps elements and relations in a literal or conceptual world to elements and relations on paper. Because graphics rely on human ability to make spatial inferences and on widespread spatial metaphors, they are relatively easy to understand and use.

Four spaces with distinctive patterns of cognition and action have been considered. Are there other functionally distinctive spaces? Of course. The brain, for example, makes far more distinctions, the space around the face, that around the hand, and more (e.g., Gross & Graziano, 1995). Have we even exhausted the four spaces reviewed? No, there is far more to do and say about cognition and action in each of these spaces.

Each space reviewed subserves different functions involving different spatial elements and reference frames. Consequently, each has a different mental structure. The elements and spatial relations abstracted for each space are determined by the activities prevalent in that space. Schematization reduces memory load, facilitates information processing, and allows integration of disparate bits of information; however, it also introduces bias and error.

REFERENCES

- Beniger, J. R., & Robyn, D. L. (1978). Quantitative graphics in statistics. *The American Statistician*, 32, 1-11.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-145.
- Brown, C. H. (1976). General principles of human anatomical paratomy and speculations on the growth of partonomic nomenclature. *American Ethnologist*, 3, 400-424.
- Brown, L. (1977). *The story of maps*. New York: Dover.
- Bryant, D. J., & Tversky, B. (1999). Mental representations of spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 137-156.

- Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language*, 31, 74-98.
- Bryant, D. J., Tversky, B., & Lanza, M. (2001). Retrieving spatial relations from observation and memory. In E. van der Zee & U. Niskanen (Eds.), *Conceptual structure and its interfaces with other modules of representation* (pp. 116-139). Oxford, UK: Oxford University Press.
- Coulton, J. J. (1977). *Ancient Greek architects at work: Problems of structure and design*. Ithaca, NY: Cornell University Press.
- Denis, M. (1997). The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive*, 16, 409-458.
- 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 and Cognition*, 20, 507-518.
- Gross, C. G., & Graziano, M. S. A. (1995). Multiple representations of space in the brain. *The Neuroscientist*, 1, 43-50.
- Heiser, J., & Tversky, B. (2002). *Mental models of complex systems: Structure and function*. Manuscript submitted for publication.
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, 18, 65-96.
- Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology*, 14, 328-352.
- Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? *Memory and Cognition*, 4, 291-297.
- Morrison, J. B., & Tversky, B. (1997). Body schemas. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the meetings of the Cognitive Science Society* (pp. 525-529). Hillsdale, NJ: Lawrence Erlbaum.
- Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 516-528.
- Suwa, M., Tversky, B., Gero, J., & Purcell, T. (2001). Seeing into sketches: Regrouping parts encourages new interpretations. In J. S. Gero, B. Tversky, & T. Purcell (Eds.), *Visual and spatial reasoning in design* (pp. 207-219). Sydney: Key Centre of Design Computing and Cognition.
- Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, 35, 371-391.
- Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, 13, 407-433.
- Tversky, B. (1991). Spatial mental models. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 27, pp. 109-145). San Diego: Academic Press.
- 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). New York/Berlin: Springer-Verlag.
- Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), *Understanding images* (pp. 29-53). New York/Berlin: Springer-Verlag.
- Tversky, B. (2000a). Levels and structure of cognitive mapping. In R. Kitchin & S. M. Freundschuh (Eds.), *Cognitive mapping: Past, present and future* (pp. 24-43). London: Routledge.
- Tversky, B. (2000b). Remembering spaces. In E. Tulving & F. I. M. Craik (Eds.), *Handbook of memory* (pp. 363-378). New York: Oxford University Press.
- Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79-111). Cambridge, MA: MIT Press.