

Using Testing to Improve Learning and Memory

Henry L. Roediger, III, Kathleen B. McDermott and Mark A. McDaniel

Washington University in St. Louis

Address correspondence to:

Henry L. Roediger, III
Department of Psychology – Box 1125
Washington University in St. Louis
One Brookings Drive
St. Louis, MO 63130-4899

e-mail: roediger@wustl.edu

Telephone: 314-935-4307

Fax: 314-935-7588

How do you prepare for a test? If you are like most students, rereading your textbook – accompanied by underlining or highlighting – is your preferred study strategy. Perhaps you underline or highlight important points the first time you read a chapter, and maybe you reread critical parts that you marked earlier. Likely you reread your class notes, too. If these study strategies seem familiar, they are not unreasonable. Repetition often aids memory. But, unfortunately, the strategies are not particularly effective. In fact, research shows that reading has less influence on later memory than you might expect. In some studies, rereading was *no more effective* in improving learning than a single, initial reading (Callender & McDaniel, 2009).

Research by cognitive and educational psychologists has shown that effective study strategies exist, although most require more effort than simple rereading. Relating new information to what we already know, a strategy called *elaboration*, can make the material meaningful and, therefore, memorable. Similarly, when we outline or take notes, we recode the information – putting it in our own words – and much research shows that such active generation improves retention. Converting verbal information into vivid mental images can also improve learning. Spacing study sessions rather than massing them together also aids long-term retention. Research evidence strongly supports the effectiveness of these and many other active inquiry strategies in mastering information (McDaniel & Callendar, 2008).

Research in our laboratories at Washington University in St. Louis has led us to advocate an additional strategy, which we call *test-enhanced learning*. As you read text material, make up questions (or use the ones supplied at the end of the chapter) and then later test yourself. If you can retrieve the information from memory, great; having

retrieved it once, you will remember it even better. If you can't retrieve the answer, study the material again and retest yourself until you're sure you know it. But even if you were able to retrieve the answer once, don't stop. Test yourself repeatedly and keep retrieving answers. Repeated retrieval is the key to long term retention (Karpicke & Roediger, 2007).

Our advice regarding test-enhanced learning represents the bottom line of this essay. In the next few pages, we'll describe some evidence that supports our advice (McDaniel, Roediger & McDermott, 2007; Roediger & Karpicke, 2006a, b).

Test-Enhanced Learning

Most students and teachers think of tests as assessment devices, designed to measure what students have learned and to provide a basis for assigning grades. In many college courses, especially large introductory courses, tests are infrequent – occurring perhaps 2 or 3 times a semester. Assessment is a perfectly valid reason for testing, but, as we have said, it can also improve learning.

Consider the powerful *testing effect* we recently demonstrated in an experiment comparing three groups who were asked to use different methods to study brief prose passages such as what you might encounter in a textbook. All participants read the passage once. Then one group (SSSS) studied the passage three more times; a second group (SSST) studied the passage two more times and then took one test; a third group (STTT) took tests on the passage three successive times after reading it once. When tested, participants simply recalled as much as they could, but because the tests were relatively immediate and the passages were reasonably short, they were able to recall

about 70 percent of the information. Importantly, no feedback was given on test results (Roediger & Karpicke, 2006b).

Then participants took a break. For some participants, the break lasted only five minutes; others were asked to come back to the laboratory after a week. How did the different study conditions affect later recall? As you can see in Figure 1, when tested after 5 minutes, the more students had studied, the better they did – evidence that that cramming just before the exam can actually work! However, after a week, the outcome was completely reversed. Now students who had studied only once but recalled the material three times did best; those who had studied it three times and were tested once were next best; and those who had studied material the most but were never tested performed worst. The clear implication is that for long-term retention – the kind educators hope to foster and students seek – repeated retrieval of material is better than reading it over and over.

The finding that the act of taking a test can enhance learning may seem surprising, but convincing supporting evidence has been reported in many different experiments over the years (see Roediger & Karpicke, 2006a for a review of the literature). In fact, you can probably find confirmation in your own elementary school experience. Think back: what strategies were suggested for learning multiplication tables, foreign language vocabulary, or other systems of material? You probably used some type of flash card with the problem or word visible and the answer hidden. Did you practice over and over until the information was so well learned that answers would immediately spring to mind? Such self-testing can be adopted for almost any type of material and can create

effective means for learning large amounts of information and retaining it for long periods of time.

The Importance of Feedback

Participants in the test-enhanced learning experiment did not receive feedback after taking the test. As we noted, however, the first tests were given under conditions in which memory performance was reasonably good. Thus, we wondered whether the testing effect would occur under conditions where memory performance is mediocre or poor. Interestingly, the answer is *no*, or at least *not always*. If students are given a test on a large amount of material or if they are given a test long after the material has been studied, the testing effect disappears (Kang, McDermott, & Roediger, 2007). In fact, if performance on the test is very poor, the test offers no benefit at all.

Suppose a test, given under conditions of poor recall, is followed by feedback (that is, correct answers)? How will performance on a retest compare with a condition in which untested participants simply read correct answers in advance? To answer that question, we asked students to read four brief articles about topics in psychology. Each article had a different follow-up: a multiple-choice test, a short-answer test, a reading task involving some of the material that was tested in other conditions; or no activity (the control condition). The design of the study is shown in Figure 2. Because they were unaware there would be follow-up, participants were presumed to have studied the articles in the same way (Kang et al., 2007, Experiment 2). To further account for unexpected influences, information in the questions was carefully controlled. For example, one short-answer question was “What is hostile media bias?” The multiple-choice question was the same, but four possible answers were provided. The reread-

condition statement was “Hostile media bias refers to the phenomenon in which people on both sides of a controversy perceive the media as being hostile to their group.” Students who took the short-answer or multiple-choice test received feedback that consisted of statements like those presented in the reread condition. Students in the reread condition read statements for a third time to equalize exposure to the material. All participants took a final short-answer test (sort of like a final exam) three days later.

As you can see in Figure 3, a short-answer test resulted in the best performance, with the multiple-choice test next best, although the difference between simply reading statements and taking a multiple-choice test was not very great. Not surprisingly, the no-review condition resulted in the worst performance. However, the relatively good performance in the rereading condition may reflect a flaw in the comparison condition: Participants read only statements that would later appear on the test – as if, for example, they had gotten hold of a copy of the final exam. Nevertheless, after a delay, testing still led to better performance than rereading, just as it had before. Performance on the short-answer test with feedback was better than on the multiple-choice test with feedback, perhaps because of the extra thinking involved in generating answers.

You may be thinking, “But wait! Maybe if a multiple-choice test had been the final test, then taking a prior multiple-choice test would have produced better performance than taking a short answer test.” We thought of that too, and so we included in our experiment conditions in which multiple-choice items made up the final test. We found that an intervening short-answer test still produced the best performance (Kang et al., 2007). The greater retrieval effort required to produce information on a short answer

test (relative to recognizing it on a multiple choice test) probably confers the greater benefit.

In other studies, we showed that delaying feedback a bit after a question is more effective than giving it immediately (Butler, Karpicke & Roediger, 2008). Most teachers who provide feedback probably do this anyway. Quizzing can also improve retention of related but untested information (Chan, McDermott & Roediger, 2006). If verified, this finding has important implications for learning: Because only a small amount of material can be quizzed, a large part of the content of college courses is not tested. The possibility that self-testing might aid in the retention of material instructors don't get around to quizzing is a comforting notion.

Does Testing Work in the Classroom?

You may be wondering whether our laboratory research evidence has real-world relevance. After all, college students study for varying amounts of time and in different ways, and the time between classroom quizzes and a final test are much longer than they are in typical lab studies. We wondered, too. And the answer is *yes*, at least in the first studies that have asked this question.

A Web-based course on brain and behavior at the University of New Mexico was the site of one of these experiments (McDaniel et al., 2007). Students in the course, who were assumed to have completed 40-page weekly reading assignments, were assigned to the now-familiar four conditions: short-answer quizzes, multiple-choice quizzes; the opportunity to reread, and no intervention at all (control condition). Because this was a real class, of course, some students may have read the assignment multiple times, some once, and some not at all. Feedback was given after the quizzes and again for the third

(read facts) condition in which students read the facts a second time to equate for exposure.

After three assignments had been completed, students took a multiple-choice unit test. The content was the same but the wording of the questions was changed, thus requiring different answers. For example, in the multiple-choice quiz students were asked, “All preganglionic axons, whether sympathetic or parasympathetic, release _____ as a neurotransmitter.” (The correct answer is acetylcholine). On the unit exam, the question was “All _____ axons, whether sympathetic or parasympathetic, release acetylcholine as a neurotransmitter.” Thus, students could not memorize answers from the quiz, but had to know the whole fact. Students in the two groups that had been quizzed did better on the unit exams (54% correct) than those in the other two groups (46% correct). Again, the short-answer format offered a better opportunity for retention than did the multiple-choice format, with the latter format leading to only slightly better retention than rereading, and rereading proving superior to no re-exposure at all. In short, the results of the classroom study agree remarkably well with those of the laboratory study, as well as with other laboratory results (e.g., Butler & Roediger, 2007). Keep in mind that the lab experiment used short-answer items as the final test, whereas the classroom experiment used multiple-choice questions. We find it reassuring that the same pattern of results holds with both types of test.

In current research in three middle-school classrooms, we have also found strong testing effects. Students in social studies, for example, achieved scores about ten points higher on material on which they had been quizzed than on material that was not quizzed (Roediger et al, 2007). Quizzing helped raise grades a letter, from B- to A- levels.

Practical Implications for How to Study

One of us (KBM) teaches an undergraduate course on human memory, which meets twice a week for 1.5 hours. As in most classes, readings are assigned for that day, and the lecture builds on and supplements the readings. However, unlike most classes, the last 5-10 minutes of the class are devoted to a quiz, with questions drawn from everything assigned as a reading for that day and from the lecture. Students also take three regular tests and a final exam. The news about the quiz schedule always leads a few faint-hearted students to drop the course. However, by the end of the semester, students' reactions are mostly quite positive.

Of course, as a student, you are in no position to tell the instructor how many tests to give, and it would take an unusually brave student to even broach the idea in class. How can you use the information we have provided to craft more effective study strategies for yourself? In providing advice, we draw on research not discussed in this essay (see McDaniel & Callendar, 2008, for a summary) to provide advice on how to read your textbook and get the most from it, how to best use lectures and lecture notes, and how to review prior to tests.

Reading a textbook may seem obvious and easy, but effective reading requires a strategy sometimes called the PQ4R technique: *Preview, Question, Read, Reflect, Recite, Review*.

- Preview. Skim through the chapter looking at the headings and the figures to see what topics will be covered. If there are summaries, read them first to see where you will be heading. Also, see if there is a set of key terms and review questions at the end of the chapter, and possibly other study aids.

- Question, Read. If there were no key terms and no questions, make them up as you go along. List the important terms and write the definition.
- Reflect. Read with an inquiring mind: Try to convert headings into questions that will be answered. So, if one heading is “Classical conditioning” ask yourself “What is classical conditioning?” and your reading will answer that question. Jot down these questions to help you with later study and review. Some students find it worthwhile to outline chapters as they are reading, which forces them to think hard about the relationship among section of a chapter and produce it in their own words.
- Recite, Review. Reading carefully takes longer than simply blasting through the chapter with a highlighter, but studies show that students remember the material much better when they have carefully considered it this way. At the end of the chapter, review the key terms (either provided in the book or collected by you). Look up and restudy any terms about which you are shaky. If you can’t remember it just after you have read it, what chance will you have on a test later? Similarly, review the study questions and try to answer them. Again, review (or give yourself feedback) on questions you cannot answer.

Paying careful attention and taking good notes in class is an art that some students master and others do not. We suggest regular class attendance, focused listening, copious note-taking, and careful attention to the instructor’s suggestions. These steps may seem obvious, but, surprisingly, many students view them as optional. After a class, read over your notes carefully. You were writing fast, at least some of the time, so you should go

over the notes while the lecture is still fresh in mind and flesh them out. Or, if taking notes on your laptop, you may need to edit them to extract key points.

Treat your notes the same way that you read the text: cull out key terms and think of possible review questions. What the teacher covers in class is often what she or he considers most important and is therefore likely to include on the test. For some parts of the lecture that you might have found confusing, consult your notes, the Internet, or your instructor during office hours.

If you have thoughtfully studied your text and class notes, preparing for exams should be relatively easy. But, as should be clear by now, repeatedly rereading your text and notes is not a particularly effective strategy. Pre-exam time is the moment to put into practice our *test-enhanced learning* method. Test yourself on key terms and definitions. Examine the review questions in the book or the ones you created. If you are unsure of the meaning of a term or can't answer a question, review the book, your lecture notes, or both. Keep questioning yourself until you have the information down. By testing yourself, you can help gauge whether or not you know the material.

Repeated rereading can make information seem quite familiar without making it easy to retrieve when you need it – on the test. Repeated retrieval is the key to long-term retention (Karpicke & Roediger, 2007). So, even if you think you know the answers to questions, if you keep practicing retrieving them at spaced intervals, you will remember them better over the long term. Self-testing not only aids learning, but it also lets you know what you do not know and where you need to focus your efforts.

Conclusion

Repeated retrieval makes information more likely to be accessible in the future. Testing in the classroom accomplishes this goal, but so does studying by self-test. As William James wrote (1890, p. 646), “A curious peculiarity of our memory is that things are impressed better by active than by passive repetition. I mean that in learning (by heart, for example), when we almost know the piece, it pays better to wait and recollect by an effort from within, than to look at the book again. If we recover the words in the former way, we shall probably know them the next time; if in the latter way, we shall very likely need the book once more.”

Acknowledgements

The research reported here was supported by a James S. McDonnell Foundation 21st Century Science Initiative in Bridging Brain, Mind and Behavior/Collaborative Award and by the Institute of Education Sciences, U.S. Department of Education, through Grant #305H060080 to Washington University in St. Louis. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

- Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*, 514-527.
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). 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.
- Callender A. A. and McDaniel M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology, 34*, 30-41.
- Chan, C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval induced facilitation: Initially nontested material can benefit from prior testing. *Journal of Experimental Psychology: General, 135*, 533-571.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- Kang, S., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modulate the effect of testing on memory retention. *The European Journal of Cognitive Psychology, 19*, 528-558.
- Karpicke, J. D. & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*, 151-162.

- McDaniel M.A., Anderson J.L., Derbish M.H., and Morrisette N. (2007) Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*, 494-513.
- McDaniel, M. A., & Callender, A. A. (2008). Cognition, memory, and education. (pp. 819-843). In H. L. Roediger (Ed.), *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference*, 4 vols. Oxford:Elsevier.
- McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*, 200-206.
- Roediger, H. L., & 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., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249-255.
- Roediger, H.L., McDaniel, M.A., McDermott, K.B., & Agarwal, P.K. (2007, November). Testing-enhanced learning in the classroom: The Columbia Middle School project. Poster presented at the Psychonomic Society meeting, Long Beach, CA.

Figure Captions

1. Proportion of ideas recalled after 5 minutes and after one week as a function of three experimental conditions. The data on the left side show recall on the final test given after 5 minutes; the data on the right, results after a week.
2. The design of the Kang, McDermott and Roediger (2007) experiment. Students studied a passage and then either took a short-answer test, took a multiple choice test, reread certain facts, or had no activity related to the passage. They took a final test three days later.
3. Performance on the final short-answer test is shown as a function of the activity participants performed after reading. The immediate short-answer test with feedback produced the best performance on the final test. Data are from Kang et al. (2007).

Figure 1

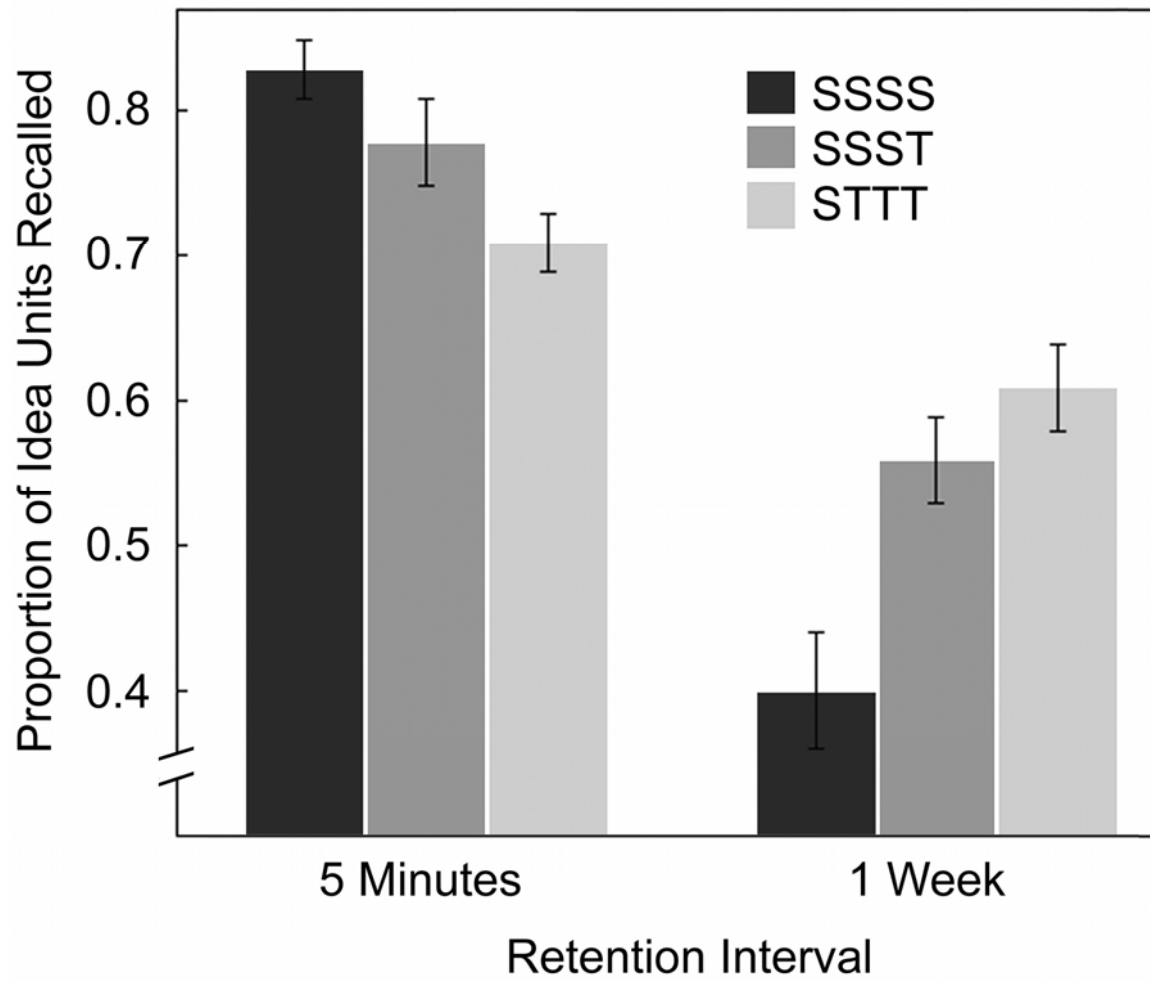


Figure 2

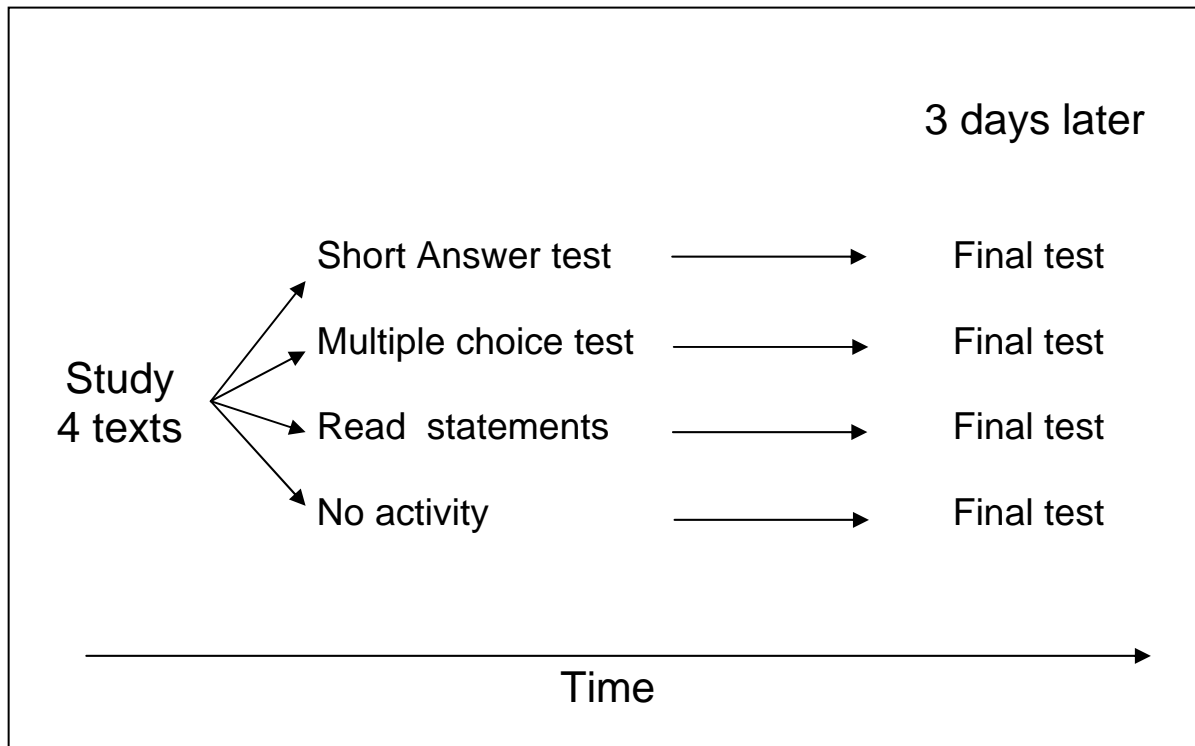


Figure 3

