Chapter 6

The Science of Learning and Its Applications

Mark A. McDaniel and Cynthia Wooldridge

In this chapter, we appeal to basic findings in cognitive psychology and theoretical advances in learning sciences to provide guidelines for organizing instruction to improve student learning. Our development of the foundations will not be extensive (see also McDaniel & Callender, 2008; Sawyer, 2006). Our objective is to highlight instructional techniques that are grounded in basic principles and to support these techniques with brief summaries of the pertinent studies conducted with authentic educational materials and in authentic educational contexts (e.g., classrooms). We have chosen to focus on techniques that are either not currently widespread in standard practice or are generally accepted but perhaps not instantiated optimally.

Test-Enhanced Learning

We begin with a cornerstone of educational practice, *testing*, that is significantly underappreciated as an effective learning tool, and is certainly underutilized in higher education as a technique to promote learning. Testing is typically used by educators to evaluate students and assign grades. Yet testing is not neutral; it also modifies learning. Accordingly, we suggest that low- or no-stakes quizzing (i.e., retrieval of target content) can be a key component for assisting students in learning target content.

Quizzing (with feedback) mobilizes at least three direct benefits for learning. First, quizzing promotes active retrieval of information from memory. A body of basic experimental evidence has established that active retrieval produces a powerful positive effect on later retention (Roediger & Karpicke, 2006) and transfer (Butler, 2010; McDaniel, Howard, & Einstein, 2009). Second, feedback that is provided after quizzing may be especially potent for stimulating learning. Memory research suggests that failing to answer a test question can potentiate learning for the correct answer when it is later provided (Kornell, Hays, & Bjork, 2009). When students answer a question incorrectly but with high confidence, the test-potentiated learning is especially great (Butterfield & Metcalfe, 2001). Experiments conducted with educationally relevant material consistently confirm that feedback (providing the correct answer) produces significant learning gains (Butler & Roediger, 2008). A third key outcome of quizzing is improvement in metacognition. Basic research suggests that learners generally cannot judge how well they will remember previously studied information (Dunlosky & Nelson, 1994). These poor metacognitive judgments in turn negatively affect the efficacy of student-directed study activities. Theoretically, then, interventions that improve metacognition should result in more effective student-directed studying. Quizzing also has a number of positive indirect effects. These include encouraging students to keep up with material (Leeming, 2002), possibly lowering test anxiety, and alerting students to adopt self-quizzing as a learning tool (Karpicke, 2009).

Evidence

Studies conducted in college courses (Daniel & Broida, 2004; Lyle & Crawford, 2011) and medical schools (Larsen, Butler, & Roediger, 2009) have demonstrated that low- or no-stakes quizzing improves performance on subsequent class examinations. One potential criticism of relying on quizzing to assist learning is that it is only useful for learning “inert” facts that will not transfer to other uses. Recent laboratory evidence disfavors this hypothesis. Active processing via retrieval creates knowledge that can also be retrieved in other contexts (see, e.g., Butler, 2010; McDaniel et al., 2009). However, the evidence just cited relied on initial tests (quizzes) that required recall; in contrast, most of the published experiments in authentic classrooms have relied on multiple-choice quizzes, which tend to require recognition rather than recall processes.

In the classroom context, initial findings with multiple-choice quizzes are mixed, with one experimental study finding that multiple-choice quizzing limited the potency of the quizzing benefits (McDaniel et al., 2009).

In contrast, in a web-based college brain and behavior course, taking a no-stakes online multiple-choice quiz repeatedly (four times) produced benefits on exam questions that were related (not identical) to quizzed content, benefits that were as robust as those produced by repeated short-answer quizzing (McDaniel, Wildman, & Anderson, 2010). Also, in a college educational psychology course, multiple-choice quiz questions followed by class discussion about the reasoning supporting the answers significantly improved course exam performances (which had both similar and dissimilar questions to those given on the quiz) relative to no quizzing or to a condition in which the multiple-choice quizzes were not accompanied by class discussion (Mayer et al., 2009). On balance, the available experimental evidence suggests that even multiple-choice quizzing, if administered with appropriate parameters (e.g., perhaps repetition of quizzes, discussion of quiz answers) can stimulate learning that leads to flexible use of target material.

Spacing

In education, target information may often be presented several times. Also, homework and workbooks often mass practice on one particular kind of item, instead of spacing. To the extent that repeated presentation of material can be spaced rather than massed, much laboratory work indicates that learning should be more efficient and retention should be improved with spaced presentation (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). The idea is that repetition of target content and practice of cognitive skills is more effective when repetition is spaced rather than massed. Unfortunately, in college instruction, key concepts covered in one massed lesson are often not considered again during the course. Spacing the coverage of these key concepts throughout the course would be expected to significantly improve retention of course material. Several studies conducted in college classrooms support the idea that spacing produces better retention of educational content than massed repetition (Rohrer & Taylor, 2006).

Desirable Difficulties

A common assumption is that instruction that enhances performance during *learning* also produces superior long-term retention and transfer. However, performance during learning can be a poor indicator of whether that knowledge or skill will be accessible (or available) in the future (Bjork, 1994).

The counterintuitive idea that has emerged from basic memory and skill learning literatures is that introducing difficulties and challenges during learning has desirable outcomes, such as promoting retention and transfer of learned material and more accurate metacognition regarding the degree of learning (Bjork, 1994; McDaniel & Butler, in press). These challenges might include interleaving of content rather than blocking content (e.g., Kornell & Bjork, 2008), spacing of content rather than massing it, and generation of content rather than reception of content (e.g., McDaniel, Waddill, & Einstein, 1988).

Evidence

There is a body of basic research supporting the idea that introducing difficulties during learning can increase long-term retention and transfer (Bjork, 1994; McDaniel & Butler, in press). Additionally, difficulties that create disfluency (e.g., presenting target materials in a font that is difficult to read) have been shown to prompt individuals to engage more controlled problem-solving strategies and to think more abstractly (Alter & Oppenheimer, 2008). For instance, when people were presented with math problems for which a reflexive answer is incorrect, presentation in disfluent type fonts produced more correct answers (Alter, Oppenheimer, Epley, & Eyre, 2007). However, only a handful of research-based efforts have been directed at developing and evaluating educationally relevant desirable difficulties.

Rohrer and Taylor (2007) found that interleaving instruction and subsequent practice on different types of mathematics problems (computing volumes of different solids) produced better application (retention) of the solution procedures to new problems than did blocking the problems, even though initial performance on the practice problems was superior when the practice was blocked relative to interleaving (see Kornell & Bjork, 2008, for parallel effects with learning about artists' painting styles). Note that the standard practice for arranging practice problems in textbooks is to block practice problems by topic. More desirable from the perspective of promoting retention would be to mix and distribute practice problems from different procedures.

Regarding authentic classroom contexts, in an experiment using passages from an introductory psychology text, key terminology was remembered better when college students were required to generate the terms in the context of the assigned reading (from word fragments) relative to when students read the terms (DeWinstanley & Bjork, 2004). Once students had experienced better memory performance after generating than reading, on subsequent paragraphs these students remembered read terms as well as they remembered generated terms. This result may suggest that students acquire better strategies for encoding material once they have been forced to generate.

In science classes, students can be required to generate predicted outcomes prior to classroom demonstrations, rather than being told about the expected result. In the domain of physics, research has established that demonstrations typically do not enhance learning (Crouch, Fagen, Callan, & Mazur, 2004). To investigate the benefits of generating predictions, Crouch et al. conducted an experiment in a college physics class that contrasted end-of-semester exam performances relating to demonstrations for which students generated predictions relative to exam performances after standard demonstrations (no predictions were required). Confirming previous findings, relative to a no-demonstration condition, observation of the demonstration promoted no significant improvement in students' ability to explain the outcome of related (to the demonstration) physical situations on a test at the end of the semester. When students were required to generate predictions prior to the demonstrations, their exam performances improved significantly. Interestingly, a third condition in which students were required to discuss and evaluate the generated predictions after the demonstrations did not produce significant increases in exam performances relative to the generate-only condition. This pattern suggests that generation of predictions may be sufficient to stimulate students to ponder and evaluate the demonstration with regard to the targeted conceptual information.

Several prominent challenges exist for effectively implementing desirable difficulties in the classroom. One challenge is to identify presentation formats or tasks that are tractable and acceptable in school settings, but that nevertheless create some difficulty for the student in initial processing or learning of the target material. Another challenge is that the desirability of any particular difficulty will depend on a number of factors that vary in the educational environment (see McDaniel & Butler, in press). For instance, difficulty will not be desirable when the summative assessments are not sensitive to the processing stimulated by difficulty (Thomas & McDaniel, 2007) or when the learners' cognitive skills (and prior knowledge) are overly challenged by the difficulty (McDaniel, Hines, & Guynn, 2002). Accordingly, successful implementation of desirable difficulties may depend in part on instructors' sensitivities to whether students' skills are sufficient to accommodate difficulties that are introduced and whether the summative assessments reflect the learning that is enhanced by a particular difficulty.

Interleave Example Solutions and Problem-Solving Exercises

Typically, in mathematics and science courses, instructors present a slew of example problems, and then students are required to solve a set of related problems. Yet, experimental evidence shows that student learning is markedly enhanced when worked example solutions are alternated with problems that the student is asked to solve. For example, Sweller and Cooper (1985) found that eighth- and ninth-grade students who solved eight algebra problems (as students might have to do in a homework assignment) took more time to complete the problems and performed worse on a post-test than students who were given pairs of a solution example followed by a problem. Note that this interleaved condition required less generation of solutions or active solving of solutions than did the condition in which students solved eight problems, but the worked example–problem pairing nevertheless led to superior performance (see also Cooper & Sweller, 1987; Renkl, 2002).

The implication of this research is that the standard practice of a teacher presenting a set of solved problems followed by homework assignments on a set of problems could be improved. Instead, teachers might present one worked example and then have students (either in small groups or as individuals) solve a problem on their own. Then the teacher could orient the class to another worked example and give the students a second problem to solve.

Forge Understanding

Understanding is the foundation for assimilating new information, remembering that information, and applying it. Next, we mention several techniques that evidence shows will assist in stimulating greater understanding.

Ask Deep-Level Questions

Encouraging students to engage in self-explanation, often by posing "why" questions, promotes deep understanding (see, e.g., McDaniel & Donnelly, 1996). Once students have acquired basic knowledge about a topic of study, deeper explanations and understanding of the key concepts can be facilitated by questions that prompt deep explanations of the target concepts. These questions often involve asking "what if" and "how does X compare to Y"; they are intended to prompt "deep" explanations that relate causes and consequences, motivation of people (e.g., involved in historical events), and scientific evidence for particular theories. The questions and explanations can occur in the context of classroom instruction, discussion, and independent study.

Use Graphics With Verbal Descriptions

Augmenting text and verbal descriptions with relevant graphical presentations that illustrate key processes and concepts facilitates student understanding. For instance, scientific processes and how things work (such as disk brakes, volcanic eruptions, bicycle pumps) can be visually illustrated through diagrams and schematics. Experimental studies demonstrate that such schematics improve learning, including problem solving and application of target constructs (Mayer, 2009). These visual representations may help students construct a mental model that effectively supports deep understanding of the content. It is worth noting that the available evidence suggests that pictures or series of pictures can be as effective in promoting learning as animated narratives.

Abstract and Concrete Representations of Concepts

In introducing students to a concept, teachers can focus on concrete realizations of the concept or can render a more abstract representation of the concept. Researchers suggest limitations of relying exclusively on either approach. Learning with concrete objects facilitates initial understanding but does not necessarily foster transfer to related contexts (Resnick & Omanson, 1987), whereas introducing the concept at an abstract level may slow students' mastery, though application to novel contexts may be facilitated. Current approaches suggest incorporating both concrete and abstract representations of target concepts during instruction. One critical feature of this approach is that teachers guide students toward the relevant and shared components of the concrete and abstract representations. Another particular approach is that of "concreteness fading," wherein initial learning is supported with a concrete representation that is gradually replaced with a more abstract representation.

Analogy

A key component of understanding is activating and focusing relevant prior knowledge on new material. Theorists have suggested that use of analogy to activate familiar concepts (prior knowledge) in the service of understanding new concepts may facilitate classroom learning (e.g., Halpern, 1987). For instance, if students are taught that "memory operates like a library," then the aspects of a library that are familiar to students can be activated to understand that organization of memory is essential for efficiency of memory, and that locating information in memory may be a process of restricting search to a general topic (e.g., a floor within a library) and then individuating particular information within that topic (a particular book or set of books on that floor). Using educationally relevant content, laboratory experiments with college students have shown benefits of analogy for learning astrophysics concepts, especially when summative tests focus on inference-level responses (Donnelly & McDaniel, 1993).

Appropriate Summative Testing

The benefits of the instructional techniques outlined in the previous paragraphs (or for that matter any instructional technique) will hinge in part on the nature of the exams that are constructed to evaluate students' mastery of the material. In the basic memory literature, a well-established principle is that particular encoding (study) activities effectively enhance memory performance to the extent that the criterial task depends on the information/processing engaged during encoding (e.g., McDaniel, Friedman, & Bourne, 1978). For instance, a generation task that focuses the learner on interrelations among the target concepts in a text will produce benefits on a test of relational information but not on a test of details (relative to a no-generation control); by contrast, a generation task that focuses the learner on the details in a text will benefit the detail test but not the relational test (Thomas & McDaniel, 2007).

The implications for educational practice are straightforward: Exams need to be constructed to reflect the kind of skill and knowledge that are targeted in the instructional goals and activities. In practice, however, this transfer-appropriate principle is often not appreciated. Teachers may design study activities that engage their students in analysis and synthesis of the core materials but then unwittingly undermine the effectiveness of these activities by giving exams that focus on individual details (see McDaniel, 2007, for an authentic example). In a related vein, in science and math domains, in many cases exams focus on students recording correct answers (i.e., a student's score depends on how many correct answers were provided). More effective for evaluating understanding and transfer of knowledge is to also focus on the thinking processes and approach that students use in arriving at their answers. The idea is to have students "show their work" (externalize thinking processes) and to give them feedback on the validity of the approach, not just on their final answer.

Another important consideration in effective use of exams is whether to include cumulative testing. Theoretically, cumulative testing is beneficial because it produces spacing of the material and it provides additional opportunities for active retrieval of target material (relative to giving a single unit exam on target content), both of which should contribute to long-term retention of key material. Some researchers have noted that when students are only given unit exams, students often comment after an exam that they no longer have to worry about that material. In some sense, the material may be treated like that in laboratory experiments in which subjects are directed to forget some studied items, which results in poorer retention for those items (e.g., Szpunar, McDermott, & Roediger, 2007).

Conclusions

Research in psychology and education has pointed to many different means of designing courses to maximize student learning. There is not sufficient evidence at this time to claim that a particular instructional technique is "best," and so the challenge is for instructors to choose from among these many techniques. Some of the earlier recommendations require small changes to existing courses. Some researchers have added small changes such as generating predictions prior to demonstrations (Crouch et al., 2004) or various hands-on activities (Cobern et al., 2010) and found that such modest changes may be sufficient to significantly increase learning and transfer. Accordingly, we believe that it may be possible for instructors to foster significant gains in student learning without dramatic changes in their current teaching methods. Incorporating the relatively modest kinds of changes like those suggested in this chapter may be sufficient to stimulate and enhance student learning.

References

Alter, A. L., & Oppenheimer, D. M. (2008). Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). *Psychological Science, 19*, 161–167.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*, 569–576.

Bjork, R. A. (1994). 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.

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

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

Butterfield, B., & Metcalfe, J. (2001). Errors made with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1491–1494.

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.

Cobern, W. W., Schuster, D., Adams, B., Applegate, B., Skjold, B., Undreiu, A., et al. (2010). Experimental comparison of inquiry and direct instruction in science. *Research in Science & Technological Education, 28*(1), 81–96.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79,* 347–362.

Crouch, C. H., Fagen, A. P., Callan, J. P., & Mazur, E. (2004). Classroom demonstrations: Learning tools or entertainment? *American Journal of Physics, 72*, 835–838.

Daniel, D. B., & Broida, J. (2004). Using web-based quizzing to improve exam performance: Lessons learned. *Teaching of Psychology, 31*, 207–208.

DeWinstanley, P. A., & Bjork, R. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory and Cognition, 32*, 945–955.

Donnelly, C. M., & McDaniel, M. A. (1993). Use of analogy in learning scientific concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 975–987.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33*, 545–565.

Halpern, D. F. (1987). Analogies as a critical thinking skill. In D. E. Berger, K. Bezdek, & W. P. Banks (Eds.), *Applications of cognitive psychology: Problem solving, education, and computing* (pp. 75–86). Hillsdale, NJ: Lawrence Erlbaum.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*, 469–486.

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

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 989–998.

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

Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, *29*, 210–212.

Lyle, K. B., & Crawford, N. A. (2011). Retrieving essential material at the end of the lecture improves performance on statistics exams. *Teaching of Psychology, 38,* 94–97.

Mayer, R. E. (2009). *Multimedia learning.* (2nd ed.) New York: Cambridge University Press.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., et al. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34*, 51–57.

McDaniel, M. A. (2007). Rediscovering transfer as a central concept. In H. L. Roediger, Y. Dudai, & S. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 267–270). New York: Oxford University Press.

McDaniel, M. A., & Butler, A. C. (in press). A contextual framework for understanding when difficulties are desirable. In A. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork*. New York: Taylor & Francis.

McDaniel, M. A., & Callender, A. A. (2008). Cognition, memory, and education. In J. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 819–844). Oxford, UK: Elsevier.

McDaniel, M. A., & Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology*, *88*, 508–519.

McDaniel, M. A., Friedman, A., & Bourne, L. E., Jr. (1978). Remembering the levels of information in words. *Memory & Cognition, 6,* 156–164.

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

McDaniel, M. A., Waddill, P. J., & Einstein, G. O. (1988). A contextual account of the generation effect: A three-factor theory. *Journal of Memory and Language*, *27,* 521–536.

McDaniel, M. A., Wildman, K. M., & Anderson, J. L. (2010). *Using quizzes to enhance summative-assessment performance in a web-based class: An experimental study.* Manuscript under review.

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction, 12,* 529–556.

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.). *Advances in instructional psychology* (Vol. 3, pp. 41–95). Hillsdale, NJ: Lawrence Erlbaum.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science, 17*, 249–255.

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20*, 1209–1224.

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35*, 481–498.

Sawyer, R. K. (Ed.). (2006). *Cambridge handbook of the learning sciences*. New York: Cambridge University Press.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*, 59–89.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory & Cognition, 35*, 1007–1013.

Thomas, A. K., & McDaniel, M. A. (2007). Metacomprehension for educationally relevant materials: Dramatic effects of encoding-retrieval interactions. *Psychonomic Bulletin & Review*, *14*, 212–218.