Spatial Representations and Imagery in Learning

Daniel L. Schwartz and Julie Heiser

Stanford University

To appear in the Handbook of the Learning Sciences. K. Sawyer (Ed.), Cambridge University Press

Spatial representations, when used well, support learning in reading, mathematics, and science. They also enable mental simulations and visualizations that prompt innovation and scientific discovery. Spatial representations, both external drawings and internal images, exploit people's sophisticated perceptual-motor system. The embodiment of thought in perceptual processes has promising implications for learning. In this chapter, we emphasize spatial representations that people construct and transform in their mind's eye. The process of working with these mental spatial representations is called, "imagery."

Spatial representation is different from other forms of cognitive representation studied by learning scientists—linguistic, conceptual, logical—because spatial representations partake of perceptual processes and experiences. Neurological evidence, for example, indicates that perceptual regions of the brain activate when people imagine movement (Kosslyn, 1994). Yet, spatial representations are not mere echoes of perception. They can integrate non-perceptual knowledge that allows people to imagine things they have not seen. Spatial representations have four key properties that determine their unique value for education. We begin with a brief review early psychological research on spatial memory, and then describe the four key properties. We show how these properties can be used to help people learn about the world through mental models and simulations. Afterwards, we discuss ways to help people leverage imagery to innovate new ideas and scientific insight.

1.0 Spatial representations become special

Much of the inspiration for examining spatial representations for learning began with studies of pictorial memory. People have expansive memories for pictures. In a heroic study, Standing (1973) showed people 10,000 pictures over five days. (This is the equivalent of one picture every 15 seconds for eight straight hours on five consecutive days!) Afterwards, people saw a combination of new pictures and original pictures. People correctly recognized the original pictures at a rate of 83%. Vivid pictures were recognized even more frequently. Standing extrapolated that if people saw a million vivid images, they would retain 986,300 in the near term, and would recognize 731,400 after a year. This estimate is well beyond memory for words and sentences (Shepard, 1967). Visuals can be an excellent way to help students remember.

One explanation for this impressive memory is Paivio's (1986) "dual coding" hypothesis. When people see a visual scene, they also explain its content to themselves. Pictures yield a perceptual code and a verbal code in memory, which doubles the chances of retrieval. Additionally, the perceptual and verbal representations can commingle. For example, the ancient Greeks invented the Method of Loci. Orators memorized long speeches by associating elements of their speech with objects along a standard path through a cathedral. Retracing the objects on the path cued people's memory for the speech. More recently, Bower, Karlin and Dueck (1975) gave people nonsensical drawings or "droodles" with or without descriptive captions (e.g. "a midget playing a trombone in a telephone booth"). Participants who received the descriptive captions better recalled and reproduced the droodles. The dual coding hypothesis implies that mental spatial representations are not "pictures in the head." Instead, they can be changed by other mental processes. For example, Stevens and Coupe (1978) found that people who lived in San Diego judged that Reno was to their east. This is wrong, and people made the error because they had schematized San Diego as a coastal city and Reno as an inland city. On the positive side, the ability to integrate perceptual and verbal information is extremely valuable for learning, for example, when reading an evocative story.

The last 30 years of research have revealed several distinct features of spatial representations in relation to learning, starting with huge memory effects and their permeability to verbal information. More recently and as we will describe later in this chapter, spatial representations and imagery also support understanding. Four distinct qualities of perception make spatial representations special for thinking and learning.

2.0 Unique features of spatial representation

Representations of spatial information, whether internal or external, capitalize on the perceptual system. The perceptual system has a structure that enforces and enables specific spatial computations. Imagery is a representation of perceptual experience, and therefore, it inherits the structure of perception to complete computations that are difficult to perform linguistically. Knowing which computations imagery handles particularly well can help one decide when to use spatial materials and processes in education. The four properties of perception with special relevance to education are effortless structure, determinism, action coupling, and pre-interpretation.

These four properties of perception are ubiquitous, and they should be distinguished from perceptual experiences that need to be learned. For example, all people experience a world in depth, but not all people have learned to see a penny. Nickerson and Adams (1979) found that only 42% of their subjects could select the correctly drawn penny from of a set of 14 variants. These Americans had handled thousands of pennies, but they never learned to see them well.

One of the most beguiling aspects of perception is that people easily see what they have learned, yet they can completely overlook what they have not. Consequently, people often believe they perceive all there is to be seen in a situation. This is highly relevant to issues of learning. Educators often provide explanations of phenomena that students have not learned to perceive, and therefore, students do not realize they are missing something. Bransford, Franks, Vye, and Sherwood (1989) described clinical psychology students who learned to diagnose symptoms from print materials. When the students began their internships, they were unable to perceive the symptoms in patients and had difficulty making diagnoses.

In a seminal paper, Gibson and Gibson (1955) argued that perceptual learning involves the increased discernment or pick-up of information; for example, the ability to differentiate an edible and poisonous mushroom. So rather than describing learning as a constructive process of creating more abstract mental representations, they argued that learning gets people closer to the world by improving their abilities to perceive information that has always been there. Gibson and Gibson suggested the use of contrasting cases, much like tasting wines side-by-side. By juxtaposing cases that are similar in many ways, people can begin to discern or notice what makes the cases distinctive. For example, Howard Gardner (1982) describes an art exhibit that juxtaposed original paintings and forgeries. At first people could not tell the difference, but over time, they began to perceive the features that differentiated the originals from the copies (cf. Eisner, 1972; Goodwin, 1994). It is an important lesson for educators that what they perceive may not be the same thing as their students (Nathan & Koedinger, 2000), and it takes special steps, like contrasting cases, to help students see what is important.

2.1 Four qualities of perception that make imagery special

When students can see what is important, they can recruit the special qualities of the perceptual system for thinking with imagery. We now describe four features of perception used by imagery.

2.1.1 Effortless Structure

The function of perception is to provide a cohesive, stable experience that permits action. Perception differs from sensation. Sensation provides spatial information through modalities like sound, sight, and touch. It is perception’s task to integrate and structure that information. When people handle a piece of typing paper, for example, they perceive a white rectangle. However, their fingers sense edges and corners, not a complete object. Their retinas sense a trapezoid due to foreshortening. Moreover, the shape of the trapezoid changes as the paper gets moved around. Nevertheless, perception delivers the unitary experience of a constant rectangle.

Perception packages sensation with little discernible effort, because evolution has conferred specialized abilities that are well matched to recurrent structures in the spatial world including shape, motion, and color. The early perceptual research by the Gestalt psychologists attempted to catalog environmental structures that make “good form” for perception (Wertheimer, 1938). One indicator that perception includes specialized abilities is that different types of visual information – color, motion, brightness, what, where – are processed in separate brain regions. A benefit of these evolved specializations is that cognition gets structural output from perception “for free.” This is relevant to imagery and learning because it suggests that spatial representations can be useful when they provide structures that do not require intensive cognitive effort to manage. For example, students may more easily grasp the structure of a factorial experiment when they can see the factors crossed in a matrix.

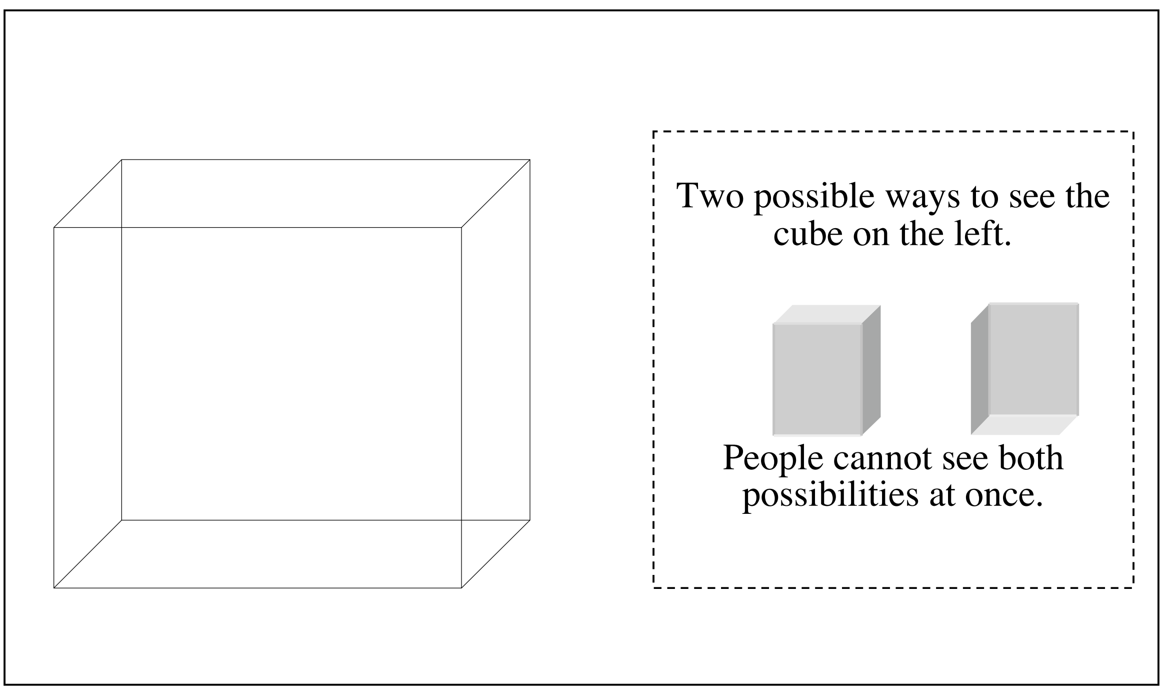


Figure 1. The Necker cube demonstrates determinate perception.

2.1.2 Determinism

Perceptual structure is deterministic – at any given moment, people only see one set of structures. In Figure 1, there are two ways to perceive the orientation of the cube, but perception limits people to one version at a time, and they cannot see mutually inconsistent consequences.

The determinism of spatial perception can be contrasted with language. People can say "the tree is next to the bush" and be vague about where. In sight, the tree needs to be in front, behind, left, or right of the bush (or some variant in between). The spatial relations are determined and cannot be easily omitted as in language. This is relevant to imagery, because it implies that people will represent specific situations in their imagination. Mani and Johnson-Laird (1982), for example, provided people sentences like "the cross is above the square, and the triangle is next to the square." People then saw pictures and were asked if they matched the sentence. Some people rejected pictures that showed the triangle to the left of the square, and other people rejected pictures that showed the triangle to the right of the square. They had taken the indeterminate sentence and made it spatially determinate when they constructed their image. Determinism can be valuable for learning, because it prevents people from being vague.

2.1.3 Perception-action coupling

The third relevant property of perception involves the motor system. Visual perception is much more than watching. People do quietly watch, but perception is tightly coupled with people's abilities to take action. Perception guides motor action. Also, motor actions guide perception. People, for example, move their heads to get a better look at something, and active touch (Gibson, 1962) helps people figure out the shape of an object. People are constantly learning perceptual-motor couplings so that their perceptions and actions are coordinated. People quickly learn to adjust their swing when they change tennis rackets. This recalibration is handled by the perceptual-motor system and involves minimal cognitive effort.

The coupling between perception and action is relevant to imagery, because people need to imagine the consequences of an action. For example, a person might imagine what happens if they try to jump across a wide chasm. To conduct this imagery, people need to animate their image. The animations that people can complete are related to the types of actions they can take, including changing perspective, moving an object, using a tool, walking down a path, and so forth. The relevance of the coupling between action and perception for imagery is exemplified in a study by Parsons (1987). Participants sat at a computer keyboard that showed a hand on the monitor. They had to determine whether it was a left or right hand. The time people took to make the decision was related to how difficult it would be to move their own hand from the keyboard into that position, even though people did not actually move their hands. One implication for education is that it is fruitless to ask people to imagine changes that humans could never possibly precipitate, for example, rotating four-dimensional structures.

2.1.4 Pre-interpretation

The final quality is that perception is largely "bottom-up." Perception often occurs prior to one's beliefs or knowledge about a situation, and one's beliefs cannot easily override a perception. As mentioned, some aspects of perception are learned, particular those aspects that vary across many environments (e.g., the flavor of wine). Those things that are learned are more prone to interpretation than those things built into the system. Many researchers have argued about the line between those aspects of perception that can and cannot be influenced by experience. For example, one hypothesis proposed that people who live in environments without buildings cannot see straight lines, because 'straightness' is a learned property of a "carpentered world" (Gregory, 1966). There will always be arguments about which aspects of perception can be influenced by beliefs and culture. Regardless, perception can occur independently of beliefs. Optical illusions are a good example. Sometimes the world violates the expectations of evolution, and the automatic processes of perception get fooled. Figure 2 displays an illusion, where the two line segments look like they have different lengths. Believing that the lines are the same length does not help people see them as the same length (though people can always assert that they are).

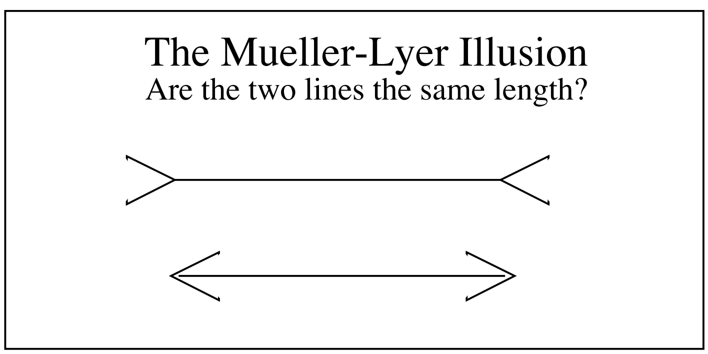


Figure 2. Knowing that the lines are the same length does not block the illusion.

When perception operates independently of beliefs, it is pre-interpretive. Pre-interpretation is an important quality for imagery. People can do imagery work before they get "locked" into a particular interpretation. With language, people need to interpret words, and this shapes the kinds of conclusions they will reach. Pre-interpretive images, on the other hand, allow people to manipulate images to see what forms emerge. We discuss the relevance of pre-interpretation to learning, along with the other three properties, next.

3.0 Imagery for getting closer to the world: Mental models and simulations

People can create images from visual input. For example, people slowly build a mental map while navigating a city. People can also create images from touch. Shelton and McNamara (2001) asked people to arrange a number of objects without looking at them. This improved people's abilities to recognize the layout of the objects when viewed from novel orientations. People can form images based on sound and language; for example, when hearing a book. Sometimes, when people need to construct an image that conflicts with concurrent visual input (e.g., words on a page), it interferes with their imagery (Brooks, 1968). People can also create images in the absence of any immediate input, for example, when anticipating an encounter.

Determining whether people are using imagery is difficult. The fact that someone is working with spatial information does not imply the use of imagery. For example, Schwartz and Black (1996) found that people initially solve gear problems by imagining their movement, but over time they learn to use a quick verbal rule even though the problem is spatial (e.g., adjacent gears turn opposite directions). For most learning scientists, it is more important to design tasks that harness imagery than prove its existence. Therefore, we emphasize the four perceptual properties of spatial representations that educators can recruit to improve thinking and learning.

In this section, we emphasize research where imagery helps people anticipate or learn how a possible world might appear. We begin with examples of people constructing mental models to understand a situation, and we emphasize the relevance of effortless structure and determinism. We then present examples of people using simulations to draw inferences, and we emphasize action-coupling and pre-interpretation. To help highlight the four perceptual properties in imagery, we emphasize them separately, but all four at play in each example.

3.1 Constructing Mental Models

One role for imagery in education is to help people make sense of things they have not experienced first hand and can only hear or read about. Zwann (2004) describes language comprehension as "guided experience." Guided experience, of course, is not as vivid as direct experience, but it engages common mechanisms. These common mechanisms permit people to construct spatial "mental models" and draw inferences, almost as though they were there. Mental models are internal representations where changes within the model into changes in the world (as opposed, for example, to manipulating an algebra formula). For example, Morrow, Bower, and Greenspan (1987) demonstrated that, when reading, people track the spatial location of the characters in a mental model; people can quickly answer questions about a spatial location if the character recently moved to that location, which is what would occur if people were actually walking about the space themselves (see Zwaan & Radvansky, 1988 for a review). Mental models can greatly enhance people's abilities to learn from what they read. To help early readers, it is important (a) to provide information that supports the effortless structure of imagery, and (b) help students learn to construct determinate mental models.

3.1.1 Effortless Structure

Good readers often rely on imagery when trying to comprehend discourse. People, for instance, can generate metric structures in imagery. Metric structures include the intervals between positions and object boundaries. For example, a metric image that portrays the distance between New York and Los Angeles would include the space in between (scaled, of course). This is different from stating that New York and Los Angeles are 3000 miles apart without representing the intermediate space in between. The effortlessness of creating spatial structures often yields cases where people spontaneously create metric structures that go beyond a text. Morrow and Clark (1988), for example, asked participants to read “The tractor is approaching the fence,” or “The mouse is approaching the fence.” Afterwards, when asked to estimate the distance between the fence and the mouse or the tractor, they estimated a larger distance between the tractor and the fence. The sentences were silent about this distance. More generally, one benefit of the effortless structure of imagery is that it permits people to generate images easily so they can scan them for interesting relations.

However, people can have difficulty recruiting the effortless structure of imagery, when the available information does not provide structural cues. Rock and Di Vita (1987), for example, showed people wire figures, like a twisted coat hanger. People could not recognize identical wire figures when they were shown at a different orientation. People were quite bad at this task, because there was so little structural information in the figures. Farah, Rochlin and Klein (1994) replicated the study, but they molded clay into the wire so the stimuli now looked like complicated potato chips. In this case, people were quite good at the task, because the clay provided more cues to the overall structure of each object. Imagery can only be as good as the structural information people have to begin with, and designers of information displays and texts need to keep this in mind.

3.1.2 Determinism

Good readers spontaneously construct deterministic structures of what they are reading. For example, here are two sentences, "A turtle is on a log," and, "A fish swims under the log." Most people easily infer the fish swam under the turtle. People construct an image that determines this relation, and they can see it in their image. Given a brief delay, people are even likely to think they read the sentence, "The fish swam under the turtle" (Bransford, Barclay & Franks, 1972). Early readers, however, do not always construct determinate models of what they read, and therefore, they may leave ideas vague. Schoenfeld (1992) describes examples of students blindly solving math word problems where the situation was impossible, but the students never knew because they did not try to model the problem. Although imagery can provide determinate structures, people need to learn that they should construct images. Imagery does not always arise spontaneously.

It is particularly important to encourage early readers to imagine narratives so they can better understand. Young children have difficulties with imagery compared to adults, and therefore, they need special support (Reiser, Garing, & Young, 1994).

Glenberg, Gutierrez, Levin, Japuntich, and Kaschak (2004) describe an experiment where young children read a passage and manipulated figurines so they portrayed the actions in the passage (e.g., the farmer walked in the barn). After some practice, children were asked to simply imagine manipulating the figurines. As a posttest, the children read a final passage without any prompting. Children who completed this sequence were better at remembering and drawing inferences about the new passage compared to (a) children who received no training, (b) children who were only instructed to imagine the passage; and importantly, (c) children who manipulated the figurines without the intermediate instructions to imagine manipulating. Encouraging imagery through the initial use of physical modeling helped the children learn a strategy to make more determinate relations in their understanding of a text, and this improved their comprehension.

3.2 Running Mental Simulations

A second application of imagery is to help people imagine changes through a mental simulation. Shepard and colleagues (see Shepard & Cooper, 1986) performed ground-breaking studies that proved people can simulate the movement of objects. They called it "analog imagery" to emphasize that people were imagining <u>continuous</u> movements through metric space. Participants saw two objects at different angles. Figure 3 provides an example where the angular disparity between the objects is about 90^{o}. People had to decide if the objects were the same shape. Shepard varied the angular disparity between the objects (15^{o}, 45^{o}, etc.) and measured how long it took people to decide if they matched. Shepard reasoned that people would take longer to complete the task when the two objects had a greater angular disparity, because people would have to mentally rotate one of the objects farther so they could see if it matched.

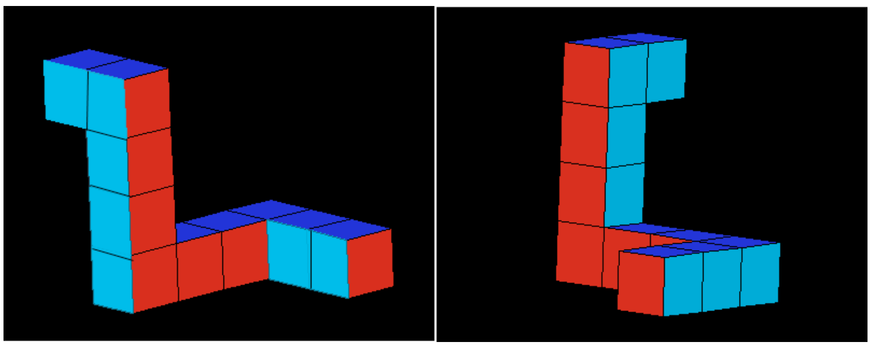


Figure 3. People imagine rotating the objects to solve the problem.

Figure 4 shows one set of results. People exhibited a strong linear relation between the angular disparities of the objects and how long it took them to solve the problems. If one looks closely, there are two clusters of times to solve the problem at 225^{o}. This is because some people rotated the object the shorter direction (135^{o}), so it took them less time than participants who rotated the object 225^{o}. Simulations like these can greatly help students draw inferences, and like mental models, students need to learn to conduct these simulations. This requires learning (a) how to "set-up" simulations so they match different possible actions; and, (b) checking their answers lest they rely on faulty pre-interpretations.

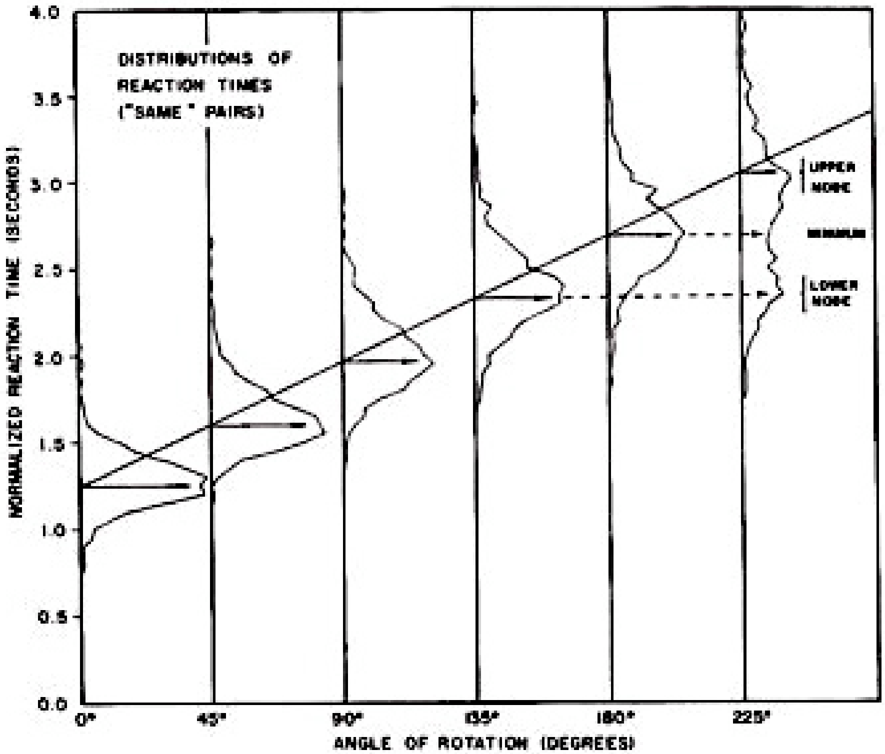


Figure 4. Time to decide if two objects are the same as a function of the angular disparity between the objects. At each angle of disparity, the graph also shows the distributions of response times. From Shepard & Cooper (1986), pg. 58.

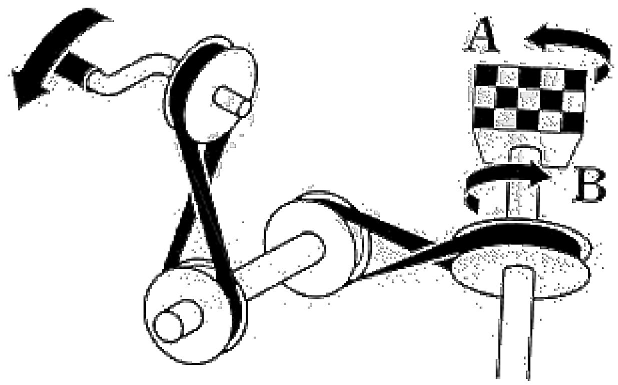
3.2.1 Action-Coupling

Simulations anticipate the consequences of possible motor actions (e.g., turning a block, drawing a line). Simulations are possible because of perception-action coupling. Relevant evidence comes from studies where people accomplish imagery tasks more quickly and accurately when they take movement. Simons and Wang (1998), for example, showed that people easily imagine looking at a set of objects from a new perspective, when they are allowed to walk to that perspective, even if they keep their eyes closed! Wohlschlåger and Wohlshlåger (1988), Wexler and Klam (2001) showed that people imagine object rotations more quickly when they can physically rotate an object with their hands, even if they are not directly touching the object (e.g., Schwartz &

Holton, 2000 had people pull a string that turned a table on which an object rested). People often think of simulations as guiding possible actions. However, in these cases, actions guided imagery.

The coupling of imagery to the motor system is relevant to reading. Klatzky, Pellegrino, McCloskey, and Doherty (1989), for example, told people to form their hand into an open grip or a flat palm. They were then asked, "Is it possible to squeeze a tomato?" People were faster to say "yes" when they had made the open grip. So, not only does embodying words through imagery and action improve comprehension (as described above), it also improves access to relevant possible actions and their likely consequences.

The coupling of imagery and the motor system enables people to anticipate possible changes, and this can help students solve problems, for example, when rotating molecular models to see if they can bond with each other. At the same time, imagery is often limited to the complexity of action available to the body. People can only image so much in a single simulation. Hegarty (1992, 2000) examined people's abilities to "mentally animate" complex pulley systems as shown in Figure 6. Hegarty found that good problem-solvers simulated one interaction at a time and in the right order. For example, they imagined how the belt would rotate if they turned the crank. They would then propagate the result by imagining how the moving belt would turn the next pulley in the series. Spatial representations interact with other forms of knowledge, in this case problem decomposition, and therefore, it is important to help students to develop the skills necessary to decompose a problem into a form that is amenable to imagery.

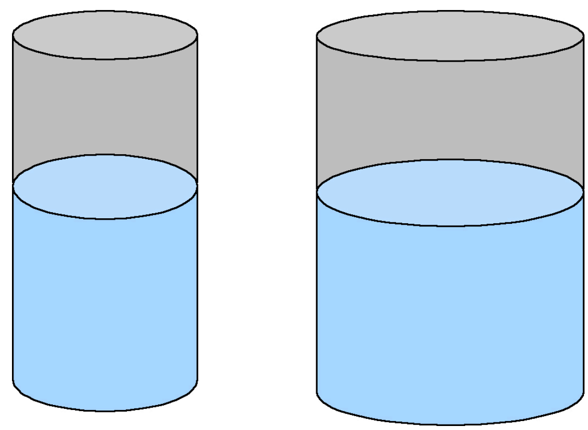


If the handle is turned, will the box turn direction A or B?

Figure 7. Gear and belt task used to examine complex imagery. From Hegarty (2000).

3.2.2 Pre-interpretation

Pre-interpretation is relevant to simulations of interacting objects. When people simulate interacting objects, they need to coordinate their relative rates of movement. For example, imagine a large gear driving a small gear. The speed of the small gear will depend on the large gear, and it will be faster than the large gear. To simulate this scenario, people need to coordinate their relative rates. The same is true of most simulations involving multiple components, for example, scissors with long blades but short handles. Pre-interpretive expertise helps people coordinate movements in complex simulations.



Does one glass start pouring water at a lesser angle of tilt?

Figure 5. A problem where pre-interpretive knowledge helps solve a problem that people typically cannot solve verbally. Adapted from Schwartz and Black (1999).

The role of pre-interpretive expertise in imagery was demonstrated by Schwartz and Black (1999). Participants received the problem in Figure 5. A wide and narrow glass of the same heights are filled to identical levels of water. Will the two glasses start spilling water at the same angle of tilt? Very few people gave the right answer when they could only look at the glasses. In a second condition, participants were handed one glass at a time. There was no water in the glass, just a black strip indicating the “pretend” water level. People closed their eyes and tilted the glass until they thought the water would just start to spill out. They repeated the process with the second glass. In this case, every individual correctly tilted the narrow glass farther than the wider glass. Thus, people had pre-interpretive, action-dependent expertise that enabled them to simulate the water-glass movement interactions. Scientists often use simulations to tap their intuitions (Clement, 1994), and asking learners to imagine “what would happen if,” can help them ground explicit understanding with their own pre-interpretive expertise. At the same time, educators need to watch for student misconceptions that are borne of pre-interpretive expertise. Pre-interpretive expertise is for pragmatic action, and therefore, it does not always match scientific truths and can interfere with science learning (e.g., naïve physics, McCloskey, 1983).

Pre-interpretive expertise often appears in tasks that involve motor activity, especially for young children. Krist, Fieberg, and Wilkening (1993) asked children to push a ball off a ledge so it would land on a target on the floor. The researchers varied the height of the ledge, and children accurately modified their pushing force. However, when asked explicitly how hard they needed to push the ball, the children assumed a higher ledge called for a great push because the ball was further away from the target. Of particular relevance to educators, Alibali, Church, and Goldin-Meadow (1993) found that children who are in transition between developmental stages demonstrate competence through the motor system before they have achieved explicit understanding. For example, when pouring the same amount of water in a narrow and wide glass, 6-year-old children often say that the narrower glass has more water, because the water is reaches a higher level. But, when the children are close to appreciating the true answer, they often tacitly indicate the width of the class with their hands, even though they are discussing the water's height. Alibali and colleagues have shown that teachers who attend to these types of speech-gesture mismatches (which also appear in mathematics) can better deliver "just-in-time" instruction that moves children to explicit understanding. It is an interesting question whether training teachers to look at gestures is a tractable method to increase their abilities to tailor instruction.

4.0 Imagery for going beyond the world: Emergence and covariance

Here, we consider imagery that goes beyond experience and creates innovations in thought. Human imagination may never be as creative as nature; for example, who could imagine a rainbow had they never seen or heard of one. Still, human imagination is an impressive affair, and spatial representations are implicated. The history of science is filled with eminent scientists who claim imagery helped them infer structures hidden from perception.

Thus far, we have described how people use imagery to help anticipate what they might be likely to perceive. Imagery helped people be more efficient in connecting them to the empirical world of experience, and measures of performance used speed and accuracy. When imagery is used for innovation and going beyond experience, the measures of success are found in the novelty and appropriateness of the structures that people create. Innovation is critical to learning, because the goal of education is not only to make people more proficient at what they already know; it is also to help them develop new structures for thinking (Schwartz, Bransford, & Sears, 2005). In this section, we describe two uses of imagery for innovation: emergence and covariant representations.

4.1 Emergence

The effortless structure and simulations of spatial representations are important contributors to innovation and discovery. People imagine changes to see whether new patterns emerge. First consider visible action. Martin and Schwartz (in press) asked 9-year-olds to solve equivalent fraction problems, for example, "What is one-fourth of eight?" Children received eight small plastic tiles and had to indicate their answers. When children could only look at the pieces without touching, they tended to indicate one and/or four pieces as the answer. They mapped the pieces in one-to-one fashion to the numerals “1” and “4” in the fraction 1/4. However, when children could push the pieces around, they were nearly four times more successful. By moving the pieces, the children began to notice grouping structures. They discovered it is possible to make “4” groups, and “1” group has two pieces.

Imagery, like visible action, supports innovation because it helps people to move shapes into new configurations relatively effortlessly. Finke (1990), for example, asked people to imagine the letter ‘C’ and the letter ‘J’. He asked people to imagine rotating the ‘C’ so it was on top of the ‘J’. He asked people what they saw, and they often said an umbrella. Though people had not started with the idea of an umbrella, it emerged from the reconfiguration of the shapes in their imagination. In another set of studies, people were not told what geometric pattern to construct, but they still exhibited emergence. People saw a number of abstract shapes (e.g., square, circle, line). Finke asked the participants to close their eyes and imagine, for example, how the shapes could create a piece of equipment for cleaning gutters. They produced many creative solutions.

Asking students to imagine may also help them innovate emergent solutions to important problems, but simply telling them to close their eyes and be creative is probably insufficient. A major challenge of innovation is that explicit interpretations can interfere with developing new ones. For example, Chambers and Reisberg (1985) asked people to look at Figure 7, which can be interpreted as a duck or a rabbit. Once people had an interpretation, they closed their eyes. Chambers and Reisberg asked if they could come up with a second interpretation; could they overcome their original interpretation (e.g., duck) and see a second interpretation (e.g., rabbit)? Not a single participant over several studies could do the re-interpretation. Explicit interpretations interfered with people's abilities to see an alternative. More generally, the great challenge of innovation is "breaking set." Once reasoning begins with an interpretation, it is hard to shake free. In Finke's studies, where new interpretations did emerge, people did not begin with a pre-interpretation. They simply encoded a set of geometric shapes. Because imagery can transform shapes without interpretation, people could move the shapes in their imagination and continue to make determinate structures until they saw an interesting pattern and develop an interpretation of it.

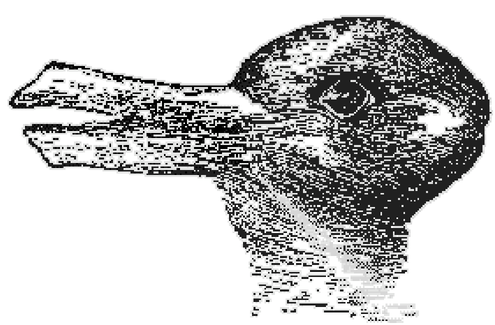


Figure 7. Duck/rabbit image used in many experiments (e.g. Chambers & Reisberg, 1985).

To help students use their imaginations to innovate new (for them) ideas, it is important to remember that innovation favors the prepared mind. Keulke, who famously imagined the benzene ring, knew that he needed a structure that could join multiple atoms, and he was prepared to recognize the significance of what he imagined. To recognize an emergent structure, students need a strong understanding of the constraints and problems the structure must handle, and this is something that appropriate instruction can do. Yet, at the same time, students must avoid interpretations that pre-figure their solutions, because this will interfere with their ability to develop new interpretations. Thus, innovation through imagery requires a delicate balance. Compared to the massive efficiency literature on how to make people faster and more accurate, the learning sciences could use more research on techniques for fostering innovation (Schwartz, Bransford, & Sears, 2005).

4.2 Covariant Representations

Perhaps the most impressive instances of spatial innovation involve covariant representations. Covariant spatial representations do not resemble their referents. For example, a speedometer does not look like a speeding car and a clock does not look like time. However, changes in speed or time map neatly onto the changes displayed in the dials. Among their strengths, covariant representations make it possible to represent non-spatial phenomena in spatial form. Venn Diagrams, for example, can represent a space of personality traits. There are a many visual representations that people have invented to support reasoning about both spatial and non-spatial matters.

Creating a covariant spatial representation is an impressive feat of creativity. Galileo has been credited with inventing the first covariant representation to make an argument (Cummins, 1989). He used area to represent distance. Students can learn to invent covariant representations given appropriate educational support. Schwartz (1995) found that very few adolescents spontaneously construct visualizations to solve problems. However, once they were encouraged to invent their own representations, and they experienced their exceptional benefits for problem solving, the students spontaneously started to invent their own forms several weeks later on novel problems. Bamberger (1991) describes how children, given prompting, will invent increasingly precise visual representations of musical form that includes pitch and duration. There have been promising educational efforts that capitalize on children's facility with covariant representations. diSessa, Hammer, Sherin, & Kolpakowski (1991) describe children's "meta-representational expertise" as they are encouraged to progressively create more refined visual representations of motion.

One of the benefits of covariant representations is that they distill the most critical aspects of a situation into a form that enables people to bring to bear their spatial abilities for working with structure. Larkin and Simon (1987), for example, investigated why "a picture is worth 10,000 words." They found that the structure of spatial representations makes them very easy to search. A matrix that uses rows for trip locations and columns for costs permits a person to index their search by cost or location. Covariant representations also create gestalts that guide the perception of important forms. Descartes' invention of the X, Y coordinate system permits the easy detection of linear and curvilinear patterns. When experts externalize and communicate covariant representations, they can help learners see how they structure their thoughts.

One scaffold for helping students build covariant representations comes from work on Teachable Agents (Schwartz et al., in press). Students teach a computer agent by using predefined forms to build important covariant representations. In turn, the agent can show how it reasons based on the structure the student creates. Figure 8 shows Betty, whom students teach by creating a directed graph that uses links like "increases" and “decreases.” Betty can then animate her reasoning by tracing through the links to answer questions. Students learn to structure their thought visually and trace relations in their own thoughts. For example, students who teach Betty are more likely to reason about multiple causal pathways (Biswas, Schwartz, Bransford, & TAG-V, 2001).

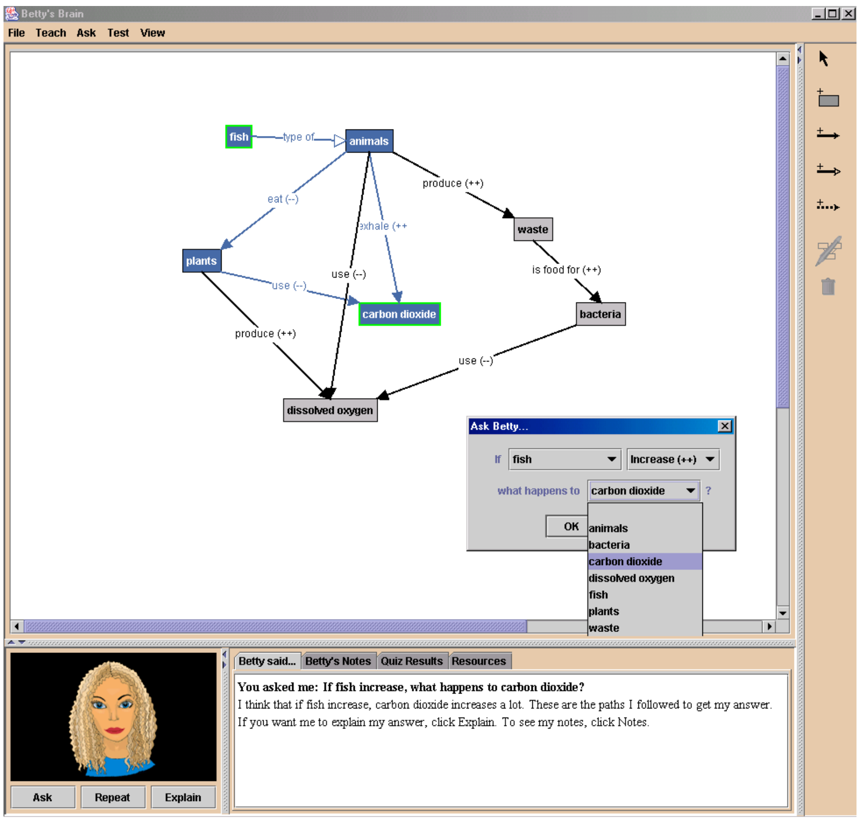


Figure 8. A Teachable Agent.

5.0 Conclusions

A search of the learning-science relevant databases indicates that human spatial competence has received well below half the attention given to either language or social behavior. This is surprising given human abilities to wield space to create art, visualizations, and multipart tools. Hopefully, this imbalance will change as new technologies permit more spatial methods of interacting with information. Because much of the research that does exist is about subtle individual differences in spatial abilities, people overlook how capable all humans are with spatial representations. In some cases, it leads to the belief that "spatial" students should receive special instruction that emphasizes imagery, but that non-spatial students should receive some other form of instruction. It may be a mistake to assume that only people with high imagery abilities should be presented with visual information. For example, they may find visual tasks too easy, and therefore, they may not elaborate as deeply. Additionally, there is evidence that spatial visualization ability develops with experience (Baenninger & Newcombe, 1989) so denying low-spatial individuals the chance to work with visual information may limit their development. Finally, although subtle differences in spatial ability appear when spatial information is presented in difficult ways, they usually disappear when effective visual representations are designed to scaffold spatial reasoning (Heiser, 2004). Assembly diagrams that are step-by-step and present consistent structural cues erase the differences between low and high spatial individuals on assembly tasks. The learning sciences should focus on learning rather than assumptions about ability, and it should examine how to design visual environments that will benefit all learners.

Our review of imagery has been unorthodox. Most work on imagery has emphasized its introspective vividness or its geometric properties (e.g., does it use a viewpoint or object-centered coordinate system). In contrast, we argued that it is important to understand the relation of imagery to other sources of knowledge, and therefore, we identified four properties that can supplement non-spatial forms of reasoning: effortless structure, determinism, perception-action coupling, and pre-interpretation. When people learn calculus, they need to combine linguistic, mathematical, and spatial processes, and therefore, it is important to investigate the strengths and weaknesses that imagery can bring to a larger learning endeavor.

We also made a distinction between uses of imagery that mimic perceptual experience and uses of imagery that go beyond experience. This is an important distinction, because the ways that we assess the educational benefits of imagery will be quite different for each (e.g., speed and accuracy versus novel structure). We have primarily discussed internal spatial representations, but this distinction cuts across both internal and external spatial representations. External spatial representations, such as maps and diagrams can serve to bring you closer to the world and also farther from the world. For example, when architects use their sketches as abstractions to help them innovate new forms, they omit details that would have to be determinate and therefore constrain their thinking. Further along in the design process, these sketches are used to represent details and make all structures determinate (Suwa & Tversky, 1997). Similarly, people can use imagery to work with abstract shapes to see what structures emerge, and they can use imagery to construct specific models and run simulations to comprehend and predict in more detail. The fundamental challenge for the learning sciences is not to determine whether people use imagery or whether imagery is “good.” Rather the challenge is more precise – when and how can people use which function of imagery to support learning, creativity, and reasoning.

6.0 Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant Nos. BCS-0214548 and REC-0231946. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

7.0 References

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. Cognitive Psychology, 25, 468-523.

Baenninger, M. & Newcombe, N. (1989). The Role of Experience in Spatial Test Performance: A Meta-Analysis. Sex Roles, 20, 327-344.

Bamberger, J. (1991). The mind behind the musical ear: How children develop musical intelligence. Cambridge: MA: Harvard University Press.

Biswas, G., Schwartz, D. L., Bransford, J. D., & TAG-V. (2001). Technology support for complex problem solving: From SAD Environments to AI. In K. Forbus & P. Feltovich (Eds.), Smart machines in education (pp. 71-98). Menlo Park, CA: AAAI/MIT Press.

Bower, G. H., Karlin, M. B., & Dueck, A. (1975). Comprehension and memory for pictures. Memory & Cognition, 3, 216-220.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

Bransford, J. D., Franks, J. J., Vye, N. J. & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 470-497). NY: Cambridge University Press.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 22, 349-368.

Chambers, D., & Reisberg, D. (1985). Can Mental Images Be Ambiguous? Journal of Experimental Psychology: Human Perception and Performance, 11, 317-28.

Clement, J. (1994). Use of physical intuition and imagistic simulation in expert problem solving. In D. Tirosh (Eds.), Implicit and explicit knowledge. Norwood, NJ: Ablex.

Cummins, R. (1989). Meaning and mental representation. Cambridge, MA: MIT Press.

diSessa, A., Hammer, D., Sherin, B. & Kolpakowski, T. (1991). Inventing graphing: Metarepresentational expertise in children. Journal of Mathematical Behavior, 10, 117-16.

Eisner, E.W. (1972). Educating artistic vision. NY: The Macmillan Company.

Farah, M., Rochlin, R., Klein, K. (1994). Orientation Invariance and Geometric Primitives in Shape Recognition. Cognitive Science, 18, 325-344.

Finke, R. (1990). Creative Imagery: Discoveries and Inventions in Visualization. Hillsdale, NJ: Erlbaum.

Gardner, H. (1982). Art, Mind and Brain. NY: Basic Books.

Gibson, J.J. (1962). Observations on active touch. Psychological Review, 69, 477-491.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. Psychological Review, 62, 32-41.

Glenberg, A. M., Gutierrez, T,, Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. Journal of Educational Psychology, 96, 424-436.

Goodwin, Charles (1994). Professional Vision. American Anthropologist, 96, 606-633.

Gregory, R.L. (1966) Eye and Brain. London: Weidenfeld and Nicholson.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. Journal of Experimental Psychology: Learning, Memory and Cognition, 18, 1084-1102.

Hegarty, M. (2000). Capacity Limits in Diagrammatic Reasoning. In M. Anderson, P. Cheng and V. Haarslev (Eds.). Theory and Application of Diagrams. Berlin: Springer.

Heiser, J. (2004). External Representations as Insights to Cognition: Production and Comprehension of Text and Diagrams in Instructions. Unpublished Doctoral Dissertation, Stanford University.

Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. Journal of Memory & Language, 28, 56-77.

Kosslyn, S. M. (1994) Image and Brain: The Resolution of the Imagery Debate. Cambridge, MA: MIT Press.

Krist, H., Fieberg, E. & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 952-966.

Larkin, J. & Simon, H. (1987) Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

Mani, K. & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. Memory & Cognition 10, 181-187.

Martin, T., & Schwartz, D.L. (in press). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. Cognitive Science.

Morrow, D. G., & Clark, H. H. (1988). Interpreting words in spatial descriptions. Language and Cognitive Processes, 3, 275-292.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and Language, 26, 165-187.

McCloskey, M. (1983) Naive theories of motion. In D. Gentner and A. Stevens (Eds.), Mental Models, (pp. 229-324). Hillsdale, NJ: Erlbaum.

Nathan, M. J., and Koedinger, K. R. (2000). An investigation of teachers' beliefs of students' algebra development. Cognition and Instruction, 18, 209-237.

Nickerson, R.S., and Adams, J.J. (1979). Long-term memory for a common object. Cognitive Psychology, 11, 287-307.

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford, England: Oxford University Press.

Parsons L.M. (1987). Imagined spatial transformations of one's hands and feet. Cognitive Psychology 19, 178–24.

Reiser, J. J., Garing, A. E., & Young, M. E. (1994). Imagery, action, and young children's spatial orientation. It's not being there that counts, it's what one has in mind. Child Development, 65, 1262-1278.

Rock, I., & Di Vita, J. (1987). A case of viewer-centered object perception. Cognitive Psychology, 19, 280-293.

Schoenfeld, A. (1992) Learning to think mathematically: Problem solving, metacognition and sense making in mathematics. In D. Grouws (Ed.), Handbook of Research on Mathematics Thinking and Learning (pp. 334-370). New York: MacMillan.

Schwartz, D. L. (1993). The construction and analogical transfer of symbolic visualizations. Journal of Research in Science Teaching, 30, 1309-1325.

Schwartz, D. L. & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. Cognitive Science, 20, 457-497.

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 116-136.

Schwartz, D. L., Blair, K. P., Biswas, G., Leelawong, K., Davis, J. (in press). Animations of thought: Interactivity in the teachable agent paradigm. To appear in R. Lowe & W. Schnotz (Eds), Learning with Animation: Research and Implications for Design. UK: Cambridge University Press.

Schwartz, D. L., Bransford, J. D., Sears, D. L. (2005). Efficiency and innovation in transfer. (in press). In J. Mestre (Ed.), Transfer of learning from a modern multidisciplinary persepctive. CT: Information Age Publishing.

Schwartz, D. L. & Holton, D. (2000). Tool use and the effect of action on the imagination. Journal of Experimental Psychology: Learning, Cognition, and Memory, 26, 1655-1665.

Shelton, A. L., & McNamara, T. P. (2001). Visual memories from non-visual experiences. Psychological Science, 12, 343-347.

Shepard, R. N. (1967). Recognition memory for words, sentences and pictures. Journal of Verbal Learning and Verbal Behavior, 6, 156-163.

Shepard, R. N., & Cooper, L. A. (1986). Mental images and their transformation. Cambridge, MA: MIT Press.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. Psychological Science, 9, 315-320.

Standing, L. (1973). Learning 10,000 pictures. Quarterly Journal of Experimental Psychology, 25, 207-222.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. Cognitive Psychology, 10, 422-437.

Suwa, M. and Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. Design Studies, 18, 385-403.

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. B. Ellis (Ed.), A Sourcebook of Gestalt Psychology, (pp. 71-88). Harcourt, Brace and Company, 1938.

Wexler, M., & Klam, F. (2001). Movement prediction and movement production. Journal of Experimental Psychology: Human Perception and Performance, 27, 48-64.

Wohlschläger, A., & Wolshläger, A. (1998). Mental and manual rotation. Journal of Experimental Psychology: Human Perception and Performance, 24, 397-412.

Zwaan, R.A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In B.H. Ross (Ed.), The Psychology of Learning and Motivation, Vol. 43 (pp. 35-62). NY: Academic Press.

Zwaan R.A, & Radvansky G.A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162-185.