Chapter 11:

The Science of Learning:

Mechanisms and Principles

Stephen M. Kosslyn

Abstract

An enormous amount is now known about how humans process and store information. This chapter organizes that knowledge into a set of 16 principles, which we have used systematically to help students master the material they are taught. The principles are organized under two very general maxims: *Think it Through* (the more you think something through, paying attention to what you are doing, the more likely you are later to remember it) and *Make and Use Associations* (associations not only help us organize material so that it is easy to store in memory but also give us the hooks that will allow us later to dig the material out of memory, to recall it). The 16 principles can be invoked in different combinations by using different application techniques, such as mnemonics and explaining material to yourself.

The "science of learning" comprises findings in a wide range of areas, including the study of memory, perception, comprehension, learning, and reasoning. We've learned a tremendous amount about how humans process and store information, and that knowledge can be used systematically in education to help students master the material they are taught.

Oddly, although the science of learning matured decades ago, it is rarely used to facilitate teaching. Instead, most classes are taught using methods that were developed over a thousand years ago. Walk into any university and you are more likely than not to see a "sage on the stage": a faculty member at the front of the class, and rows of students dutifully putting in their time by sitting in class (some listening, some taking notes – but many doing email, monitoring Twitter, or surfing the web). To my knowledge, Minerva is the only institution to use the science of learning systematically in all aspects of the curriculum.

Lectures are a very common way of teaching, but we need to distinguish between teaching and learning. Teaching focuses on information *transmission*; learning is about knowledge *acquisition*. On the face of things, the two activities should be completely aligned. But, typically, they are not. Teaching is often done in a way that is convenient and efficient for the professor, with little thought to how best to facilitate student learning. Lectures are a superb way to teach: A single instructor can lecture to 10,000 people as easily as 10. But study after study has documented that lectures are a terrible way to acquire information (let alone to acquire deep knowledge, which requires not just learning information, but also gaining an understanding of its broader context and utility).

Consider, for example, a large review and analysis of 225 studies of how well students learn from lectures versus active-learning seminars (Freeman et al., 2014). This review was restricted to STEM (science, technology, engineering, and mathematics) courses, which presumably are among the most challenging offered to undergraduates. The results were dramatic. The authors report the following:

> "The studies analyzed here document that active learning leads to increases in examination performance that would raise average grades by a half a letter, and that failure rates under traditional lecturing increase by 55% over the rates observed under active learning. The analysis supports theory claiming that calls to increase the number of students receiving STEM degrees could be answered, at least in part, by abandoning traditional lecturing in favor of active learning…. Finally, the data suggest that STEM instructors may begin to question the continued use of traditional lecturing in everyday practice, especially in light of recent work indicating that active learning confers disproportionate benefits for STEM students from disadvantaged backgrounds and for female students in male-dominated fields..."

This is from just one meta-analysis, but the same conclusion is offered repeatedly—even in non-STEM courses. Clearly, active learning is better than passively listening to lectures.

So why are lectures still the dominant mode of teaching in most universities? Part of the problem may be that faculty don't understand enough about the science of learning in order to take advantage of it. This is not to suggest that other problems are not also prevalent (such as the economics of universities, incentive structures, institutional rigidity)--but certainly most faculty who care about being effective instructors would benefit from a more thorough understanding of the science of learning.

In this chapter, I provide a very brief overview of the key principles that I organized from the empirical literature. I will summarize 2 overarching principles (which I'll call "maxims") and then consider 16 specific principles that fall under them. At the outset, I must note that different reviewers have organized the literature differently, producing different numbers of principles. For example, Graesser, Halpern, & Hakel (2008) identify 25 principles whereas Willingham (2010) identifies only 9. The differences appear to arise primarily from what principles are considered "special cases" or variants of other principles. In what follows, I've chosen a level of granularity that easily maps into active learning exercises that can be used in the classroom (chapter 12).

Two Maxims

Before we start, I need to make several distinctions clear:

Important distinctions

First, as noted above, "learning" is the process of acquiring information, of picking it up and storing it mentally. In contrast, "memory" refers to stored information, and the term typically includes the processes of retaining the information and then subsequently digging it out of storage for use. Learning and memory are different sides of the same coin: If you don't acquire it, it can't later be used; and if you can't locate it or dig it out, it may as well not exist.

Second, it is useful to distinguish between two different sorts of memories: *Dynamic* memories exist only as long as they are actively maintained. In contrast, *structural* memories persist even when they aren't being actively considered. Here's a metaphor: Imagine that someone is trying to remember the shape of a four-sided geometric form. To do so, she walks along a path she creates on a lawn, in the shape of the form. She walks this path over and over. While she is walking, the representation of the shape is dynamic; it depends on her continued movement. If she stops walking, the representation is lost. But after a while, she wears a dirt path through the lawn. After this happens, it no longer matters whether she keeps walking. She can stop slogging along the path and the shape persists. At this point the representation has transitioned from being dynamic to being structural.

Third, it is useful to distinguish between two different types of learning. On the one hand, we learn *declarative* information—such as vocabulary words, addresses, concepts and theories. On the other hand, we learn *procedural* information—such as how to drive a car, how to negotiate, debate, and use the rules of grammar to speak a second language. The following principles typically should be applied in different orders when learning the different types of information. For example, having students master underlying principles may be more important when acquiring procedural knowledge than when acquiring declarative knowledge, and hence should be introduced early to help students form a "mental model." But in all cases, I claim, the underlying principles of learning are the same.

Fourth, the principles can be invoked by a large number of "application techniques." For example, reflecting on how you would explain something to yourself will help you learn. But this activity itself is not a separate principle. Rather, it is a way to draw on a set of the principles, bringing them to bear in learning. We will return to this point after we review the principles.

Purposes of the principles

The principles I describe below are intended to accomplish three aims:

First, many of them can lead students to learn even if the student does not intend to do so. This is a remarkable discovery: You often learn not through intention, but simply as a consequence of using information. Think about what you remember at the end of the day about what you did from the time you woke up that morning (e.g., conversations you had, details of a newspaper article you read, etc.). How much of that material did you consciously *try* to memorize? Very little, I would wager.

Second, by using the principles I summarize below repeatedly, you can transition from doing something consciously to doing it automatically. For example, consider what happened when you learned to drive a car. At first, it was extremely laborious. Your driving instructor told you what to do and you did your best to follow instructions. But with practice, you soon could do those things without consciously thinking about them.

And third, these principles can help learners apply what they learned to all relevant contexts. This is a challenge because people transfer what they learned in one context to a novel context only with effort. And the more dissimilar the new context is to the one in which the material was originally learned, the harder it is to make the shift. “Far transfer” occurs when one uses learned material in very novel contexts (which, on the surface, do not resemble the circumstances in which the material was learned) and uses it well after the material was learned (Barnett & Ceci, 2002).

Maxim I: Think It Through

The first maxim is *Think it Through*. The key idea is very simple: The more you think something through (“turn it over in your mind”), paying attention to what you are doing, the more likely you are later to remember it.

This maxim is at the core of how you can recall facts and figures from a newspaper article you read, even though you didn’t try to memorize it. You stored the material in memory simply because you paid attention and thought it through. *Incidental learning* is learning that occurs without consciously trying to acquire the knowledge; it occurs as a byproduct of cognitive processing that is used to understand or analyze.

As you go through the list of specific principles below, you may notice something that is conspicuously absent—mention of “motivation” as a principle. I’ve often encountered admonitions to "find out what the students are interested in, and play to those interests." But the evidence suggests that the key is to get the students engaged. They might get engaged because they are motivated to do so or simply because the situation requires it—as far as I can tell, the reasons why they are engaged will make little difference. The key is to lead them to perform the relevant processing and to pay attention while they are doing so.

This is the essence of the first maxim, *Think It Through*.

Maxim II: Make and Use Associations

The second maxim is *Make and Use Associations*. Associations not only help us organize material so that it is easy to store in memory but also give us the hooks that will allow us later to dig the material out of memory, to recall it.

A dramatic demonstration of the power of using associations to organize material was reported by Ericsson, Chase & Faloon (1980). They asked an undergraduate student to commit to coming into the lab at least 3 times per week, and he did this for about a year-and-a-half. At each session, the researchers simply read him a sequence of random digits, one digit per second, and asked him to repeat them back. They started with a single digit, which he correctly recalled. They then gave him two other randomly selected digits, which he recalled, and then three, and so on, increasing the size of each new list until he failed to recall the entire sequence (8 digits, on that first day). Each session began where the previous one had left off, with a new list of that length (with a new combination of random digits). Every set consisted of a new set of digits; he wasn't given practice learning the same set over and over. When the study finally ended, this participant could recall a list of 79 random digits!

How did he do this? As it happened, the participant in the study was a long-distance runner who had run numerous marathons. He associated the random digits with times for particular segments of races. For example, if he heard "3, 4, 9, 2" he might associate these digits with the time, "3 minutes, 49.2 seconds." Thus, four digits were converted to a single "chunk" (i.e., an organized unit) using associations. He eventually devised other strategies for making such associations, such as relating digits to specific people's ages or specific notable dates.

Associations are not only important for helping us enter new information into memory, so that it is stored effectively, but also for helping us later to retrieve this information. Associations can serve as cues and reminders. The game of Charades illustrates this process in slow motion. For example, say that the presenter gets down on all fours and moves around like an animal. The players might shout "a cat," "a dog," "a donkey." The position is a cue, which activates these concepts. The presenter then sticks her fingers up from her forehead, mimicking horns. Someone says "a deer," someone else says "a goat." The presenter then stands up and mimics having a cape with an animal running by—leading the viewers to shout out "a bull"! Each cue evokes specific associations, which in turn retrieve certain information from memory. This process of being cued to recall specific information happens all of the time, every day, virtually every time we recall something.

Sixteen Specific Principles

Now to the specific principles that underlie these two overarching maxims. I used three criteria to select and formulate these principles: First, the principle could not be explained by appeal to other principles; it has to tap a distinct type of process. Second, the principle had to have been derived from highly replicable studies that demonstrated large effects on learning. Third, the principle had to have straightforward implications for instruction; it had to be clear how to implement it in practical situations. In what follows, I provide very brief descriptions of each principle.

Principles that underlie Think It Through

First, six principles fall under the umbrella of the maxim *Think It Through*.

Deep processing. The more mental operations one performs while paying attention to such operations, the more likely it is that one will later recall that information (Craik & Lockhart, 1972; Craik et al., 2006). This is the most obvious implication of the maxim *Think It Through*. For example, if you formulate an example of how every one of these principles can be used in a specific situation, you will remember them much better than if you simply read and understand them.

Desirable difficulty. We can think of this as the Goldilocks Rule (not too hot, not too cold—just right!). Learning is best when the task is not so easy as to be boring but not so hard as to be over the learner's head (Bjork, 1988, 1999; VanLehn et al., 2007). To get the most out of thinking it through, the person needs to be as engaged as possible—no more, no less. For example, if you are good at math, you will need more challenging examples of new concepts to stay engaged than would someone who has less knowledge.

Generation effect. Simply recalling information—especially when effort is required—strengthens memory for that piece of information; the mere act of digging information out of memory reconstructs and strengthens the mental representation of the information. For example, a consequence of this principle is that frequent testing can enhance learning if it leads learners to recall relevant information (Butler & Roediger, in press; Roediger & Karpicke, 2006).

These first three principles all focus strongly on the fact that more processing of the relevant information will produce better memory. The next three principles focus on ways to induce people to engage in additional processing. (Note: My interrupting here and pointing this out should allow you to create two large groups for this set of principles: The first three and the second three – which respects the fact that we can easily store no more than four units in a "chunk," as discussed below.)

Interleaving. Instead of just focusing on one type of problem (e.g., in math), it's best to intermix different types of problems. The same principle implies (but to my knowledge has not yet been investigated) that when learning French, it's best to do a bit of studying French, then some history, then some math, and then back to French. This makes sense because it's easier to pay attention to something new than to sustain paying attention to the same material, extended over time. For example, you would probably learn this material more effectively if you do something else after you finish this section, and return to the second set of principles later.

Dual codes. If I give you a short paragraph to remember, you will recall it better if I also included some illustrations. In general, presenting both verbal and visual material enhances memory. In this case, the brain stores multiple representations in memory (some verbal, some visual—which are stored in different parts of the brain), which gives you multiple shots at later digging the information out of storage (Kosslyn, 1994; Mayer, 2001; Moreno & Valdez, 2005). For example, if you are only given a name or verbal description to remember, your memory will be vastly improved if you can visualize (i.e., form a mental image of) the named object or scene: Not only will you create a second type of representation (in addition to the verbal material itself), but also the mere effort of visualizing the described object or scene will enhance subsequent memory.

Emotion. Leading someone to feel emotion when experiencing an event generally will enable him or her to recall that event more effectively. Emotion focuses attention and also causes the brain to devote extra resources to storing the information. Negative emotions, in particular, narrow attention and focus one on details. For such negative emotions, beta-blockers will remove this extra boost, which provides hints about the underlying pharmacological events that produce this extra processing (Erk et al., 2003; Levine & Pizarro, 2004; McGaugh, 2003, 2004). For example, if you are anxious about how an interview will go (and didn't take beta blockers!), you probably will remember more details about the interview than if you are not anxious.

To summarize, we've just reviewed six principles, all of which are special cases of the maxim "Think It Through."

Principles that underlie Make and Use Associations

The second overarching maxim is *Make and Use Associations.* It is useful to distinguish two general classes of these principles.

Structure information by using associations

The first class is "Structure information by using associations." Six principles fall in this category:

Chunking. As we saw in the case of the marathon runner who could memorize staggering numbers of randomly selected digits (Ericsson et al., 1980), you can use associations you already have in your memory in order to organize material into relatively few chunks (organized units). People can easily store in memory 3 or 4 organized chunks – and, remarkably, each of these units can contain 3 or 4 chunks. For example, if you want to learn a list of 16 principles, figure out ways to organize them into 4 or more groups. Organizing material into manageable units clearly facilitates learning (e.g., Brown, Roediger & McDaniel, 2014; Mayer & Moreno, 2003).

Build on prior associations. When learning something new, the more associations you can find to information already stored in memory, the better (e.g., Bransford, Brown & Cocking, 2000; Glenberg & Robertson, 1999; Mayer, 2001). For example, when meeting a new person, you can remember his name by associating his face with someone else you already know who has the same name. One way to do this is to visualize the face of the person you already know and then morph that mental image into the face of the new person (Kosslyn, 1994). If you do this a few times, you will associate the new person's face with that of the familiar person. And the familiar person's face is already associated with the appropriate name.

The fact that prior associations can be used to learn new information resolves an old conundrum: At one time researchers worried about a "paradox of the expert," which hinged on the fact that the more you know, the easier it is to learn even more (Reder & Anderson, 1980; Smith, Adams & Schorr, 1978). The intuition was that the more you know, the "fuller" memory should be—and hence it should be harder, not easier, to store new information. However, researchers have learned that the more information you already know, the more existing associations you can use to store new information. The more branches you have, the more leaves and fruit can be hung on this structure. Hence, there's no actual paradox.

Foundational learning. When acquiring complex information, learning is enhanced when a teacher takes advantage of existing associations to provide the most basic material first, and then to integrate new material a bit at a time (Bransford et al., 2000; Wandersee, Mintzes & Novak, 1994). Presenting foundational material first provides a "backbone" for additional information, allowing an organized mental structure to be built up over time. For example, presenting the two general "Maxims" first should have given you a structure for understanding the specific principles.

Deliberate practice. Building up a structured representation is more efficient if you receive feedback along the way, so that you can correct aspects of the representation when it isn't optimal (Brown, Roediger & McDaniel, 2014; Ericsson, Krampe &

Tesch-Romer, 1993). For example, when learning French, it's good to have a native speaker listen to you and carefully correct your pronunciation. Such feedback is most effective when learners use "deliberate practice": Deliberate practice occurs when you pay careful attention to mistakes and use the ways that an error differs from the correct performance to correct subsequent performance. Note, however, that this principle alone is not enough to make you into an expert (Hambrick et al., 2014).

The following two principles build on the previous ones, but focus specifically on the relationship between examples and underlying principles. (Note again: By interrupting here and pointing this out, I'm hoping to help you create two large chunks – which should help you get your mental arms around this material.)

Appropriate examples. Abstract ideas cannot be fully understood without examples. But examples must be memorable, in part by being associated with prior information. Multiple examples of the same concept, method or theory must be associated with each other, so that they form a cluster that is associated with the to-be-learned material. For example, when learning the concept of "far transfer," it's not enough for me to provide the example that debate techniques learned in class should then be used months later when arguing politics with friends. You would need a few different examples of "far transfer," and I would need to make sure that you have associated them with each other—in spite of the fact that on the surface they will appear very different (Hakel & Halpern, 2005).

Principles, not rote. Learning a concept, method, or theory requires not just learning examples of its applications, but also the underlying principles that organize and integrate examples (Kozma & Russell, 1997; Bransford et al., 2000). For example, the key to far transfer is to distinguish between surface characteristics (the particular example) and underlying deep characteristics (which tell you which knowledge should be transferred to the present case). For instance, the principles of debate can also be used in teaching, but that doesn't require becoming confrontational (a surface characteristic of debate)—but rather being sensitive to the other person's goals and perspectives (a deep characteristic). The principles must be associated with the examples. In general, focusing on how information relates abstractly (at what is called a "deep structural level" in the literature) to other information enhances memory (Chi & VanLehn, 2012).

Create rich retrieval cues

The other class of principles that falls under the maxim *Make and Use Associations* is "Create rich retrieval cues." The key idea here is that you need to associate distinctive information with what you learn so that you later can be effectively reminded of it when you want to recall it.

Dynamic representations arise from recent experiences or thoughts, and often are conscious. Thus, they are easy to recall. Structural representations, in contrast, are like the crates and boxes in that giant warehouse in the last scene of *Raiders of the Lost Ark*. We can retain uncounted numbers of such representations, and they often aren't well organized. We access these representations by using cues and reminders. For example, you might associate one of those crates with a coffin, and hence seeing that shape would remind you of it (and you then could search for other such shapes, if the initial one turned out to be incorrect). Thus, to be easily recalled later, it's crucial that structural representations include characteristics that makes them easily cued later, which can include being associated with a distinctive time and place. The following principles can produce such cues.

Associative chaining (aka story telling). Stories are built on a series of interlocking causes and effects—this is the essence of a plot. Creating an interlocking sequence of associations that have a narrative arc (i.e., a story) to integrate material will not only help you create larger chunks (stories are one way to build associations to create chunks), but—more than that—you also can use each part of the story to cue the next part when you later recall the material. Such cueing can greatly facilitate later recall of the information incorporated into the story (Bower & Clark, 1969; Graesser, Olde & Klettke, 2002). For example, to learn the principles under the *Think It Through* maxim, you could create a story about a friend who uses each of these principles in an effort to learn the computer programming language Python, adopting a new principle when the previous one proves inadequate.

Spaced practice. Cramming may be an efficient way to study, but it's a bad way to learn. Here's an analogy: When I was young, I had a black wooden desk. I thought it would look much better if I painted it white. Being in a hurry, I ignored the advice to use several thin coats of paint rather than one thick coat. I poured on a single thick coat. At first, it seemed just fine; the black paint was covered up. But in practically no time, the paint began to chip—and soon the desk was a blotchy mess, much uglier than it was in the first place. Something similar happens with memory: Trying to store information in one fell swoop leaves it vulnerable to being lost. One reason for this is that if you cram you will have only one set of retrieval cues, the associations set up the one time you stored the information. If you instead spread studying out over time, you will associate the material with lots of different cues (such as cues in the room or rooms where you study, your feelings at the time, and thoughts you have while considering the information). It is much better to use information repeatedly over a relatively long span of time in the course of learning it (Brown, Roediger, & McDaniel, 2014; Cepeda et al., 2006, 2008; Cull, 2000). For example, when learning this material you might want to read it once, and then go back to it a few times to review.

Different contexts. Far transfer is the holy grail of learning. As noted earlier, far transfer occurs when one can retrieve information learned in one context (e.g., a classroom) and apply it in a very different context (e.g., to a seemingly unrelated problem in a work environment, years later). Far transfer appears in part to be possible because one has learned a group of varied examples and has a firm grasp of the principles that underlie the relevant concept, method or theory (Hakel & Halpern, 2005; Van Merrienboer et al., 2006). But it also depends critically on knowing when learned information is relevant. To facilitate this, one should associate the material with numerous different contexts. For example, studying in different places will enhance your ability later to use the information in different contexts.

Avoiding interference. Distinctive retrieval cues are crucial in part because they can help the learner avoid interference from other information (Adams, 1967; Anderson & Neely, 1996). Psychologists have documented two types of interference: *Proactive interference* occurs when material you learned previously interferes with learning new information. For example, if you learned Spanish, you might have a problem learning that "de" is pronounced "duh" in French, not "day" as it is in Spanish. *Retroactive interference* occurs when learning new material impairs your ability to recall previously learned material. In the language example, once you learn the French pronunciation, you might have difficulty recalling the Spanish one. Creating distinctive retrieval cues can help you avoid both types of interference (e.g., you could associate the French pronunciation with an image of a French person having difficulty understanding why a learner is having this problem, perhaps dismissively saying "duh," and a Spanish person taking a siesta in the middle of the "day").

Using the Principles

The principles just summarized are a "base set" of processes that underlie all learning. They are like letters in an alphabet; different combinations of the same principles are used in types of learning.

Many "application techniques" have been developed to evoke different combinations of processes to produce effective learning. For example, researchers have shown that people learn effectively by explaining things to themselves (e.g., Chi et al., 1994). Creating an explanation is a special case of the generation effect, and checking to ensure that it is correct hinges on deliberate practice. Thus, the method is effective not because it introduces a new kind of process, but because it effectively recruits combinations of specific underlying principles. Similarly, consider mnemonics, which can be a very effective way to learn. In fact, I relied on two of them in the above (the descriptions of learning new names and of avoiding interference both involved mnemonics). But there's nothing special here: Mnemonic techniques involve combinations of specific processes, such as deep processing and drawing on previous associations to form new ones.

The sixteen principles reviewed here underlie all forms of learning, ranging from learning a golf swing to learning copyright law to learning about the principles of learning. In many cases, you initially learn a set of rules or instructions that must be consciously mediated (via what Kahneman, 2011, calls "System 2"), and only after practice does the material become automatic (i.e., can be accomplished by what Kahneman calls "System 1").

We at Minerva designed our curriculum from scratch and could systematic and principled in doing so. We decided to take advantage of the science of learning, and designed (and are continuing to design) every one of our classes to rely on application techniques: Every class is built around active learning, and every one of our active learning exercises draws on combinations of the principles just described. If these principles are respected in how material is presented and used, students will learn without necessarily trying to learn.

Acknowledgements

Brianna Smrke helped me clarify earlier versions of many of the ideas presented here, Behnam Arzaghi did a great job catching errors, and Laurence Holt provided astute feedback that shaped the presentation. I also wish to thank Diane Halpern and Daniel Levitin for useful conversations on this and related topics.

References

Adams, J. A. (1967). *Human memory*. New York: McGraw-Hill.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer *Psychological bulletin, 128*, 612-637.

Butler, A. C., & Roediger H.L., III. (in press). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology.*

Bjork, R. A. (1988). Retrieval practice and maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.). *Practical aspects of memory: Current research and issues, Vol 1.* (pp. 396-401). NY: Wiley.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.). *Attention and performance XVII: Cognitive regulation of performance: interaction of theory and application*. (pp. 435-459). Cambridge, MA: MIT press.

Anderson, M.C. & Neely, J.H. (1996). Interference and inhibition in memory retrieval. E.L. Bjork & R.A. Bjork (Eds.), *Memory. Handbook of perception and cognition (2nd ed.)*. (pp. 237-313). San Diego, CA: Academic Press.

Bower, G.H., & Clark, M.C. (1969). Narrative stories as mediators for serial learning. P*sychonomic Science, 14*, 181-182.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn* (expanded ed.). Washington, D.C.: National Academy Press.

Brown, P. C., Roediger, H. L. III, and McDaniel, M. A. (2014). *Make it stick: The science of successful learning.* New York: Belknap Press.

Cepeda, N. J.; Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354-380.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*, 1095-1102.

Chi, M. T. H., de Leeuw, N., Chiu, M-H, & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439-477.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist, 47*, 177-188.

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

Craig, S. D., Sullins, J., Witherspoon, A., & Gholson, B. (2006). The deep-level reasoning effect: The role of dialogue and deep-level-reasoning questions during vicarious learning. *Cognition and Instruction, 24*, 565-591.

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology, 14,* 215-235.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*, 1181-1182.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100,* 363-406.

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *NeuroImage, 18*, 439–447.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111,* 8410-8415.

Glenberg, A.M., & Robertson, D.A. (1999). Indexical understanding of instructions. *Discourse Processes, 28*, 1-26.

Graesser, A.C., Halpern, D.F., & Hakel, M. (2008). *25 Principles of Learning*. Task Force on Lifelong Learning at Work and at Home, http://www.psyc.memphis.edu/learning/whatweknow/index.shtml (for a summary, see also Graesser, A.C. (2009). Inaugural Editorial for *Journal of Educational Psychology*, *Journal of Educational Psychology*, 2009, 1010, 259-261.)

Graesser, A. C., Olde, B., and Klettke, B. (2002). How does the mind construct and represent stories? In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative* *impact: Social and cognitive foundations.* (pp. 231-263). Mahwah NJ: Lawrence Erlbaum Associates.

Hakel, M., & Halpern, D. F. (2005). How far can transfer go? Making transfer happen across physical, temporal, and conceptual space. In J. Mestre (Ed.), *Transfer of learning: From a modern multidisciplinary perspective.* (pp.357-370). Greenwich, CT: Information Age Publishing.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34-45.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *43*, 949-968.

Levine, L.J., & Pizarro D.A. (2004) Emotion and memory research: A grumpy overview. *Social Cognition*, *22,* 530-554.

Mayer, R. E. (2001). *Multimedia learning.* NY: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43-52.

McGaugh, J.L. (2003). *Memory and emotion: The making of lasting memories.* New York: Columbia University Press.

McGaugh, J.L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience, 27,* 1-28.

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development, 53*, 35-45.

Reder, L.M. & Anderson, J.R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology, 12,* 447-472.

Roediger, H. L. III., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Psychological Science, 1,* 181-210.

Smith, E. E., Adams, N., & Schorr, D. (1978). Fact retrieval and the paradox of interference. *Cognitive Psychology, 10,* 438-464.

Van Merrienboer, J., Jeroen, J. G., Kester, L., & Pass, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20,* 343-352.

VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., & Rose, C.P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31,* 3-62.

Wandersee, J.H., Mintzes, J.J. & Novak, J.D. (1994). Research on alternative conceptions in science. In: D. L. Gabel (Ed.). *Handbook of research on science teaching and learning.* (pp. 177-210). New York: MacMillan.

Willingham, D. (2009). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom.* New York: Jossey-Bass.