Part II

Functional Processes for Learning

5

Learning as Coordination

Cognitive Psychology and Education

Daniel L. Schwartz
Stanford University

Robert Goldstone
Indiana University

It is human nature to create dichotomies—mine versus yours, hot versus cold. Dichotomies usefully structure and simplify the world. They can also lead people astray. Aesop's fable of *The Satyr and the Man* captures this risk:

> A Man was walking in the woods on a very cold night. A Satyr came up to him. The Man raised both hands to his mouth and kept on blowing at them.
>
> "What do you do that for?" asked the Satyr.
>
> "My hands are numb with the cold," said the Man, "and my breath warms them."
>
> Later, the Satyr saw the Man again. The Man had a bowl of steaming soup. The Man raised a spoon of soup to his mouth. He began blowing upon it.
>
> "And what do you do that for?" asked the Satyr.
>
> The Man said, "The soup is too hot, and my breath will cool it."
>
> The Satyr shouted, "The Man blows hot and cold with the same breath!"
>
> The Satyr ran away. He was afraid the Man was a demon.

Each pole of the dichotomy contains a truth—the man's breath warmed his hands and cooled his soup. The problem is that the satyr treated the categories of hot and cold as mutually exclusive and did not seek a deeper analysis. Instead, he became agitated and fled the possibility of a unifying explanation.

Education has produced its share of dichotomies: abstract versus concrete, memorizing versus understanding, teacher-centered versus student-centered, authentic tasks versus decomposed practice, efficiency versus innovation, and many more. Often the categories of such dichotomies become mutually exclusive alternatives, and people advocate for one versus the other. Since at least the behaviorism of B. F. Skinner (1986), scholars have argued whether discovery or entrainment is better for learning (e.g., Kirschner, Sweller, & Clark, 2006; Tobias & Duffy, 2009). The so-called math and reading wars are strong examples of heated polarization in education.

More prevalent and less extreme than heated debates, people simply accept the definition of one category as the negation of another, as in the case of active versus passive learning. People do not flee like the satyr, but they do not seek a deeper analysis either. On deeper analysis, familiar categories of learning, often taken as mutually exclusive, have underlying mechanisms that can make them complementary. So rather than choosing one or the other, the best strategy is to choose both.

The chapter follows a central thesis: A major task of teaching and instruction is to help learners coordinate categories of cognitive processes, capabilities, and representations. While nature confers basic abilities, education synthesizes them to suit the demands of contemporary culture. So, rather than treating categories of learning and instruction as an either–or problem, the problem is how to coordinate learning processes so they can do more together than they can alone. This thesis, which proposes a systems level analysis, is not the norm when thinking about teaching and learning. More common is the belief that learning involves strengthening select cognitive processes rather than coordination across processes. Our chapter, therefore, needs to develop the argument for learning as coordination. To do so, we introduce findings from the field of cognitive psychology.

Cognitive psychology focuses on the mechanisms of mind and brain that determine when and how people solve problems, make decisions, interpret situations, remember, learn, and adapt. There are many reviews of cognitive psychology as it relates to education (e.g., Koedinger, Booth, & Klahr, 2013; Pashler et al., 2007). There are also cognitively minded books for education (Bransford, Brown, & Cocking, 2000; Mayer, 1987), cognitive psychology textbooks (Anderson, 2000), and excellent free online resources (www.learnlab.org/research/wiki/index.php/Main_Page). These all introduce the central constructs of cognitive psychology, including attention, different forms of memory, expertise, problem-solving strategies, schemas, and more. Many topics originally investigated by cognitive psychology have matured to the point that they now have their own chapters in this *Handbook* and do not need further coverage here (e.g., see Chapters 9 and 15). Therefore, the goal of the present chapter is not to provide an encyclopedic review. Instead, the primary goal is to provide framing and examples for how to view learning from a cognitive perspective that is relevant to questions of teaching and instruction ranging from reading to math. A second goal is to introduce cognitive neuroscience, which is increasingly a part of the cognitive psychology tool kit. We show where neuroscience can complement behavioral analyses.

The first section of the chapter considers the natural human tendency towards categorization with a special focus on reconsidering one of the most influential categorical frameworks in education—Bloom's taxonomy. The next section presents a view of the mind and brain that helps to indicate why mutually exclusive categories of learning are problematic. The third section presents the heart of the thesis: A major goal of school-based instruction is to help learners coordinate different cognitive processes in the service of cultural goals such as being able to read. The section is populated with examples from research on the teaching and learning of math and reading. The remaining sections provide two examples of common dichotomies, including memorization versus understanding and concreteness versus abstraction. The examples provide a glimpse into how cognitive processes that putatively occupy the poles of a dichotomy can work in concert. The conclusion considers *dichotomania* more generally and offers a tentative prescription.

Categorical Thinking and Education

Before developing our alternative to dichotomous thinking, it is worth understanding the power of categories and boundaries, which make dichotomies possible. Boundaries appear throughout cognition. At the lowest levels, vision has dedicated neural circuitry that detects the edges that separate one object from another. Rainbows present to us a continuous range of wavelengths, yet we tend to see rainbows as consisting of seven distinct bands of color. At the highest levels, people intentionally impose boundaries. Political systems depend on fabricated social boundaries that often become physical ones. Creating boundaries is fundamental to the human experience (Medin, Lynch, & Solomon, 2000) and reaches from basic perception to cultural organization.

Categories follow from boundaries; they collect those things that fall within a physical or conceptual boundary. Categories simplify and stabilize an otherwise ever-changing world. The category of "self" applies during dinner and when waking up, even though one is quite different at those two time points. Without categories, experience would be a flow of inchoate sensations without organizational structure. Once categories are fixed mentally, people de-emphasize differences among members of a category, and accentuate differences across categories (Goldstone & Hendrickson, 2010; Harnad, 1997).

Language is an important contributor to category formation (Borodistky, 2001; Lupyan, 2008). When speaking, it is impossible to convey the totality of experience and all the subtle variations one might be experiencing right now at this very second. Language fixes the flow of experience into categories. Through language, people can reflect upon and communicate categories. Lawyers' carefully worded statements, political platforms, and the movement toward non-sexist and non-discriminatory language are all motivated by the realization that the words we use do not just label our experiences, but also shape and warp these experiences. Being labeled as a member of a category, by a stereotype for instance, can have large effects on how people experience and perform in the world (Steele, 1997). Humans create categories, and categories create humans (McDermott, 1996).

Given their centrality in human thought, categorization schemes can be extremely powerful. An important goal of education is to help students learn cultural and scientific categorization schemes (e.g., republics, taxonomies). Categories, even imperfect ones, can advance science. They offer initial hypotheses that drive research that may even eventually replace the original categories. On the negative side, once a categorization scheme is in place, it can be difficult to displace. It took over a thousand years to overhaul the categories of Aristotelian physics with the modern conception of force. People still spontaneously develop Aristotelian categories to understand physical phenomena, and it takes substantial instruction to displace those naïve misconceptions (Hestenes, Wells, & Swackhamer, 1992).

Bloom's taxonomy of educational outcomes (Figure 5.1) provides an example of the strengths and weaknesses of categorization schemes (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). On the positive side, the taxonomy was a brilliant effort to create an assessment framework. It helped educators focus on a more differentiated set of outcomes than the coarse observation that a student "learned." The taxonomy describes a pyramid of the following order, going from bottom to top: memory (called "knowledge" back then), comprehension, application, analysis, synthesis, and evaluation. More recently, some scholars have put in a new top layer, labeled creativity. The pyramid was seminal in pointing out that there are learning outcomes that go beyond the repetition of behavior, which was the prevailing behaviorist perspective at the time.

On the negative side, many people interpret the categories as forming a prerequisite structure. Students must first learn the lower-order skills at the bottom of the pyramid (memory), before engaging in the skills at the top of the pyramid (evaluation). This interpretation fuels a back-to-basics mentality, so that students should memorize before trying to apply their learning usefully. However, the science

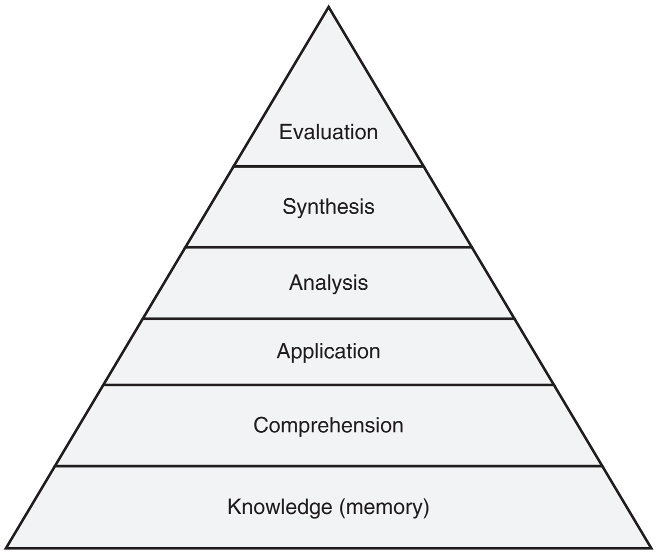

Figure 5.1 Bloom's taxonomy of cognitive outcomes is a framework for analyzing learning outcomes from 70 years ago (Bloom et al., 1956). Contemporary research does not support the implied ordering that people should learn the bottom of the pyramid before engaging the top of the pyramid.

of learning does not support this interpretation. For example, comprehension occurs above memory in the taxonomy, but people can remember ideas better when they comprehend them (Bransford & Johnson, 1972). Making memories a prerequisite for comprehension does not work very well. Similarly, having students learn a new topic in an application context is a useful way to help them simultaneously learn the facts and evaluate their applications (Barron et al., 1998).

Bloom's taxonomy neatly captures the strengths and weaknesses of categorizations in education. It is a compelling and intuitive categorization scheme, and as such, it has had tremendous influence on practitioners and scientists alike. At the same time, it has been difficult to change, despite 70 years of subsequent research that challenges the pyramidal structure. Moreover, people use the categories in ways that violate their intent. Bloom's taxonomy is an assessment framework for evaluating instructional outcomes. It is not a framework for learning or designing instruction, but people still use it that way.

The Distributed Nature of Cognition

One of the important qualities of cognition is that different categories of thinking comprise distributed and overlapping subprocesses at another lower level of description (Rumelhart, McClelland, & the PDP Research Group, 1986). For instance, subtraction and multiplication are separate categories of mathematical operation, and each requires its own set of mental steps to compute an answer. It seems safe to say that when people are doing subtraction, they have "shut off" multiplication. However, at a lower level of analysis, they are engaging many of the same underlying processes for both types of computation. What appears to be different at one level of analysis is not so different at another. We provide an example by introducing brain research that uses functional magnetic resonance imaging (fMRI).

Brain cells are alive and therefore always active to some degree. If one simply looked at the activation of the brain for any category of thought, all the cells would be active. The constant activation of the brain makes for an interesting methodological problem, because it is not possible to say that one cognitive process (set of cells) is on, and another cognitive process is off. To solve this problem, brain research examines relative changes to levels of activation.

The MRI machine used for brain research is the same machine that doctors can use to collect images of soft tissues, such as a torn knee cartilage. For knee injuries, the machine records structural data on the shape and density of tissue. When used for fMRI, the scanner can detect changes in blood flow within the brain. When people complete a task, some of the brain cells do more work than others do. These working cells need to be replenished with oxygenated blood, and the fMRI picks up the changes in the blood flow. fMRI does not capture the firing of the neurons when people are completing the task, but rather the increase in blood flow after the task (about 2 seconds later).

fMRI research depends on comparing the amount of local blood flow for different tasks. When a region of the brain receives more blood, scientists infer it has been more active. A study by Lee (2000) demonstrates a typical research strategy. People completed subtraction tasks and multiplication tasks. The fMRI recorded brain activity during the two tasks. The investigators studied the average activation patterns across the brain for the subtraction tasks, as well as the activation patterns for the multiplication task. The whole brain is active for both tasks, but the researcher wanted to find out which brain regions are selectively more active for one task versus the other. To find out, the investigator took the activation patterns for subtraction and removed the activation patterns in common with multiplication. In other words, the scientist statistically removed all the activation for subtraction that was common with multiplication. The leftover activation indicates which parts of the brain are involved preferentially in subtraction compared to multiplication. The researchers then flipped the comparison. They took the brain activity for the multiplication task, and removed activity that was in common with the subtraction task. Figure 5.2 shows the results. The black regions indicate areas that are more active for multiplication than subtraction, and the white areas show the areas that are more active for subtraction than multiplication.

The circled intraparietal sulcus (IPS) region was more active for subtraction. The IPS is also involved in various spatial attention tasks and judgments about the size of things (Uddin et al., 2010). One interpretation is that people are consulting some form of spatial representation—a mental number line—when doing subtraction (Dehaene, Piazza, Pinel, & Cohen, 2003). They are making sense of the relative magnitudes of the numbers while completing the subtraction on symbolic digits. In contrast, the area indicated as AG (angular gyrus) is more active for multiplication. This region

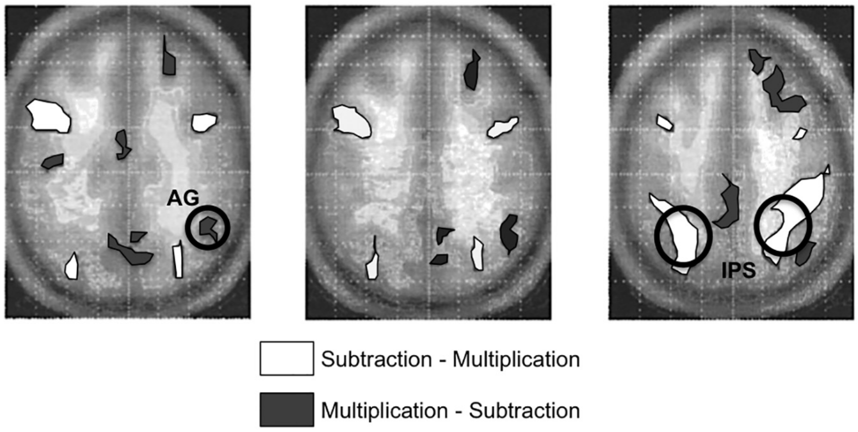

Figure 5.2 Regions of the brain are selectively active for different mathematics tasks. The figure shows three different slices of the brain. Areas in white are more active for subtraction than multiplication. Areas in black are more active for multiplication than subtraction. The circled region labeled AG indicates the rough location of the angular gyrus. The circled regions labeled IPS indicate the rough location of the intraparietal sulcus. (Adapted from Lee, K. (2000). Cortical areas differentially involved in multiplication and subtraction: A functional magnetic resonance imaging study and correlation with a case of selective acalculia. *Annals of Neurology*, 48(4), 657–661.)

is also involved in the retrieval of factual, verbal memories. Thus, people seem to rely on the semantics of quantity (e.g., size and order) for subtraction, and they appear to rely on rote memory for multiplication, consistent with the idea they are consulting memorized multiplication tables.

Given these results, it may be easy to feel the pull of dichotomous thinking. For instance, one might want to conclude that when people are doing multiplication, their sense of magnitude (the IPS) is shut down. One might even go further to make the reckless conclusion that the proper method of multiplication instruction is to emphasize the memorization of verbal math facts without regard for a sense of magnitude. We take a closer look at IPS activation to see why these are mistaken conclusions.

Cochon, Cohen, van de Moortele, & Dehaene (1999) compared brain activation during multiplication with activation when staring at a small cross on the screen. Figure 5.3 indicates the IPS is very active for multiplication compared to doing a non-mathematical task. When interpreting this new result, one might now conclude that it is important for multiplication facts to be tightly connected with one's sense of magnitude. This is a very different conclusion from a dichotomous interpretation of the results in Figure 5.2.

In summary, cognitive processes are always "on" to some degree. It is a mistake to view them as dichotomous, where one process excludes another. It is tempting to do so, because dichotomous categories simplify the world. On closer inspection, however, exclusive categories often hide a deeper truth about cognition, much as the satyr's assumptions about hot and cold hid a deeper truth.

Nature Confers Cognitive Processes; Education Coordinates Them

A major goal of typical education is to coordinate evolutionarily conferred abilities into ensembles that can achieve culturally relevant goals. Whereas most everyone learns to speak and interpret language, education coordinates our evolutionarily bestowed linguistic capacity with the visual system so that people can also read language. Similarly, education coordinates the IPS, largely implicated in spatial attention, so it can contribute to mathematical thinking. Evolution bestowed humans with the ability to coordinate and re-coordinate cognitive and neural processes.

A nice example comes from a study by Mackey, Miller Singley, and Bunge (2013). The authors compared brain changes among students who did or did not take a course that provided training for the Law School Admission Test (LSAT) exam (the entrance test for law schools). The LSAT is rich in hypothetical thinking, which requires one to set aside the facts that one knows, and instead, draw logical conclusions based on the stated premises in the problem. That is why it is called "hypothetical reasoning." The effect of the LSAT training was to coordinate the prefrontal and parietal regions. One interpretation is that the prefrontal regions learned to suppress spontaneous memory intrusions from the parietal regions, so people would rely on the premises and logic rather than their memories. Learning to deactivate memory retrieval is useful for doing the types of tasks that appear in the LSAT, a cultural invention.

The reader may have entertained the analogy that learning is like strengthening a muscle. A better analogy would be learning to dance. Dancing requires the coordination of many muscles, as well as the strengthening of the muscles in response to one another. Strengthening without coordination is ineffective. Woltz, Gardner, and Bell (2000), for instance, found that if people already know how to do one set of computation steps very well, they may display more errors when performing a new, related computation compared to a person who has less initial experience. Even though the seasoned subjects had strengthened some relevant computation "cognitive muscles," the coordination was wrong.

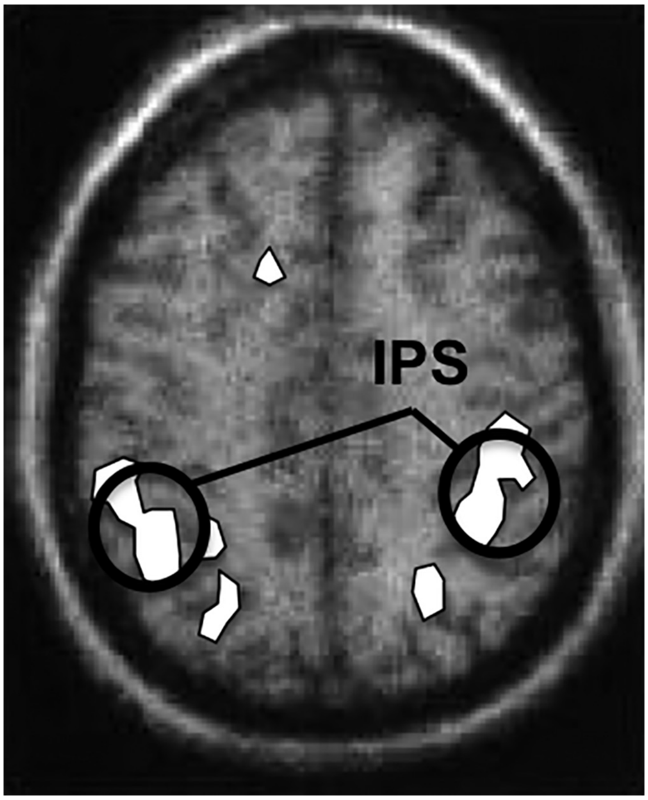

Figure 5.3 The intraparietal sulcus (IPS) is active for multiplication tasks relative to staring at a fixation cross. (Adapted from Cochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience, 11*(6), 617–630. Reprinted by permission of MIT Press Journals. Copyright 1999 Massachusetts Institute of Technology.)

Consider the case of learning to type. One approach might be to have people strike a key faster and faster when they see the relevant letter. For instance, one sees the letter “” appear on a screen and then types the letter “” as quickly as possible. It is not hard to imagine a fun little computer game that could train this kind of response. It fits the muscle analogy, where one emphasizes the strengthening of an isolated skill. Typing programs, however, do not take this approach. Instead of helping people learn how to type each letter as quickly as possible, these programs help people coordinate multiple keystrokes. The bottleneck in typing is how well people can coordinate their fingers to handle collections of letters (i.e., words). Moreover, people also need to coordinate the movements of their eyes with their hands, if they are typing from a document. They need to look ahead by just the right amount to anticipate how to coordinate their fingers for the transition from one word to the next. This fits the dance analogy, where one emphasizes the coordination of activity. *Education is more about teaching the brain to dance than teaching it to lift weights.*

Dichotomous thinking brings with it a focus on single cognitive processes, often to the exclusion of others. This can lead to tenacious misconceptions. One major misconception may be the belief in learning styles. The belief is that different people have different favored cognitive abilities, and therefore, instruction should match a person’s favored cognitive ability. To be sure, there are individual differences in some foundational capacities. For instance, some people are better at mentally manipulating spatial information than others (Hegarty & Waller, 2005), and there are researchers who work on strengthening these very specific skills (Feng, Spence, & Pratt, 2007). However, this does not support the claim that, therefore, people with high spatial ability should receive instruction spatially, which is the immediate implication of some of the research on learning styles. Despite a thriving belief in learning styles, their effects must be small, because there is surprisingly little evidence to support the idea that people with different native strengths should receive different types of instruction (Pashler, McDaniel, Rohrer, & Bjork, 2009). When people claim they are visual learners, they may be claiming that they can interpret spatial information more easily, or perhaps, they are saying that they do not like to read very much, which is a motivation issue. Regardless, when one thinks of learning the important content taught in schools, it often depends on the coordination of the linguistic, spatial, conceptual, attention, memory, and other systems.

Examples of Learning as Coordination

Learning to coordinate is foundational to the biology of the brain as it adapts to new information. At the cellular level, brain cells “learn” to coordinate their signals with one another. All learning requires coordination at the cellular level. The neurons need to communicate to accomplish work. Learning comprises an increase and decrease in the number and strength of connections among neurons, so they can coordinate their communication more effectively for specific tasks. Of course, knowing this fact does not get one very far in thinking about the macro-level of learning that teachers handle in classroom instruction. Therefore, in this section we provide some examples of coordination for the types of tasks and learning found in schools. The examples come from reading, mathematics, and conceptual change.

Learning to Read

A crisp example of the role of learning as coordination involves reading. By the time that children are learning to read, they have extensive vocabularies. They can detect words in sound, and they can use these sound-based words to retrieve their meanings from memory. Arrow 1 in Figure 5.4 indicates this coordination of hearing and memory. With reading, children now have the challenge of hooking up their visual system to their auditory system, as indicated by arrow 2. They need to learn that the look of a set of letters (a written word) corresponds to a sound. Establishing this coordination takes time, because the children need to learn how to see and hear the letters. Over time and with many hundreds of hours of practice, people begin to establish coordination between sight and meaning. They develop a link directly between the look of a word and its meaning, as indicated by arrow 3.

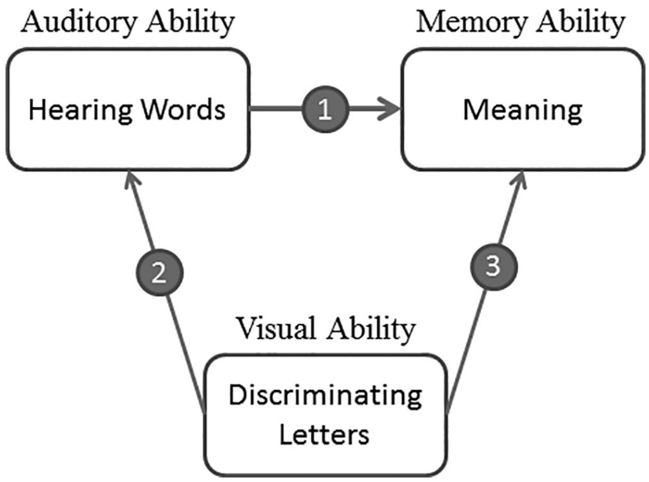

Figure 5.4 Circuits that enable reading. Evolutionarily conferred language circuits map between word sounds and meaning. It requires explicit education to coordinate the activity of the visual system so that reading language also becomes possible.

The link between vision and word meaning becomes automatic with practice. When seeing a word, it is hard to ignore its meaning, as shown by the Stroop task (Stroop, 1935) in Figure 5.5. Moreover, people can read without having to sound out words, which enables them to read much faster. An interesting fact is that once sight and meaning have been coordinated, people do not lose the coordination between sight and sound. For instance, if you run into a word that you do not know immediately, you may notice that you subvocalize that word—you are sounding it out in your head in the hopes that arrow 1 will help find the meaning, because you cannot find the direct link from sight to meaning. It is informative to note that population variability in the ability to speak language is much lower than the variability found in reading. This is because reading depends on the special coordinating arrangements of culture and school, whereas speaking and understanding oral language are conferred by nature.

Approximate Addition

The significance of well-coordinated processes also appears in mathematics tasks. Tsang, Dougherty, Deutsch, Wandell, and Ben-Shachar (2009) investigated children's abilities to do approximate addition. In approximate addition, people receive an addition problem and have to choose which of two answers is closer without computing the answer exactly. Given 27 + 14, is 40 or 60 closer to the answer? The task is an experimental version of the standard "estimate the answer" assignment in school. Tsang had children complete numerous problems and found that there were reliable individual differences in children's performance.

Figure 5.5 The Stroop effect. The task is to say the color of the word (black), but people automatically read and retrieve the meaning of the word (white), which slows down their time to complete the task of saying black.

The researchers then took measures of the brain's white matter using MRI. The white matter consists of fibers or tracts that connect regions of gray matter that reside on the surface of the brain. The gray matter is responsible for different types of computations, whereas the white matter helps distal brain regions communicate. Figure 5.6 shows the brains of two children and the white-matter tract of interest (anterior superior longitudinal fasciculus: aSLF) for the approximate addition task. Children who had a more coherent tract connecting the two areas of the brain were also the ones who did better on the approximate addition task. The implication is that they were better able to coordinate the computations between different brain regions.

At this fine level of granularity, the coordination of different processes appears as biological, and one can ask whether and what types of educational experiences might improve the structure of these specific biological pathways. The researchers did not address this question. A likely hypothesis is that the children need to engage in tasks that co-activate and force the coordination of the two areas of gray matter to drive changes in the connective white matter (see Scholz, Klein, Behrens, & Johansen-Berg, 2009).

Conceptual Change in Mathematics

Conceptual change refers to major shifts in how people think of a situation or problem (see also Chapter 18, this volume). For instance, young children change from a conception of a flat earth to a round one (Vosniadou & Brewer, 1992). Conceptual change in mathematics provides a strong example of learning as coordination. To an adult, the digit "5" coordinates multiple quantitative meanings seamlessly. For instance, 5 can refer to cardinality—five total things. It can refer to ordinality—fifth in a series. It can also refer to magnitude—5 is bigger than 3. Infants, and many animals, have innate abilities for each of these separate meanings of number. They can differentiate between two and three objects at a rapid glance; they can tell whether something comes before or after something else; and they can judge larger and smaller. The task of instruction is to coordinate these different abilities to make an integrated concept of number. For instance, Griffin, Case, and Siegler (1994) created a kindergarten curriculum that involved board games where students had to translate between the different meanings. They might roll a die and count the total number of dots (cardinality). They would then move their character on the game board the same number of spaces forward, thereby translating between ordinality and cardinality. They might then have to decide who has more total spaces so far, translating between

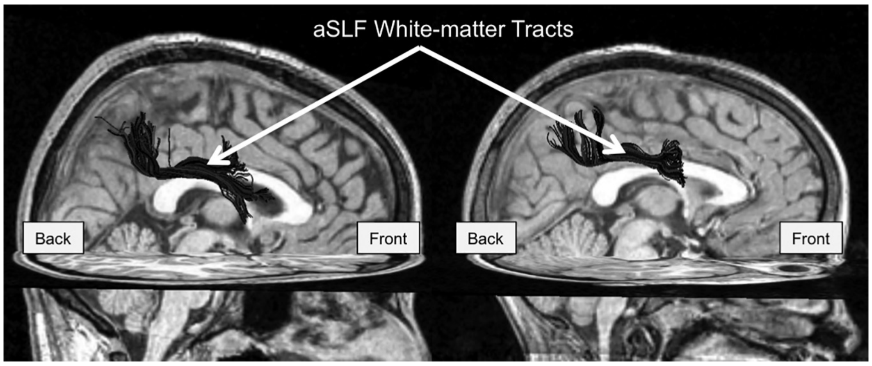

Figure 5.6 White-matter tracts connect distant surfaces of the brain. These two brains show differences in the anterior superior longitudinal fasciculus (aSLF) white-matter tracts that connect two regions of the brain that coordinate to complete approximate addition tasks. (For visual clarity, the aSLF tracts are shown in black and the many other white-matter tracts of the brain have been removed from the image.) (Courtesy of Dr. Jessica Tsang, based on data collected in Tsang et al., 2009.)

magnitude and ordinality. These researchers found that children who played the coordinating games did better in first grade the following year compared to children who played games that tried to improve each sense of quantity independently (e.g., just counting dots to find cardinality without translating the results into ordinal position or to make a magnitude comparison).

When people learn fundamentally new concepts, they need to re-coordinate the relations between evolutionarily old neural circuits. Dehaene and Cohen (2007) proposed that people "exapt" neural circuits for cultural purposes through a process of cortical recycling. Exapt means that a structure originally evolved to serve one function is borrowed to serve another. For example, the visual circuits responsible for fine discrimination of natural phenomena become repurposed to identify symbolic letters. Blair, Tsang, and Schwartz (2013) looked for evidence of borrowing primitive perceptual computations in the context of a mathematical conceptual change; namely, learning the integers.

The integers introduce the negative numbers and zero to the natural numbers. The understanding of negative numbers is unlikely to have been conferred by nature, given that they are a recent invention (Varma & Schwartz, 2011). The integers also depend on the introduction of new mathematical structure in the form of the additive inverse: $X + -X = 0$. The authors asked what innate abilities were exapted to handle the additional structure of the negative numbers. Adults had to decide the mid-point of two digits, for instance, 2 and 10 (answer: 6), –6 and 2 (answer: –2). As the digits became more symmetric about zero, people answered more quickly. For instance, people could solve –5 and 7 faster than –3 and 9. This was true even if people heard the digits rather than seeing them on a screen. Interestingly, they also found that brain regions associated with detecting visual symmetry (e.g., visual area V5) became more active for the more conceptually symmetric problems. Based on this evidence, it appears that people exapt their abilities to detect symmetry to help make sense of the integers, which can be conceptualized as symmetric about zero.

The authors went a step further to determine if this finding had implications for instruction. They created a curriculum for fourth graders that emphasized symmetry (Figure 5.7) so that students could coordinate their innate abilities with symmetric structures to understand the negative numbers. They found that this curriculum led to superior abilities to solve novel integer problems compared to current instructional models, which do not help students coordinate their knowledge of natural numbers and symmetry to build an understanding of integers.

Memory and Understanding

We now turn to the discussion of dichotomies that may be familiar to the reader. The poles of these dichotomies reflect important cognitive processes and outcomes. The risk is that people treat the poles as mutually exclusive and argue for one over the other. We begin with the distinction between memory and understanding. These processes need each other. For instance, a common technique in classrooms around

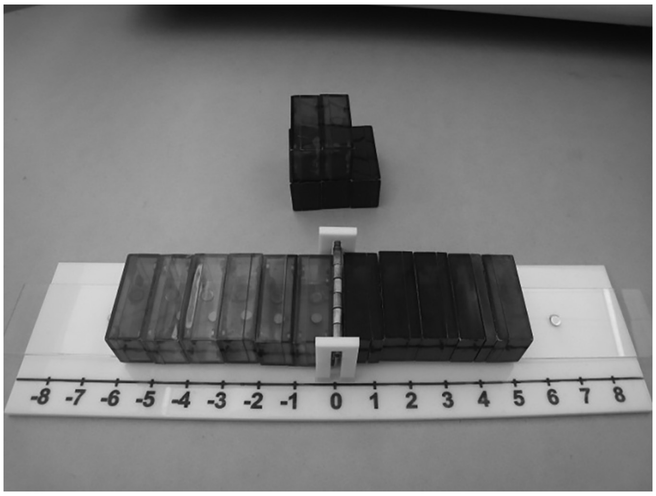

Figure 5.7 Hands-on materials created to emphasize the symmetry of positive and negative materials. Children received integer addition problems (e.g., 5 + –6) that they modeled by setting out positive and negative blocks about the zero point. To find the answer, they clapped the blocks together, folding up from the zero point. The number of extra blocks on either side gives the answer.

the world is to have students activate their prior knowledge before a lesson. "Do you remember hugging your dog? Did you notice the warmth? That is because a dog is a mammal, and mammals are warm-blooded." Activating prior knowledge is an example of retrieving memories to help one understand new ideas better. In this case, understanding depends on memory. A second common instructional technique is to ask people to make sentences out of new words. Making a meaningful sentence with a new word will help people remember the word. In this case, memory depends on understanding. Despite the obvious interdependence of memory and understanding, they are often placed into the following exaggerated opposition:

Rote memorization ↔ Deep understanding

People need to memorize important recurrent facts. Knowing the fact families in math is a great asset for solving problems that depend on factoring. Remembering is faster than problem solving, and for many problems, speed matters. Being able to remember an answer also frees up cognitive resources useful for understanding broader aspects of a problem. Similarly, people need understanding. If one truly understands, then one can recreate what may be forgotten. There are important differences between memorization and understanding, but as fits our argument, they work better in coordination than in isolation. We begin with a brief review of the memory literature, and then the literature on understanding. We then consider why the coordination of memory and understanding is important for the transfer of learning from one setting to another.

Memory

Memory is one of the most intensely studied and theorized domains within cognitive psychology. How can people gain memories without limit, yet still remember the right memory at the right time and at blazing speeds? For example, here is a random word—"peach." You probably recalled the right fruit in about 0.6–0.75 seconds. Given how many memories you have about so many different things, it is a stunning achievement.

Humans have many distinct memory systems, each specializing in a different type of information. At an extreme, one can consider the immune system to be a type of memory. When people receive bone marrow transplants, doctors kill the existing marrow and then replace it. As a result, the immune system "forgets" all the diseases it has encountered and it needs to relearn. For cognitive phenomena, there are multiple memory systems, and recent evidence suggests that each requires separate sleep cycles to help consolidate the memories of the day (Stickgold, 2005). For instance, given a typing lesson, people will type faster after sleeping on the lesson than they did at the end of the typing lesson. If people's sleep is interrupted during the specific cycle associated with this form of procedural memory, they will not perform better in the morning.

Gaining a memory depends on two processes. One is encoding the memory, or "getting it in there." The other is retrieval, or "getting it back out." We consider each briefly.

Encoding involves laying down the initial trace of a memory. Ideally, the way one encodes a memory will improve the chances of remembering it later, and this is an important emphasis of good instruction. There are a number of study techniques for improving memory. One class of strategies relies on the meaning of what one is trying to learn. For instance, connecting a new idea to a pre-existing idea improves encoding. If you are trying to learn a new phone number, it helps to find familiar mathematical patterns. Given 422-8888, one might improve the encoding of the phone number by thinking, "4 divided by 2 makes 2, and adding them up makes 8 of which there are 4 again." This works much better than just repeating the phone number, which is a recipe for forgetting as soon as one stops repeating the digits. In general, the depth of processing (Craik and Lockhart, 1972) and the relevance of elaboration (Stein & Bransford, 1979) predict the success of memory encoding. The more you think about a new idea and relate it to other ideas in meaningful ways, the better the chances of remembering it. It is as if you are laying down lots of neural roads, so it is easier to get back to the idea from other ideas. A second class of general encoding strategy—spaced practice—works regardless of the content of what one is learning (Cepeda et al., 2009). If one plans to work on memorizing words for a total of 10 minutes, it is better to use five separate sessions of 2 minutes each rather than one big session of 10 minutes. Cramming for a test is a bad way to create memories for a lifetime.

The second process of memory is retrieval, which involves bringing the memory back out. Retrieving a memory increases the chances of being able to retrieve it again later. A seminal demonstration comes from the "generation" effect (Slamecka & Graf, 1978). People received word pairs in one of two conditions. In the read condition, the words were presented completely, for instance, FAST : RAPID. In the generate condition, the words were presented as FAST: R_P_D. People knew the words had to be synonyms, and they could easily generate the missing letters to generate "rapid." People read or generated very many words. A short time later, they recalled as many of the words as possible. The generate condition remembered more of the words. One possible explanation is that the generate task required working a little harder to remember the word "rapid" during the task, which made subsequent retrieval a little easier. The importance of retrieval practice has resurfaced recently as the testing effect (Karpicke & Blunt, 2011). Taking a test, which requires retrieving memories, improves the chances of retrieving those memories later, for example, on a future test. Of course, the implication is not necessarily that students should take repeated tests, but rather, they should practice remembering what they know. If one wants to learn using flashcards, it is better to try to remember what is on the other side of the flashcard than just turning it over to see the answer.

With practice, the neural coordination of memories changes (McClelland, McNaughton, & O'Reilly, 1995). An example comes from children solving simple mental addition and subtraction problems. Behaviorally, children are very accurate at all ages, but they answer more quickly as they develop more experience with math facts. Figure 5.8 shows changes in brain activity with development (Rivera, Reiss, Eckert, & Menon, 2005). The bottom panel highlights areas that decrease activity. As children gain experience, they do not rely on the prefrontal areas of the brain as much. Among other things, the prefrontal area is responsible for the deliberate control of processing. With experience, the children do not need to do as much deliberate control to help them put their memories together to come up with the answer. The bottom panel shows areas of the brain that become more active for the arithmetic tasks as children gain more experience. With experience, these parietal areas become responsible for holding the relevant memories, and children can access them directly with little deliberate effort. Tasks that once required flexible but cognitively costly executive control come to be executed by quickly retrieving stored memories (see also Chapter 19, this volume).

At the behavioral level, a clear case of memory transformation involves skill acquisition (Anderson, 1982). Acquiring skills involves a transition from declarative to procedural memory. Declarative memory refers to things you can say, and procedural memory refers to things that you do. Imagine that you are learning to change lanes while driving. At first, you followed declarative instructions—"check your blind spot, turn on your blinker, check your blind spot, turn the wheel . . . " With practice, you no longer needed to

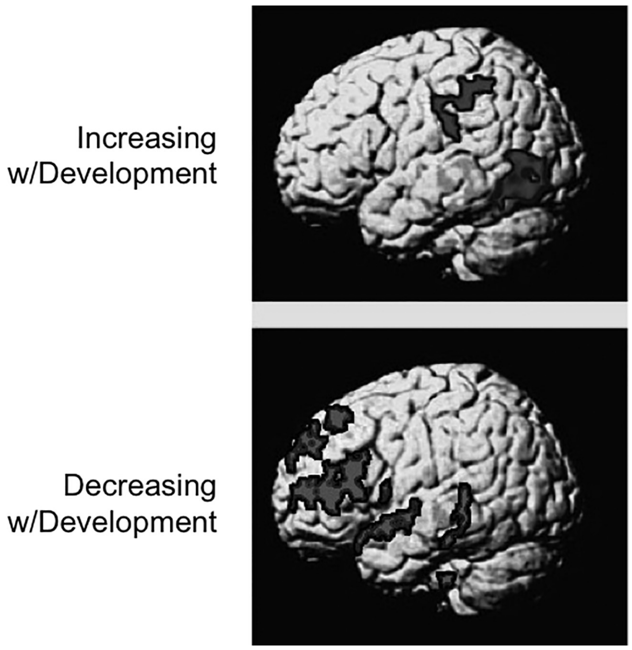

Figure 5.8 Changes in how the brain solves mental addition problems as children develop. The top panel shows that children increasingly rely on parietal regions to solve addition problems, and the bottom panel shows that they decreasingly rely on executive control from the prefrontal cortex to accomplish the tasks. (Adapted from Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, 15(11), 1779–1790. By permission of Oxford University Press.)

rely on these verbal memories. Instead, you developed procedural memory. You can tell because you do not need to talk to guide yourself through the steps. Instead, you can just execute them. After even more practice, these skills become automated. They require very little cognitive control or attention to execute. For instance, you can change lanes while also talking to your passenger. Because you do not need to pay attention to the skill execution, you can pay attention to talking. The way this works is that all the steps become "chunked" together so that one quickly leads to the next, and it requires little cognitive control to coordinate the transition from one step to another. They become one big step.

The transition from declarative to procedural memory has been an important guide for the design of many curricula. One of the most notable involves computerized "cognitive tutors" (Anderson, Corbett, Koedinger, & Pelletier, 1995). These intelligent computer programs track a student's progress. By monitoring how well the learner is performing on various tasks, it can infer whether the child has developed chunked procedural knowledge. If not, the program can back up to provide the student with relevant practice.

Understanding and Analogy

A common objection to memory-focused models of instruction is that students may learn to recall or execute a skill, but they may not understand it. For instance, students who memorize math steps may not really *understand* what those math steps mean. But, what counts as "understanding?"

The definition of "understanding" has been a subject of philosophical investigation since at least the time of Socrates. In cognitive psychology, different investigators choose different ways to operationalize understanding that are most relevant to their topic of study. (The term "operationalize" means that one indicates which measureable behaviors provide evidence for a given mental state or process.) For instance, a researcher who studies mathematics learning may operationalize understanding as the ability to verbally justify the generality of a particular claim (does the operation of addition always make a greater total quantity?). A person who studies language acquisition may operationalize understanding as the ability to identify the referent of a word (given the word dog, point to the right thing). Thus, there is no single definition of understanding. Nevertheless, it has been possible to make important empirical advances.

One major advance has been the distinction between surface features and deep structures. A deep structure is a set of necessary relations that characterize what is the same across many instances (e.g., mammal: warm-blooded, hair). A surface feature is a property that may or may not be important (e.g., red hair). A classic example comes from Chi, Feltovich, and Glaser (1981). They had novices and experts categorize physics problems. Experts grouped spring and inclined-plane problems together, whereas novices did not. The experts identified that the problems shared the deep structure of being about potential energy. To the novices, these situations seemed completely different, because one involved springs and one involved inclined planes, which do not look alike.

In mathematics and science education, helping students learn the deep structure is important. A phenomenon's deep structure is typically what verbal principles and formulas describe.

Oftentimes, people rely on surface features while learning, and this causes them to miss the deep structure. In a telling study, students learned the probability formulas for computing combinations and permutations (Ross, 1987). (As a reminder, imagine pulling two chips from a bag of red and blue chips. There are three possible combinations: two reds, two blues, or one red and one blue. Permutations further consider the possible orderings: red → red, blue → blue, red → blue, blue → red.) In the study, the students learned to compute the number of combinations using marbles as the example, and they learned permutations using cars as the example. On the posttest, students received combination and permutation problems. They used the combination formula for problems about marbles—regardless of whether the problem called for finding combinations or permutations of marbles. Similarly, if a problem involved cars, the students used the permutation formula whether it was appropriate or not. The students had memorized the formulas just fine. The problem was that they relied on the surface features of the problems (e.g., cars or marbles) to decide which formula to use. They did not learn to recognize the deep structure of combinations and permutations, which holds up regardless of cars or marbles.

Analogies capitalize on the distinction between deep structure and surface features. Consider the abbreviated test question:

Deluge is to Droplet as:

(a) Landslide is to Pebble
(b) Cloudburst is to Puddle

Many people choose (b) as the answer, because it shares the surface feature of being about water. Answer (a) can be construed as a better answer because it shares the same deep structure as the prompt, which might be summarized as "many harmless events can accumulate into a disaster." Being able to work with the deep structure of a situation is one useful operationalization of understanding.

Transfer and Induction—Where Understanding and Memory Work Together

Research on transfer highlights the importance of coordinating memory and understanding. Transfer refers to the use of prior learning in a new situation. Liberal education is predicated on the notion of transfer, because students learn in school, but they need to use this learning outside of school. In contrast, training-oriented instruction often does not need to consider issues of transfer. The application context typically shares the same surface and deep features as the original learning conditions. Training airline pilots in a simulator does not raise large transfer challenges, because the simulated cockpit is very similar to the cockpit of the plane; they share the same surface and deep features.

Transfer depends on the coordination of memory and understanding. For instance, in the preceding study, the students remembered the procedures for computing permutations and combinations. This was insufficient for effective transfer, however. They did not understand the deep structure of situations that call for the use of one or the other formula. Of course, had the students never memorized the procedures, they would not have had a formula to transfer either. How can we help students both memorize and understand?

One solution is to rely on inductive learning (Holland, Holyoak, Nisbett, & Thagard, 1986). Induction refers to the process by which people use multiple instances to create a new category or rule. (It contrasts with deduction, where people start with a rule or category, and determine what instances are possible.) Through induction, people may find the deep structure that unifies discrete memories and generalizes to new situations.

Discrete memories do not transfer well, because they typically apply to a single situation. For instance, memorizing 3 + 1 = 4 will not help solve 4 + 1. However, if people memorize that 3 + 1 = 4, 4 + 1 = 5, 5 + 1 = 6, and so forth, they might induce the rule that "any number plus 1 equals the next number in order." Induction is an important way that people generalize from the instances they have encountered and go beyond the information given (Bruner, 1957).

People are always inducing patterns from their memories and experience. However, they may not induce what we consider most important. For instance, given the series of +1 problems above, a student might correctly but inappropriately induce, "this teacher really likes to give problems with a 1 in them." Through education, we want people to induce particular patterns that generalize well, not idiosyncratic ones. Learning through analogy is a powerful way to help people induce targeted understanding from a set of instances.

A classic study on learning from analogy clarifies the role of induction in coordinating memory and understanding for transfer. Gick and Holyoak (1983) tried to determine what would help people solve Duncker's radiation problem, short of giving them the answer. Here is the problem:

> A patient has a tumor that needs to be irradiated. If the doctor uses a beam that is powerful enough to kill the tumor, it will kill healthy cells as it passes on the way to the tumor. If the doctor uses a radiation beam that is weak enough that it will not hurt healthy cells, then it will not kill the tumor. What can the doctor do?
>
> *The answer: The doctor can use multiple weak beams from different angles that simultaneously converge on the tumor.*

To see what would help people solve this problem, the researchers constructed several analogs to the radiation problem. For example, in one analog, a general wanted to attack a fortress and had to split up his troops to converge from different angles so they would not be too heavy for any one bridge. In another, firefighters needed to use multiple bucket brigades to douse a fire. In some cases, the researchers also described the general principle, "Split up forces to converge on a central target." Given these elements, the researchers tried different combinations to see which ones would support transfer to the radiation problem.

Table 5.1 The Effects of Induction and Explanation on Transfer Data from Gick and Holyoak (1983)

Percent who Solved the Radiation Problem	Read Principle	Did not Read Principle
Received no analog	28%	18%
One analog	32%	29%
Two analogs	62%	52%

College students were randomly assigned to one of several conditions. One factor was the number of analogies included in the packet: zero, one, or two of the analogs (fortress and fire problems). A second factor was whether or not the packet included a statement of the principle. On the last page of all the packets was the radiation problem. Table 5.1 shows the percentage of students who solved the radiation problem at the end of the packet. (Students who received neither the analogs nor the principle received filler materials in their packet and served as the control condition.)

The most notable findings involve the first column. Students who were told the correct principle without receiving any examples did not transfer very well. One interpretation of this result might be that the students did not understand the principle without an example. However, students who received one analog (an example) plus the statement of the principle did not do much better. Why would a single example with a principle be ineffective for transfer? The principle indicated the deep structure of the problem, and the students had an example to help make sense of the principle. It cannot simply be that the students did not know the deep structure.

The students who received two examples (analogs) did much better, with or without a statement of the principle. (The students who received the two examples without the principle were often able to induce the principle from the two analogs, so they did not need to read the principle.) One possible reason that the single analog and the principle did not work very well is that students did not learn the range of variation that might appear for this particular principle. For example, those students who only learned about the story of attacking the fortress from multiple bridges, even with the statement of the principle, had no way of knowing that it can apply to lots of situations. The surface features, while incidental to the deep structure, are still important for transfer. Remembering variability of the surface features allows people to appreciate that the deep structure they understand can apply to many situations. Thus, memory of several instances and understanding work together, and one without the other does not work very well for transfer.

The value of coordinating memory and understanding for transfer yields some simple instructional prescriptions. Loewenstein, Thompson, and Gentner (2003) describe a study where they found that asking business students to find the analogous structure between case studies led to superior learning compared to a condition where students handled each case separately without looking for the common structure. Providing two analogous examples works well for transfer, but students need encouragement to induce the common deep structure that unifies otherwise discrete examples (Schwartz, Chase, Oppezzo, & Chin, 2011).

Despite the simplicity of helping students induce the deep structure across instances, there is an instructional tendency to use on single examples plus a statement of the rule. Felder and Silverman (1988) noted that almost all engineering professors claim to use deductive instruction methods when teaching others—going from general rule statements to specific instantiations—even though they often themselves use inductive learning methods, proceeding from particulars to generalities. Additionally, people may neglect the potential of using analogies to support induction. In a review that compared instruction cross-nationally, Richland, Zur, and Holyoak (2007) found that U.S. teachers tended not to capitalize on the use of analogy compared to teachers in Hong Kong and Japan. Perhaps, by understanding how memory and understanding coordinate, educators will take more advantage of analogical induction.

Concrete and Abstract

A long-standing distinction, at least since the time of Plato, is the dichotomy between concrete and abstract mental representations. The idea is that concrete thinking is tied to the perceptual-motor particulars of a situation, whereas abstract operations rise above immediate experience to use logical relations and hypothetical thinking. The dichotomy is relevant to instruction, because people often favor one over the other mode of operation. For example, in California, the state's science curriculum commission proposed legislation that would limit hands-on learning to "no more than 20 to 25 percent" of instructional time. This resulted in an outcry from educators and business people, and the final legislation reversed the proposal to "at least 20 to 25 percent" of science instruction using hands-on material (emphasis added, 2004, www.cascience.org/csta/leg_criteria.asp).

A similar distinction between perception and abstraction occurs in intuitive frameworks for thinking about the brain: The brain is an information-processing system that takes in information from the senses and transforms it in different ways along a series of processing stages. The early stages of information processing are "low-level." For vision, these early stages would include extracting edges from a scene, creating contours for objects, determining stable colors not influenced by lighting conditions, and segregating objects from their backgrounds. "High-level" stages are further "downstream" in the flow of information processing, executed only after a considerable amount of sensory and perceptual processing has completed. High-level processes would involve cognitive actions such as inferring likely career choices of a friend, deciding where to search for an answer, and creating a new diagram for representing relations among cognitive functions.

The intuitive separation of low-level and high-level operations fuels many of the dichotomies found in education, such as Bloom's taxonomy. One key dichotomy might be represented as follows:

Perception ↔ Conception

As with all dichotomies, there is some truth to it. The brain is not a homogeneous lump. It is organized into spatially distinct modules with specialized functions. Some of the best-articulated modules are those dealing with perception, and these perceptual modules are often the first to be activated in response to a stimulus. In contrast, concepts do not need to be stimulus-driven. One can bring to mind the concept of "dog" without seeing or hearing a dog. Moreover, concepts have abstract and logical relations to other concepts, such as "not a cat."

Despite the important distinctions between concepts and percepts, they are not dichotomous. Gross anatomical considerations indicate a high degree of coordination between low-level perceptual processes and high-level conceptual processes. First, there are strong down-stream connections from perception to conception, so that stimulation of the perceptual system gives rise to relevant concepts. Second, there are also reciprocal up-stream connections from conceptual levels of analysis to the perceptual system (Lamme & Roelfsema, 2000). The tight integration is manifest when considering the time course of identifying an object such as a dog's tail. As stimulus information arrives from the senses, populations of neurons complete low-level analysis, such as edge detection, to extract important perceptual features. At 120 milliseconds, these neural populations operate similarly, regardless of whether one is explicitly paying attention to the object, and whether one has knowledge that there is a dog attached to the object. In this time window, the neurons are engaged in pre-attentive, unconscious perceptual processing. Yet, at 160 milliseconds, the same neural populations will respond to high-level information about where to pay attention and the knowledge of the context of the object (it is a dog). It is this latter activity that people consciously experience—"it is the tail of a dog" (Fahrenfort, Scholte, & Lamme, 2007). Thus, if one wants to know whether a neuron is "strictly perceptual," the answer depends on when one asks the question as much as it depends on which neuron is involved.

These broad considerations of neural architecture echo in the functional behavior of people. People adjust the processing of lower-level regions so that they are better adapted to the needs of higher-level cognition. Our perceptual systems are surprisingly adaptive, even late in life, and they typically adapt so that we can perform high-demand tasks more effectively (Fahle & Poggio, 2002).

The renowned philosopher Quine (1977) argued that advanced scientific thought must dispense with notions of perceptual similarity as the basis for its categories (see also Chapter 1, this volume). The argument seems plausible at first sight, and it has some similarity to the idea that understanding depends on finding deep relational structures rather than relying on surface features. Our perceptual systems might mislead us into believing that samples of fool's gold (pyrite) are true gold. Better to rely on the periodic table and the physical chemistry of the elements. There is a good deal of appeal to this argument. However, if one used the possibility of error as a reason to discard perception from scientific thinking, one would also have to throw out conception, because people's concepts are frequently wrong as well (e.g., McCloskey & Kohl, 1983). *It is the possibility of error that creates the possibility of learning and discovery.*

Fittingly, people can learn to perceive. Learning is not confined to abstract matters such as $F = ma$. Perception can be educated and augmented so it complements conceptual thinking. People can look for subtle properties that distinguish fool's gold from the real thing. People can supplement their biological perceptual apparatus with tools such as microscopes, hardness scales, and quantitative measurements of malleability.

Education and experience change the processing of the perceptual system (Goldstone, 1998). Some of these changes can occur in perceptual areas relatively early in the brain's information-processing stream. For instance, consider expert perception. Researchers measured the electrophysiological activity of dog and bird experts while they looked at pictures of dogs and birds (Gauthier, Tarr, & Bubb, 2010). For dog experts, enhanced electrical activity occurred 164 milliseconds after the presentation of dog, but not for a bird. Reciprocally, bird experts showed quick activation for bird pictures but not for photographs of dogs. This is an impressively fast processing effect of expertise given that transmitting a simple electrical signal from one end of a neuron to the other requires about 10 milliseconds. For brain evidence of experience-driven changes to perception, Furmanski, Schluppeck, and Engel (2004) used fMRI to measure brain activity before and after 1 month of practice with detecting hard-to-see lines. Practice increased responses in the primary visual cortex (area V1) and the degree of change correlated with detection performance. Bao, Yang, Rios, and Engel (2010) found changes in electrical activity as fast as 50–70 milliseconds after stimulus onset. Perception, even at the first stages of information uptake, can be educated. In fact, there is evidence that auditory training can produce differential responses in sensory receptors, such as the cochlea (Puel, Bonfils, & Pujol 1988), a sensory organ just inside of the eardrum. Perceptual changes are found at many different neural loci and a general rule seems to be that brain regions associated with early perceptual analysis are implicated in finer, more detailed, and generally less transferable knowledge (Ahissar & Hochstein, 1997).

Even learning abstract topics such as algebra can be improved by harnessing perceptual learning in instruction. A nice example comes from Kellman, Massey, and Son's (2010) perceptual learning modules. The algebra learning modules have high feedback and minimal explicit instruction. They try to develop students' sensitivity at noticing preserved structures in equations across algebraic transformations. For example, students are given trials on which they must determine that $6y - 17 = 32 - 5x$ is a valid transformation of $6y + 5x - 17 = 32$, but that neither $6y - 17 = 32 + 5x$ nor $6y - 17 = 32 - x - 5$ are. Although this kind of training might seem like "mere symbol pushing," the argument from perceptual learning is that by training students to see contrasts between valid and invalid algebraic transformations, they come to naturally perceive or induce the underlying structure of algebra.

Lawrence Barsalou has presented a particularly influential account of the grounding of conception in perception in the form of perceptual symbols theory (Barsalou, 1999).

By this account, conceptual knowledge involves activating brain areas dedicated for perceptual processing. When a concept is brought to mind, sensorimotor areas of the brain are reactivated. Even abstract concepts, such as truth and negation, may be grounded in complex perceptual simulations of combined physical and introspective events. Interestingly, Barsalou's research shows that when people engage in perceptual simulations, their understandings of a concept are likely to be richer and more flexible compared to when they do not. Reasoning based on perception is the "smart," not "stupid" stuff. This result is echoed by studies showing that students who show greater mathematical competence are more, not less, likely to engage in perceptual solutions to algebraic tasks (Goldstone, Landy, & Son, 2010). For example, students who exhibit relatively good mastery of mathematics are more likely to solve problems such as $x - 2 = 7$, by imagining the 2 moving from the left side of the equation to the right side, turning into a +2 as it does so. Rather than viewing perceptual processes as antagonistic to proper formal thought, it is precisely by properly executing these perceptual processes that formally sanctioned reasoning is achieved effectively.

A major challenge of school-based instruction is helping students coordinate the abstract, symbolic representations of culture with the perceptual world of experience. Glenberg, Gutierrez, Japuntich, and Kaschak (2004), for instance, noted that young readers often do not construct a mental model of what they are reading, but instead, they are just saying the words aloud. To help, the researchers had young children manipulate figurines to correspond with each sentence they read (e.g., "the man went into the barn"). They then told the children they should do this in their head when reading. This improved reading comprehension later, even when the children no longer manipulated figurines physically. An important challenge for an educationally relevant cognitive psychology is to develop new theories and evidence that helps guide fresh instructional efforts to coordinate perception, action, and conception (Goldin-Meadow & Beilock, 2010). Simply juxtaposing a concrete and abstract representation may not be sufficient for people to learn to coordinate their perceptual-motor abilities with their symbolic ones.

Dichotomania

We have sampled a pair of familiar dichotomies. There are others. For instance, a common dichotomy is the distinction between passive versus active learning, which appears in the college instruction literature (e.g., Prince, 2004). Passive learning largely refers to sitting in a large lecture listening to a professor's exposition, whereas active learning refers to being engaged in problem solving during class. Dichotomies with family resemblances include learning by doing versus being told, as well as discovery learning versus direct instruction. The intuition that students can learn more effectively when they are experientially engaged, or at least not being crushed by tedious exposition, is worthwhile. At the same time, a tremendous amount of learning occurs through reading and hearing explanations, for example through mass media, the internet, and books. Experience and explanation each has its place. So again, the task is how to coordinate these different types of learning. Experiential activities can provide direct engagement of a phenomenon or problem, whereas lectures and readings can provide explanations of those experiences in ways that students are unlikely to discover on their own. On this model, one way to coordinate active and passive learning is to use active experience to create a time for telling (Schwartz & Bransford, 1998). For instance, Arena and Schwartz (2014) had students play a modified version of the arcade game Space Invaders that prepared them to then learn a formal treatment on statistical distributions. By itself, the game showed little direct benefit for learning, but when combined with a formal exposition, students learned more from the exposition than otherwise equivalent students who had not played the game.

What can we do about all these dichotomies? It may be useful to notice that many of the dichotomies make one of the poles of the dichotomy something construable as "true understanding." Rote memorization was contrasted with conceptual understanding. Attention to surface features was contrasted with attention to deep principles. Low-level perception was contrasted with high-level abstract reasoning. It is a recurrent motif to contrast the upper reaches of human thought with the lower capabilities shared by animals.

Overcoming dichotomania requires a more humble mindset. First, as we have proposed, what separates humans from animals is the ability to coordinate cognitive processes in concert with cultural demands and opportunities. Through this process of coordination, both the "bottom" and the "top" of cognition refashion one another. Humans are adaptive, and this should be the emphasis of our thinking about learning, not how one type of thinking is superior to another.

Second, it is important to appreciate that even true understandings are always partial and fragmentary. A noteworthy attitude shared by many accomplished scholars is their insistence on how much they, and we, do not yet understand. The dichotomous endpoint of "true understanding" is illusory, and a realization of this may yield a less disparaging attitude toward the purported opposite pole. True and complete understanding certainly has its attractions over more brittle and biologically constrained forms of intelligence, but the latter have the distinct advantage of actually existing.

As part of a more humble attitude towards posing dichotomies, one also needs an attitude towards becoming more knowledgeable. The acceptance of dichotomies presupposes fixed poles, when they may not be fixed but rather grow with respect to one another. Creating dichotomous categories may be an important first step in making intellectual advances; it is native to human thought (Smith & Sera, 1992). Nevertheless, one should avoid becoming a satyr and running away from opportunities to grow beyond the opposition.

In an analogous case, Carol Dweck (2012) has observed that people differ in their implicit views—their mindsets—about the origins of human ability. Some people with a "fixed" theory believe that ability is largely innate. In contrast, those with a "growth" theory believe that ability results from hard work. A "fixed" mindset has an analogous structure to dichotomania. A fixed mindset presupposes there are poles of "smart" and "not-so-smart" people, and there is no path from one to another. For dichotomania, one may feel that there are mutually exclusive cognitive processes, some being better than others, and with no bridge between.

Just as people who adopt a "growth" theory are more likely to achieve actual success, so our understanding of learning may be more successful if we adopt a growth theory. Such a perspective does not focus on the wide gap between the endpoints of putative dichotomies, but rather considers how different processes can be placed into productive relations. The point is to reflect not only on our lofty positions as intelligences capable of infinite flexibility, but also on how we can get to that point using finite means. By this account, properly harnessed and coordinated memory, perception, action, habit formation, and attention processes can grow into a well-organized system that we take as showing improved educational outcomes.

Acknowledgment

The writing of this chapter was supported by the National Science Foundation under grant nos. SMA-0835854, EHR-0910218, as well as the Institute of Education Sciences, U.S, Department of Education R305A1100060. 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 granting agencies.

References

Ahissar, M.& Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*, 401–406.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). New York: Worth.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of Learning Sciences*, *4*, 167–207.

Arena, D. A., & Schwartz, D. L. (2014). Experience and explanation: Using videogames to prepare students for formal instruction in statistics. *Journal of Science Education and Technology, 23,* 538–548.

Bao, M., Yang, L., Rios, C., and Engel, S. A. (2010). Perceptual learning increases the strength of the earliest signals in visual cortex. *Journal of Neuroscience, 30*, 15080–15084.

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., Bransford, J. D., & CTGV. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *Journal of the Learning Sciences, 7*, 271–312.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Blair, K. P., Tsang, J. M., & Schwartz, D. L. (2013). The bundling hypothesis: How perception and culture give rise to abstract mathematical concepts in individuals. In S. Vosniadou (Ed.), *International handbook of research on conceptual change II* (pp. 322–340). New York: Taylor & Francis.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York: David McKay.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22.

Bransford, J. D., Brown, A. L., & Cocking, R .R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior, 11,* 717–726.

Bruner, J. S. (1957). Going beyond the information given. In H. Gruber, G. Terrell, & M. Wertheimer (Eds.), *Contemporary approaches to cognition* (pp. 258–290). Cambridge, MA: Harvard University Press.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice. *Experimental Psychology*, *56*(4), 236–246.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Cochon, F., Cohen, L, van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience, 11*(6), 617–630.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3/4/5/6), 487–506.

Dweck, C. S. (2012). *Mindset: How you can fulfill your potential*. London: Constable & Robinson.

Fahle, M., & Poggio, T. A. (2002). *Perceptual learning*. Cambridge, MA: MIT Press.

Fahrenfort, J., Scholte, H., & Lamme, V. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19*(9), 1488–1497.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education*. Engineering Education, 78*, 674–681.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*(10), 850–855.

Furmanski, C. S., Schluppeck, D., and Engel, S. A. (2004). Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology, 14*, 573–578.

Gauthier, I. Tarr, M. J., and Bubb, D. (Eds.) (2010). *Perceptual expertise: Bridging brain and behavior.* Oxford: Oxford University Press.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.

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.

Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science, 5,* 664–674.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Interdisciplinary Reviews: Cognitive Science, 1,* 65–78.

Goldstone, R. L., Landy, D. H., & Son, J. Y. (2010). The education of perception. *Topics in Cognitive Science, 2,* 265–284.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25–49). Cambridge, MA: MIT Press.

Harnad, S. (1997). Categorical perception. Cambridge: Cambridge University Press.

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. *The Cambridge Handbook of Visuospatial Thinking* (pp. 121–169). New York: Cambridge University Press.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141–158.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Karpicke, J. D. & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborate studying with concept mapping. *Science, 331,* 772–775.

Kellman, P. J., Massey, C. M., & Son, J. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, *2*, 285–305.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Koedinger, K., Booth, J., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342* (6161), 935–937.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, *23*, 571–579.

Lee, K. (2000). Cortical areas differentially involved in multiplication and subtraction: A functional magnetic resonance imaging study and correlation with a case of selective acalculia. *Annals of Neurology*, *48*(4), 657–661.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education, 2*(2), 119–127.

Lupyan, G. (2008). From chair to "Chair": a representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General*, *137*, 348–369.

Mackey, A. P., Miller Singley, A. T., & Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience, 33*(11), 4796–4803.

Mayer, R. E. (1987). *Educational psychology: A cognitive approach*. Boston, MA: Little, Brown.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419.

McCloskey, M., & Kohl, D. (1983). Naive physics: the curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 146.

McDermott, R. P. (1996). The acquisition of a child by a learning disability. *Understanding Practice: Perspectives on Activity and Context*, 269.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51,* 121–147.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning (NCER 2007-2004).* Washington, DC: National Center for Education Research, Institute of Education Sciences, and U.S. Department of Education.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles. *Psychological Science in the Public Interest*, *9*, 105–119.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231.

Puel, J. L., Bonfils, P., & Pujol, R. (1988). Selective attention modifies the active micromechanical properties of the cochlea. *Brain Research,* 447, 380–383.

Quine, W. V. (1977). Natural kinds. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds* (pp. 155–175). Ithaca, NY: Cornell University Press.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, *316*, 1128–1129.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased specialization in the left inferior parietal cortex. *Cerebral Cortex*, *15*(5), 1779–1790.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 629–639.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (Eds.) (1986). *Parallel distributed computing: Explorations in the microstructure of cognition*. *Volume 1: Foundations*. Cambridge, MA: MIT Press.

Scholz, J., Klein, M. C., Behrens, T. E., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience*, *12*(11), 1370–1371.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition & Instruction*, *16*, 475–522.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759–775.

Skinner, B. F. (1986). Programmed instruction revisited. *Phi Delta Kappan, 68*, 103–110.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 592–604.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24*, 99–142.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.

Stein, B. S., & Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 769–777.

Stickgold, R. (2005). Sleep dependent memory consolidation. *Nature, 437*, 1272–1278.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662.

Tobias, S., & Duffy, T. (Eds.) (2009). *Constructivist instruction: Success or failure*. New York: Routledge, Taylor & Francis.

Tsang, J. M., Dougherty, R. F., Deutsch, G. K., Wandell, B. A., & Ben-Shachar, M. (2009). Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. *Proceedings of the National Academy of Science, 106*, 22546–22551.

Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. *Cerebral Cortex, 20*(11), 2636–2646.

Varma, S., & Schwartz, D. L. (2011). The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts. *Cognition*, *121,* 363–385.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24,* 535–585.

Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 601.