Oxford Handbooks Online

Memory in Educational Settings

Elizabeth J. Marsh and Andrew C. Butler

The Oxford Handbook of Cognitive Psychology
Edited by Daniel Reisberg

Print Publication Date: Mar 2013 Subject: Psychology, Cognitive Psychology
Online Publication Date: Jun 2013 DOI: 10.1093/oxfordhb/9780195376746.013.0020

Abstract and Keywords

We describe three theoretical principles from cognitive science that have implications for educational practice: introducing desirable difficulties during learning, processing information to extract meaning, and the importance of a match between the processing that occurs during initial learning and the processing required by the criterial task. We use these principles to evaluate the effectiveness of three strategies typically used to guide the initial learning of material (advance organizers, underlining and highlighting, and note taking) and three strategies for poststudy processing (retrieval practice, feedback processing, and spacing of practice). Finally, we consider the issues of long-term retention, transfer, and the relativity of memory phenomena, all of which often constrain the applicability of basic research to educational settings.

Keywords: memory, education, learning, long-term retention, transfer-appropriate processing, desirable difficulties, levels of processing

Cognitive psychologists and educators share a common goal: Both want to understand how to promote long-term learning and memory. Both are interested in the answers to questions like “How should people study material in order to remember it after time has passed?” and “What causes people to forget material they once knew?” However, cognitive psychologists tend to answer these questions using convenience samples (i.e., college undergraduates), relatively short delays between study and test, laboratory experiments, and simple stimuli such as word lists. But the educator is interested in specific student populations, very long-term memory, applications in the classroom, and complex educational materials such as lectures and textbooks. These differences in approaches have limited conversation between cognitive psychologists and educators, which is unfortunate as there is great potential for both groups to benefit from interacting. This chapter reflects a recent interest in the field to shrink this divide and summarizes a growing body of work aimed at connecting basic cognitive research to educational situations.

We begin this chapter with some guiding theoretical principles and then turn to specific applications. We are interested in general cognitive principles that can be applied to education, regardless of the subject matter or particular classroom. To this end, we will describe some basic research from cognitive psychology that is important for understanding the theory, and we will consider the implications for educational practice. We will then discuss some specific strategies (including ones implemented during the initial learning of material as well as some implemented after material has been studied) and provide an example of how such strategies can be combined to optimize learning. Next we will explore some issues with applying these guiding principles and learning strategies in educational settings. Finally, we will close by considering the relativity of memory and pointing out some future directions for research.

(p. 300) Guiding Theoretical Principles

Introducing "Desirable Difficulties" During Learning

In education, techniques that speed up the process of learning are commonplace—from language-learning programs like Rosetta Stone (a commercial product) to speed reading strategies like skimming. The logic behind such techniques seems reasonable: Why would one want to spend a longer period of time mastering material when it is possible to learn it faster? The problem with this idea is the implicit assumption that the level of performance attained during learning will be maintained over the long run. However, this assumption is false—learning material to some criterion does not guarantee that it will be remembered in the future. Performance during learning is a poor predictor of future performance because it reflects the momentary accessibility of knowledge (i.e., retrieval strength) rather than how well it has been stored in memory (i.e., storage strength) (see Bjork & Bjork, 1992).

To understand the distinction between retrieval strength and storage strength, try answering the following question: What did you eat for dinner last night? You probably retrieved the answer very quickly and easily because retrieval strength is high for that particular memory (since it is very recent and you have not eaten any other dinners in the interim). However, a month from now you would be unlikely to remember what you ate last night because retrieval strength dissipates over time. Instead, remembering your dinner a month from now would depend upon the storage strength of the memory of that particular dinner. Storage strength increases when a memory is retrieved or the event is reexperienced. If last night's dinner is like most dinners, the memory will not accumulate much storage strength because you are unlikely to think about it or tell someone else about it (i.e., retrieve or reexperience it). Returning to education, if knowledge is to be retained over long periods of time, then the goal of learning must be to increase storage strength, not momentary accessibility. A student being able to quickly retrieve an answer at one time does not guarantee long-term learning.

Based on the distinction between retrieval strength and storage strength, R. A. Bjork and colleagues developed the concept of "desirable difficulties" in learning (Bjork, 1994a, 1994b; Christina & Bjork, 1991; Schmidt & Bjork, 1992). The main idea is that introducing difficulties during learning will result in superior long-term retention because the greatest gains in storage strength occur when retrieval strength is low. For example, consider the practice of using flash cards to study vocabulary words. If you study a word and then try to remember it immediately, then the gain in storage strength will be relatively low because it is so easy to retrieve the word right away (retrieval strength is high). However, if you wait 5 minutes before attempting to retrieve the word (when retrieval strength will be lower), then the gain in storage strength will be larger. The implication for educational practice is that instead of arranging the conditions of learning to be easier and faster for the learner, educators should introduce difficulties into the learning process in order to promote long-term retention of knowledge.

In short, the theory of desirable difficulties encourages us to think about introducing conditions that will slow the learner down, and make learning a bit harder—with benefits for long-term learning. As will be developed later in the chapter, however, this theory is not to be interpreted as "make learning impossible"—the learner must be able to overcome the difficulty for it to be a desirable one.

Processing Information to Extract Meaning

A second relevant theoretical idea involves how people simultaneously process information on many different levels. At the most basic level, incoming information is processed by the nervous system in order to organize and understand sensory input. At higher levels, the information is processed with respect to existing knowledge in order to extract meaning. According to the levels-of-processing framework (Craik & Lockhart, 1972; Craik & Tulving, 1975), much of this processing occurs automatically, but attention can be directed toward any given level. Critically, directing attention at a lower or "shallow" level of processing (e.g., focusing on the orthography of words while reading) disrupts higher or "deeper" levels of processing (e.g., determining what those words mean). As a result, the type of processing in which one consciously engages determines what information will be encoded into memory and retained. The type of processing is more important than the intent to learn (e.g., Craik & Tulving, 1975); the implication is that a student who deliberately prepares for a test but who does not engage in deep processing will not do as well as the student who processes the material deeply, even if the latter student is not deliberately trying to learn the material.

(p. 301) For the purposes of the present chapter, we are more concerned with types of processing in which meaning is extracted from information, since most educational tasks emphasize meaning over perceptual information. With that in mind, it is important to point out that there are many different ways in which information can be processed at this "deeper" level (e.g., Packman & Battig, 1978). One helpful distinction involves *item-specific processing* versus *relational processing* (Hunt & Einstein, 1981). Item-specific processing involves encoding the various characteristics or properties of a particular piece of information. For example, judging the pleasantness of a word, filling in missing letters in a text, and creating a mental image of each step in a science experiment all focus the learner on a single to-be-remembered item.

In contrast, relational processing refers to the encoding of similarities and differences across pieces of information. For example, sorting words into categories, ordering sentences to create a coherent text, and explaining why each subsequent step in a science experiment follows the preceding step all involve comparing to-be-remembered events to each other. In short, both item-specific and relational processing can involve meaning extraction, but they direct the learner to different aspects of the to-be-remembered events.

To summarize, it is key to consider what type of processing different learning strategies encourage. For most educational tasks, students will benefit from strategies that encourage them to extract the meaning of to-be-remembered information. However, as will be described in the next section, it is also important to think about what type of processing will be required later on to succeed on the criterial task.

Importance of Match Between Processing at Encoding and Retrieval

Although processing information at "deeper" levels involving meaning extraction generally results in superior memory performance (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975), the way in which information is processed is not the sole determinate of what will be remembered. Rather, memory performance is the joint product of the way in which the memory was encoded (i.e., the memory trace that is stored) and the way in which it is retrieved (i.e., the cues provided) (e.g., Tulving & Pearlstone, 1966; Tulving & Osler, 1968). This idea is codified in the theory of transfer-appropriate processing, which states that memory performance will be enhanced to the extent that the processes engaged during initial learning match the processes required for the criterial task (Morris, Bransford, & Franks, 1977; for a review see Roediger, Weldon, & Challis, 1989). Thus, "goodness" of encoding is important for establishing the potential for good memory performance, but the match between processing at encoding and retrieval is critical to realizing that potential (e.g., Fisher & Craik, 1977; Moscovitch & Craik, 1976).

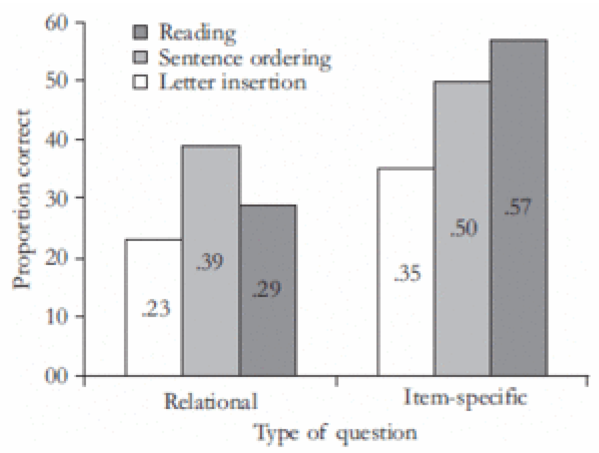


Figure 20.1 Proportion correct on the final test as a function of encoding task (reading, sentence ordering, letter insertion) and question type (relational vs. item-specific). (Data from McDaniel, Hines, Waddill, & Einstein, 1994, Experiment 1.)

To illustrate the concept of transfer-appropriate processing, consider an experiment by McDaniel and colleagues (McDaniel, Hines, Waddill, & Einstein, 1994, Experiment 1). Subjects read a text that described baseball and performed one of three tasks: generating words in the text by inserting deleted letters (an item-specific processing condition), ordering scrambled sentences to form a coherent text (a relational processing condition), or simply reading the text (a control condition). Later, subjects took a short-answer test that probed item-specific information (e.g., the name of the baseball stadium) and relational information (e.g., the sequence of events during a major play in the game). Having inserted letters or ordered the sentences led to better performance on both types of final questions, relative to simply reading the text (see Fig. 20.1). However, more importantly, subjects who had inserted missing letters performed the best on the item-specific questions, whereas subjects who had ordered sentences performed the best on the relational questions. Memory performance on the final test was enhanced when there was a match between processing at encoding and retrieval.

(p. 302) In short, there is not always a single right answer about which study strategy will be best. Rather, as reflected in ideas about transfer-appropriate processing, the ideal study strategy depends upon what the student will need to do later.

Summary

We have presented three general cognitive principles that are critical for determining memory performance in educational settings: introducing desirable difficulties during learning, engaging in processes that emphasize meaning extraction, and matching learning processes to the processes needed to excel on the final criterial task (i.e., transfer-appropriate processing). We turn now to applying these theoretical ideas to specific educational activities.

Learning Strategies for Studying Material

Much research has focused on strategies for learning material, and to review all possible strategies is beyond the scope of this chapter. Instead, we focus on a few common strategies and apply the guiding theoretical principles described earlier to help us understand how each strategy can be used to optimize learning. We chose one strategy (advance organizers) that is typically implemented by the instructor or in course materials (e.g., textbooks) and two strategies (highlighting/underlining and note taking) that are normally under student control.

Receiving Advance Organizers

It is important that the student approach the learning situation with the information he or she needs to appropriately process the to-be-remembered information. Without the relevant background knowledge, the student will not be able to extract the desired meaning from a text or lecture (in other words, deep processing will be impossible). Thus, the first study strategy we will consider, advance organizers, involves giving readers information (normally in paragraph form) that helps them to orient to and understand the important information in a to-be-read text, as well as connect it to their prior knowledge. Critically, advance organizers are aimed at providing a larger conceptual framework for the to-be-learned material; they help the reader to understand the upcoming material but do not contain the exact same information. For example, in one of the earliest studies (Ausubel, 1960), students in the experimental group read an advance organizer explaining the relationship between metals and alloys before reading the target passage on the metallurgy of steel. In contrast, students in the control group also read a paragraph before reading the critical passage on steel, but they read about the history of steel-making methods (which would not help them understand the science of steel). Both groups read their introductory passages 48 hours before reading the target passage on steel; 3 days after that, they answered a series of multiple-choice questions about the steel passage. Subjects in the experimental group, who received the advance organizer, outperformed subjects in the control group (*M*s of 46% and 39% correct, respectively).

Numerous studies have replicated this benefit of advance organizers (see Corkill, 1992, and Mayer, 1979, for reviews). Advance organizers can help the student to be more efficient; for example, students who read an advance organizer did not need to listen to a lecture on 35 mm cameras more than once to reach the same level of learning as observed in control subjects who repeatedly listened to it (Bromage & Mayer, 1986; see also Mayer, 1983). However, advance organizers are not always helpful, and understanding when they do versus do not help the reader is informative about how they work. Advance organizers work by providing a framework or schema for integrating the incoming information—thus, if the framework is already activated, advance organizers will not provide any additional benefit. Consequently, advance organizers help students learn disorganized material but not organized material (Mayer, 1978). Advance organizers are also less effective when students are likely already familiar with their content. For example, Ausubel and Fitzgerald (1961) examined the usefulness of an advance organizer comparing Christianity and Buddhism, prior to reading a text about Buddhism. The advance organizer was helpful, but only for subjects who scored low on a test about Christianity. Students who knew a lot about Christianity presumably did not need that schema activated prior to passage reading. Overall, advance organizer studies often use math or science materials, as students are unlikely to be familiar with the overarching conceptual frameworks for such topics (Mayer, 1979).

Advance organizers work primarily by improving the encoding of the to-be-learned material. Standard manipulations known to increase encoding have their expected effects when applied to advance organizers. The benefits of advance organizers increase following deep processing (by requiring students to paraphrase the organizer; Dinnel & Glover, 1985) and with spacing (reading the organizer ahead of time is better than reading it immediately before (p. 303) the critical text; Glover, Bullock, & Dietzer, 1990). Because advance organizers are typically aimed at overarching conceptual frameworks, they help students to process the target relationally (according to the relevant schema) as opposed to focusing on individual details (item-specific processing). Consequently, advance organizers promote transfer of knowledge. For example, Mayer (1975, 1976) examined how advance organizers affected the ability to apply new knowledge about a computer language called FORTRAN. In these studies, the advance organizer explained how a computer works using concrete analogies, such as comparing a computer's output to a telephone note pad. Subjects in both the advance organizer and control groups learned computer commands such as READ, WRITE, GO TO, and IF, but only the advance organizer subjects had learned an overarching model of the computer before encountering the specific programming commands. When asked to recall what they had learned, subjects who had received an advance organizer were more likely to recall conceptual information and to make appropriate inferences (Mayer, 1976). Subjects who had received advance organizers were also better able to write and interpret programming statements than were control subjects (Mayer, 1975).

While typically treated as an encoding phenomenon, advance organizers can also influence retrieval. Although the benefits of advance organizers depend upon reading them before encoding the target text (readers do not benefit from reading an organizer after encoding; Mayer & Bromage, 1980), rereading the organizer before a *delayed* test can also help performance (Corkill, Glover, Bruning, & Krug, 1988). Rereading the organizer right before test likely reactivates the original encoding context, increasing the overlap between encoding and retrieval, as well as providing a spaced rehearsal of the crucial conceptual information.

In short, advance organizers can be connected to the theoretical ideas laid out earlier in this chapter. Most obviously, they encourage deep processing of unfamiliar material by giving students the knowledge they need to process the material meaningfully. Advance organizers are less useful when read immediately before the target text; it is a desirable difficulty to insert a delay between these two phases and require readers to retrieve the information from the advance organizers from long-term memory. Finally, advance organizers work by helping the learner to link to-be-remembered information to a larger conceptual framework, meaning the benefits are most visible on tests that tap relational information.

Highlighting and Underlining

Students often face a large amount of information and need to figure out how to best allocate their study time. As described in the last section, the teacher can help students to appropriately direct their attention with tools such as advance organizers. However, students need to be able to identify important concepts on their own, and one common technique is to highlight or underline information that needs to be remembered. The one study we found that directly compared highlighting and underlining found no difference between these conceptually similar strategies (Fowler & Barker, 1974), so we will discuss these techniques together.

Highlighting and underlining have potential as learning strategies, because it has been demonstrated that students are more likely to remember text that appears in a distinctive format. For example, students who read *Scientific American* articles with five statements underlined were more likely to later correctly answer questions about those statements than were students who read the same texts without any underlining (Cashen & Leicht, 1970). In such cases, underlining or highlighting presumably captures the reader's attention, increasing encoding of the distinctive text. Several studies support the idea that readers benefit from attentional cues in texts that include underlining, highlighting, capitalization, and other ways of making text distinctive (e.g., Lorch, Lorch, & Klusewitz, 1995).

However, a second step is involved when the student uses highlighting or underlining as a study strategy: He or she must decide which text to mark. No benefit has been observed in several studies that allowed students to underline as much or as little of texts as desired (e.g., Idstein & Jenkins, 1972; Stordahl & Christensen, 1956). Two issues may drive such results. First, even if marking text does not lead to an overall boost in recall, it may promote memory for the parts of the text that were highlighted—the issue is that this benefit is potentially at the expense of other nonhighlighted text (e.g., Fowler & Barker, 1974). The literature is unclear about the effects of marking part of a text on other nearby but unmarked sentences, and whether it should hurt or help memory for the unmarked sentences (e.g., Cashen & Leicht, 1970). Second, students often highlight too much of a text, and people are less likely to remember marked text if it is more prevalent (Lorch et al., 1995). In addition to reducing the distinctiveness of the marked text, over-highlighting and over-underlining may reduce deep processing. That is, it may take more cognitive effort to **(p. 304)** figure out which one or two points in a paragraph are most important, and thus selective highlighting may be better than highlighting the bulk of a text. Some data to support this idea comes from Fowler and Barker (1974), who examined the relationship between the amount of *Scientific American* texts that were highlighted and performance on later multiple-choice items. Students who highlighted more text performed worse on the later test (r = –.287), although the small sample size (n = 19 in the relevant condition) meant the result was not significant. Supporting the idea that less highlighting might be better, Rickards and August (1975) found that students asked to underline just one sentence per paragraph later recalled more of a science text than did a no-underlining control group.

In short, highlighting (or underlining) as typically implemented is not a strong study strategy, even though it is a popular one. One issue is that overuse of this technique may reduce deep processing. A better strategy is to select just one sentence to highlight per paragraph, as this likely increases both deep processing and relational processing as sentences are compared to one another. In addition, highlighting in this way represents a desirable difficulty because of the effort required to identify the key sentence in a given paragraph.

Note Taking

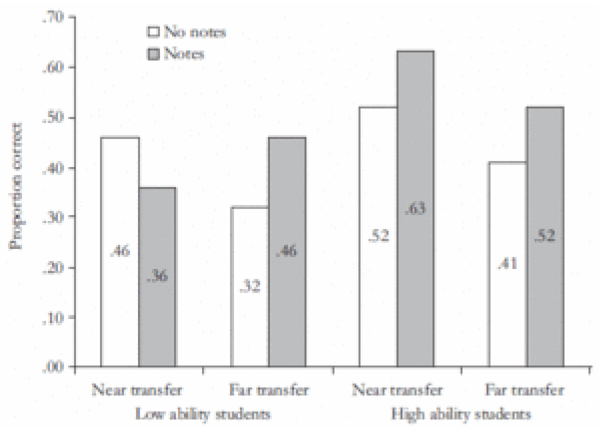
Note taking is also a common strategy that students report using when listening to lectures or while reading texts (e.g., Palmatier & Bennett, 1974). Note taking may affect memory through two different routes: one involving initial learning (encoding) and the other involving storage. Our focus will be on encoding; that is, how does note taking change the way the to-be-remembered information is processed, with consequences for later memory? However, this focus is not meant to minimize the importance of reviewing notes after encoding; an important consequence of note taking is that it yields external storage of the to-be-remembered material, which can then be used for poststudy review (see Kiewra, 1983, for an overview of the review function of note taking, and see the next section of this chapter for discussion of other poststudy activities).

Overall, research supports the common belief that taking notes promotes memory, with the benefits of note taking observed both in the classroom (e.g., Crawford, 1925a) and the lab (e.g., Crawford, 1925b), and with both text (e.g., Bretzing & Kulhavy, 1981) and lecture materials (e.g., Di Vesta & Gray, 1972). Some of the earliest evidence comes from Crawford (1925a), who examined the relationship between students' notes and performance on quizzes in seven college classes. Across studies, there was a median correlation of .50 between the number of correct points in a student's notes and his or her quiz score. In contrast, there was no relationship between the number of vague points noted and later quiz performance. This early study highlights two themes in the note-taking literature: first, the quantity of notes matters (with more complete notes generally yielding greater memorial benefits) and, second, the quality of the notes matters (with precise notes being better than vague ones).

Consistent with these two themes, instructor- provided support that improves the quality of notes benefits memory. For example, Moore (1968) held up red/green cards to cue students on whether to write down lecture points or not; students liked this technique and it improved later performance relative to a control group that took notes without cues. Similarly, students can benefit from receiving partial notes or handouts from the instructor to guide note taking (e.g., Austin, Lee, Thibeault, Carr, & Bailey, 2002). On the other hand, if the educational situation will not permit the student to take complete and/or quality notes, the student will not benefit from note taking. In fact, note taking may even hurt performance in such situations because students must pay attention to both the lecture and note taking, and dividing attention during encoding impairs memory (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). The student listening to a fast-paced lecture does not have the resources to take good notes and thus may be worse off if she tries to take notes while listening (e.g., Peters, 1972).

Understanding *why* note taking benefits memory requires an analysis of the processing involved. Two findings suggest that note taking may help students to extract the "big picture" of a text or lecture. The first piece of support comes from detailed analyses of the content of notes and later recall of the target material, to see whether listeners recalled all lecture propositions equally (Einstein, Morris, & Smith, 1985). In this study, an independent group of subjects rated the importance of 126 propositions in a lecture on the history of individual differences, and these ratings were used to classify the importance of what note takers recorded and recalled. Critically, note takers remembered more important propositions than did subjects who only listened to the lecture. It was *not* that note takers recalled more of everything. Instead, note takers only recalled more of (p. 305) the most important propositions, a finding Einstein et al. (1985) argued was consistent with organizational processing because determining importance requires comparisons across to-be-remembered items.

Even more critically, note taking promotes transfer (e.g., Peper & Mayer, 1986). For example, in one experiment (Peper & Mayer, 1978, Experiment 2) subjects learned about the chi-square statistic, either via a video lecture or via text. A later test consisted of both near- and far-transfer questions; near-transfer questions were straight-forward applications of the formula, whereas far-transfer questions went beyond the material covered in the experiment (e.g., requiring the subject to say the situation did not allow the use of the chi-square statistic). Note takers and control subjects performed similarly on the near-transfer problems, but note takers outperformed the controls on the far-transfer problems. This benefit of note taking occurred regardless of whether students heard the lecture or read the text, but it was particularly salient for low-ability subjects who had not scored as well on the SAT math test (see Fig. 20.2).



Click to view larger

Figure 20.2 Proportion correct on the final test as a function of note taking (notes vs. no notes), question type (near transfer vs. far transfer), and student ability (high vs. low). (Data from Peper & Mayer, 1978, Experiment 2.)

Although note taking yields memorial benefits, the type of notes naturally taken may not be the most powerful mnemonic possible. One issue involves how much note taking involves going beyond the to-be-remembered information (versus simply copying; see Marsh & Sink, 2010), including paraphrasing the to-be-remembered information and connecting it to stored knowledge. Returning to the levels of processing framework, the key issue is the depth of processing note taking naturally affords. We have just reviewed evidence that note taking encourages relational processing that affords transfer. Nevertheless, additional benefit may come from techniques that encourage the reader or listener to process the material even more deeply. For example, King (1992) trained students to summarize material, specifically how to identify and encapsulate the main idea. This group of students was compared to another group who took notes naturally, and who later had a chance to review those notes. Students who summarized the lecture performed better on both immediate and delayed comprehension tests than students who took notes (see Bretzing & Kulhavy, 1979, for similar results). Similarly, the note-taking group did not do as well as a group of students trained to ask themselves (and answer) questions about the material. Students might benefit from incorporating some of these deep processing techniques into their notes; in other words, training might help students to take notes that include more of the generative processing thought to be key for transfer (e.g., Peper & Mayer, 1978).

In summary, the literature on note taking is consistent with the theoretical ideas laid out at the beginning of this chapter. Note taking promotes transfer, suggesting that it does involve deep processing of the to-be-remembered material. However, it is important that the situation not be one that makes note taking too hard; the difficulty must be desirable and not impossible to overcome.

(p. 306) Learning Strategies for Poststudy

After the initial processing of material is complete, there are many ways in which that information can be further processed (such as reviewing one’s notes, as already discussed). One of the most popular strategies involves repeatedly reading the material (e.g., Karpicke, Butler, & Roediger, 2009; see too Kornell & Bjork, 2007). As the reader can likely surmise from the concept of desirable difficulties, this strategy tends to be relatively ineffective (see Callender & McDaniel, 2009). However, there are many strategies for reprocessing information that produce substantial mnemonic benefits. We will describe three such strategies: engaging in retrieval practice, processing feedback to correct errors, and spacing out practice over time.

Retrieval Practice

In most educational settings, testing is used as an assessment tool. As such, it serves two purposes: (1) evaluating student learning (i.e., summative assessment) and (2) providing feedback to guide future learning activities (i.e., formative assessment). Much has been written about each of these purposes (for an excellent review, see Black & Wiliam, 1998), and thus we will focus on a less appreciated reason for testing: the use of tests as a learning tool. In both educational and psychological theories of learning, testing is often assumed to be a neutral event, much like measuring someone's weight. Just as stepping on a scale does not change how much someone weighs, testing is assumed to measure the contents of memory but leave them unchanged. However, memory research has shown that retrieving information from memory actually changes memory (e.g., Bjork, 1975), improving long-term retention of the material (e.g. Butler & Roediger, 2007; Carrier & Pashler, 1992; Karpicke & Roediger, 2008).

The finding that retrieval practice produces superior long-term retention has been termed *the testing effect* (for review, see Roediger & Karpicke, 2006a). Memory researchers working at the intersection of psychology and education were the first to demonstrate the mnemonic benefits of testing (Abbott, 1909; Gates, 1917; Jones, 1923–1924; Spitzer, 1939). However, despite being over 100 years old, the testing effect remained relatively underappreciated until a recent resurgence of interest in the phenomenon (see Marsh, Roediger, Bjork, & Bjork, 2007; McDaniel, Roediger, & McDermott, 2007; Pashler, Rohrer, Cepeda, & Carpenter, 2007). Importantly, retrieval practice improves retention even in the absence of feedback or additional study opportunities (e.g., Glover, 1989). In addition, the benefits of retrieval practice cannot be explained by reexposure to the material, because students remember more after retrieval practice than if they restudy the original material for an equivalent amount of time (Roediger & Karpicke, 2006b).

Retrieval practice (i.e., testing) represents a strategy for reprocessing material that yields significant memorial benefits in educational settings (e.g., Carpenter, Pashler, & Cepeda, 2009; McDaniel, Agarwal, Huelser, McDermott, & Roediger, in press; McDaniel, Anderson, Derbish, & Morrisette, 2007; Metcalfe, Kornell, & Son, 2007; for review, see Bangert-Drowns, Kulik, & Kulik, 1991). For example, Larsen, Butler, and Roediger (2009) investigated whether retrieval practice enhances long-term retention of medical information that was taught during a weekly didactic conference for residents on a neurology rotation. A group of pediatric and emergency medicine residents participated in an interactive teaching session that covered two topics: (1) status epilepticus and (2) myasthenia gravis (two types of neurological disorders). For one of the topics, the residents took a series of three short-answer tests on the material with each test spaced about 2 weeks apart. For the other topic, the residents studied a review sheet at the same intervals. After 6 months, a final test was given for both topics. Performance on the final test showed that repeated testing on the material led to much better retention relative to repeated studying—and the difference in performance between the two groups was almost one standard deviation (effect size = .91).

When testing is used as a learning tool in educational settings, there are several ways in which its efficacy can be enhanced. First, tests that require students to produce a response, such as short-answer and essay tests, generally lead to better retention than tests that simply require the selection of the correct response, such as multiple-choice tests and true/false tests (e.g., Butler & Roediger, 2007; Kang, McDermott, & Roediger, 2007). One explanation for this finding, consistent with the idea of desirable difficulties, is that retrieval practice requires greater effort and therefore produces better retention (e.g., Carpenter & DeLosh, 2006; Pyc & Rawson, 2009). Second, taking multiple tests results in better retention than taking a single test (Bahrick & Hall, 2005; Wheeler & Roediger, 1992), so it is beneficial to repeat questions on quizzes and give cumulative exams. Third, successful retrieval is the key to learning from tests, so providing feedback after the test (p. 307) is essential, especially if test-takers do not retrieve many correct responses (e.g., Kang et al., 2007).

To summarize, retrieval practice represents an effective method for promoting long-term retention. This strategy constitutes a desirable difficulty because testing requires the learner to effortfully retrieve information from memory (particularly when a response must be produced by the learner) and may initially slow learning relative to comparable tasks like repeated study. Retrieval practice also provides the learner with an opportunity to reprocess the material in a way that results in the extraction of additional meaning (i.e., above and beyond that gleaned from initial study). We now turn to describing when and how to give feedback.

Processing Feedback to Correct Errors

Providing feedback after an attempt to retrieve information from memory is critical, because it helps to correct memory errors (e.g., Pashler, Cepeda, Wixted, & Rohrer, 2005) and maintain correct responses (e.g., Butler, Karpicke, & Roediger, 2008; Fazio, Huelser, Johnson, & Marsh, 2010). An extensive literature exists to support the efficacy of feedback in both controlled laboratory experiments and applied studies in educational settings (for reviews, see Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Butler & Winne, 1995; Hattie & Timperley, 2007; Kulhavy & Stock, 1989). Researchers have identified a variety of factors that determine the effectiveness of feedback, such as the nature of the to-be-learned material, the amount of prior learning before retrieval is attempted, and the retention interval between feedback and the subsequent use of that knowledge (e.g., retrieval on a final test). However, given the limited space we can devote to this topic, we will focus on two of the more important factors: the content of the feedback message and the timing of feedback.

Perhaps the most important aspect of feedback is the nature of the information provided. At the minimum, the feedback message needs to inform the learner whether one's response is correct or incorrect. However, there are many different ways of elaborating the feedback message beyond this basic form (for a taxonomy of feedback messages, see Kulhavy & Stock, 1989), and numerous studies have explored the efficacy of including additional information in the message. The most consistent result is that providing learners with the correct answer in the feedback message produces better subsequent performance than simply indicating whether an answer is correct or incorrect (e.g., Pashler et al., 2005; for a meta-analysis, see Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). This finding makes sense because informing the learner that a given response is incorrect will not help the learner to correct the error if the learner does not have any recourse to learn the correct answer.

There is little consensus as to whether feedback should contain additional information beyond the correct answer. Some studies have found that elaborating the basic feedback message has a positive effect on learning; for example, learning can be improved by providing students with the original study materials so that they can self-grade their tests (Andre & Thieman, 1988) or requiring them to generate the feedback by rearranging letters to form the correct answer (Lhyle & Kulhavy, 1987). However, many other elaborations have shown no effect or, occasionally, a negative effect on learning (see Kulhavy & Stock, 1989). One reason for these disparate findings may be the way in which learning is assessed on the final test. For example, studies on explanation feedback (i.e., providing an explanation of why the learner's response is correct or incorrect) often find no benefit relative to correct answer feedback (e.g., Kulhavy, White, Topp, Chan, & Adams, 1985; McDaniel & Fisher, 1991). However, these studies generally assess learning with verbatim repetitions of the same questions, so that the learner merely has to memorize the correct responses rather than understand the concepts. Explanation feedback may be more effective for promoting transfer to situations in which learners must apply their knowledge in a new context (Butler, Godbole, & Marsh, 2010).

Another key factor involves the timing of the feedback message. Studies have produced contradictory results on feedback timing. Some studies have found it is better to give feedback immediately, while others have found it is better to give it after a delay (for a meta-analysis, see Kulik & Kulik, 1988). The literature on this topic is complicated by differences in how "immediate" and "delayed" feedback are defined across studies, serious methodological flaws (e.g., confounding the timing of feedback with the type of feedback; Angell, 1949), and the deeply entrenched notion that feedback must be given immediately to be effective (e.g., Skinner, 1954). Although some reviewers have concluded that feedback should be given as soon as possible to be effective (e.g., Mory, 2004; Kulik & Kulik, 1988), the weight of evidence seems to suggest that delayed feedback produces superior retention. In fact, many studies have shown that even short (p. 308) delays produce better retention, such as waiting until the end of the test rather than giving feedback after each question (e.g., Butler & Roediger, 2008). Of course, the key assumption in comparing the effectiveness of immediate and delayed feedback is that feedback is processed in the same way regardless of its timing (see Butler, Karpicke, & Roediger, 2007). For example, if students are motivated to study all the feedback when it is given immediately, but only study some of it if it is given a week later, then immediate feedback may be more effective.

In summary, processing feedback helps to correct errors and maintain correct responses. Consistent with the notion of processing material to extract meaning, feedback messages that are elaborated beyond the basic form of verification (i.e., correct/incorrect) may promote deeper learning and therefore produce superior transfer to new tasks. All things being equal, delayed feedback seems to be more effective than immediate feedback. Delaying feedback represents a desirable difficulty because it generally takes more effort to reinstate the initial learning context when processing feedback after a delay. However, it is important to note that if full processing of the delayed feedback cannot be guaranteed, then giving immediate feedback is probably the better choice.

Spacing Out Practice Over Time

Students often put off studying until the last possible moment (oftentimes the night before an exam). Procrastination is a behavior that is manifested at all levels of education (even teachers wait until the night before to compose a test). In the literature, this pattern of behavior has been termed the “procrastination scallop” (Michael, 1991). Study behavior remains relatively infrequent over time, but then it increases rapidly as an exam or some other reason to learn the material approaches. For instance, Mawhinney and colleagues (Mawhinney, Bostow, Laws, Blumenfeld, & Hopkins, 1971) conducted a classroom experiment in which they tested students either every day or every 3 weeks. When students were tested daily, they studied for 60 to 80 minutes every day. In contrast, students who were tested every 3 weeks studied about 20 minutes a day at first and then increased the amount of time that they studied until they reached 120 minutes a day right before the test (i.e., the classic procrastination scallop pattern).

Why is procrastination so common in educational settings? One likely factor is that cramming, which results from procrastination, is a very effective strategy in the short term. When students cram or “mass” practice, they probably do fairly well on an immediate test, leading to the illusion that they know the material. However, as we discussed in the section on “Introducing ‘Desirable Difficulties’ During Learning,” performance on an immediate test is not a good predictor of long-term retention. As any student will readily attest, most of the information that is learned from cramming is quickly forgotten. Thus, cramming may be useful under some circumstances, but it is not an effective way of producing knowledge that will be retained over long periods of time.

Many studies have shown that spacing or distributing practice over time produces better long-term retention of material than massing practice (i.e., cramming), a finding called *the spacing effect* (e.g., Glenberg, 1976; Melton, 1970). Ebbinghaus (1885/1967) described the spacing effect in his pioneering investigation of human memory, making this finding one of the oldest in the literature. After over a century of investigation, the spacing effect has been shown to be extremely robust and easy to replicate (for a review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Dempster, 1989). Based on the ubiquity of the spacing effect, the Institute of Education Sciences recommended spacing as a strategy to promote retention in a practice guide entitled *Organizing Instruction and Study to Improve Student Learning* (Pashler, Bain, et al., 2007; see recommended readings).

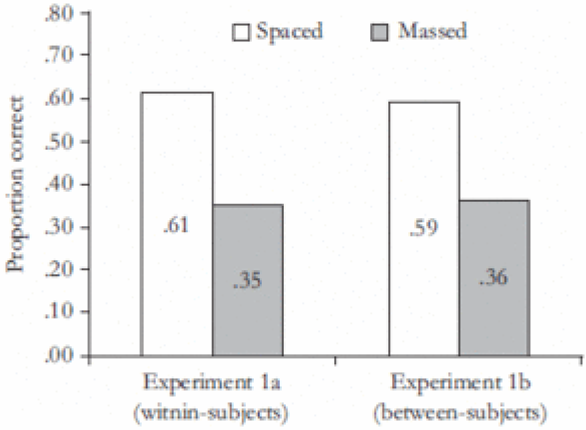


Figure 20.3 Proportion of artists selected correctly on the multiple-choice test given after the first block of trials as a function of learning condition (spacing vs. massing) for Experiments 1a (left side) and 1b (right side). (Data from Kornell & Bjork, 2008, Experiments 1a and 1b.)

Although many studies of the spacing effect have used lists of words and other simple information, the finding has also been demonstrated with more complex sets of material. For example, Kornell and Bjork (2008; Experiments 1a and 1b) investigated whether spacing would lead to better inductive learning of the styles of different painters in a pair of experiments. They had subjects study six paintings by 10 different artists in either a massed fashion (all the paintings from a given artist were presented consecutively) or spaced out in time (the paintings from each artist were interleaved with paintings from other artists).[1] Spacing was manipulated within subjects in Experiment 1a and between subjects in Experiment 1b. After each presentation block, subjects took a test in which they saw new paintings by the same artists, and they had to indicate which artists had painted them. Figure 20.3 shows the results of the test that was given after the first block (i.e., before subjects received any feedback). As the results clearly show, subjects in both experiments were better able to correctly identify the artist (p. 309) when they had studied the paintings under spaced conditions rather than massed conditions.

Is there an optimal interval for spacing practice? Studies have shown benefits of distributing practice over intervals that range from a few seconds (see Underwood, 1961) to many years (e.g. Bahrick, Bahrick, Bahrick, & Bahrick, 1993), so there is not just one optimal interval. Rather, it seems that the optimal interval depends on how long the knowledge needs to be retained after the last practice (i.e., the retention interval). Cepeda and colleagues (2006; see also Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008) performed a meta-analysis that included 317 experiments from 184 articles on the spacing effect. They found that the optimal spacing interval is approximately 10%–20% of the retention interval. Thus, if the goal is to retain the material for 5 days, then practice should be spaced over intervals between 12 and 24 hours. However, if the desired retention interval is 10 years, then the optimal interval between practices is about 1 to 2 years.

Overall, spacing out practice over time is a powerful method for promoting long-term retention. Reprocessing material after a delay may lead to the extraction of different information (and different meaning, increasing deep processing), which can facilitate transfer in certain situations (e.g., Kornell & Bjork, 2008). Spacing constitutes a desirable difficulty in that it may initially slow learning as people have a harder time remembering previously learned material after a delay (and probably forget some of it); however, these initial difficulties will pay dividends over the long run because spacing will produce better retention relative to massing practice. In terms of identifying the spacing schedule that will optimize retention, the ratio described earlier may sound complicated; however, it can be simplified by stating that spacing over greater intervals generally leads to retention over longer periods of time.

Combining Techniques

Successful students are unlikely to limit themselves to a single technique, so future research should assess how to combine the kinds of techniques reviewed here. We alluded earlier to one combination known to benefit memory: reviewing notes (a poststudy strategy) taken during a lecture (a study strategy). A second example involves the 3R technique, a self-controlled method of learning that McDaniel, Howard, and Einstein (2009) created by shortening an older technique (Robinson, 1941). The 3R technique involves reading the text, reciting aloud all the information that can be remembered without looking at the text, and then rereading the text (i.e., read-recite-review—hence the 3R name). The 3R technique incorporates two of the poststudy learning strategies discussed earlier: retrieval practice (the recite part of the technique) and processing feedback (obtained when the text is reviewed).

In one experiment, McDaniel et al. (2009, Experiment 2) compared the 3R technique with two other common techniques: rereading and note taking (the latter of which we reviewed earlier as an effective strategy). Subjects read engineering texts while using one of the three techniques (3R, rereading, or note taking). Learning was assessed via a series of tests given either immediately or after a 1-week delay. A free-recall test measured memory for the entire text, a multiple-choice test measured retention of facts from the text, and a short answer test required inferences to be made based on information from the text. Subjects who used the 3R technique recalled more of the texts than subjects who either reread the texts or took notes, and this result held both immediately and a week later. Using the 3R technique also increased performance on both fact-based, multiple-choice questions and inference-based, short-answer questions, relative to rereading. The 3R technique and note taking led to similar performance on the multiple-choice and short-answer tests, but the 3R technique required less time during the initial study phase (meaning it was a more efficient way to learn the material).

The 3R technique is probably effective because it requires significant effort on the part of the learner—recalling a text is much harder to do than simply rereading it, and actively reviewing a text (p. 310) to obtain feedback is more difficult than passively receiving feedback. Furthermore, additional meaning is extracted when the student practices retrieval and then reviews each text for feedback. That is, when students reread the text *after* having attempted retrieval, the processing will change, leading to the encoding of new information as well as reencoding learned information in new ways.

In short, strategies that combine techniques have the potential to be even more effective than a single approach. However, more is not always better; a complex strategy that combines ineffective techniques may not be helpful and may even hurt the learner. The guiding principles can be used to evaluate the components of more complex strategies to yield predictions about whether techniques will help the learner.

Applying the Guiding Principles and Learning Strategies in Educational Settings

Now that we have described several guiding principles and learning strategies, it is important to think about their application in the classroom. We note that most of the strategies reviewed here have been studied in some form in the classroom (e.g., Crawford, 1925a; Larsen et al., 2009). However, when thinking about the transition between the laboratory and the classroom, two obstacles must be considered: Experimenters often use short delays and simple materials, whereas educators are interested in long-term learning and mastery of more complex materials. In this section, we will highlight the advances that have been made toward solving each of these problems.

Long-Term Learning

First, consider the issue of long-term learning. Cognitive psychologists often think they are using long delays when they examine memory over days or even weeks (in fact, delays as short as 20 minutes are often labeled “long-term memory”). Even in the majority of studies cited in this chapter, many of which explicitly claim to have educational applications, learning is often assessed after a few days. However, the goal of education is learning that will be retained for years. This is difficult to study in the laboratory, for obvious reasons. However, Harry Bahrick brilliantly tackled this challenge using naturalistic studies; for example, he examined people’s memory for Spanish vocabulary and grammar learned during high school (Bahrick, 1984). As will be described, these studies have yielded a unique finding: the concept of a *permastore*, which refers to relatively permanent retention of knowledge over long periods of time, even if it is not used in the interim. Bahrick’s study on Spanish included subjects of all ages, so that the amount of time since initial learning varied widely—from high school students who were currently enrolled in Spanish classes to older adults who had taken it 50 years earlier. Retention of high school Spanish dropped off sharply over the first 5 years, but then stayed stable over the next 30 years before showing a final decline (which was probably related to cognitive aging). Interestingly, a portion of the knowledge remained accessible up to 50 years later, even though it had not been used since high school, leading Bahrick to propose the concept of the permastore.

Of interest is how the types of learning strategies reviewed in this chapter may contribute to the type of long-term learning studied by Bahrick and emphasized by educators. Many studies have shown that knowledge acquired in the classroom is retained over long periods of time (e.g., Conway, Cohen, & Stanhope, 1991; Landauer & Ainsle, 1975; Semb, Ellis, & Araujo, 1993; for review, see Semb & Ellis, 1994), and our focus is on the techniques that promote this long-term retention. While noting that many of these studies are naturalistic in their design (i.e., no independent variables, no random assignment to groups, many uncontrolled variables, etc.), and thus the conclusions drawn from them must be interpreted with some caution, it is important that many of these studies yield results that converge with the results of careful laboratory studies. For example, Bahrick and colleagues provided evidence that spaced poststudy sessions promote long-term memory for foreign language vocabulary (Bahrick et al., 1993). Similarly, taking a test on material helps students to remember the information a year later (e.g., Landauer & Ainslie, 1975; see also Semb & Ellis, 1994).

Intriguingly, some of the strategies reviewed here may also be useful for recovering knowledge, long after it has been forgotten. That is, just because students learned something at one point in time does not mean they will be able to retrieve it in the future. Bahrick referred to this type of knowledge as *marginal knowledge*: information that was stored but cannot currently be retrieved (this distinction is very similar to Tulving’s distinction between *availability* and *accessibility* of information in memory; see Tulving & Pearlstone, 1966). In one study, the power of feedback to reactivate marginal knowledge was examined, using general-knowledge materials (Berger, Hall, & Bahrick, 1999). Subjects received (p. 311) 5 seconds of answer feedback after they failed to correctly answer each of a series of short-answer questions (e.g., they saw "Einstein" if they failed to correctly answer the question, "What is the last name of the person who proposed the theory of relativity?"). Critically, half of the items referenced real facts (like the Einstein fact just shown) and half were yoked controls that referenced fictional entities (e.g., "What is the last name of the person who proposed the theory of maladaptivity?" Answer: Alfred). The fictional items were important as a control for new learning; the feedback could only activate preexisting marginal knowledge for real items and not for fictional items. After various retention intervals of up to 9 days, subjects were asked the critical questions again. Feedback boosted performance on the real items, even after 9 days, whereas performance on the fictional items quickly dropped to zero after a delay, since there was no marginal knowledge to activate for these items.

In short, strategies like retrieval practice, spacing, and feedback have the potential to enhance learning over the long run, as desired by educators. One of the major challenges for future research is to further connect the learning strategies reviewed in this chapter to the types of retention intervals important in education.

Learning Beyond Facts

Experimentalists have a tendency to simplify the materials used, in order to control for as many potentially confounding factors as possible. This is not possible in actual classrooms, nor is it desirable. Rather, in the classroom, educators have many different goals for their students, and these goals vary as a function of the level of education, the type of course, and the time frame given for learning, among many other factors. One way of categorizing these goals is through Bloom's (1956) taxonomy of educational objectives, which conceptualizes learning as a hierarchy in which the various levels must be mastered in sequential order. The cognitive domain is comprised of six levels (from lowest to highest): knowledge (e.g., learning facts, concepts, etc.), comprehension (e.g., understanding the relationship between ideas), application (e.g., using knowledge to solve new problems), analysis (e.g., finding evidence to support a hypothesis), synthesis (e.g., combining different accounts of an event to understand what occurred), and evaluation (e.g., assessing the validity of an idea according to certain criteria). Although Bloom's taxonomy has been criticized over the years (e.g., Moore, 1982), it has proven to be a valuable tool for both researchers and educators. In fact, its massive influence on the fields of education and educational psychology recently led to a major effort to update the taxonomy so that it could continue to be used (Anderson et al., 2001).

Cognitive psychologists often claim to be using educationally relevant materials when they are using general-knowledge facts (e.g., Lima is the capital of Peru) or simple texts. However, such learning is classified as Level 1 in Bloom's taxonomy, reflecting the learning of simple facts and concepts. To the extent that research on learning strategies will be translated to the classroom, it must move beyond the learning of facts and concepts to higher levels of knowledge (e.g., evaluation). Throughout this chapter we have highlighted instances where learning strategies afford transfer; for example, as described earlier, receiving advance organizers promoted interpretation of new programming statements, note taking improved performance on conceptual questions about the chi-square statistic, and reading explanation feedback enhanced transfer of knowledge about complex concepts such as diffusion.

Recent work suggests that retrieval practice may also facilitate transfer. In one study, Butler (2010, Experiment 2) had subjects study a set of prose passages, each of which contained several critical concepts (e.g., how bats use echolocation to navigate and find prey in the dark). A concept was defined as information that had to be extracted from multiple sentences. After reading each passage, the subjects either repeatedly restudied the entire passage, repeatedly restudied isolated sentences from the passage that contained the critical concepts, or repeatedly answered test questions about the critical concepts and received feedback on their answers (e.g., Some bats use echolocation to navigate the environment and locate prey. How does echolocation help bats to determine the distance and size of objects?). The feedback provided after each test question was essentially the same information as that presented in the condition in which subjects restudied the isolated sentences from the passage.

After a 1-week delay, students took a transfer test that consisted of questions that required the application of the critical concepts from the passages (e.g., “An insect is moving toward a bat. Using the process of echolocation, how does the bat determine that the insect is moving toward it [i.e., rather than away from it]?”). When subjects were repeatedly tested on the concepts, they performed significantly better on the transfer test relative to when they repeatedly studied the passages or repeatedly (p. 312) studied the isolated sentences. The comparison between repeated testing and repeated study of the isolated sentences is particularly important, because these two conditions were well matched in terms of exposure to the material and total time spent learning. In fact, the only real difference between these conditions was that the subjects attempted to retrieve the critical concepts in one condition and not in the other.

The finding that testing promoted transfer of knowledge is an important one (see also Johnson & Mayer, 2009; Rohrer, Taylor, & Sholar, 2010). However, encouraging as these results are, the tasks involved were still quite simple ones, and they did not require the use of knowledge as proposed in the highest levels in Bloom’s taxonomy. Future research should continue to investigate transfer and should tackle goals such as evaluation and synthesis.

The Relativity of Memory

A recurring theme in the last section involved concerns that basic research may not generalize to actual educational situations; we discussed whether the learning strategies reviewed in this chapter are likely to boost memory over long delays, with more complex materials, and for more sophisticated learning goals involving transfer of knowledge. In effect, such questions emphasize the importance of considering memory in particular contexts (e.g., McKeachie, 1974; see Roediger, 2008). This emphasis on the relativity of memory differs from the emphasis of the first half-century of cognitive research, when memory researchers focused on discovering general laws that governed how we remember and forget (e.g., Jost, 1897; Ribot, 1881; Thorndike, 1911). The switch to emphasizing the relativity of memory means that when questions like "Does X lead to better memory?" are posed, the answer is almost always "Well, it depends…."

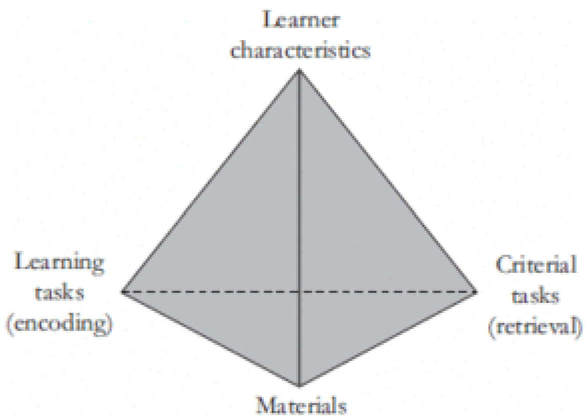


Figure 20.4 The tetrahedral model of memory experiments. (Adapted from Jenkins, 1979.)

To better understand why memory effects tend to be relative, it is helpful to consider the classes of variables that can influence memory performance in any given circumstance. Figure 20.4 presents a tetrahedral model of memory that was created by James Jenkins (1979). In the model, each vertex of the tetrahedron represents a class of variables: the learning or encoding tasks, the to-be-learned materials, the characteristics of the learner, and the criterial tasks that assess memory (see Roediger, 2008 for a possible fifth vertex). Cognitive psychologists often carefully control for these different factors by specifying how the learner should process to-be-remembered information (e.g., by judging its pleasantness) and using words for materials, college students for learners, and standard tests such as recall to measure retrieval. In contrast, educators are often at the other extreme, arguing for interventions aimed at specific populations in specific contexts, as opposed to a "one-size-fits-all" approach. The best approach likely falls in the middle, with the goal being as general of advice as possible, while considering how the advice to an educator might change depending on factors such as learner characteristics or test type.

We have already discussed how the interaction of encoding and retrieval tasks can have a profound effect on memory performance. In passing we also mentioned some of the different types of materials that researchers have used, as well as some individual differences. However, we did not have the opportunity to review research on how certain types of materials naturally afford specific types of processing (e.g., the narrative structure of fairytales invites relational processing). Research has shown that learning is better when the processing induced by the materials is complementary to the processing that is induced by the learning task (see McDaniel & Einstein, 1989, 2005). For example, performing a relational processing task (e.g., sorting mixed-up sentences into the correct order) on a fairytale would not be an effective way of learning because the processing induced by the task is redundant with the processing induced by the materials. In contrast, performing such a relational processing task on an expository text (which does not naturally induce relational processing) would be effective, because it would be complementary to the processing naturally afforded by the text.

Likewise, characteristics of the learner will impact whether the processing induced by a particular task enhances memory. If the learner has (p. 313) difficulty with a particular type of processing or tends to naturally engage in a certain type of processing, then giving that learner a task that induces that same type of processing will be detrimental and redundant, respectively. For example, readers who score poorly on the Nelson-Denny Reading Test (Brown, Nelson, & Denny, 1973) do not benefit from processing a text using a letter-insertion task like the one described earlier (e.g., McDaniel, Hines, & Guynn, 2002). The Nelson-Denny Reading Test is a measure of reading ability that reflects, among other things, people's proficiency at decoding words. If people have trouble decoding words, then performing a task that requires generating words from fragments will not be helpful because it places additional stress on word decoding.

The relativity of memory means that desirable difficulties must be considered within the context of the other three vertices of Jenkins's model (for discussion, see McDaniel & Butler, 2010). The desirability of any difficulty will be determined by three factors. First, the particular type of processing in the learning task must match the type of processing in the retrieval task (e.g., deWinstanley, Bjork, & Bjork, 1996). Second, the type of processing invited by the materials must be complementary to that induced by the learning task (e.g., McDaniel & Einstein, 1989). Third, and finally, the abilities of the learner must allow the difficulty induced by the learning task to be overcome (e.g., McDaniel et al., 2002).

Conclusions and Future Directions

In this chapter, we described three cognitive theories that can guide educational practice: desirable difficulties (and how it can be a good thing to make learning harder), meaning extraction (and how it is important that learning tasks encourage the student to process the material at "deeper" levels), and transfer-appropriate processing (and how the strategy should make the learner attend to whatever it is he or she will need to succeed on the final criterial task). We used these three theories to evaluate the effectiveness of strategies that guide the initial learning of material (advance organizers, underlining/highlighting, and note taking) as well as strategies that guide poststudy processing (retrieval practice, feedback, and spacing). We could have sampled additional strategies, such as the use of adjunct questions during reading (a strategy involving retrieval practice), summarization exercises (which can be completed during learning or poststudy, encouraging meaning extraction), and the use of imagery (which is often involved in mnemonics created during study), to name just a few. The application of the theories is not limited to the examples we chose, and future research will continue to connect the theories to additional encoding and poststudy strategies.

When applying cognitive psychology to education, it is critical to think about long-term retention and transfer of knowledge (two important directions for future research). The results of experiments are more likely to generalize to the classroom if researchers keep Jenkins's tetrahedral model in mind, by thinking about how results might change for different students, materials, strategies, and criterial tasks. Earlier in this chapter we alluded to the practice guide published by the Institute of Education Sciences (Pashler, Bain, et al., 2007; see suggested readings); this guide made recommendations for the classroom based on basic cognitive science research. One key criterion for inclusion in the guide was that a principle had to have been shown in both the classroom and the laboratory. Successful translation of cognitive research will likely require explicit classroom demonstrations, as educators are unlikely to be convinced by assumptions of generality. Moving into the classroom may also benefit cognitive psychologists because effects that persist in the classroom are likely to be robust (see Rubin, 1989) and careful observation of the classroom may reveal interesting new phenomena to study. By increasing the collaboration and communication between cognitive psychologists and educators, we will enhance our understanding of how to optimize learning to promote long-term retention and transfer of knowledge.

Acknowledgments

The writing of this chapter was supported by a collaborative activity award from the James S. McDonnell Foundation. The opinions expressed are those of the authors and do not represent the views of the James S. McDonnell Foundation. We thank Anna Goswick for her assistance with research and editing.

References

Abbott E. E. (1909). On the analysis of the factors of recall in the learning process. Psychological Monographs, 11, 159–177.

Anderson L. W., Krathwohl D. R., Airasian P. W., Cruikshank K. A., Mayer R. E., Pintrich P. R.,…Wittrock M. C. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives. New York : Longman.

Andre T., & Thieman A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. Contemporary Educational Psychology, 13, 296–307.

Angell G. W. (1949). The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. Journal of Educational Research, 42, 391–394.

Austin J. L., Lee M. G., Thibeault M. D., Carr J. E., & Bailey J. S. (2002). Effects of guided notes on university students’ responding and recall of information. Journal of Behavioral Education, 11, 243–254.

Ausubel D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology, 51, 267–272.

Ausubel D. P., & Fitzgerald D. (1961). The role of discriminability in meaningful verbal learning and retention. Journal of Educational Psychology, 52, 266–274.

Bahrick H. P. (1984). Semantic memory content in permastore: 50 years of memory for Spanish learned in school. Journal of Experimental Psychology: General, 113, 1–29.

Bahrick H. P., Bahrick L. E., Bahrick A. S., & Bahrick P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. Psychological Science, 4, 316–321.

Bahrick H. P., & Hall L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. Journal of Memory and Language, 52, 566–577.

Bangert-Drowns R. L., Kulik J. A., & Kulik C. C. (1991). Effects of frequent classroom testing. Journal of Educational Research, 85, 89–99.

Bangert-Drowns R. L., Kulik C. C., Kulik J. A., & Morgan M. (1991). The instructional effect of feedback in test-like events. Review of Educational Research, 61, 213–238.

Berger S. A., Hall L. K., & Bahrick H. P. (1999). Stabilizing access to marginal and submarginal knowledge. Journal of Experimental Psychology: Applied, 5, 438–447.

Bjork R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), Information processing and cognition: The Loyola Symposium (pp. 123–144). Hillsdale, NJ : Erlbaum.

Bjork R. A. (1994a). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). Cambridge, MA : MIT Press.

Bjork R. A. (1994b). Institutional impediments to effective training. In D. Druckman & R. A. Bjork (Eds.), Learning, remembering, believing: Enhancing human performance (pp. 295–306). Washington, DC : National Academy Press.

Bjork R. A., & Bjork E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), From learning processes to cognitive processes: Essays in honor of William K. Estes (Vol. 2, pp. 35–67). Hillsdale, NJ : Erlbaum.

Black P., & Wiliam D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy, and Practice, 5, 7–74.

Bloom B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Essex, England : Harlow.

Bretzing B. H., & Kulhavy R. W. (1979). Notetaking and depth of processing. Contemporary Educational Psychology, 4, 145–153.

Bretzing B. H., & Kulhavy R. W. (1981). Note-taking and passage style. Journal of Educational Psychology, 73, 242–250.

Bromage B. K., & Mayer R. E. (1986). Quantitative and qualitative effects of repetition from learning from technical text. Journal of Educational Psychology, 78, 271–278.

Brown J. I., Nelson M. J. B., & Denny E. C. (1973). The Nelson-Denny reading test. Boston, MA : Houghton Mifflin.

Butler A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 1118–1133.

Butler A. C., Godbole N., & Marsh E. J. (in press). Explanation feedback is better than correct answer feedback for promoting transfer of learning. Journal of Educational Psychology.

Butler A. C., Karpicke J. D., & Roediger H. L., III. (2007) The effect of type and timing of feedback on learning from multiple-choice tests. Journal of Experimental Psychology: Applied, 13, 273–281.

Butler A. C., Karpicke J. D., & Roediger H. L., III. (2008). Correcting a meta-cognitive error: Feedback enhances retention of low confidence correct responses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34, 918–928.

Butler A. C., & Roediger H. L., III. (2007). Testing improves long-term retention in a simulated classroom setting. European Journal of Cognitive Psychology, 19, 514–527.

Butler A. C., & Roediger H. L., III. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. Memory and Cognition, 36, 604–616.

Butler D. L., & Winne P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. Review of Educational Research, 65, 245–281.

Callender A. A., & McDaniel M. A. (2009). The limited benefits of rereading educational texts. Contemporary Educational Psychology, 34, 30–41.

Carpenter S. K., & DeLosh E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. Memory and Cognition, 34, 268–276.

Carpenter S. K., Pashler H., & Cepeda N. J. (2009). Using tests to enhance 8th grade students' retention of U. S. history facts. Applied Cognitive Psychology, 23, 760–771.

Carrier M., & Pashler H. (1992). The influence of retrieval on retention. Memory and Cognition, 20, 632–642.

Cashen V. M., & Leicht K. L. (1970). Role of the isolation effect in a formal educational setting. Journal of Educational Psychology, 61, 484–486.

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 effect in learning: A temporal ridgeline of optimal retention. Psychological Science, 19, 1095–1102.

Christina R. W., & Bjork R. A. (1991). Optimizing long-term retention and transfer. In D. Druckman & R. A. Bjork (p. 315) (Eds.), In the mind's eye: Enhancing human performance (pp. 23–56). Washington, DC : National Academy Press.

Conway M. A., Cohen G., & Stanhope N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. Journal of Experimental Psychology: General, 120, 395–409.

Corkill A. J. (1992). Advance organizers: Facilitators of recall. Educational Psychology Review, 4, 33–67.

Corkill A. J., Glover J. A., Bruning R. H., & Krug D. (1988). Advance organizers: Retrieval context hypothesis. Journal of Educational Psychology, 80, 304–311.

Craik F. I. M., Govoni R., Naveh-Benjamin M., & Anderson N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. Journal of Experimental Psychology: General, 125, 159–180.

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

Craik F. I. M., & Tulving E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268–294.

Crawford C. C. (1925a). The correlation between college lecture notes and quiz papers. Journal of Educational Research, 12, 282–291.

Crawford C. C. (1925b). Some experimental studies of the results of college note-taking. Journal of Educational Research, 12, 379–386.

Dempster F. N. (1989). Spacing effects and their implications for theory and practice. Educational Psychology Review, 1, 309–330.

deWinstanley P. A., Bjork E. L., & Bjork R. A. (1996). Generation effects and the lack thereof: The role of transfer-appropriate processing. Memory, 4, 31–48.

Dinnel D., & Glover J. A. (1985). Advance organizers: Encoding manipulations. Journal of Educational Psychology, 77, 514–521.

Di Vesta F. J., & Gray G. S. (1972). Listening and note taking. Journal of Educational Psychology, 63, 8–14.

Ebbinghaus H. (1967). Memory: A contribution to experimental psychology (H. A. Ruger & C. E. Bussenius Trans.). New York : Dover. (Original work published 1885).

Einstein G. O., Morris J., & Smith S. (1985). Note-taking, individual differences, and memory for lecture information. Journal of Educational Psychology, 77, 522–532.

Fazio L. K., Huelser B. J., Johnson A., & Marsh E. J. (2010). Receiving right/wrong feedback: Consequences for learning. Memory, 18, 335–350.

Fisher R. P., & Craik F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning and Memory, 3, 701–711.

Fowler R. L., & Barker A. S. (1974). Effectiveness of highlighting for retention of text material. Journal of Applied Psychology, 59, 358–364.

Gates A. I. (1917). Recitation as a factor in memorizing. Archives of Psychology, 6 (40), 1–104.

Glenberg A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. Journal of Verbal Learning and Verbal Behavior, 15, 1–16.

Glover J. A. (1989). The "testing" phenomenon: Not gone but nearly forgotten. Journal of Educational Psychology, 81, 392–399.

Glover J. A., Bullock R. G., & Dietzer M. L. (1990). Advance organizers: Delay hypotheses. Journal of Educational Psychology, 82, 291–297.

Hattie J., & Timperley H. (2007). The power of feedback. Review of Educational Research, 77, 81–112.

Hunt R. R., & Einstein G. O. (1981). Relational and item-specific information in memory. Journal of Verbal Learning and Verbal Behavior, 20, 497–514.

Idstein P., & Jenkins J. R. (1972). Underlining versus repetitive reading. Journal of Educational Research, 65, 321–323.

Jenkins J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 429–446). Hillsdale, NJ : Erlbaum.

Johnson C. I., & Mayer R. E. (2009). A testing effect with multimedia learning. Journal of Educational Psychology, 101, 621–629.

Jones H. E. (1923–1924). The effects of examination on the performance of learning. Archives of Psychology, 10, 1–70.

Jost A. (1897). Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen [The strength of associations in their dependence on the distribution of repetitions]. Zeitschrift fur Psychologie und Physiologie der Sinnesorgane, 16, 436–472.

Kang S. H. K., Mc Dermott K. B. & Roediger H. L., III. (2007). Test format and corrective feedback modulate the effect of testing on memory retention. European Journal of Cognitive Psychology, 19, 528–558.

Karpicke J. D., Butler A. C., & Roediger H. L., III. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? Memory, 17, 471–479.

Karpicke J. D., & Roediger H. L., III. (2008). The critical importance of retrieval for learning. Science, 15, 966–968.

Kiewra K. A. (1983). The process of review: A levels of processing approach. Contemporary Educational Psychology, 8, 366–374.

King A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. American Educational Research Journal, 29, 303–323.

Kornell N., & Bjork R. A. (2007). The promise and perils of self-regulated study. Psychonomic Bulletin and Review, 14, 219–224.

Kornell N., & Bjork R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? Psychological Science, 19, 585–592

Kulhavy R. W., & Stock W. A. (1989). Feedback in written instruction: The place of response certitude. Educational Psychology Review, 1, 279–308.

Kulhavy R. W., White M. T., Topp B. W., Chan A. L., & Adams J. (1985). Feedback complexity and corrective efficiency. Contemporary Educational Pscyhology, 10, 285–291.

Kulik J. A., & Kulik C. C. (1988). Timing of feedback and verbal learning. Review of Educational Research, 58, 79–97.

Landauer T. K., & Ainslie K. I. (1975). Exams and use as preservatives of course-acquired knowledge. Journal of Educational Research, 69, 99–104.

Larsen D. P., Butler A. C., & Roediger H. L., III. (2009). Repeated testing improves long-term retention relative to repeated study: A randomized, controlled trial. Medical Education, 43, 1174–1181. (p. 316)

Lhyle K. G., & Kulhavy R. W. (1987). Feedback processing and error correction. Journal of Educational Psychology, 79, 320–322.

Lorch R. F., Lorch E. P., & Klusewitz M. A. (1995). Effects of typographical cues on reading and recall of text. Contemporary Educational Psychology, 20, 51–64.

Marsh E. J., Roediger H. L., III, Bjork R. A., & Bjork E. L. (2007). The memorial consequences of multiple-choice testing. Psychonomic Bulletin and Review, 14, 194–199.

Marsh E. J., & Sink H. E. (2010). Access to handouts of presentation slides during lecture: Consequences for learning. Applied Cognitive Psychology, 24, 691–706.

Mawhinney V. T., Bostow D. E., Laws D. R., Blumenfeld G. J., & Hopkins B. L. (1971). A comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules. Journal of Applied Behavior Analysis, 4, 257–264.

Mayer R. E. (1975). Different problem-solving competencies established in learning computer programming with and without meaningful models. Journal of Educational Psychology, 67, 725–734.

Mayer R. E. (1976). Some conditions of meaningful learning for computer programming: Advance organizers and subject control of frame order. Journal of Educational Psychology, 68, 143–150.

Mayer R. E. (March 27–3l, 1978). Can advance organizers counter the effects of text organization? Paper presented at the Annual meeting of the American Educational Research Association, Toronto, Canada.

Mayer R. E. (1979). Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. Instructional Science, 8, 133–167.

Mayer R. E. (1983). Can you repeat that: Qualitative effects of repetition and advance organizers on learning from science prose. Journal of Educational Psychology, 75, 40–49.

Mayer R. E., & Bromage B. K. (1980). Different recall protocols for technical texts due to advance organizers. Journal of Educational Psychology, 72, 209–225.

McDaniel M. A., Agarwal P. K., Huelser B. J., Mc Dermott K. B., & Roediger H. L., III. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. Journal of Educational Psychology, *103*, 399–414.

McDaniel M. A., Anderson J. L., Derbish M. H., & Morrisette N. (2007). Testing the testing effect in the classroom. European Journal of Cognitive Psychology, 19, 494–513.

Mc Daniel M. A., & Butler A. C. (2010). A contextual framework for understanding when difficulties are desirable. In A. S. Benjamin (Ed.), Successful remembering and successful forgetting: Essays in honor of Robert A. Bjork (pp. 175–199). New York : Psychology Press.

McDaniel M. A., & Einstein G. O. (1989). Material appropriate processing: A contextualist approach to reading and studying strategies. Educational Psychology Review, 1, 113–145.

McDaniel M. A., & Einstein G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), Experimental cognitive psychology and its applications (pp. 73–85). Washington, DC : American Psychological Association.

McDaniel M. A., & Fisher R. P. (1991). Tests and test feedback as learning sources. Contemporary Educational Psychology, 16, 192–201.

McDaniel M. A., Hines R. J., & Guynn M. J. (2002). When text difficulty benefits less-skilled readers. Journal of Memory and Language, 46, 544–561.

McDaniel M. A., Hines R. J., Waddill P. J., & Einstein G. O. (1994). What makes folk tales unique: Content familiarity, causal structure, scripts, or superstructures? Journal of Experimental Psychology: Learning, Memory, and Cognition, 2 *0*, 169–184.

McDaniel M. A., Howard D. C., & Einstein G. O. (2009). The read-recite-review study strategy: Effective and portable. Psychological Science, 20, 516–522.

McDaniel M. A., Roediger H. L., III, & Mc Dermott K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. Psychonomic Bulletin and Review, 14, 200–206.

McKeachie W. J. (1974). Instructional psychology. Annual Review of Psychology, 25, 161–193.

Melton A. W. (1970). The situation with respect to the spacing of repetitions and memory. Journal of Verbal Learning and Verbal Behavior, 9, 596–606.

Metcalfe J., Kornell N., & Son L. K. (2007). A cognitive-science based program to enhance study efficacy in a high and low-risk setting. European Journal of Cognitive Psychology, 19, 743–768.

Michael J. (1991). A behavioral perspective on college teaching. Behavior Analyst, 14, 229–239.

Moore D. S. (1982). Reconsidering Bloom's taxonomy of educational objectives, cognitive domain. Educational Theory, 32, 29–34.

Moore J. C. (1968). Cueing for selective note-taking. Journal of Experimental Education, 36, 69–72.

Morris C. D., Bransford J. D., & Franks J. J. (1977). Levels of processing versus transfer-appropriate processing. Journal of Verbal Learning and Verbal Behavior, 16, 519–533.

Mory E. H. (2004). Feedback research review. In D. Jonassen (Ed.), Handbook of research on educational communications and technology (pp. 745–783). Mahwah, NJ : Erlbaum.

Moscovitch M., & Craik F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. Journal of Verbal Learning and Verbal Behavior, 15, 447–458.

Packman J. L., & Battig W. F. (1978). Effects of different kinds of semantic processing on memory for words. Memory and Cognition, 6, 502–508.

Palmatier R. A., & Bennett J. M. (1974). Notetaking habits of college students. Journal of Reading, 18, 215–218.

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

Pashler H., Cepeda N. J., Wixted J. T., & Rohrer D. (2005). When does feedback facilitate learning of words? Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 3–8.

Pashler H., Rohrer D., Cepeda N. J., & Carpenter S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. Psychonomic Bulletin and Review, 14, 187–193.

Peper R. J., & Mayer R. E. (1978). Note taking as a generative activity. Journal of Educational Psychology, 70, 514–522. (p. 317)

Peper R. J., & Mayer R. E. (1986). Generative effects of note-taking during science lectures. Journal of Educational Psychology, 78, 34–38.

Peters D. L. (1972). Effects of note taking and rate of presentation on short-term objective test performance. Journal of Educational Psychology, 63, 276–280.

Pyc M. A., & Rawson K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? Journal of Memory and Language, 60, 437–447.

Ribot T. (1881). Les maladies de la memoire [Diseases of memory]. Paris: Germer Bailliere.

Rickards J. P., & August G. J. (1975). Generative underlining strategies in prose recall. Journal of Educational Psychology, 67, 800–805.

Robinson F. P. (1941). Effective study. New York : Harper and Brothers.

Roediger H. L., III. (2008). Relativity of remembering: Why the laws of memory vanished. Annual Review of Psychology, 59, 225–254.

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

Roediger H. L., III, & Karpicke J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17, 249–255.

Roediger H. L., III, Weldon M. S., & Challis B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, III, & F. I. M. Craik (Eds.), Varieties of memory and consciousness: Essays in honor of Endel Tulving (pp. 3–41). Hillsdale, NJ : Erlbaum.

Rohrer D., Taylor K., & Sholar B. (2010). Tests enhance the transfer of learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 233–239.

Rubin D. C. (1989). Issues of regularity and control: Confessions of a regularity freak. In L. W. Poon, D. C. Rubin, and B. A. Wilson (Eds.), Everyday cognition in adulthood and late life. Cambridge, England : Cambridge University Press.

Schmidt R. A., & Bjork R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science, 3, 207–217.

Semb G. B., & Ellis J. A. (1994). Knowledge taught in school: What is remembered? Review of Educational Research, 64, 253–286.

Semb G. B., Ellis J. A., & Araujo J. (1993). Long-term memory for knowledge learned in school. Journal of Educational Psychology, 85, 305–316.

Skinner B. F. (1954). The science of learning and the art of teaching. Harvard Educational Review, 24, 86–97.

Spitzer H. F. (1939). Studies in retention. Journal of Educational Psychology, 30, 641–656.

Stordahl K. E., & Christensen C. M. (1956). The effect of study techniques on comprehension and recall. Journal of Educational Research, 49, 561–570.

Thorndike E. (1911). Animal intelligence: Experimental studies. Oxford, England : Macmillan.

Tulving E., & Pearlstone Z. (1966). Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381–391.

Tulving E., & Osler S. (1968). Effectiveness of retrieval cues in memory for words. Journal of Experimental Psychology, 77, 593–601.

Underwood B. J. (1961). Ten years of massed practice on distributed practice. Psychological Review, 68, 229–247.

Wheeler M. A., & Roediger H. L., III. (1992). Disparate effects of repeated testing: Reconciling Ballard’s (1913) and Bartlett’s (1932) results. Psychological Science, 3, 240–245.

Further Reading

McDaniel M., & Einstein G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), Experimental cognitive psychology and its applications (pp. 73–86). Washington, DC: APA.

Pashler H., Bain P. M., Bottge B. A., Graesser A., Koedinger K., Mc Daniel 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, U.S. Department of Education.

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

Willingham D. T. (2009). Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom. San Francisco: Jossey-Bass. (p. 318)

Notes:

(1.) The careful reader may have noticed that the example just described involves spacing during *initial learning*, as opposed to spacing of *poststudy* activities. To be clear, spacing is important for both behaviors, and we could have included it in either section of the chapter.

Elizabeth J. Marsh

Elizabeth J. Marsh, Department and Psychology and Neuroscience, Duke University, Durham, NC

Andrew C. Butler

Andrew C. Butler, Department of Psychology and Neuroscience, Duke University, Durham, NC