1

Functional Significance of Visuospatial Representations

Barbara Tversky

abstract

Mental spaces are not unitary. Rather, people conceive of different spaces differently, depending on the functions they serve. Four such spaces are considered here. The space of the body subserves proprioception and action; it is divided by body parts, with perceptually salient and functionally significant parts more accessible than others. The space around the body subserves immediate perception and action; it is conceived of in three dimensions in terms of relations of objects to the six sides of the body: front/back, head/feet, left/right. The space of navigation subserves that; it is constructed in memory from multimodal pieces, typically as a plane. The reconstruction generates systematic errors. The space of external representations, of pictures, maps, charts, and diagrams, serves as cognitive aids to memory and information processing. To serve those ends, graphics schematize and may distort information.

introduction: four functional spaces

When physicists or surveyors exercise their trades, aspects of space are foreground, and the things in space background. Things are located in space by means of an extrinsic reference system, in terms of metric measurement. Within the reference system, aspects of the space, whether large or small, distal or proximal, for entities small or large, are uniform. Surveyors laying out a road, for example, need to know the exact distance from point A to point B, the exact curvature of the terrain, the exact locations of other objects, natural and built. In other words, they need first to measure aspects of the space as accurately as possible. For human cognition, the void of space is treated as background, and the things in space as foreground. They are located in space with respect to a reference frame or reference objects that vary with the role of the space in thought or behavior.

Which things, which references, which perspective depend on the function of those entities in context, on the task at hand. In human cognition, the spatial relations are typically qualitative, approximate, categorical, or topological rather than metric or analog. They may even be incoherent, that is, people may hold beliefs that cannot be reconciled in canonical three-dimensional space. Human directions to get from A to B, for example, are typically a string of actions at turning points, denoted by landmarks, as in "go down Main to the Post Office, take a right on Oak." The directions are given in terms of entities in the space, paths, and landmarks, and in approximate terms, right, left, straight (Denis, 1997; Tversky & Lee, 1998). In addition, for human cognition there are many spaces, differing in the roles they play in our lives. Those considered here are the space of the body, the space surrounding the body, the space of navigation, and the space of external representations, such as diagrams and graphs. These mental spaces do not seem to be simple internalizations of external spaces like images (e.g., Kosslyn, 1980, 1994b; Shepard, 1994; Shepard & Podgorny, 1978); rather, they are selective reconstructions, designed for certain ends.

What are the different functions that space serves us? The space of the body, the space around the body, the space of exploration, and a uniquely human space, the space of depictions, serve different functions in human activity and hence in human cognition. Things in space impinge on our bodies, and our bodies act and move in space. In order to interpret those impingements, we need knowledge of the receptive surfaces on the body. In order to coordinate those actions, we need knowledge of what the body can do and feedback on what the body has done. The space of the body has a perceptual side, the sensations from outside and inside the body, and a behavioral side, the actions the body performs. Proprioception tells one about the other. Representations of the space of the body allow us to know what the parts of our bodies can do, where they are, what is impinging on them, and, importantly, how to interpret the bodies of others. Actions of others may have consequences for ourselves, so we need to anticipate those by interpreting others' intentions. The space around the body is the space in which it is possible to act or see without changing places, by rotating in place. It includes the surrounding objects that might get acted on or need to be avoided. The space around the body represents the space that can immediately affect us and that we can immediately affect. Both these spaces are experienced volumetrically, although the space of the body is decomposed into its natural parts and the space around the body is decomposed into the six regions projecting from the six surfaces of the body. The space of navigation is the space of potential travel. It is too large to be seen at once, so it is pieced together from a variety of kinds of experiences: perceptual, from actual navigation, or cognitive, from maps or descriptions. In contrast to the space of the body and the space around the body, it is known primarily from memory, not from concurrent perception. It is typically conceived of as primarily flat. Finally, the space of external representations considered here is typically space on paper meant to represent an actual space, as in a map or architectural drawing, or to represent a metaphoric space, as in a diagram or graph. External representations are creations of people to aid cognition. They can be directly perceived, but they themselves are representations of something else. This is a capsule of what is yet to come.

The Space of the Body

Through our bodies, we perceive and act on the world around us and learn about the consequences of our actions. One way that we view and think about bodies is as objects. Common objects can be referred to at several levels of abstraction. What I am wearing on my feet can be called clothing or shoes or running shoes. What I am sitting on can be referred to as furniture or a chair or a desk chair. Despite those possibilities, there is a preferred level of reference, a most common way of talking in everyday speech, the level of shoe or chair, over a broad range of contexts. This level has been termed the *basic* level (Rosch, 1978). The basic level has a special status in many aspects of human cognition. Central to recognition and to categorization of objects at the basic level is contour or shape. Underlying shape for most objects are parts in the proper configuration (cf. Biederman, 1987; Hoffman & Richards, 1984; Tversky & Hemenway, 1984). Although objects have many features, parts constitute the features most diagnostic of the basic level of categorization. Many other cognitive tasks converge on the basic level. For example, it is the highest level for which people can form a general image; people report that forming images of shoes or chairs is not difficult, but forming single images of clothing or furniture is not possible. It is the highest level for which action patterns are similar. The same behaviors are appropriate to different kinds of shoes and different kinds of chairs, but not toward different pieces of clothing or furniture. The basic level is also the highest level for which a general image, one that encompasses the category, can be formed, the highest level for which action patterns are similar, the fastest level to identify, the earliest level acquired by children and introduced to language, and more (Rosch, 1978).

Thus, the basic level has a special status in perception, action, and language. Parts may be critical to the basic level because they form a link from perception or appearance of an object to its function. Parts that are perceptually salient tend to be functionally significant as well; moreover, the shapes of parts give clues to their functions (Tversky & Hemenway, 1984). Think of arms, legs, and backs of chairs, and of course, of people. What is especially intriguing for the parts of the human body is that the size of the brain representations are not proportional to the physical size of the parts themselves. The brain has twin representations of the body, on either side of the sensorimotor cortex, one for the sensory part, one for the motor part. In both cases, certain parts, like lips and hands, have larger than expected amounts of cortex devoted to them, and other parts, like backs, have smaller than expected amounts of cortex devoted to them.

Bodies are a privileged object for humans. Unlike other objects, they are experienced from inside as well as from outside. People determine the actions of their own bodies and those actions provide sensory feedback. Insider knowledge of the body seems to affect how bodies are perceived. Consider an interesting phenomenon in apparent motion. Apparent motion occurs when two similar integrative stimuli occur in rapid succession. Instead of perceiving two static images, people perceive a single image that is moving. Apparent motion is the basis for movies and for the lights on movie marquees. The motion is normally seen at the shortest path. However, when the shortest path for apparent motion violates the ways that bodies can move, a longer motion path is seen for intermediate interstimulus intervals (Heptulla-Chatterjee, Shiffrar, & Freyd, 1996). Thus, when a photo of an arm in front of the body and an arm behind the body are played in rapid succession (but not too rapid), viewers see the elbow jutting out rather than passing through the body. The shortest path is preferred for objects, even when it violates a physical property of the world, that one solid object cannot pass through another solid object, suggesting that knowledge of the body is privileged for perception. In other experiments, people were asked to judge whether two photos of humans in contorted positions of the body were the same or different. Observers were more accurate when they actually moved the limbs, arms or legs, whose positions were changed in the photos, provided the movements were random (Reed & Farah, 1995). Neuroscience literature also indicates privileged areas of the brain for representing the body; when those areas, primarily in parietal cortex, are damaged, there can be disruption of identification or location of body parts (e.g., Berlucchi & Aglioti, 1997; Gross & Graziano, 1995). Moreover, sections of the lateral occipital temporal cortex are selectively responsive to the sight of human bodies (Downing, Jiang, Shuman, & Kanwisher, 2001).

Insider knowledge of the body seems to affect mental representations of the space of the body as well, as revealed in the speed with which different body parts are identified. Despite diversity in languages, certain body parts are named across most of them: head, arm, hand, leg, foot, chest, and back (e.g., Andersen, 1978). These parts differ in many ways, including size, contour distinctiveness, and function. In detecting parts in imagery, size is critical; larger parts are verified faster than smaller ones (Kosslyn, 1980). In object recognition, parts that are distinct from their contours, parts that stick out, are critical (Biederman, 1987; Hoffman & Richards, 1984; Tversky & Hemenway, 1984). Finally, although the functional significance of parts is correlated with contour distinctiveness, the correlation is not perfect. Is one of these factors (size, perceptual salience, or functional significance) more critical to mental conceptions of the body than others? In a series of experiments, participants saw either the name of one of the frequently named body parts or a depiction of a side view of a body with one of the parts highlighted (Morrison & Tversky, 1997; Tversky, Morrison, & Zacks, 2002). They compared this to a depiction of a side view of a body with a part highlighted, responding same or different depending on whether the parts matched. Neither of the comparisons, name–body or body–body, revealed an advantage for large parts; on the contrary, large parts were slower to verify than small ones. For both comparisons, verification times were faster for parts that were high on contour distinctiveness and functional significance. Functional significance was roughly indicated by relative size in sensorimotor cortex. For body–body comparisons, verification times were more highly correlated with contour distinctiveness; these comparisons can be quickly made just on the basis of visual appearance, without processing the body as a body or the parts as actual parts. That is, the two pictures can be treated as meaningless visual stimuli for the comparison entailed. In contrast, for name–body comparisons, verification times were more highly correlated with functional significance. In order to compare a name with a depiction, at least some aspects of meaning must be activated. Names are powerful. In this case, it appears that names activate aspects of meaning of body parts that are closely tied to function.

People move the separate parts of their bodies in specific ways in order to accomplish the chores and enjoy the pleasures of life. They get up and dress, walk to work (or to their cars), pick up mail, open doors, purchase tickets, operate telephones, eat food, hug friends and family. The space of the body functions to achieve these ends. Different body parts are involved in different sorts of goals and functions, the feet and legs in navigating the world, the hands and arms in manipulating the objects that serve us. Mental representations of the space of the body reflect the functions of the body parts.

THE SPACE AROUND THE BODY

The space around the body is the arena for learning about the world and for taking actions and accomplishing goals in it. The proximal space from which the world can be perceived and in which action can readily be taken is a second natural delineation of space by function. One effective way to study the cognition of space, the space around the body and other spaces as well, is through narrative descriptions of space. When descriptions of space are limited and coherent, people are able to construct mental models of them (e.g., Ehrlich & Johnson-Laird, 1982; Franklin & Tversky, 1990; Glenberg, Meyer, & Lindem, 1987; Mani & Johnson-Laird, 1982; Morrow, Greenspan, & Bower, 1989; Rinck, Chapter 9; Rinck, Hahnel, Bower, & Glowalla, 1997; Taylor & Tversky, 1992b; Tversky, 1991). The mental spatial models are mental representations that preserve information about objects and the spatial relations among them and are updated as new information comes in. They allow rapid inferences of spatial elements, locations, distances, and relations from new viewpoints.

Narratives have been used to establish mental models of the space around the body (e.g., Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990; Franklin, Tversky, & Coon, 1992; Tversky, Kim, & Cohen, 1999). Participants studied narratives that addressed them as "you," and placed them in an environment such as a hotel lobby, a museum, or a barn, surrounded by objects at all six sides of their bodies, front, back, head, feet, left, and right. Thus, the narratives described the world from the point of view of the observer (you), in terms of directions from the observer. After learning the environment from narratives, participants were reoriented to face a new object, and probed with direction terms for the objects currently in those directions. Several theories predicting the relative times to retrieve objects at the various directions around the body were evaluated (Franklin & Tversky, 1990). The data did not fit the Equiavailability Theory, according to which all objects should be equally accessible because none is privileged in any way. The data also did not conform to a pattern predicted from an Imagery Theory, according to which observers would imagine themselves in a scene and then imagine themselves examining each direction for the relevant object; an imagery account predicts slower times to retrieve objects in back of the observer than to left and right, counter to the data. The pattern of retrieval times fit the Spatial Framework Theory best. According to this theory, people remember locations of objects around the body by constructing a mental spatial framework consisting of extensions of the axes of the body, head/feet, front/back, and left/right, and attaching the objects to them. Accessibility of directions depends on asymmetries of the body and asymmetries of the world. The only asymmetric axis of the world is the up/down axis created by gravity. Gravity of course has broad effects on the way the world appears and the way we can act in it. For the upright observer, this axis coincides with the asymmetric head/feet axis of the body. Times to retrieve objects at head and feet are in fact, fastest. The front/back axis is also asymmetric but does not coincide with any asymmetric axis of the world. The front/back axis separates the world that can be readily perceived and acted on from the world behind the back, difficult for both perception and action. Finally, the left/right axis lacks any salient asymmetries, and is, in fact, slowest.

The spatial situation can be varied in many ways: by altering the orientation of the observer (Franklin & Tversky, 1990), by adding more observers (Franklin et al., 1992), by putting the array in front of the observer instead of surrounding the observer (Bryant et al., 1992), by having the environment rotate around the observer instead of having the observer turn to reorient in the environment (Tversky et al., 1999). These variants in the situation lead to consequent variants in the retrieval times that can be accounted for by extensions of the Spatial Framework Theory. When the observer is described as reclining, and turning from side to front to back to side, no body axis correlates with gravity. Retrieval times in this case depend only on body asymmetries. The front/back axis of the body seems to be the most salient as it separates the world that can be readily perceived and manipulated from the world behind the back. Along this axis, front has a special status, as it is the direction of orientation, of better perception, of potential movement. In fact, for the reclining case, times to retrieve objects in front and back are faster than times to retrieve objects at head and feet, and times to front faster than those to back (Bryant et al., 1992; Franklin & Tversky, 1990). What about narratives describing two characters, for example, in different scenes? In that case, the viewpoints of each character in each scene are taken in turn; in other words, participants construct and use separate mental models for each situation, yielding the spatial framework pattern of data. However, when two characters are integrated into a single scene, participants seem to construct a single mental model that incorporates both characters, and take a single, oblique point of view on them and the objects surrounding them (Franklin et al., 1992). In this case, they do not take the point of view of either of the characters so their bodies are not aligned with any of them. Thus no area of space is privileged for the participant, and in fact, reaction times are the same for all directions for both characters. How about when narratives describe the environment as rotating rather than the observer as turning? In the case of the rotating environment, participants take twice as much time to reorient as when narratives describe the observer as reorienting. In the world we inhabit, people move, not environments, so although people can perform mental feats that the world does not, it takes longer to imagine impossible than possible, normal, mundane interactions with the world (Tversky et al., 1999).

Not only can the spatial situation be varied, the mode of acquisition can be varied; the space around the body can be acquired from narrative, from diagrams, from models, and from experience (Bryant & Tversky, 1999; Bryant, Tversky, & Lanca, 2001; Franklin & Tversky, 1990). As long as retrieval is from memory rather than perception, the Spatial Framework pattern of retrieval times obtains (Bryant et al., 2001). When responding is from perception, then patterns closer to the Imagery model obtain. This is because it in fact takes longer to look behind than to look left or right. Surprisingly, as participants learn the environments, they cease looking, so that although the information is available from perception, they respond from memory. As a consequence, the retrieval times come to correspond to the Spatial Framework model. Although diagrams and models are both external spatial representations of the scenes, they instill slightly different mental models (Bryant & Tversky, 1999). The models were foot-high dolls with depictions of objects hung in the appropriate directions around the doll. When learning from models, participants adopt the embedded point of view of the doll, and, just as from the original narratives, they imagine themselves reorienting in the scene. The diagrams depicted stick figures with circles at the appropriate directions from the body; the circles contained the names of the objects. When learning from diagrams, participants adopt an outside point of view and imagine the scene rotating in front of them, as in classic studies of mental rotation (e.g., Shepard & Cooper, 1982). We speculated that the three-dimensional models encouraged participants to take the internal viewpoint of the doll whereas the flat and flattened space of the diagram encouraged participants to treat the diagram as an object, in other words, to mentally manipulate the external representation instead of using it to induce an internal perspective. These perspectives, however, are flexible; when directed to do so, participants used the diagram to take an internal viewpoint or used the model to adopt an external one. The two perspectives and the mental transformations of them, viewing an object from outside versus viewing a surrounding environment from inside, appear in other analogous tasks, and are subserved by different neural substrates (e.g., Zacks, Rypma, Gabrieli, Tversky, & Glover, 1999). They reflect the two dominant perspectives people take on space, an external view, prototypically the view people have on objects that they observe and manipulate, and an internal view, prototypically the view people have on environments that they explore. One remarkable feature of human cognition is that it allows both viewpoints on both kinds of external realities.

The space around the body, that is, the space immediately surrounding us, the space that functions for direct perception and potential action, is conceptualized in three dimensions constructed out of the axes of the body or the world. Objects are localized within that framework and their relative locations are updated as the spatial situation changes. The mental spatial framework created out of the body axes underlies perspective-taking, allows updating across rotation and translation, and may act to establish allocentric or perspective-free representations of the world from egocentric experience.

The Space of Navigation

The space of navigation serves to guide us as we walk, drive, fly about in the world. Constituents of the space of navigation include places, which may be buildings or parks or piazzas or rivers or mountains, as well as countries or planets or stars, on yet larger scales. Places are interrelated in terms of paths or directions in a reference frame. The space of navigation is too large to perceive from one place so it must be integrated from different pieces of information that are not immediately comparable. Like the space around the body, it can be acquired from descriptions and from diagrams, notably maps, as well as from direct experience. One remarkable feature of the human mind is the ability to conceive of spaces that are too large to be perceived from one place as integral wholes. In order to conceive of spaces of navigation as wholes, we need to paste, link, join, superimpose, or otherwise integrate separate pieces of information. In addition to being separate, that information may be in different formats or different scales or different perspectives; it may contain different objects, landmarks, paths, or other details. Linking disparate pieces of information can be accomplished through spatial inferences anchored in common reference objects, reference frames, and perspectives. The linkage is necessarily approximate, leading to consistent errors, as shall be seen in the section on cognitive maps (see also Montello, Chapter 7, for a more detailed discussion of navigation and Taylor, Chapter 8, for a more detailed discussion of cognitive maps as well as externally presented maps).

Places

Many navigable environments can be loosely schematized as landmarks and links, places and paths. Places that is, configurations of objects such as walls and furniture, buildings, streets, and trees, selectively activate regions of the parahippocampus, part of the network of brain structures activated in imagining travel. Not only is this area selectively active under viewing of scenes, but also patients with damage to this area experience severe difficulties acquiring spatial knowledge of new places (e.g., Aguirre & D'Esposito, 1999; Cave & Squire, 1991; De Renzi, 1982; Epstein & Kanwisher, 1998; Rosenbaum et al., 2000). The brain has areas selectively sensitive to only a small number of kinds of things, places, faces, objects, and bodies, suggesting both that these entities have special significance to human existence and that they are at least somewhat computationally distinct.

Perspective of Acquisition

Descriptions of the space of navigation locate places with respect to one another and a reference frame, from a perspective. They typically use one of two perspectives or a mixture of both (Taylor & Tversky, 1992a, 1996). In a *route* perspective, the narrative takes a changing point of view within an environment, addressing the reader or listener as "you," describing you navigating through an environment, locating landmarks relative to your changing position in terms of your left, right, front, and back. For example, "As you drive down Main Street, you will pass the bank on your right and the post office on your left. Turn right on Cedar, and the restaurant will be on your left." In a *survey* perspective, the narrative takes a stationary viewpoint above the environment, locating landmarks relative to each other in terms of an extrinsic frame of reference, typically, north-south-east-west. For example, "The bank is east of the post office and the restaurant is north of the post office." The components of a perspective, then, are a landmark to be located, a referent, a frame of reference, a viewpoint, and terms of reference. In both speech and writing, perspectives are