Psychonomic Bulletin & Review
2007, 14 (2), 187-193

Enhancing learning and retarding forgetting: Choices and consequences

Harold Pashler
University of California, San Diego, La Jolla, California

Doug Rohrer
University of South Florida, Tampa, Florida

Nicholas J. Cepeda
University of California, San Diego, La Jolla, California
and University of Colorado, Boulder, Colorado

and

Shana K. Carpenter
University of California, San Diego, La Jolla, California

Our research on learning enhancement has been focusing on the consequences for learning and forgetting of some of the more obvious and concrete choices that arise in instruction, including questions such as these: How does spacing of practice affect retention of information over significant retention intervals (up to 1 year)? Do spacing effects generalize beyond recall of verbal materials? Is feedback needed to promote learning, and must it be immediate? Although retrieval practice has been found to enhance learning in comparison with additional study, does it actually reduce the rate of forgetting? Can retrieval practice effects be extended to nonverbal materials? We suggest that as we begin to find answers to these questions, it should become possible for cognitive psychology to offer nonobvious advice that can be applied in a variety of instructional contexts to facilitate learning and reduce forgetting.

Scope of Our Research Program

The potential of research in learning and memory, and in cognitive psychology generally, to improve instructional techniques has been discussed for decades. However, it is rather disconcerting to note how few examples exist of actual translation from cognitive science research into classrooms or learning technologies. Why is this? One factor may be pernicious intellectual fashions within the field of education, where empirical testing is sometimes regarded as "naive positivism" rather than an essential precondition for rational practice (Carnine, 2000). Nonetheless, before blaming practitioners, it might be reasonable to begin with a question closer to home: Has memory and learning research provided many results that have nonobvious and concrete implications for instructional procedures?

A brief perusal of cognitive psychology textbooks might leave one unsure. The finding that seems to be most widely cited as having practical relevance to instruction is the benefit of *elaborative encoding* on long-term memory storage (see, e.g., Hyde & Jenkins, 1973). Although the validity of this principle is not in doubt, it seems not to have provided much nonobvious or concrete guidance for practitioners. Along with other writers represented in this special issue, we have been seeking to add to the stock of useful information. Our strategy is to look for key *choices* that arise in designing instructional procedures—choices that might well affect the success and durability of learning, but whose impact is not intuitively obvious. Interestingly, this search often leads us to questions that drew more attention during an earlier era of psychology (see, e.g., Starch, 1927) than they have in recent years (even though, we contend, some of them have implications for issues of much current theoretical interest; see, e.g., Mozer, Howe, & Pashler, 2004).

The present article gives an overview of our main results to date, focusing on four broad themes: the effects of temporal distribution of learning (spacing), the form and timing of feedback, the effects of testing (retrieval practice), and the consequences of guessing when a learner is not sure.

Spacing of Practice: Temporal Variables

One of the most basic choices that faces a learner or instructor is when study should take place. Research on the temporal distribution of practice goes back at least as far as Ebbinghaus (1885/1964) and is the subject of hundreds of

H. Pashler, hpashler@ucsd.edu

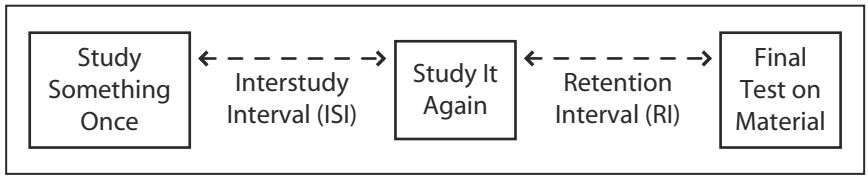

Figure 1. The basic design of a spacing experiment. Subjects have two opportunities to learn the same material, separated by an ISI. After an RI that is measured from the second learning episode, a final test is given. A spacing experiment most typically has one RI and several values of ISI.

articles. One might even assume the topic had been "studied to death." If so, practical payoffs have been strangely elusive. In 1988, Frank Dempster published an article on spacing in *American Psychologist* subtitled "A Case Study in the Failure to Apply the Results of Psychological Research," a description that remains apt to this day. Whether one looks in classrooms, instructional design texts, or at current instructional software, one finds little evidence that anyone is paying attention to the temporal distribution of study. Moreover, programs that deliberately compress learning into short time spans (immersion learning, summer boot camps) seem to be flourishing.

But exactly what practical advice about spacing can be given to practitioners on the basis of findings from the memory lab? Our research group recently performed a meta-analysis of the spacing literature (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006) and found that only a tiny proportion of spacing research has examined retention intervals as long as 1 day. Bahrick (e.g., in Bahrick, Bahrick, Bahrick, & Bahrick, 1993) carried out pioneering studies with longer intervals, but his subjects were trained to mastery on each learning session, allowing study time to increase with spacing. Therefore, though the literature is large, it seemed to us not to provide a strong basis for prescribing the most efficient allocation of study time. For that reason, we commenced several new lines of experiments.

In discussing spacing, we refer to the basic design shown in Figure 1. Here, the learner studies the same information on two occasions (S1 and S2), separated by an interstudy interval (ISI). After an additional retention interval (RI)—measured from S2—a final test is given. The literature involving short RIs reveals that the effects of varying ISI are often nonmonotonic in character, with final-test performance rising as ISI is increased from zero, and then falling as ISI is increased beyond an optimum value (e.g., Crowder, 1976; Glenberg & Lehmann, 1980).

To maximize the likelihood of obtaining results that will generalize to practical contexts, our new studies have used materials that seem representative of at least the simpler sorts of learning tasks that people undertake in daily life—materials such as facts, vocabulary, and the like. One of our first studies (Cepeda et al., 2006) used a 10-day RI and taught subjects Swahili–English word pairs. In Session 1, subjects learned pairs to a criterion of perfect performance (on every trial, the computer displayed Swahili, and the subject typed in English and received feedback). In Session 2, a fixed number of additional learning trials were given on the same word pairs. Increasing ISI from 15 min to 1 day improved final-test recall, in line with prior results using word lists (Edwards, 1917; Glenberg & Lehmann, 1980). However, as ISI increased beyond 1 day, final-test performance declined by a small amount.

Next, we moved on to a 6-month RI, teaching subjects little-known facts as well as the names of obscure visually presented objects (Cepeda et al., 2006). Here, final-test performance increased until ISI reached about 1 *month*, with a shallow drop beyond that. Hence, the optimal ISI appears to increase as RI increases, as was found in the literature involving short time intervals. However, it appears from our data that when RI is substantial, the optimal ratio of ISI to RI is not 1:1, as some have suggested (Crowder, 1976), but rather something closer to 10%–20% (see Figure 2).

To verify these conclusions within a single experiment, we are currently carrying out a much larger Web-based study using ISIs ranging from 20 min to 15 weeks and RIs ranging from 1 to 50 weeks. Again, subjects are learning relatively unfamiliar facts. Results to date (from about 1,800 subjects) suggest that when the retention interval is 1 week, the optimal ISI is about 1 day, but for a 50-week RI, an ISI of 3 weeks is best among the values we examine. Thus, the results are consistent with our earlier findings that optimal ISI increases as RI increases.

In sum, spacing clearly does have powerful effects on memory over substantial retention intervals. Moreover, test performance after a given RI is optimized when the ISI takes some intermediate value, although a longer-than-optimal spacing is not nearly as harmful to final memory as a shorter-than-optimal spacing. Our data imply that to promote retention over years, ensuring an ISI of several months or even a few years is likely to be far more effective than using shorter intervals.

Spacing effects in math problem solving. Do these spacing principles also govern learning tasks that go beyond the recall of atomic facts or associations? To explore one aspect of this issue, we have been examining the effect of spacing of practice on retention of mathematical skills. In one recent study, college students learned a simple (but unfamiliar) principle of combinatorics: how to determine the number of different orderings of a letter sequence with at least one repeated letter (Rohrer & Taylor, 2006). The students saw a tutorial and then worked 10 practice problems that were either massed into a single session or distributed over two sessions separated by 1 week. After attempting each problem, students were shown the com-

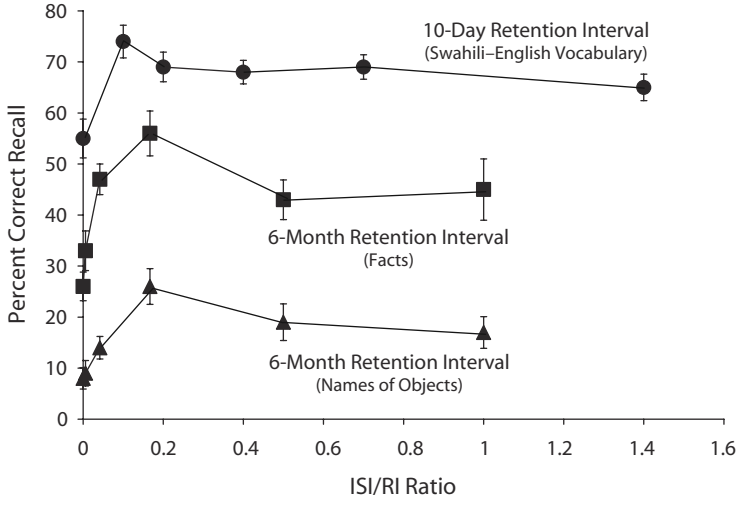

Figure 2. Results of two spacing experiments. This figure shows percentages of correct recall on the final test as a function of interstudy interval (ISI) divided by the retention interval (RI). The top line shows the first study discussed in the text, with a 10-day RI; performance peaks at a ratio of 0.1 (1-day ISI). The lower two lines show the 6-month RI studies; the peaks in both are at about a ratio of 0.2 (28-day ISI). Overall, the results suggest that for any given RI within this range, final-test memory is optimized for an ISI that is about 10%–20% of the RI. Note that, in every case, having too little spacing is worse than having too much. The error bars indicate ±1 standard error.

plete solution. A final test was given either 1 or 4 weeks after the last practice problem. Spacing had no effect at the 1-week retention interval, but it had a substantial effect at the 4-week interval (Figure 3). Spacing is evidently a potent variable for at least one form of math skill learning, and the interaction of ISI and RI seems broadly in line with the findings described earlier for fact memory. Interestingly, most current mathematics texts mass practice problems relating to a given topic into one problem set presented immediately following textual presentation of that topic. Our data suggest that—at least for promoting retention—this may be a grievous error.

Spacing in perceptual categorization learning. We have also looked at perceptual categorization learning, a task that—despite its prominence in the cognitive science literature—is almost absent from the spacing literature. In some of our studies, we taught subjects to categorize checkerboard patterns (as in Fried & Holyoak, 1984). We have observed no benefit of a 3-day ISI over a 10-min ISI for either 1 or 3 weeks' retention. We have also found no spacing benefits when subjects were taught to identify the genre and artist of relatively unfamiliar paintings (e.g., by Caravaggio, Buoninsegna, or Glackens) and were later tested on novel paintings by the same artists.

In much the same vein, with the assistance of a dermatologist, we created a Web site (www.learnmelanoma.org) that teaches people to discriminate benign from cancerous skin lesions, and within this framework we are comparing various spacing schedules. So far, 550 subjects have completed the study, and again we see little evidence of spacing effects.

In summary, spacing principles applicable to declarative memory tasks seem to extend beyond declarative memory for facts and associations to at least some forms of mathematics skill learning. However, perceptual categorization tasks seem not to show such effects, as far as we can tell. Evidently, much more research is needed to chart the boundaries of the effects.

Overlearning. Another practical choice that presents itself—closely related to spacing—is the option of *overlearning*: continuing to practice material after error-free performance is attained. Overlearning has been shown to increase later performance in comparison with smaller de-

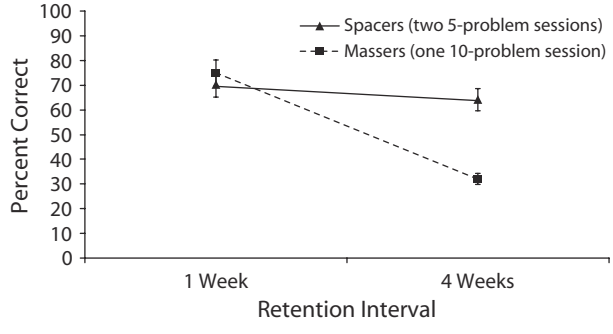

Figure 3. Spacing in mathematics learning. After learning how to solve a permutation task, students worked 10 practice problems that were either massed in one practice session or distributed across two sessions (separated by 2 weeks). The benefit of spacing grew with the retention interval. Error bars indicate ±1 standard error.

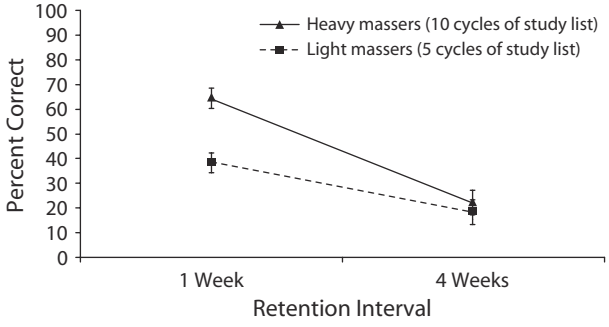

Figure 4. Overlearning vocabulary definitions. Students learned a list of word–definition pairs (e.g., *cicatrix–scar*) by cycling through the list and self-testing (as with flashcards) either 5 or 10 times. The benefit of heavily massed practice (i.e., overlearning) virtually disappeared as retention interval increased. Error bars indicate ±1 standard error.

grees of practice (see, e.g., Krueger, 1929) and has often been advocated as a generally useful learning strategy (e.g., Driskell, Willis, & Copper, 1992; Foriska, 1993). However, overlearning involves massed rather than spaced practice, which—for reasons described above—suggests that it might be an inefficient way to promote later memory.

To shed more light on this question, we assessed the gains produced by overlearning on tests given after varying retention intervals. In one study (Rohrer, Taylor, Pashler, Wixted, & Cepeda, 2005), college students learned novel vocabulary (e.g., *cicatrix–scar*), cycling through a list of word–definition pairs either 5 or 10 times. The extra 5 cycles yielded a substantial benefit after 1 week, but the gain was no longer apparent after 4 weeks (Figure 4). Likewise, in a similar experiment involving the combinatorics task described earlier, a threefold increase in the number of practice problems produced no overlearning gain at either retention interval (Figure 5; see Rohrer & Taylor, 2006). From a long-term perspective, overlearning appears to be inefficient almost to the point of wasting time.

Of course, there may sometimes be little alternative to overlearning a skill that might need to be performed at some unknown time without error (e.g., learning the Heimlich maneuver or how to land an airplane after engine failure). Furthermore, overlearning may enhance speed long after retrieval accuracy has reached ceiling (see, e.g., Logan & Klapp, 1991), and that speedup may sometimes be useful. These caveats aside, overlearning has the deficiencies of massed practice, and when the choice presents itself, our results suggest that overlearning will typically represent an inefficient use of study time.

Feedback

Another important choice faced by instructors is whether to provide feedback, and if so, when and in what form. Skinner (1968) and his followers (e.g., Vargas, 1986) argued that immediate feedback is crucial to promoting effective learning. However, in the classroom, students usually take tests and receive feedback much later, if at all. Therefore, if Skinner's hypothesis is right, the practical implications are enormous. From a very different perspective, other writers have argued that providing regular feedback may retard retention, even when it enhances performance during learning (Schmidt & Bjork, 1992).

To shed light on this issue, we had subjects learn Luganda–English word pairs (Pashler, Cepeda, Wixted, & Rohrer, 2005). An initial learning session consisted of two initial exposures to the materials, followed by several tests. The type of feedback accompanying the learning session tests was varied between subjects, and 1 week later, a final test was administered. When the subject made an error on an item during the learning session, providing feedback with the correct answer resulted in a roughly fivefold increase in the chance of successful recall on the final test. More impoverished forms of feedback, such as merely telling the subject that a response was right or wrong, accomplished little. On the other hand, when a subject correctly recalled an item during the learning session, providing or withholding feedback made essentially no difference. Oddly, even if the correct recalls were made with low confidence, withholding feedback seemed harmless. In subsequent studies, we have also looked at the effects of withholding corrective feedback from some tests of a given item, but not all. The learning curves have so far shown that withholding corrective feedback after an error is always harmful, even if done only intermittently.

What about timing of feedback? In one recent study, we had subjects learn obscure facts (e.g., *Alaska is the U.S. state with the highest percentage of people who walk to work*), followed by a test (*What is the state . . . ?*) and then feedback (*Alaska*). The test was given immediately or delayed 1 day, and the subsequent feedback was given either immediately after the test or delayed 1 day (i.e., a 2 × 2 design).

On a final test 2 weeks later (Figure 6), the groups that received delayed feedback performed better, not worse, than those that received immediate feedback, whether the test was immediate or delayed. The effect was largest for items the subjects answered correctly, but surprisingly,

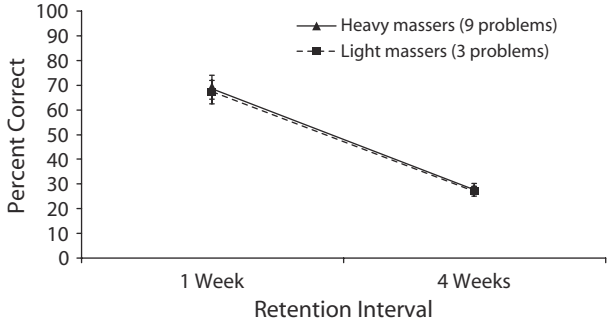

Figure 5. Overlearning mathematics. After learning how to solve a permutation task, students worked either three or nine practice problems within the same practice session. There was no benefit whatsoever of the additional massed practice (i.e., overlearning) at either the 1- or the 4-week retention interval. Error bars indicate ±1 standard error.

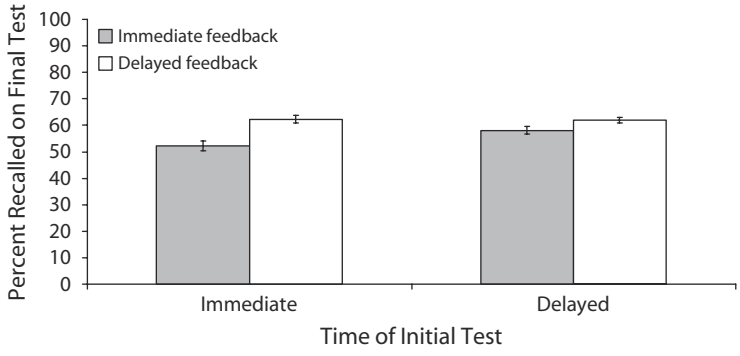

Figure 6. Effects of feedback timing and test timing on the learning of obscure facts. Subjects first read a list of facts and then were given a test of these facts either immediately or 1 day later. Feedback was given either immediately after each question on the test or 1 day later. On a final test 2 weeks later, those who received delayed feedback outperformed those with immediate feedback, regardless of whether the initial test was immediate or delayed. Far from being harmful, as one might infer from Skinnerian accounts, delaying feedback was in fact slightly helpful. Error bars indicate ±1 standard error.

similar trends were found even for errors. From these results, immediate feedback seems not at all essential for fact learning, and it may not even be optimal (presumably because delays provide spaced practice, at least after correct responses).

Retrieval Practice: Benefits From Tests

Prior research has shown that learning is often enhanced when the learner is required to recall information rather than simply restudying it (see Roediger & Karpicke, 2006a, for a review). This *testing* (or *retrieval practice*) *effect*—discussed by McDaniel, Roediger, and McDermott (2007) in the present issue—has been found in free recall (e.g., Allen, Mahler, & Estes, 1969; Carpenter & DeLosh, 2006) as well as in cued recall, including foreign language vocabulary learning (Carrier & Pashler, 1992), face–name learning (Carpenter & DeLosh, 2005), definitions (Cull, 2000), and general knowledge facts (McDaniel & Fisher, 1991).

In determining how best to exploit testing as an instructional device, one important issue that arises is whether the form of retrieval used in learning must be identical to the sort of later retrieval one hopes to promote. We started examining this question by looking at the *direction* of test in Swahili–English vocabulary learning (e.g., *kelb–dog*). Does practice recalling *dog* (after seeing *kelb* → *?*) facilitate later recall in the opposite direction (*?* → *dog*) in comparison with simply restudying the pair (*kelb–dog*)? We find that it does (Figure 7; see also Carpenter, Pashler, & Vul, 2006). We are even finding that covert retrieval practice (in which subjects are asked to retrieve without providing an observable response) suffices to enhance learning. These results encourage the idea that retrieval practice has broad practical potential.

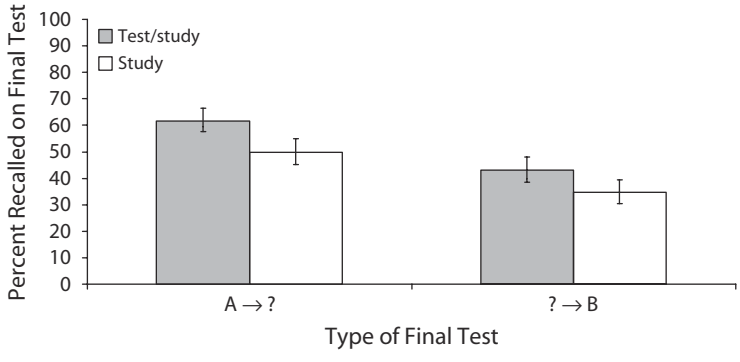

Figure 7. Is the retrieval practice benefit in cued recall (i.e., those in a testing-with-feedback condition outperforming those with pure study) confined to cases with a final test given in the same direction as the retrieval practice? Apparently, it is not. Subjects first studied word pairs (A–B) and were given either a cued-recall test (A → ?) or a restudy opportunity (A–B). The final test was cued recall in the same direction (A → ?) or the opposite direction (? → B). Recall in both directions benefited more from testing with feedback than from restudying. Error bars indicate ±1 standard error.

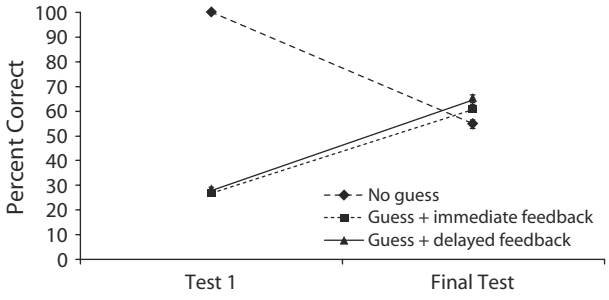

Figure 8. Effects of forced guessing on the learning of obscure facts. Subjects either read the fact along with the correct answer ("no guess"), guessed about the correct answer and were given immediate feedback, or guessed about the correct answer and were given delayed feedback. On Test 1, forced guesses are at the chance level, as expected. On a final test 1 week later, recall of the correct answers was not impaired by having guessed on those items (this held even when the initial guess was incorrect). Error bars indicate ±1 standard error.

Does retrieval practice attenuate forgetting? Some studies have found that the benefits of retrieval practice appear to grow with retention interval (e.g., Roediger & Karpicke, 2006b), possibly suggesting that retrieval practice slows the rate of forgetting (Wheeler, Ewers, & Buonanno, 2003). We have been examining this issue using a formal analysis of forgetting functions (Carpenter, Pashler, Wixted, & Vul, 2006). In one study, subjects studied obscure facts and then encountered each fact again in either a cued-recall test (with feedback) or an additional study presentation (question + answer), as in Carrier and Pashler (1992). Different items were tested after 5 min or 1, 2, 7, 14, or 42 days. The power function $y = a(bt + 1)^{-c}$ was fit to each subject's data to estimate the degree of learning (a) and the rate of forgetting (c) associated with testing versus restudying. Testing increased the degree of learning in comparison with restudying and reduced the rate of forgetting.

Retrieval practice and nonverbal tasks. Retrieval practice effects have been studied almost entirely with verbal materials. Seeking to assess the generality of these effects, we have begun investigating retrieval practice in learning of maps. In one recent study (Carpenter & Pashler, in press), subjects studied two maps (each depicting about a dozen land features—e.g., roads and rivers), using either conventional study or a covert retrieval procedure. In that procedure, subjects were repeatedly shown the same map with one land feature deleted and asked to covertly retrieve an image of the missing feature in the map. When subjects reported having done so as best they could, the computer showed them the intact map again, and the test–feedback cycle continued (always testing with a different feature). On a final test, subjects were asked to draw the full maps. Drawings were better and more complete when learned through covert retrieval rather than additional study. Thus, we are optimistic that covert retrieval may be extended to various other nonverbal learning tasks with practical significance.

Forced Guessing: Is It Harmful?

As described above, retrieval practice often seems to be a useful learning strategy. However, if one seeks to utilize retrieval practice, one must often ask the learner a question that he or she cannot answer. If the learner tries to answer and produces the wrong response, will this undermine learning, as some theorists have suggested (e.g., Guthrie, 1952)?

To assess this issue, one of our recent studies began by asking subjects to answer very difficult trivia questions (e.g., *The weight of what land mammal is equivalent to the weight of a blue whale's tongue?*) by choosing from among four plausible answers [e.g., (a) *Bengal tiger*, (b) *Grizzly bear*, (c) *Wolverine*, (d) *African elephant*]. For one-third of the questions, the correct answer (here, *African elephant*) was highlighted at the outset. For another third, subjects were required to guess and then given corrective feedback. For the remaining third, subjects guessed and were given feedback only at the end of the session (Figure 8). Even when initial guesses were wrong and feedback was delayed, forced guessing did not impair learning.

In this study and others, we were unable to find any costs associated with guessing when completely unsure. It will be interesting to see whether future research turns up exceptions to this finding.

Summary and Conclusions

Our results can be summarized as follows. We find that over substantial time periods, spacing has powerful (and typically nonmonotonic) effects on retention, with optimal memory occurring when spacing is some modest fraction of the final retention interval (perhaps about 10%–20%). These benefits seem to generalize to math skills, but not—as far as we can tell—to perceptual categorization. Retrieval practice appears to enhance initial learning and slow forgetting. Retrieval practice can also be extended well beyond overt retrieval of verbal responses. Feedback seems to be quite essential to the learning of facts—but only after errors. Furthermore, this feedback need not be immediate. Finally, guessing when the learner is completely unsure, which usually results in an error, seems quite harmless as long as feedback is eventually provided. Naturally, these conclusions are all preliminary, and it will be useful to learn more of their boundary conditions. Nonetheless, these findings—along with those described in the other articles within this special issue—seem to provide encouragement that over the next few years, as various groups join in exploring such intriguing and often neglected questions, we may finally be able to provide a fuller response to the challenge "What nonobvious advice can cognitive psychologists offer about how best to go about learning?"

AUTHOR NOTE

This work was supported by the Institute of Education Sciences (U.S. Department of Education, Grant R305H040108). We thank Kelli Taylor, John Wixted, and Ed Vul for useful discussions. Ed Vul and David Cun provided expert programming assistance in this project. Correspondence concerning this article should be addressed to H. Pashler, Department of

Psychology, 0109, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0109 (e-mail: hpashler@ucsd.edu).

REFERENCES

ALLEN, G. A., MAHLER, W. A., & ESTES, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning & Verbal Behavior*, **8**, 463-470.

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.

CARNINE, D. (2000). Why education experts resist effective practices (and what it would take to make education more like medicine). Thomas B. Fordham Foundation, April 1, 2000. Available at www.edexcellence.net/foundation/publication/publication.cfm?id=46.

CARPENTER, S. K., & DELOSH, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, **19**, 619-636.

CARPENTER, S. K., & DELOSH, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, **34**, 268-276.

CARPENTER, S. K., & PASHLER, H. (in press). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*.

CARPENTER, S. K., PASHLER, H., & VUL, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review*, **13**, 826-830.

CARPENTER, S. K., PASHLER, H., WIXTED, J. T., & VUL, E. (2006). *The effects of tests on learning and forgetting*. Manuscript submitted for publication.

CARRIER, M., & PASHLER, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, **20**, 633-642.

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.

CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

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

DEMPSTER, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, **43**, 627-634.

DRISKELL, J. E., WILLIS, R. P., & COPPER, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, **77**, 615-622.

EBBINGHAUS, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885)

EDWARDS, A. S. (1917). The distribution of time in learning small amounts of material. In W. B. Pillsbury, J. W. Baird, & M. F. Washburn (Eds.), *Studies in psychology: Contributed by colleagues and former students of Edward Bradford Titchener*. Worcester, MA: Louis N. Wilson.

FORISKA, T. J. (1993). What every educator should know about learning. *Schools in the Middle*, **3**, 39-44.

FRIED, L. S., & HOLYOAK, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 234-257.

GLENBERG, A. M., & LEHMANN, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition*, **8**, 528-538.

GUTHRIE, E. R. (1952). *The psychology of learning* (Rev. ed.). New York: Harper.

HYDE, T. S., & JENKINS, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning & Verbal Behavior*, **12**, 471-480.

KRUEGER, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology*, **12**, 71-78.

LOGAN, G. D., & KLAPP, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 179-195.

MCDANIEL, M. A., & FISHER, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology*, **16**, 192-201.

MCDANIEL, M. A., ROEDIGER, H. L., III, & MCDERMOTT, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, **14**, 200-206.

MOZER, M. C., HOWE, M., & PASHLER, H. (2004). Using testing to enhance learning: A comparison of two hypotheses. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 975-980). Mahwah, NJ: Erlbaum.

PASHLER, H., CEPEDA, N. J., WIXTED, J. T., & ROHRER, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 3-8.

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.

ROHRER, D., & TAYLOR, K. (2006). The effects of overlearning and distributed practise on the retention of mathematics knowledge. *Applied Cognitive Psychology*, **20**, 1209-1224.

ROHRER, D., TAYLOR, K., PASHLER, H., WIXTED, J. T, & CEPEDA, N. J. (2005). The effect of overlearning on long-term retention. *Applied Cognitive Psychology*, **19**, 361-374.

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.

SKINNER, B. F. (1968). *The technology of teaching*. New York: Prentice Hall.

STARCH, D. (1927). *Educational psychology*. New York: Macmillan.

VARGAS, J. S. (1986). Instructional design flaws in computer-assisted instruction. *Phi Delta Kappan*, **67**, 738-744.

WHEELER, M. A., EWERS, M., & BUONANNO, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory*, **11**, 571-580.