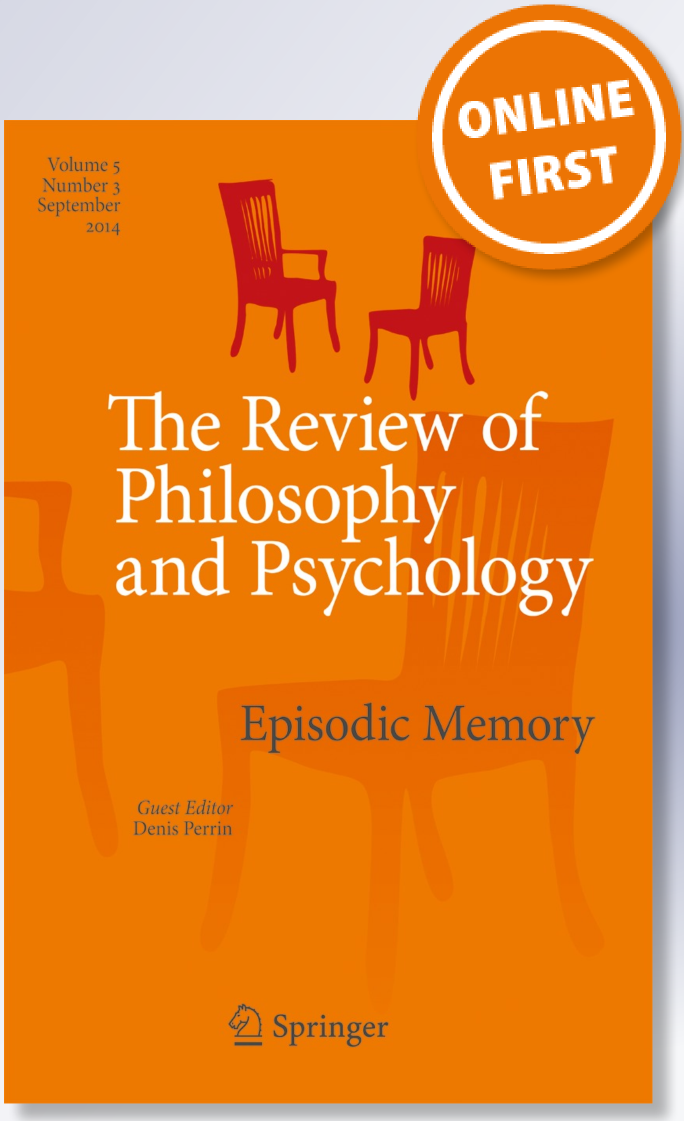
The Cognitive Design of Tools of Thought

Barbara Tversky

Review of Philosophy and Psychology

ISSN 1878-5158

Rev.Phil.Psych.
DOI 10.1007/s13164-014-0214-3



Springer

Rev.Phil.Psych.
DOI 10.1007/s13164-014-0214-3

The Cognitive Design of Tools of Thought

Barbara Tversky

Abstract When thought overwhelms the mind, the mind puts it into the world, notably in diagrams and gestures.Both use space and arrays of elements, depictive and non-depictive, to convey ideas, concrete and abstract,clear and sketchy. The arrays and the non-depictive elements like boxes and arrows serve to showrelationships and organizations, thematic, categorical, and more. on paper, in the air, in the diagrammedworld. Human actions organize space to convey abstractions: spraction.

1 Putting Ideas into the World

Thought can quickly outgrow the mind. To augment the mind, people use anything at hand, their hands, their bodies, arrangements of sticks and stones and coffee cups, sketches in the sand, scribbles on paper napkins, and more. Some ancient outpourings of thought have survived the ravages of time, evocative paintings of animals on the walls of caves (Altamira in Spain, Chauvet in France, Apollo 11 in Namibia), maps incised in stone (Bedolina map in Val Camonica, Italy), histories carved in columns (Trajan) and arches (Titus), explanations and journeys painted on walls of tombs (Valley of the Kings, Egypt), myths glazed on vases (Greece), tallies inscribed in bones (Ishango bone, Republic of Congo).

These external representations of thought are designed to serve thought, for self or for others, for many ends. Here we explore how they are designed, how they use marks and arrangements of marks in space to reflect and affect thought. They are variously referred to as diagrams, sketches, graphics, visualizations, models, material, cognitive tools, cognitive artifacts, and more, and their roles in the enrichment of collective and individual minds cannot be underestimated (e. g., Donald 1991; Hutchins 1995; Kirsh 1995; Malafouris 2013; Norman 1993; Stjernfelt 2011; Suchman 1987; Tversky 2001). With notable exceptions (such as abaci), they arrange marks or things on the space of a surface we will call a page. Importantly, the arrangements typically differ from the arrangements in the natural world, that is, they are designed. External representations are selective; they schematize, including some information, omitting other information, and often simplifying and exaggerating the information included. Many, like maps,

B. Tversky (✉)
Columbia Teachers College and Stanford University, 525 West 120th St, New York, NY 10027, USA
e-mail: btversky@stanford.edu

Published online: 26 October 2014

have been refined through use in communities, a kind of informal user-testing. As such, they provide insight into the structure of the thought that generated them as well as into the correspondences between thought and the visual-spatial expressions of thought.

The approach taken here is empirical, using two research paradigms, one from linguistics and one from psychology. Evidence in linguistics often relies on the analysis of illustrative examples. We analyze examples of cognitive artifacts produced over time and space that schematize in similar ways and that design marks in space in similar ways, revealing those similarities (and ignoring the differences). Evidence in psychology often relies on experimental manipulations. We have developed a program that might be called empirical semiotics, experiments examining regularities in the ways people create, interpret, and use various external representations of meaning. In short, we analyze external expressions of thought, in the wild and in the laboratory, extracting from both the visual and spatial features that characterize a wide range of diagrams and underlie their communicative foundation. We draw connections between material representations of thought and ethereal representations of thought created by gestures. Both are formed by the actions of hands in space, so parallels abound. We integrate action in space with the abstractions they create on a page, in the air, and in the world, developing the concept of *spraction*. We begin historically, with external representations of things in the wild.

2 Diagramming Things and Events: Analyzing Examples in the Wild and Domesticated

Maps In examining diagrams of things, we will collect their characteristics, implicitly or explicitly contrasting the ways representations of things differ from the world they represent and differ from another prevalent way of representing things, words and sentences. We begin with maps, representations of places in spaces, pervasive throughout the world. The Babylonian clay map depicted below can be seen at the British Museum. The British Museum dates it 700–500 BC; it is regarded as the oldest preserved map in the world (Brown 1979). This map, however, is not strictly geographic; rather it mixes space and time, geography with history, including myth, a not uncommon mixture in ancient maps, for example, in Mexico (Boone 2010) Fig. 1.

Maps map; they represent landmarks and the spatial relations among them in the real world using marks and spatial relations on a page. Like modern maps, this ancient map uses blobs or point-like marks to represent places or landmarks and lines to represent paths between places more or less geographically arrayed. The same information could be, and often is, represented by words, but maps have rich correspondences (or second-order isomorphisms, Shepard 1984) with the worlds they represent and words for the most part do not. Because of these spatial correspondences, literal and metaphoric, visual-spatial representations are more direct than representations in words and sentences.

That said, words and sentences and their array on a page can and do correspond to meanings on occasion. The letters of one word are spatially separated from those of others; new ideas are indicated by paragraph or chapter breaks. Events are often, though by no means always, expressed in temporal order. Interestingly, in Greek mathematics, the sentences of logical syllogisms were ordered similarly and line up as if in a table, an arrangement that may have facilitated the development of logic and logical reasoning (Netz 1999). Note that these are spatial devices of the sort that characterize diagrams.

Fig. 1 Babylonian clay map. British Museum. Courtesy of Wikipedia commons

Some sound-meaning correspondences have also been found for specific words (e.g., Jurafsky 2013).

Maps (and diagrams) have richer correspondences with the worlds they represent. Yet, as this ancient map and contemporary maps illustrate, maps do not merely shrink the world, as much is omitted and much of what is there is distorted, coastlines are smoothed and roads are often enlarged or they wouldn't be seen. Like the map above, modern maps are typically annotated. Maps serve multiple purposes. They provide a larger view of the world than a viewer can see from any particular place (save a spaceship). Maps can show how the space is used, and how geography interacts with human activity. They facilitate planning routes, and serve city planning and design. What ends they serve depend on their design, on what information is preserved or omitted, and how it is arranged.

Rock Art Graffiti is (are) not a recent urban phenomenon. The oldest surviving graffiti is rock art of various forms, carvings in stone, paintings in caves, images in tombs, found all over the world, and dating back at least as early as the Late Stone Age, 40,000 years ago (e.g., Pike et al. (2012). These images intrigue, fascinate, mystify, and puzzle their viewers, and lead to endless speculation about their meanings and perhaps spiritual significance. Easier to discern is their content and how they are arrayed, and both are revealing. Although appearing at unrelated sites all over the world, many early cave paintings and petroglyphs are remarkably similar in form and content. The image on the following page is from the Chauvet Cave in Southern France, thought to date 30–32,000 years ago (Sadier et al. 2012). Interestingly, the unique depiction of overlapping heads in the middle has been interpreted as an early attempt at animation (Azema, and Rivere 2012), foreshadowing Duchamp's well-known *Nude Descending a Staircase* as well as many comics.

Fig. 2 Image from the Chauvet Cave. Courtesy of Wikipedia Commons

What is curious about these evocative visual relics (Figs. 2, 3 and 4) is what they are not. Most do not seem to depict individuals. Rather than conveying specifics, they seem to be conveying generalities. Some depict rudimentary scenes, notably stampedes or a hunter and his prey, but more frequently, they depict conglomerations of single isolated images, taken out of their natural settings, animals, stick figures, geometric figures or symbols. The animals and stick figures are readily identified, again because of correspondences, here of parts, configurations of parts, and shapes between the depictions and the actual animals. Whether crude, as in petroglyphs or more refined, as in cave paintings, the forms themselves are idealized, prototypic forms, rather than forms with specific, individualized identities. The views are stereotyped canonical views, front or side views of stick or blobby humans, sometimes ornamented, and side views of animals, sometimes in parades of the same type. The sizes of the figures are approximately the same, not proportional to their actual sizes. They are generally upright and oriented similarly. The images are often grouped by type, animals organized categorically, humans together elsewhere, usually not arranged in familiar scenes or themes. Except for the hunter-prey pairs and stampedes, the forms are often loosely lined up as if on a bulletin board, in sloppy rows and sometimes columns, perhaps to be contemplated, juxtaposed, and compared. It is important to remember that it is likely that rock art had many authors over time, so unlike the other external representations discussed, these works were not likely to be designed as wholes.

Fig. 3 Prehistoric petroglyph from Val Camonica, Italy. http://www.reidsitaly.com/destinations/lombardy/lake_iseo/valcamonica.html

Diagrams Despite the multiple authorship of rock art, its spatial and visual characteristics–arrays of like groups of similarly sized and oriented canonical images spatially separated loosely into groups, aligned loosely into rows and columns–are characteristics associated with designed diagrams of things, in contrast to depictions of action scenes. In their essay *The Culture of Diagram*, Bender and Marrinan (2010) argue that diagrams have an intermediate status between depictions, which, according to Goodman (1976) are "dense," and descriptions, which are "articulated." Here I borrow from but extend Bender and Marrinan's analyses. They offer the illustrations in Diderot and D'Alembert's magnum opus, *L'Encyclopedie* as examples. Figure 5 shows a typical example, one of the pages showing the Pinmaker's Factory. The layout and juxtaposition make the contrast between depictions of scenes and diagrams explicit, by showing both on the same page.

Fig. 4 Ancient petroglyph from Newspaper Rock State Historical Monument Utah. From www.davejenkins.com Creative Commons GFDL

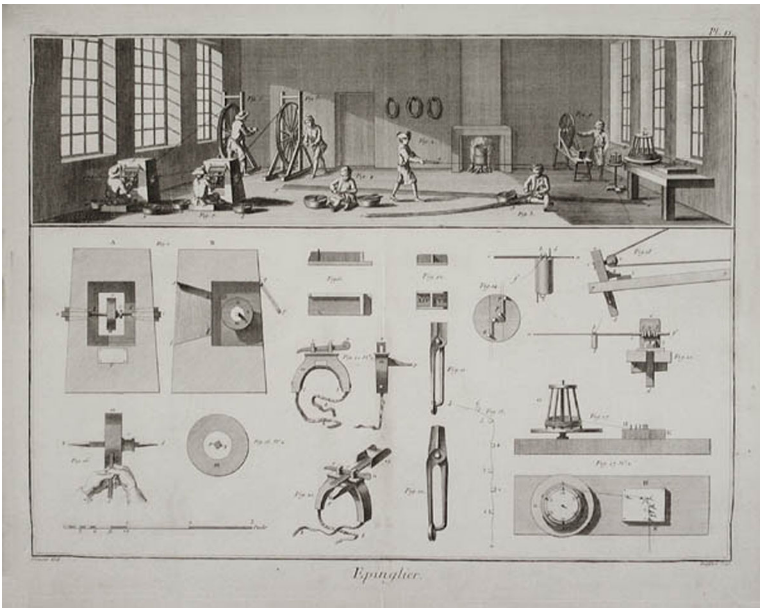


Fig. 5 Pinmaker's Factory from Diderot and D'Alembert, *L'Encyclopedie*. Courtesy of Wikipedia commons

The implication is that each individual image at the bottom of the page is in itself a depiction, but the array is not, the array is articulated like a description. Bender and Marrinan note several diagrammatic features of the bottom half of the page: the sizes of the images, not proportional; the shadows, meant to accentuate the features rather than natural light; the annotation, that is, labels and numbers; the white space separating the images, meant to point to certain *correlations* among the objects. *Correlation* is Bender and Marrinan's translation of Diderot and D'Alembert's *rapport*; perhaps a better term for this community would be *correspondence*. The intermixing of dense depictive images in an articulated, description-like display signals, according to Bender and Marrinan, the presence of a diagram. Some diagrams, we will argue, but not all. Along with many others, Bender and Marrinan regard diagrams not as in and of themselves, but as "objects to be worked with," presumably to be used for understanding and inference. Going beyond their analysis and foreshadowing later discussion, the characteristics of the array, which can be varied, enable and promote different understandings and inferences.

Carrying their analysis further, the scene at the top of the page is depicted "realistically," that is, people at work using implements in their natural setting, a room designed for that activity with walls, windows, floor, a door, even pictures and a fireplace to warm the room. The workers are shown engaged with their tools, and both workers and objects are arrayed in a space that supports their activities. The light source is natural, and indicated by appropriate shading and shadows. The white space is actual empty space. The scene at the top is active; the array at the bottom is frozen, static. The array of the scene at the top is constrained and determined by gravity, by the architecture of the room, by the characteristics and activities of the workers and their implements. By contrast, the array of tools and parts of tools at the bottom of the page is constrained and determined by kinds, objects that share appearance and function. The image sizes of the objects are similar, and the objects are angled to show the critical features of the parts and tools. Shadows and shading are used to accentuate the structures of the tools. The tools and tool-parts are grouped by category and arrayed in columns and rows, separated by approximately equal spaces, as in the Gestalt principle of grouping by proximity. Both parts of the page, the action scene at the top and the diagram of the tools at the bottom are contained in frames or boxes. The frames are descriptive, not depictive; they explicitly separate one form of representation from the other. Thus, depictions of scenes and diagrams have strikingly different spatial as well as visual characteristics, and presumably of communicative intent. At the top of the page, the objects are contextualized; at the bottom, they are decontextualized, extracted from their context of use, and recontextualized into kinds in the array at the bottom of the page, forming what today might be termed a catalog. At the same time, the kinds of diagram exemplified in *L'Encyclopedie* and in the anatomical and mechanical diagrams common at the time are not purely descriptive as they maintain depictive properties. Here, we suggest that this kind of diagram, a diagram of things, is derived from depictions, by distillation and rearrangement. Intriguingly, the distillations and rearrangements mirror both the ways people think about the world and the ways they design and arrange things in the world.

These ways of depicting and grouping, scenes of action and classes of things, map onto two common useful ways that people simplify and comprehend the world, by grouping multitudes of things together: partonomic and taxonomic (e. g., Miller and Johnson-Laird 1976; Tversky 1990). As their names suggest, partonomies are hierarchical organizations of parts; examples include the parts of the body, the parts of objects, the parts of a scene, as well as parts of more abstract entities, the parts of government. Taxonomies are hierarchical organizations of kinds; examples include kinds of animals, kinds of plants, kinds of tools, kinds of musical instruments, that is, sets of things that share appearance and function. People's partonomies of scenes include the people, things, and activities that typically appear in them (Tversky and Hemenway 1983). Within scenes, people, things, and activities often come in organized triplets or even quadruplets: in the Pinmaker's Factory above, one worker is spinning thread on a wheel, in another scene, a baker is stirring batter with a spoon; in yet another, a tailor is cutting fabric with scissors. In many other cases, pairs and groups of people work cooperatively to attain a goal. Those organizations of knowledge, grouping objects that people commonly use interactively, such as paper and scissor and hammer and nail and knife and bread, have been called thematic. Adding an actor to a thematic grouping of objects creates a rudimentary scene. Interestingly, one of the few repeated thematic organizations that appear with some frequency in ancient petroglyphs is hunter-prey, with the hunter often holding a hunting tool.

One of the formative developments in the long period between rock art and *L'Encyclopedie* is the page, which both frames and bounds one set of images, and separates that set of images from other sets of images. Each two-page spread, of *L'Encyclopedie* was designed as a whole, to integrate information relevant to a single topic. The integration is accomplished, at least in part, visually and spatially. The two-page spread is subdivided into pages, and each page is further subdivided, more implicitly by white space, more explicitly by lines and frames. Going upwards, the pages are grouped into chapters and volumes, signified by white space. Thus the visual and spatial design of *L'Encyclopedie* is hierarchical, and reflects and supports the hierarchical organization of the information. These visual and spatial features also serve written texts; writing is organized by visual and spatial features like spaces between words and sentences, indentation for paragraphs, and new pages for new chapters. Written language has diagrammatic features. We will return to these diagrammatic visual and spatial devices, lines, frames, and spacing.

Bender and Marrinan point to another way that depictions have been used in diagrammatic arrays, to portray events in time, different from kinds and parts, and demanding different arrays of different kinds of depictions. Bundling sequences of interrelated actions into meaningful *events* is yet another common way to organize experience (e. g., Casati and Varzi 1996; Zacks and Tversky 2001). Bender and Marrinan are referring in particular to the sequences of photographs by Muybridge and Marey, such as that in Fig. 6.

These time-lapse photos show successive stills of human and animal bodies in motion. They are arrayed horizontally in reading order and they set the stage for motion pictures. However, successive stills that constitute a larger event also have ancient roots, below an example from the tomb of the 15th c BC astronomer Nakht in Thebes, Egypt. In this case, as evident from Fig. 7, the events are horizontally arrayed, but, unusually, ordered from bottom to top.

Visual stories, histories, and explanations in tombs (above), columns, arches, friezes, and scrolls, are also ancient, if later than rock art. In contrast to petroglyphs and cave paintings, these images are integrated series of related actions in time, in linear arrays. Each image is typically embedded in a comprehensive and coherent spatial scene, so their scale and scope are far greater than the second by second portrayals of movements of bodies of Muybridge and Marey, though both freeze moments in time. Like the diagrams of the *Encyclopedie,* they are highly selective. They segment, separate, and

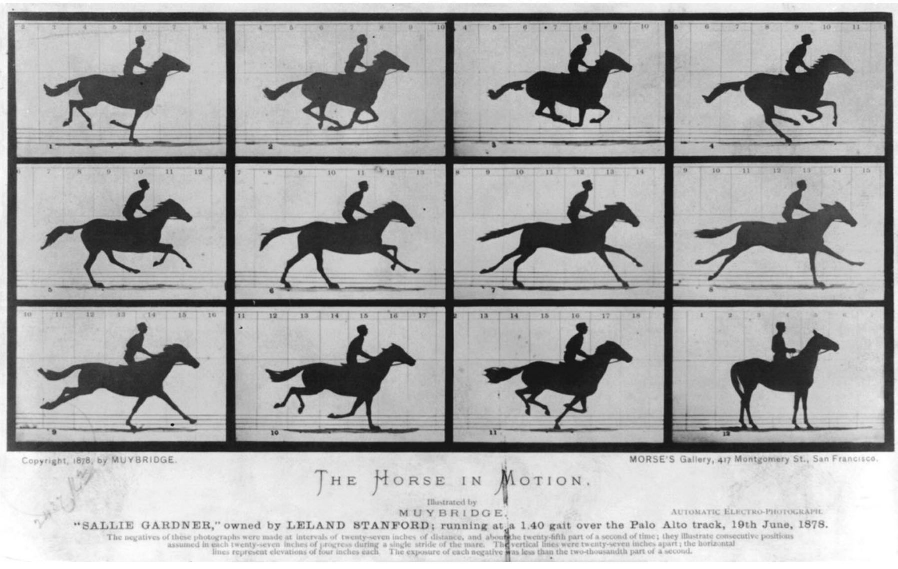


Fig. 6 Time-lapse photographs by Muybridge in part to settle a bet on whether all four legs of a galloping horse were in the air at the same time. Image courtesy of Wikipedia Commons

Fig. 7 Wall painting showing planting and harvesting under an overseer from the tomb of the 15th c BC astronomer Nakht, Thebes, Egypt. Image courtesy of Wikipedia Commons

frame the continuous flow of human activity by natural joints, typically, key events in history or key steps in explanations (e. g., Zacks and Tversky 2001; Tversky and Zacks 2013). As in many cave paintings and petroglyphs, the images are typically arrayed in horizontal rows or vertical columns. The arrays are frequently, but not always, in temporal order, and the images are typically, but not always, the same size. Some classical art deviates from that; for example, in some Greek, Etruscan, or Roman vases or frescoes, the array of figures is determined by importance rather than temporal sequence, with the most important image in the middle, and larger, and the less important smaller, to either side (Small 1999). Sometimes, the segmentation is implicit, accomplished with gestalt, perceptual features such as grouping by proximity. Other times, the grouping is accomplished by the spatial array itself, such as groups of people facing each other or focusing on the same object or action. In other cases, even ancient ones, segmentation of a larger event into frozen critical moments is explicit, accomplished by adding non-depictive glyphs, lines or frames or boxes that both enclose one event and separate it from the others. Diagrams that show a sequence of events over time have many historical and contemporary variants. Stained glass windows are one example. The visual instructions that come with a camera or knock-down furniture, the visual explanations common in science textbooks, the visual stories in comics are other examples. As noted, such diagrams typically contain not only depictive elements but also words and symbols as well as graphic elements that are non-depictive, such as frames. Graphic novels and comics bring the communication of events in time to an art form (McCloud 1993).

These kinds of diagrams could be called bottom-up diagrams. They are distillations and rearrangements of depictions of aspects of the world, scenes and events. They seem to begin with that complex world, select and refine depictions of parts of it, and then rearrange those to articulate conceptual correspondences, using visual devices like points, blobs, lines, and frames and spatial Gestalt devices like grouping, orders, rows and columns. To expand the repertory, we turn now to diagrams of ideas.

3 Diagramming Things and Ideas: Laboratory Experiments

The diagrams analyzed so far, maps, rock art of animals, tomb paintings of agriculture, and engravings of manufacturing refine, rearrange, and reorganize things, scenes, and events in the world on a page to show certain correspondences and relations among those things and the spaces that encompass them. We turn now to experiments in which people produce, understand, and use diagrams with varying meanings to see the ways those meanings are represented and organized. In particular, experiments on production and comprehension converge to show that simple geometric forms, dots, lines, boxes, and arrows, carry context-dependent meanings related to their geometric forms.

Dots and Lines: Route Maps and Directions Explaining how to get from here to there is a common and ancient task. One way is to give a set of directions in language; another is to draw a sketch map. Might there be parallels between them? To see, we approached students in a dormitory around dinner time and asked them if they knew how to get to a local fast food restaurant (Tversky and Lee 1998, 1999). If they did, we asked them either to sketch a map or to write down the directions. Adapting a system developed by Denis (1997), we segmented and coded both maps and directions. Both shared a small number of semantic units. For landmarks, point-like marks on maps or names in directions. For paths, variations of lines. Both points and lines have foundations in the concrete world. Places from a distance resemble points and paths from a distance resemble lines. For turns, lines at right angles for maps and "take a," "make a," or "turn" for directions. For straight roads, straight lines in maps and "go down" in directions. For winding roads, a curved line in maps and "follow around" in directions. Thus, although maps could be analog, they weren't. They represented turns and paths categorically. Similarly neither represented distance; the presumption seemed to be that you keep going until the next landmark or intersection. Both had iterations of sequences of four units: start point, reorientation, progression along a path, and end point. However, the route directions could and did elide some of the units; for example, the previous end point is the start point for the next segment, but doesn't have to be mentioned twice. Path progression could and was omitted in the route directions. Maps cannot elide that kind of information; they need to preserve continuity. Thus, the semantic elements for route maps and route directions were the same, but the pragmatics of using that information differed.

In a subsequent experiment, participants were given the set of semantic elements either in pictures or in words and asked to use them to produce a large set of directions. They were told that the elements might not be sufficient, and they were free to add elements if they needed them. Very few people added elements, and the elements they added were things like freeway entrances. Neither medium represented exact distance or direction. Yet these maps that omit and distort are successful for navigation (e. g., Denis, Pazzaglia, Cornoldi, and Bertolo 1998): why? Maps and diagrams are meant to be used in context, by specific users for specific tasks. Maps are designed differently for hikers than for drivers and science diagrams are designed differently for novices than for experts. Route maps are meant to be used in navigation where the environment will disambiguate and clarify the missing or distorted information. If the map indicates a right angle turn, but the road goes off at a different angle, the traveler goes with the road, understanding that the right angle shouldn't be taken literally. If the distance to the next turn isn't proportional, it won't matter as the traveler will turn at the indicated landmark.

In fact, route maps distort metric information in many of the same ways that memory does (e.g., Tversky 1981). Maps for subway systems do the same. In discussing their revolutionary design of the "map" of the NYC subway system, Massimo Vignelli first refers to it as a map, and then corrects himself, "it is really a diagram," (Brew and Guerra 2013). The beauty of the Vignelli "map" is that it makes the subway routes and interconnections legible. It does so by simplifying and distorting the "true" spatial distances and directions and in so doing, it rearranges the natural space and the things in it.

Lines and Boxes: Graphs Lines and dots are the among the principle components of another frequent non-depictive diagram, that is, graphs. In graphs, dots or points represent variables. Just as lines in maps represent the paths between places, lines in graphs represent the relations between points. Graphs also use frames or boxes, notably bar graphs. In Talmy's (1983, 2000) parlance, dots represent concepts that are regarded as zero-dimension, lines represent concepts that are regarded as one-dimensional, and bars represent concepts that are regarded as two-dimensional.

Many kinds of data can be represented as bars or lines. To many (though not to statisticians) that choice seems arbitrary, a matter of taste. However, the format for representing data, bars or lines, affects their meaning, that is, how the data are interpreted. Because bars both contain and separate, they invite discrete comparisons. Because lines connect, they invite trend interpretations. Participants in fact interpreted line graphs as trends and bar graphs of the same data as discrete comparisons, even when contradicted by the dimensions underlying the data. Similarly, when people were given descriptions of trends or discrete comparisons, they produced line and bar graphs respectively (Zacks and Tversky 1999).

Arrows Another demonstration of the power of simple geometric forms to create meaning comes from a set of experiments on arrows (Heiser and Tversky 2006). If lines suggest paths or relations between points, then arrows, asymmetric lines, should suggest asymmetric relations. Arrows, too, are rooted in the concrete world; wooden arrows fly in the direction that they point, and water running downwards in sand creates arrows in the direction of movement. In an experiment on comprehension, students were asked to interpret a mechanical diagram (car brake, pulley system, or bicycle pump) with or without arrows. Those who saw diagrams without arrows interpreted them as conveying structure, the arrangement of parts. Those who saw diagrams with arrows interpreted them as conveying behavior in time, causal action. In an experiment on production, people who were given structural descriptions of one of the systems and asked to produce a diagram drew the system and labeled the parts, but did not use arrows. In contrast, people's diagrams for the causal descriptions, did not label parts but did put arrows to show the causal sequence of actions from start to finish.

Graphic Language? Dots, lines, boxes, and arrows form a small set of graphic but non-depictive elements or glyphs that serve as skeletal elements, a partial semantics, for constructing a wide range of diagrams for a broad range of domains. The senses of the glyphs change though a core sense remains. For example, a line in a map and a line in a graph represent relationships in both cases, in the former a physical relation and in the latter, a conceptual relation. Continuing the analogy to language, each domain has its own syntax and pragmatics. The rules for conjoining lines and dots is different for route maps, for line graphs, for circuit diagrams. The constraints on arranging boxes are different for the periodic table than for a demographic table. The pragmatics differ as well. Analysis of instances as well as experimentation on production and comprehension illuminate the semantics, syntax, and pragmatics of various kinds of diagrams, much as analysis of instances and experimentation illuminate understanding of the semantics, syntax, and pragmatics of language (Tversky 2004).

Space: Distance, Place and Direction Place in space also conveys meanings that find agreement in production and comprehension. Distance in space is used by people to represent distance in space in maps and to represent distance in time, preference, quantity, and more in graphs. Children spontaneously use order and distance in space to represent order and distance in time, preference, and quantity (e. g., Tversky et al. 1991). In addition to distance and order, children and adults agree on mappings of abstract concepts to direction in space. They map increased quantity and increased preference upwards. To go up means overcoming gravity, which entails strength and power and health and money. Upwards in space is typically associated with everything good, as supported by production and comprehension in children and adults (e. g., Schubert and Maass 2011; Tversky et al. 1991). These correspondences appear in language as well. We say we've grown close to someone, or far apart. We say someone's on the top of the heap or has fallen into a depression. Reading order also confers directionality to space, in particular to the horizontal, with increases in preference, quality, quantity, power, strength, and more conforming to reading order (e.g., Schubert and Maass 2011; Tversky et al. 1991). Place, proximity, and direction in space are powerful influences in graphics and in life; taller people, whether men or women, have higher-paying jobs (e. g., Egoff and Corder 1991)

3.1 Messy Diagrams: Sketches

Up till now, the discussion has focused on what might be called *orderly* diagrams. For the most part these diagrams, maps and visual explanations and graphs, are designed to clearly communicate a situation so that they are helpful to users in a range of tasks. Maps for understanding a region, for route planning, and city planning. Visual explanations for enacting the procedures or understanding the processes. Graphs for understanding the underlying phenomena and for predictions. In these cases, diagrams are especially useful when they emphasize the information likely to be productive for inference and omit the information that is irrelevant and distracting.

Yet, clarity can be fixating, freezing thought instead of freeing it (e. g., Smith 1995) and thought is not always clear enough to diagram neatly. In science, for example, the underlying phenomena generating the data are not always known. A case in point is the cholera outbreak in London in 1854, before the germ theory of disease had been developed. Dr. John Snow famously mapped each case of cholera as a dot on a city map. Studying the raw data revealed a cluster of deaths near a specific water pump. The pump was removed and the epidemic subsided (Johnson 2006). Mapping cases is still used to track and trace outbreaks of disease in order to understand their origins and to halt epidemics. Premature summaries might obscure the underlying phenomena.

Messy diagrams, then, can be crucial for thinking through problems, arousing and considering multiple possibilities. A common case is design, where designers typically begin with sketches that are vague and tentative. It’s often not clear what the lines mean, what they refer to, even to their creators. Designers report having conversations with their sketches, getting feedback from them (Schon 1983a, 1983b). They sketch for one reason, but on inspecting their sketches, see new patterns and relations and make unintended discoveries. Expert designers are better than novices at making new discoveries in their sketches; they also make more conceptual discoveries than novices, for example, realizing the effects of the layout on traffic, who primarily make perceptual inferences, for example, seeing patterns in layouts (Suwa and Tversky 1996, 2001, 2003; Suwa et al. 2001; Tversky and Suwa 2009). Artists report the same; some saying their drawing “talks to me” (Kantrowitz 2014). Designers adept at making new discoveries report that they get new ideas when they reconfigure the elements and relations in their sketches. Because ambiguity allows reconfiguration and reinterpretation, ambiguous sketches promote discoveries and inferences. Expert designers excel both at reconfiguring ambiguous sketches and at remote associations (Suwa and Tversky 2003; Tversky and Suwa 2009). This combination of perceptual and conceptual skills facilitates finding new interpretations. Expert designers use their perception actively in the service of discovery and innovation, a process we called *Constructive Perception*.

3.2 Diagrams Without Paper: Gesture

Diagrams require a surface that can be arranged, with marks or objects. Typically the surface is a page, and the marks graphic. Externalizing thought on a page has many advantages: the page is public; it is permanent so it can be reinspected, rearranged, and revised; it can be referred to and animated with gesture and speech. It stays there to be contemplated. In the absence of such tools, and before they were cheap and ubiquitous, people used the tools that are always with them, their hands, to create diagrams in the air. When architects are blindfolded as they design, they gesture copiously (Bilda and Gero 2006). Gestures also have advantages. They don't need anything extra. They are actions, and especially effective in conveying action. Like diagrams on paper, gestures use arrangements of virtual marks in space to represent the elements and relations of thought. Thus, like diagrams, gestures can and do represent information that is literally spatial and information that is metaphorically spatial. In many cases, the gesture forms are analogous to the marks on a page and the use of gestural space analogous to the use of diagrammatic space. For example, to represent environments, gestures point to places and make gestural lines for paths, analogous to dots or blobs for places and lines for paths in maps (Heiser et al. 2004; Jamalian et al. 2013). A string of integrated gestures, dots, lines, and boxes for the most part, that are organized in space can represent and convey a rich environment (Emmorey et al. 2000; Jamalian et al. 2013) or a family tree (Enfield 2003) or a temporal processes (Jamalian and Tversky 2012) or a sequence of actions (Kang et al. 2012).

Like diagrams and sketches, gestures serve the thought of self and alter the thought of others. When solving mental rotation problems, people often rotate their hands; rotating their hands and in the correct direction helps mental rotation (e. g., Chu and Kita 2008; Wexler et al. 1998; Wohlschläger and Wohlschläger 1998). When people read descriptions of rich environments in preparation for a later memory test, many gesture, and when they do so, they perform better. Their gestures represent the environments, point-like gestures for places and line-like gestures for paths between places. Participants rarely look at their hands, and the looks are brief glances. Thus, the representations that support memory and inference appear to be spatial/motor, rather than visual. A comprehensive theory of how externalizing thought aids thought must take into account facilitation by spatial/motor representations as well as by spatial/visual representations.

Gestures are actions, actions that affect objects in the mind rather than actions that affect objects in the world. As such they seem especially suited to conveying action. Action, process, behavior, and causality are especially difficult to extract from diagrams on the page, in contrast to extracting structure. Structure, the spatial or conceptual relations among the parts, is in the diagram. Action implies change, so it must be inferred from the diagram. Novices and those low in spatial ability find it hard to make inferences from diagrams (e. g., Tversky et al. 2013). A recent experiment has shown that gesture in fact is effective in conveying action. Participants were asked to learn the workings of an engine from one of two videos. Both videos had the same verbal script, which was sufficient to answer the questions, and saw the same diagram of the structure of the parts of the engine. In one video, the explainer's gestures showed the shapes of the parts of the engine in the proper configuration. In another video, the explainer's gestures enacted the actions of the parts of the engine, in the same place in gesture space. Following four viewings of the videos, participants answered questions about the structure and action of the engine, created a visual explanation of the workings of the engine, and finally made a video explaining the workings of the engine. Viewing action gestures affected all three measures. Participants answered more action questions correctly even those the questions could be answered from the verbal script. They showed more action in their visual explanations. In their video explanations, they used more action words and more action gestures (Kang et al. 2012). Thus, watching gestures that conveyed action instilled deep knowledge of the behavior of the engine, over and beyond what was conveyed by the words.

Like diagrams, gestures reflect and affect thought. They represent thought using actions that create virtual marks that stand for elements in virtual spaces that stand for the relations among elements. Gesture and diagram map meanings quite directly, in ways that words and sentences cannot (e.g., Tversky et al. 2009; Tversky and Kessell 2014). Gestures and diagrams are deeply intertwined; both are created by actions of the hands. Diagrams can be viewed as crystallized gestures.

4 Diagramming the World

The analyses of diagrams and gestures have led to a collection of features that characterize them and differentiate them, on the one hand from the things they represent and on the other hand from words and sentences. Like things are grouped by proximity and spatially separated from unlike things. Things are lined up meaningfully in rows and columns, and the rows and columns are sometimes ordered. Together, the rows and columns create tables or matrices, organized by the features of the things enclosed in the tables. The groups and orders are sometimes emphasized with lines and frames. In diagrams, these features create patterns that are regular, good Gestalts, and differ from the natural patterns in the world; they catch the eye. They invite explanations for the organization. Why are things grouped and ordered in those ways? This kind of reasoning is part of the "work" of diagrams, some of the ways that diagrams promote inference and discovery.

The designed world uses those same features to the same ends, gathering like things, separating them from unlike things, arranging things in meaningful orders, and more. Dishes, flatware, pots and pans are grouped into drawers and shelves of like things, that is, categories. They are further grouped into hierarchies of categories, plates on one shelf, bowls on another, both grouped by size. Books are grouped and ordered on shelves, perhaps by topic or size or time. These visual-spatial organizations express taxonomic categories of kinds and hierarchies of categories. The spaces designed by people reflect partonomic thematic organizations as well. The kitchen contains the things used for cooking and eating. The bathroom encloses the things used for personal care. Table settings express one-to-one correspondences. Assembly lines order in time the object-action couplings that form events, putting on clothes or preparing a meal or manufacturing automobiles (see Kirsh 1995). These visual and spatial arrangements that people impose on the world articulate conceptual correspondences, abstractions. Like diagrams, they create patterns that invite the mind to discover their meanings.

The spatial structures that cultures impose on their environments, inside and out, express more: they both affect and reflect the social, political, and economic organizations of those societies. The spatial arrangements of the remnants of ancient cultures are used by archeologists and historians as clues to the social, political, and economic organization of those societies. Is there division of activities within households, across households? Is size of space equal across family units or are some families allocated more space than others? What structures are proximal, which ones distant? Spatial size, proximity, grouping, framing with furnishings and walls all carry social-political-economic meaning. People use those visual and spatial arrangements to make inferences from the spatial to the social-political-economic in their daily lives. Spatial relations suggest social relations: people who sit or walk close to each other are more likely to be seen as a social group, and their degree of spatial closeness to reflect their social closeness.

The patterns in space that articulate abstractions are created by actions in space. Those actions are incorporated into gestures used to express acts of thought, putting, taking, lifting, sorting, pushing, separating, piling, and more. Those patterns are used to create the diagrams on paper that have more deliberate intent to communicate abstractions. Actions in space express abstractions and create abstractions, interlinked processes that have been dubbed *spraction* (e. g., Tversky 2011a, 2011b). Forming categories of kinds, parts, and events in the mind helps to simplify and organize James' proverbial "blooming buzzing confusion" of sensory input. Similarly, designing the world to reflect kinds, parts, events, and other abstractions simplifies a complex world by integrating and organizing both perception and action. The simple patterns created by design actions form good gestalts that attract the eye and inform and stimulate the mind. Those patterns are used to create the diagrams on paper that have more deliberate intent to communicate. The patterns in space are designed by action and support action, performing the basic tasks of life, navigation inside and outside, eating, sleeping, work, and play. Action. Space. Spraction. Inseparable and integrated.

Acknowledgments I am indebted to my colleagues and collaborators, especially Valeria Giardino, Jocelyn Penny Small, Azadeh Jamalian, Angela Kessell, Julie Heiser, Paul Lee, and Jeff Zacks. I am grateful the Varieties of Understanding Project at Fordham University and The John Templeton Foundation and to the following National Science Foundation grants for facilitating the research and/or preparing the manuscript: National Science Foundation HHC 0905417, IIS-0725223, IIS-0855995, and REC 0440103. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the funders.

References

Azema, M., and F. Rivere. 2012. Animation in Paleolithic art: A pre-echo of cinema. *Antiquity* 86: 316–324.

Bender, J., and M. Marrinan. 2010. *The culture of diagram*. Stanford: Stanford University Press.

Bilda, Z and Gero, J.S. (2006) Reasoning with internal and external representations: A case study with expert architects. In R Sun and N. Miyake (Editors), *Proceedings of the Cognitive Science Society,* Lawrence Erlbaum, Pp. 1020–1026.

Boone, E. H. (2010). *Stories in Red and Black: Pictorial Histories of the Aztec and Mixtec*. University of Texas Press

Brew, K. and Guerra, R. (2013). *Design is One: Lella and Massimo Vignelli*. Film. US distributor www.firstrunfeatures.com

Brown, L. 1979. *The story of maps*. NY: Dover.

Casati, R., and A. Varzi. 1996. Introduction. In *Events*, ed. R. Casati and A. Varzi. Aldershot: Dartmouth.

Chu, M., and S. Kita. 2008. Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General* 137: 706–723.
Denis, M. 1997. The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive* 16: 409–458.
Denis, M. Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1998). Spatial discourse and navigation: An analysis of route directions in the city of Venice. *Applied Cognitive Psychology, 12*
Donald, M. 1991. *Origins of the modern mind*. Cambridge: Harvard University Press.
Egoff, D.B., and L.E. Corder. 1991. Height differences of low and high job status, female and male corporate employees. *Sex Roles* 24: 365–573.
Emmorey, K., B. Tversky, and H. Taylor. 2000. Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation* 2: 157–180.
Enfield, N.J. 2003. Producing and editing diagrams using co-speech gesture: Spatializing non-spatial relations in explanations of kinship in Laos. *Journal of Linguistic Anthropology* 13(1): 7–50.
Goodman, N.A. 1976. *Languages of art: An approach to a theory of symbols*. Indianopolis: Hackett.
Heiser, J., and B. Tversky. 2006. Arrows in comprehending and producing mechanical diagrams. *Cognitive Science* 30: 581–592.
Heiser, J., B. Tversky, and M. Silverman. 2004. Sketches for and from collaboration. In *Visual and spatial reasoning in design III*, ed. J.S. Gero, B. Tversky, and T. Knight, 69–78. Sydney: Key Centre for Design Research.
Hutchins, E. 1995. *Cognition in the wild*. Cambridge: MIT Press.
Jamalian, A., and B. Tversky. 2012. Gestures alter thinking about time. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, ed. N. Miyake, D. Peebles, and R.P. Cooper, 551–557. Austin: Cognitive Science Society.
Jamalian, A., V. Giardino, and B. Tversky. 2013. Gestures for thinking. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, ed. M. Knauff, M. Pauen, N. Sabaenz, and I. Wachsmuth. Austin: Cognitive Science Society.
Johnson, S. 2006. *The ghost map*. London: Riverhead Books.
Jurafsky, D. (2013). Why ice cream sounds fat and crackers sound skinny. *Stanford Magazine*, July/August
Kang, S., B. Tversky, and J.B. Black. 2012. From hands to minds: How gestures promote action understanding. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, ed. N. Miyake, D. Peebles, and R.P. Cooper, 551–557. Austin: Cognitive Science Society.
Kang, S., Tversky, B. & Black, J. B. (2014) Gesture and speech in explanations to experts and novices.
Kantrowitz, A. (2014). A cognitive ethnographic study of improvisational drawing by eight contemporary artists. Doctoral dissertation, Columbia Teachers College
Kirsh, D. 1995. The intelligent use of space. *Artificial Intelligence* 73: 31–68.
Malafouris, L. 2013. *How things shape the mind: A theory of material engagement*. Cambridge: MIT Press.
McCloud, S. 1993. *Understanding comics*. Northhampton: Kitchen Sink.
Miller, G.A., and P.N. Johnson-Laird. 1976. *Language and perception*. NY: Belnap Press.
Netz, R. 1999. Linguistic formulae as cognitive tools. *Pragmatics and Cognition* 7: 147–176.
Norman, D.A. 1993. *Things that make us smart*. Boston: Addison-Wesley Longman.
Pike, A.W.G., D.L. Hoffmann, M. Garcia-Diez, P.B. Pettitt, J. Alcolea, R. De Balb, C. Gonzalez-Sainz, C. de las Heras, J.A. Lasheras, R. Motes, and J. Zilhao. 2012. U-Series dating of Paleolithic art in 11 caves in Spain. *Science* 336: 1409–1413.
Sadier, B., J.-J. Delannoy, L. Benedetti, D. Bourles, S. Jaillet, J.-M. Geneste, A.-E. Lebatard, and M. Arnold. 2012. Further constraints on the Chauvet cave artwork elaboration. *Proceedings of the National Academy of Sciences* 109: 8002.
Schon, D.A. 1983. *The reflective practitioner*. NY: Harper Collins.
Schubert, T., and A. Maass (eds.). 2011. *Spatial dimensions of social thought*. Amsterdam: Mouton de Gruyter.
Shepard, R.N. 1984. Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review* 91: 417–447.
Small, J.P. 1999. Time in space: Narrative in classical art. *Art Bulletin* 81: 562–575.
Smith, S.M. 1995. Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In *The nature of insight*, ed. R.J. Sternberg and J.E. Davidson, 229–251. Cambridge: MIT Press.
Stjernfelt, F. 2011. *Diagrammatology: An investigation on the borderlines of phenomenology, ontology, and semiotics*. Dordrecht: Springer.
Suchman, L. 1987. *Plans and situated actions*. Cambridge: Cambridge University Press.
Suwa, M., and B. Tversky. 1996. *What architects see in their sketches: Implications for design tools. Human factors in computing systems: Conference companion (pp. 191–192)*. NY: ACM.

Suwa, M., and B. Tversky. 2001. Constructive perception in design. In *Computational and cognitive models of creative design V*, ed. J.S. Gero and M.L. Maher, 227–239. Sydney: University of Sydney.

Suwa, M., and B. Tversky. 2003. Constructive perception: A skill for coordinating perception and conception. In *Proceedings of the Cognitive Science Society Meetings*, ed. R. Alterman and D. Kirsh. Mahwah: Erlbaum.

Suwa, M., B. Tversky, J. Gero, and T. Purcell. 2001. Seeing into sketches: Regrouping parts encourages new interpretations. In *Visual and spatial reasoning in design*, ed. J.S. Gero, B. Tversky, and T. Purcell, 207–219. Sydney: Key Centre of Design Computing and Cognition.

Talmy, L. 1983. How language structures space. In *Spatial orientation: Theory, research and application*, ed. H.L. Pick Jr. and L.P. Acredolo, 225–282. N. Y.: Plenum.

Talmy, L. 2000. *Toward a cognitive semantics. Vols 1 & 2*. Cambridge: MIT Press.

Tversky, B. 1981. Distortions in memory for maps. *Cognitive Psychology* 13: 407–433.

Tversky, B. 1990. Where partonomies and taxonomies meet. In *Meanings and prototypes: Studies on linguistic categorization*, ed. S.L. Tsohatzidis, 334–344. London: Routledge.

Tversky, B. 1995. Cognitive origins of graphic conventions. In *Understanding images*, ed. F.T. Marchese, 29–53. New York: Springer.

Tversky, B. 2001. Spatial schemas in depictions. In *Spatial schemas and abstract thought*, ed. M. Gattis, 79–111. Cambridge: MIT Press.

Tversky, B. 2004. Semantics, syntax, and pragmatics of graphics. In *Language and visualisation*, ed. K. Holmqvist and Y. Ericsson, 141–158. Lund: Lund University Press.

Tversky, B. 2005. How to get around by mind and body: Spatial thought, spatial action. In *Cognition, evolution, and rationality: A cognitive science for the XXIst century*, ed. A. Zilhao. London: Routledge.

Tversky, B. 2011a. Spatial thought, social thought. In *Spatial dimensions of social thought*, ed. T. Schubert and A. Maass. Amsterdam: Mouton de Gruyter.

Tversky, B. 2011b. Visualizations of thought. *Topics in Cognitive Science* 3: 499–535.

Tversky, B., and K. Hemenway. 1983. Categories of scenes. *Cognitive Psychology* 15: 121–149.

Tversky, B. and Kessell, A. (2014). Thinking in action. *Topics in Cognitive Science.*

Tversky, B., and P.U. Lee. 1998. How space structures language. In *Spatial Cognition: An interdisciplinary approach to representation and processing of spatial knowledge*, ed. C. Freksa, C. Habel, and K.F. Wender, 157–175. Berlin: Springer.

Tversky, B., and P.U. Lee. 1999. Pictorial and verbal tools for conveying routes. In *Spatial information theory: cognitive and computational foundations of geographic information science*, ed. C. Freksa and D.M. Mark, 51–64. Berlin: Springer.

Tversky, B., and M. Suwa. 2009. Thinking with sketches. In *Tools for innovation*, ed. A. Markman. Oxford: Oxford University Press.

Tversky, B., and J.M. Zacks. 2013. Event perception. In *Oxford handbook of cognitive psychology*, ed. D. Riesberg. Oxford: Oxford.

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

Tversky, B., J. Zacks, P.U. Lee, and J. Heiser. 2000. Lines, blobs, crosses, and arrows: Diagrammatic communication with schematic figures. In *Theory and application of diagrams*, ed. M. Anderson, P. Cheng, and V. Haarslev, 221–230. Berlin: Springer.

Tversky, B., J. Heiser, P. Lee, and M.-P. Daniel. 2009. Explanations in gesture, diagram, and word. In *Spatial language and dialogue*, ed. K.R. Coventry, T. Tenbrink, and J.A. Bateman, 119–131. Oxford: Oxford University Press.

Tversky, B., J. Heiser, and J. Morrison. 2013. Space, time, and story. In *The psychology of learning and motivation*, ed. B.H. Ross, 47–76. Oxford: Elsevier.

Wexler, M., S.M. Kosslyn, and A. Berthoz. 1998. Motor processes in mental rotation. *Cognition* 68: 77–94.

Wohlschläger, A., and A. Wohlschläger. 1998. Mental and manual rotatation. *Journal of Experimental Psychology: Human Perception and Performance* 24: 397–412.

Zacks, J., and B. Tversky. 1999. Bars and lines: A study of graphic communication. *Memory and Cognition* 27: 1073–1079.

Zacks, J., and B. Tversky. 2001. Event structure in perception and conception. *Psychological Bulletin* 127: 3–21.