

Three Spaces of Spatial Cognition

Barbara Tversky Julie Bauer Morrison Nancy Franklin David J. Bryant
Stanford University SUNY-Stony Brook Nyma Corp.

All our lives, we interact with the space around us, whether we are finding our way to a remote cabana in an exotic tropical isle or reaching for a ripe mango on the tree beside the cabana or finding a comfortable position in the hammock to snack after the journey. Each of these natural situations is experienced differently, and as a consequence, each is conceptualized differently. Our knowledge of space, unlike geometry or physical measurements of space, is constructed out of the things in space, not space itself. Mental spaces are schematized, eliminating detail and simplifying features around a framework consisting of elements and the relations among them. Our research suggests that which elements and spatial relations are included and how they are schematized varies with the space in ways that reflect our experience in the space. The space of navigation is too large to be seen from a single place (short of flying over it, but that is a different experience). To find our way in a large environment requires putting together information from different views or different sources. For the most part, the space of navigation is conceptualized as a two-dimensional plane, like a map. Maps, too, are schematized, yet they differ in significant ways from mental representations of space. The space around the body stands in contrast to the space of navigation. It can be seen from a single place, given rotation in place. It is the space of immediate action, our own or the things around us. It is also conceptualized schematically, but in three dimensions. Finally, there is the space of our own bodies. This space is the space of our own actions and our own sensations, experienced from the inside as well as the outside. It is schematized in terms of our limbs. Knowledge of these three spaces, that is, knowledge of the relative locations of the places in navigation space that are critical to our lives, knowledge of the space we are currently interacting with, and knowledge of the space of our bodies, is essential to finding our way in the world, to fulfilling our needs, and to avoiding danger, in short, necessary to survival.

The Space of Navigation.

We encounter the space of navigation in many different ways. The prototypical way is by exploring the environment. As we explore the environment, the salient features of the environment change. But if we are to learn the environment, we must find a way to link the different features in space. In addition to integrating different views from different locations, we often need to integrate different forms of information. Using a description to get from one place to another, say the cave to the hunting grounds, or your house to mine, is not just a modern problem, but is probably something our hunter-gatherer ancestors had to accomplish. Then, what needs to be integrated is the description with the environment as viewed. Nowadays, we commonly use maps to navigate, integrating the streets and perhaps landmarks in the map with the visible world. Sometimes, we use all these sources of information, our memories of the environment, descriptions, and maps. Of course, there are important differences among these three media, but there is also important common ground. Clearly, the easiest way to integrate these different representations is to use a common schematization for all. From the earliest work on environmental psychology, environments have been schematized to nodes and links, landmarks and the paths among them, elements and their spatial relations, often from a particular perspective (e. g., Denis, 199; Kuipers, 1978; Levelt, 198; Lynch, 19; Tversky, 1993; Tversky and Lee, in press). Linking different views or representations is possible when there is a common reference frame, such as the canonical directions or large environmental features, and when elements are schematized similarly.

That people's mental representations are schematized and that landmarks, paths, and reference frames are important aspects of these representations has been supported by myriad studies. Many of these have also shown that people's conceptions of space are not only schematized, but also distorted in systematic ways. The distortions arise from the fundamental components of the representations themselves, from the elements, landmarks and paths, the reference frames, and the perspectives.

Effects of Other Elements on Distance and Direction Judgements.

We do not remember the space of navigation in an absolute way. One way we remember the locations and directions of elements, is in relation to the locations and directions of other elements. Moreover, these seem to be schematized in such a way that the grouped elements are drawn closer together. Hirtle and Jonides (198) asked one group of University of Michigan students to sort buildings around campus into groups. Most students put university buildings in one group and town buildings in another. A second group of students was asked to judge the distances between pairs of the same buildings, without explicitly grouping them. Nevertheless, the second group's distance judgements were distorted by the implicit grouping into town and university structures. Distances between buildings in the same group, town or university, were underestimated relative to the distances between buildings in different groups. In addition to being grouped, elements in the environment are often related to other elements. Environmental elements are not cognitively equal; some are larger, or older, or better known, or more central to our activities. These "good" landmarks are often selected as reference elements for less salient elements. When asked where we are from, for example, we answer differently to different people, depending on what we believe they know of our environment. But in each case, we pick a nearby landmark that we think our questioner will know, and say near that (Shannon, 197). These local salient landmarks also distort our conceptions of our surroundings. People judge distances from an ordinary structure to a good landmark to be shorter than the distance from a landmark to an ordinary structure (Sadalla, 198; V and McNamara, in press). Other elements affect direction judgements as well as distance judgements. When asked to choose the correct map, more people select a map in which South America has been moved westward so that it is closer to due south of North America than the veridical map of the Americas. People also prefer a map in which Europe and Africa have been moved southwards to be aligned with the USA and South America to the true map. In addition, people's judgements of directions between pairs of cities show the same distortion (Tversky, 1981).

Effects of Frame of Reference on Distance and Direction Judgements.

A question whose answer is so striking that it has made it to Trivial Pursuit was posed by Stevens and Coupe (1978). They asked students in San Diego to indicate the direction between San Diego and Reno. Most thought, incorrectly, that Reno was east of San Diego. From this and similar examples, Stevens and Coupe argued that people's conceptions of the space of navigation are hierarchical, even though space is not. Instead of remembering the locations of all cities, they argued, people remember the relative locations of states, the larger units containing cities. They also remember which cities are in which states, and use the locations of the states to infer the locations of the cities. Evidence for hierarchical encoding of large spaces also comes from the time it takes to make direction judgements. People are faster to say that one city is east or north of another if the two cities are in different geographical entities, states or countries, than if they are in the same geographical entity (Maki, 198; Wilton, 19). Another reference frame is the canonical directions surrounding a geographical entity. That also distorts direction judgements. People see a natural orientation in shapes, including geographical entities, so, for examples, South America, Israel, and Italy seem somewhat tilted relative to north-south east-west. They are apparently remembered schematically so that their natural axes are closer to the canonical axes. When asked to orient a map of South America, people

place it more upright than it actually is (Tversky, 1981), and when asked to give directions between pairs of cities in Israel or Italy, people's judgements are distorted as if they had uprighted those countries (Cornoldi, 1997; what's his name, 199).

Effects of Perspective.

Yet another factor in organizing the space of navigation is the viewpoint adopted on a space. Holyoak and Mah (1982) asked students in Ann Arbor to imagine themselves either on the east coast or the west coast. They were then asked to judge the distances between pairs of cities more or less equidistant on an east-west trajectory across the USA. Those imagining themselves on the east coast judged the distance between New York and Pittsburgh to be larger than those imagining themselves on the west coast. Conversely, those on the west coast thought the distance between San Francisco and Salt Lake City was larger than those on the east coast. Each group judged the nearby distances to be relatively larger than the faraway distances, much like the popular posters of a New Yorker's (or San Franciscan's) view of the country.

The organizing effects of elements, reference frames, and perspectives have been revealed in many other studies as well. Together, they suggest we schematize the space of navigation in terms of elements, reference frames, and perspectives. These organizing structures not only schematize but also distort our conceptualizations of these spaces.

The Space Around the Body

As we move about the world, we seem to keep track of the objects around us effortlessly, so that even without looking we are aware of what is behind, to the sides, above, below, and in front of us. In contrast to the space of navigation, the space around the body can be viewed from a single place. Also in contrast to the space of navigation, the space around the body is conceptualized in three dimensions. How do we keep track of the space surrounding us?

Franklin and Tversky (1990) investigated people's mental models of their surroundings using descriptions of environments. In the paradigmatic situation, participants read descriptions of themselves as observers in settings, such as an opera house, museum, or barn, surrounded by objects at all six sides of the body. After learning the environments, participants were informed that they were now facing another object. They were then presented with direction terms, front, back, left, right, head (or above), feet (or below), and asked to respond with the object currently in that position. Three models were proposed to account for the times to retrieve objects in the probed directions. According to the Equiavailability Model, no one direction is privileged, so retrieval times to all directions should be identical (Levine, 198). Research on imagery (e. g., Kosslyn, 1981; Shepard and Podgorny, 1978) suggested the Mental Transformation Model, according to which participants imagine themselves in the position of the observer in the narratives, facing forward. To retrieve another direction, participants imagine themselves turning to face the probed direction to ascertain what object is located there. If so, retrieval times should increase with the angular disparity between forward and the probed direction, so that times to front should be fastest, times to objects displaced 90 degrees to left or right or head or feet should be next fastest, and objects displaced 180 degrees to the back should be slowest.

Neither the Equiavailability nor the Mental Transformation models accounted for the data. Instead, the data could be accounted for by the Spatial Framework Model developed by Franklin and Tversky (1990). They posited that in order to keep track of the objects around the body, people construct a mental spatial framework from extensions of the three axes of the body, and associate objects to it. The body has three essential axes, that formed by head and feet, that formed by front and back, and that formed by left and right.

The accessibility of objects lying in different directions depends on characteristics of the body and the world. The head/feet and front/back axes of the body are asymmetric, whereas the right/left axis lacks any salient asymmetries. Of the three axes of the world, only the up/down axis defined by gravity has a salient asymmetry. The other two axes are defined arbitrarily relative to a position. When the observer is upright, the characteristics of the body and of the world work together to make the head/feet axis the most salient, followed by the front/back, and last, the left/right. In fact, the retrieval times correspond to this pattern. When the observer reclines and turns from front to side to back, no axis of the body is correlated with gravity. In this case, the head/feet axis loses its primacy to the front/back axis, which separates the world that can be readily seen and manipulated from the world that cannot be readily seen or manipulated. In fact, for the reclining observer, retrieval times to front/back are faster than those to left/right, corresponding to this analysis (Franklin and Tversky, 1990).

In subsequent studies, the situation was varied, yielding variations of the original pattern of data. For example, the same pattern arises when participants judge spatial arrangements around a central object instead of a central person, and a similar pattern arises when participants judge spatial relations of objects arrayed in front of them, instead of surrounding them (Bryant, Tversky, and Franklin, 1992). The same pattern arises when participants judge spatial relations around two characters, as long as the characters are in separate scenes (Franklin, Tversky, and Coon, 1992). When participants learn scenes from observation rather than narrative, the spatial framework pattern emerges when they are tested from memory of the scenes (Bryant, and Tversky, in press). This is also the case when participants learn the scenes from models of them (Bryant, Tversky, and Lanca, in press).

In keeping track of the immediately surrounding space, then, people seem to construct mental spatial frameworks from the three major axes of the body, and to associate objects to them. Ascertaining the locations of surrounding objects is accomplished through this framework. However, the axes are not equally accessible. Relative accessibility is determined by enduring characteristics of the body and of the world. In the upright case, the front/back axis is most accessible, due to the confluence of body asymmetry and world gravity. Left/right is least accessible. When a person reclines and turns from side to side, no body axis is correlated with the axis of gravity, so accessibility depends solely on characteristics of the body. In this case, the perceptual and behavioral asymmetries of the front/back axis render it the most accessible. Thus, the way in which surrounding space is schematized biases the speed of information retrieval.

The Space of the Body

As we move about performing our daily activities, we keep track of the various parts of our bodies, of where they are, what they are doing, how they are feeling. Of course, the human body is a familiar one in our visual world as well, but unlike other objects in the world, we know bodies from the inside as well as from the outside. Morrison and Tversky (1997) have begun exploring mental maps of the body. For recognizing objects, contours or shapes seem to be critical (e. g., Rosch, Mervis,). Shapes, in turn, can be divided into their parts. Contour discontinuities (inflection points) seem to be used to partition objects into parts (e. g., Hoffman and Richards, 1984; Biederman, 1987). The first step in investigating mental maps of the body was to see which parts of the body were more accessible, more rapidly recognized.

Morrison and Tversky began with the body parts that span the body and that are most frequently named with single morphemes across languages (Andersen, 1978; Brown, 19). These include: head, arm, hand, chest, back, leg, and foot. The first task was a part

identification task. Participants saw the name of a body part on the screen and then a body with a part highlighted by a uniform white dot. Their job was to respond whether the highlighted part was the same or different from the named part as rapidly as possible. The bodies were realistic renderings in a variety of natural positions (and unnatural orientations).

Three theories make different predictions about the speed of identifying the body parts. A theory derived from research on imagery (e. g., Kosslyn, 1981) would predict that larger parts would be named faster than smaller ones, as they can be detected faster when scanning the object. This would mean that leg and back, for example, would be identified faster than hand and arm. From the work on contour discontinuities, it could be predicted that parts with greater contour discontinuities would be identified faster. Thus, for example, foot should be identified faster than chest. A third theory, the part significance theory, derives from the observation that bodies are experienced from the inside as well as the outside, so that information about sensations and actions constitute part of our knowledge of bodies, in addition to information about appearance. Some support for this position comes from research of Reed and Farah (1999). Their participants were asked to decide whether pairs of bodies depicted in complex positions but at different angles were in fact the same position. When participants made judgements while waving their arms in random movements, they were faster to judge differences in depicted arm positions; conversely, when participants moved their legs randomly, they were faster to judge differences in depicted leg positions. For objects, part size, salience and functional significance are highly correlated (Tversky and Hemenway, 1984). For bodies, these can be somewhat dissociated. The index we used for body part significance was relative size on sensory-motor cortex, reasoning that parts that have relatively greater sensory feedback and motor control would be relatively more significant in human interactions with the world. According to part significance, head and chest should be identified faster than leg.

In the study described, measuring times to verify named and highlighted body parts, part identification times in fact correlated best with part significance. Verification times were fastest to head, next to chest, hand, foot, and arm, and last to leg and back. Part discontinuity was second best, and in fact, the rankings for part significance and part discontinuity are themselves correlated. Nevertheless, part significance can account for the relatively rapid verification times for chest but part discontinuity cannot. There was no support for part size. In fact, some of the largest parts, leg and back, were slowest. This is somewhat surprising as the task was primarily a visual one that did not arouse any body sensations, such as those that might be aroused had participants been asked to verify touched or moved body parts. Apparently the names of the body parts call to mind the meanings of the parts, that is, their functions, actions, and sensations, as well as the visual features of parts, notably their contours. This position is supported by a second study in which participants compared two depictions of bodies with the same or different parts highlighted. In this task, which can be performed by visual comparison without naming, part discontinuity was the best predictor of reaction times, better than part significance. As before, part size correlated negatively with verification speed.

While this work is only preliminary, it does suggest that mental maps of the space of our bodies depend on both function and appearance, on sensation and action as well as on visual salience. In the case of the space of the body, the elements are the body parts, the spatial relations are their configuration, and the bias is toward parts with greater salience defined by contour discontinuity for a purely visual situation and toward parts with greater significance when function or action is aroused.

Three Functional Spaces

Knowing where we are, what surrounds us, and what our bodies are doing and feeling is essential to our very survival. These three arenas of knowledge correspond to three different arenas of interaction with the spatial world and to three different schematizations of the world. Are there other spaces of import to the human mind? It is possible. For example, neuroscientists might subdivide the space surrounding the body into the space that limbs can reach and the space that can be seen but not reached without moving, a division into the space that can be manipulated and the space that can be seen but not manipulated. This distinction may be important for the limbs but there is no evidence that it has cognitive significance, especially given the fact that space that can be viewed but is out of reach can be reached with simple locomotion. Similarly, the space of navigation can be divided into the space of experienced navigation and the space that is known only through maps (Montello, 1993). Here, too, it is not clear that these are cognitively distinguishable in the sense of requiring different sorts of mental representations. For one thing, the space of experienced navigation is often learned vicariously, from maps and descriptions, as well as from actual experience. For many (but not all) kinds of information and judgements, environments learned by maps, description, and navigation are indistinguishable. For example, the errors derived from reliance on other elements and frames of reference occur for environments acquired by actual navigation as well as those learned from maps (Tversky, 1981) suggesting that the underlying mental representations are similar in character.

Because the three spaces, the space of navigation, the space around the body, and the space of the body are experienced differently, interact with world differently, and serve different functions in lives, it makes sense that they are schematized differently, forming different kinds of mental representations. In each case, the schematization reflects the typical kinds of interactions that human beings have with their surroundings. For large scale navigation, we need to keep track of the important landmarks and routes in the environment and to patch together different views, different modalities, and different experiences into a larger whole that will allow seamless interaction among parts that have been cognitively separate. Using elements and reference frames gives us a way to patch those separate segments together. For the space around the body, we need to keep track of the objects surrounding us, objects that may be obstacles to our movements or that may be useful for our activities. Because their danger or utility depends on their location relative to our bodies, we keep track of them relative to our bodies. For this situation, our bodies are schematized by our body axes. Finally, for the space of the body, we need to keep track of our own limbs (and those of others in close proximity). Body parts that are more significant both perceptually and kinematically, are those that are more accessible.

The mental representations for each of these spaces differ in critical ways from classical views of imagery. The mental space of navigation does not seem to be Euclidean. Rather, this space seems to be represented qualitatively, in terms of elements and the coarse spatial relations among them. The space around the body is not searched in analog fashion, from viewpoint to probed direction. Instead, it is searched categorically, through the three body axes. Finally, for the space of the body, part size does not determine part verification times. Rather, when part verification requires comparing two depictions, times are determined by part salience, which is based on contour discontinuity. However, when part verification requires comparing a named part to a highlighted part, times are determined by part significance, which is based on functional significance as well as salience, as indexed by relative projection size in sensory-motor cortex. Although these spaces are conceptualized differently, all three are used, apparently seamlessly, as we interact with the world.

