Title
Enhancing learning and retarding forgetting: Choices and consequences

Journal
Psychonomic Bulletin & Review, 14

Authors
Pashler, Harold
Rohrer, Doug
Cepeda, Nicholas
et al.

Publication Date
2007

Peer reviewed

Enhancing Learning and Retarding Forgetting: Choices and Consequences

HAROLD PASHLER[1], DOUG ROHRER[2],
NICHOLAS J. CEPEDA,[1,3] AND SHANA K. CARPENTER[1]

[1]University of California, San Diego *[2]University of South Florida*
[3]University of Colorado, Boulder

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: How does spacing of practice affect retention of information over significant retention intervals (up to two years)? Do spacing effects generalize beyond recall of verbal materials? Is feedback needed to promote learning, and must it be immediate? Testing has been found to enhance learning; does it actually reduce the rate of forgetting? Can testing effects be extended to nonverbal materials? We suggest that as answers to these questions are accumulated, it should become possible for cognitive psychology to offer non-obvious advice that can be applied in a variety of instructional contexts to facilitate learning and reduce forgetting.

Introduction: 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 results that have non-obvious and concrete implications for instructional procedures?

A brief tour 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 (e.g., Hyde & Jenkins, 1973). While the validity of the principle is not in doubt, it seems not to have provided much non-obvious or concrete guidance for practitioners. The present authors, along with other writers represented in this special issue, 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 often leads us to questions that drew more attention in an earlier era of psychology (see, e.g., Starch, 1927) than in recent years (even though, we would contend, some of them have implications for issues of great theoretical interest; e.g., Mozer, Howe, and Pashler (2004).

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

Spacing of Practice: Temporal Variables

The study of temporal distribution of practice goes back at least as far as Ebbinghaus (1885/1964) and is the subject of hundreds of 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 based on 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 very few researchers have examined retention intervals even as long as one day. Bahrick (e.g., 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 along with spacing. Therefore, though the literature is large, it seemed to us not to provide an empirical basis for prescribing efficient procedures for learning of vocabulary or facts, and we commenced several new lines of experiments.

In discussing spacing, we will 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 inter-study interval (*ISI*). After an additional *retention interval (*RI) --measured from S2--a final test is given. The literature reveals that for short values of RI, effects of ISI are often non-monotonic in character, with final-test performance rising up to some optimal ISI value, then falling (e.g., Crowder, 1976; Glenberg & Lehman, 1980).

To maximize generalizability to practical contexts, our new studies have used materials that seem representative of (at least simpler) sorts of learning people actually undertake in daily life, such as facts, vocabulary, and the like. One of our first studies (Cepeda et al., submitted) involved 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 displays the Swahili, and the subject types in English word, then receiving 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 performance after the 10-day retention interval, in line with prior results using free recall (Edwards, 1917; Glenberg & Lehman, 1980). Larger ISIs produced a smaller drop.

Our next studies moved to a six-month RI, teaching subjects little-known facts as well as the names of obscure visually presented objects (Cepeda et al., submitted). Here, a one-month ISI produced much better final recall than a one-day or even a one-week ISI, with a shallow drop beyond that. The positive relationship of optimal ISI to RI fits the literature involving short time intervals, but it appears from our data that when RI is substantial, the optimal ratio of ISI to RI is not 1:1, as some had suggested (Crowder, 1976), but rather something closer to 10 - 20% (see Figure 2).

To verify this *these conclusions* within a single experiment, we are 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, we use relatively unfamiliar facts. Data collection is still ongoing, but results from about 1800 subjects suggest that when the retention interval is 1 week, optimal ISI is about 1 day, but for a 50-week RI, an ISI of 3 weeks is best among the values we examine.

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 is some intermediate value, although a longer-than-optimal ISI is better than a shorter-than-optimal ISI. Our data imply that to promote retention over years, insuring an ISI of several months or even a few years seems likely to be far more efficient than using shorter intervals.

Spacing Effects in Math Problem Solving

Do these spacing principles govern learning tasks that do not involve 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, in press). The students saw a tutorial and then worked 10 practice problems-- either massed into a single session or distributed over 2 sessions separated by 1 week. After attempting each problem, students were shown the complete solution. A final test was given 1 or 4 weeks after the last practice problem. The ISI manipulation had no effect at the 1-week retention interval, but a substantial effect after the 4-week interval (Figure 2). 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.

It is interesting to note that conventional mathematics texts normally mass practice problems relating to a topic in one problem set presented immediately following textual presentation of that topic. Our data suggest that--at least for promoting retention-- the problems related to a given topic should be distributed across many problem sets.

Spacing in Perceptual Categorization Learning

We have also looked at perceptual categorization learning, a task that--despite its prominence within cognitive science--is almost absent from the spacing literature. Some of our studies taught subjects to categorize checkerboard patterns (as in Fried & Holyoak, 1984). We have observed no benefit of a 3-day ISI as compared to a 10 minute ISI, for either 1-week or 3-week 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, Glackens), and later tested on novel paintings by the same artists.

With the assistance of a dermatologist, we have also created a website (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

We have also been looking at a related practical choice sometimes termed *overlearning*: immediate continued practice on some material after error-free performance has been attained. Overlearning has been shown to increase later recall probability as compared to smaller degrees of practice (e.g., Krueger, 1929), and has often been advocated as a generally useful learning strategy (e.g., Foriska, 1993; Driskell et al., 1992). 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 3). With the combinatorics task described earlier, the reduction in overlearning gain with retention interval was even more dramatic (Figure 4; see Rohrer & Taylor, in press).

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 procedures for landing a plane after engine failure). Furthermore, overlearning may enhance speed long after retrieval accuracy has reached ceiling (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 concrete choice faced by instructors is whether to provide feedback, and if so, at what time and of 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 learning 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. Type of feedback accompanying the tests was varied between subjects. One week later, a final test was administered. Feedback in the learning session that provided the correct answer dramatically improved performance

Submitted to Psychonomic Bulletin and Review; please do not quote without permission.

for items that elicited errors in the learning session--both within the learning session itself and on the final test. For these items, feedback offered a roughly five-fold 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. Of interest, when a subject correctly recalled an item, feedback made essentially no difference. Surprisingly, this held even for correct recalls that were made with low confidence. 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 some obscure facts (e.g., *Alaska is the U.S. state with the highest percentage of people who walk to work.*) followed by either (a) an immediate test (*What is the state…?)*, and then feedback (*Alaska)*, (b) an immediate test followed by feedback 1 day later, (c) a 1-day delayed test followed by immediate feedback, or (d) a 1-day delayed test followed by feedback after an additional day (Figure 5). On a final test 2 weeks later, the groups that received delayed feedback performed better, not worse, than those that received immediate feedback (i.e., regardless of whether the test was immediate or delayed). The effect was largest for items the subject answered correctly, but surprisingly, similar trends were found even for errors. Obviously, immediate feedback is not essential--and it may not even be optimal (presumably because delays provide spaced practice, at least after correct responses). Naturally, one cannot assume that these results will necessarily generalize to all practically important forms of learning (motor learning in particular does seem to benefit from withholding of feedback; see Schmidt & Bjork, 1992).

Retrieval Practice: Benefits from Tests

Prior research shows that learning is often enhanced when the learner is required to recall information, as compared to simply restudying it (see Roediger & Karpicke, 2006b, for review). This *testing* (or *retrieval practice*) *effect*-- discussed by McDaniel, Roediger, and colleagues in the current issue -- has been found in free recall (e.g., Allen, Mahler, & Estes, 1969; Carpenter & DeLosh, 2006) as well as cued recall, including foreign language vocabulary learning (Carrier & Pashler, 1992) face-name learning (Carpenter & DeLosh, 2005), 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 retrievals one hopes to promote. We started examining this question by looking at the *direction* of test in foreign language learning (e.g., Dog-Hund). Does practice recalling *Dog* (after seeing *Hund→?*) facilitate later recall in the opposite order (*Dog→?),* when compared to simply restudying the pair *Hund-Dog*? We find that it does (Figure 6; Carpenter, Pashler, & Vul, in press). We are also finding that even covert retrieval practice (where subjects inwardly retrieve, but make no outward response) suffices to enhance learning. The results encourage the idea that retrieval practice is has broad practical potential.

Does Retrieval Practice Attenuate Forgetting?

Some studies have found benefits of retrieval practice appear or grow with retention interval (e.g., Roediger & Karpicke, 2006a), possibly suggesting that retrieval practice actually slows the rate of forgetting (Wheeler, Ewers, & Buonanno, 2003). We have been looking at this issue using a formal analysis of forgetting functions (Carpenter, Pashler, Wixted, & Vul, in preparation). 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), rather similar to Carrier and Pashler (1992). Different items were tested after five minutes, or 1, 2, 7, 14, or 42 days (within-subject design). 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. Although testing increased the degree of learning as compared to restudying, it did not significantly reduce the rate of forgetting (see Figure 7).

Retrieval Practice and Nonverbal Tasks

Retrieval practice effects have been studied almost entirely with verbal material. Seeking to assess the generality of the 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 like 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 try to 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 on a different feature). In a final test, subjects were asked to draw the maps. Drawings were better and more complete when studied through covert retrieval. Thus, we are optimistic that retrieval practice may be extended to various other forms of nonverbal learning tasks with practical significance.

Forced Guessing: Is it Harmful?

As described above, retrieval practice can be a useful learning strategy. However, demanding retrieval makes it inevitable that students will often be asked questions to which they do not know the answer, and forced to retrieve erroneous information. Will this undermine learning, as some theories would suggest (e.g., Guthrie, 1952)?

To assess this issue, one of our recent studies began by posing very difficult trivia questions (e.g., *The weight of what land mammal is equivalent to the weight of a blue whale's tongue?*) along with four plausible answers (*a) Bengal tiger, b) Grizzly bear, c) Wolverine, d) African elephant*). For one-third of the questions, the correct answer (*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. From this and other studies we are unable to find any costs associated with guessing when completely unsure. Naturally, however, here as elsewhere there may be important boundary conditions on this result, which we regard as fairly surprising.

Summary and Conclusions

Our results can be summarized as follows. We find that over substantial time periods, spacing has powerful (and non-monotonic) 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 but does not seem to slow forgetting. Retrieval practice can be extended well beyond overt retrieval of verbal responses. Feedback seems to be quite essential to learning of facts – but only after errors. However, the timing of this feedback seems unimportant. Finallly, guessing, which usually produces errors, seems not to impair learning much at all, so long as feedback is eventually provided. These conclusions are all preliminary, naturally, but along with those described in the other papers within this special issue, we think they offer 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 answer the question "Do cognitive psychologists have any non-obvious advice they can offer about how best to go about learning things?".

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.

Atkinson, R. C. (1972). Ingredients for a theory of instruction. *American Psychologist*, 27, 921-931.

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. http://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. (in press). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review*.

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (in preparation). *The testing effect: Enhanced retention or attenuated forgetting?*

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

Cepeda, N. J., Mozer, M. C., Coburn, N., Rohrer, D., Wixted, J. T., & Pashler, H. (submitted). *Optimizing distributed practice: Theoretical analysis and practical implications*..

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: Lawrence Erlbaum Associates.

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. (1885/1964). *Memory: A contribution to experimental psychology* (H. A. Ruger, C. E. Bussenius, & E. R. Hilgard, Trans.). New York: Dover Publications, Inc.

Edwards, A. S. (1917). The distribution of time in learning small amounts of material. In *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 and Cognition, 10*, 234-257.

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

Guthrie, E. (1952). *The psychology of learning* (Rev. Ed.). New York: Harper.

Hyde, T.S., & Jenkins, J.J. (1973). Recall of words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and 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.

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.

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

Roediger, H. L. III., & Karpicke, J. D. (2006b). *The power of testing memory: Basic research and implications for educational practice*. Manuscript submitted for publication.

Rohrer, D., & Taylor, K. (in press). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology.*

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.*

Rothkopf, E.Z. (1958) Stimulus similarity and sequence of stimulus presentation in paired-associate learning, *Journal of Experimental Psychology*, *56*, 114-122.

Saxon, J. (1997). *Algebra I (3rd Ed.)*. Norman, OK: Saxon Publishers.

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.* Englewood Cliffs, NJ: Prentice-Hall.

Starch, D. (1927). *Educational Psychology.* New York: MacMillan & Co.

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

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

Author Note

Harold Pashler, Department of Psychology, University of California at San Diego.

This work was supported by the Institute of Education Sciences (US 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 Harold Pashler, Department of Psychology, 0109, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0109. Phone: (858) 534-3974, E-mail: hpashler@ucsd.edu

Figure Captions

Figure 1. The basic design of a spacing experiment. Subjects have two opportunitites to learn the same material, separated by an inter-study interval (ISI). After a retention interval (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.

Figure 2. Results of two spacing experiments. Figure shows proportion of correct recall on the final test as a function of inter-study interval divided by the retention interval. Top line shows the first study discussed in text, with a 10-day RI; performance peaks at a ratio of .10 (1-day ISI). The lower two lines show the 6-month RI studies; peaks in both are at a ratio of .20 (28-day ISI). Overall the results suggest that for any given RI within this range, final-test memory is optimized by an ISI that is about 10-20% of the

RI. Note that having too little spacing is worse than having too much spacing, in every case. Error bars reflect plus or minus one standard error.

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 two weeks). The benefit of spacing grew with the retention interval (shown on the x-axis). Error bars reflect plus or minus one 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 the heavy massed practice (i.e., overlearning) virtually disappeared as retention interval (shown on x-axis) increased. Error bars reflect plus or minus one standard error.

Figure 5. Overlearning mathematics. After learning how to solve a permutation task, students worked either 3 or 9 practice problems within the same practice session. There was no benefit whatever of the additional massed practice (i.e., overlearning) at either 1 week or 4 week retention interval. Error bars reflect plus or minus one standard error.

Figure 6. Effects of feedback timing and test timing on the learning of obscure facts. Subjects were given an immediate test (Test 1) with immediate feedback (Group 1), an immediate test with one-day delayed feedback (Group 2), a one-day delayed test with immediate feedback (Group 3), or a one-day delayed test with one-day delayed feedback (Group 4). Performance on Test 1 was better when it was immediate, naturally. However, on the Final Test two weeks later, delayed feedback outperformed immediate feedback, for both immediate and delayed tests. Far from being harmful as one might infer from Skinnerian accounts, delaying feedback was in fact slightly helpful.

Figure 7. Is retrieval practice benefit (testing-with-feedback outperforming pure study) in cued recall confined to a final test given in the same direction as the retrieval practice? Two experiments comparing forward and backward tests indicate that it is not. Subjects first studied word pairs (A – B), and were then given a cued recall test (A - ?) or a restudy opportunity (A - B). Final test was cued recall test in the same direction (A - ?) or opposite direction (? - B). Recall in both directions benefited more from testing with feedback than from restudying, and this was true for English word pairs (upper panel), as well as English-Swahili word pairs (lower panel).

Figure 8. Is retrieval practice benefit improving learning, slowing forgetting, or both? Effects of tests vs. restudy opportunities on the learning and forgetting of obscure facts. Subjects completed either a test with feedback (Test/Study) or a restudy opportunity (Study) over each fact, and then were tested over a different subset of facts from each condition after five minutes, one day, two days, seven days, 14 days, or 42 days. Parameter estimates derived from the power function $y = a(b + 1)^{-c}$ yielded a significant

advantage in the degree of learning (*a*), but no significant reduction in the rate of forgetting (*c*), for Test/Study compared to Study.

Figure 9. 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 (Guess + Immediate Feedback), or guessed about the correct answer and were given delayed feedback (Guess + Delayed Feedback). On Test 1, forced guesses are at the chance level, as expected. On a final test one week later, recall of the correct answers was not impaired by having guessed on that item (this held even if the initial guess was incorrect).

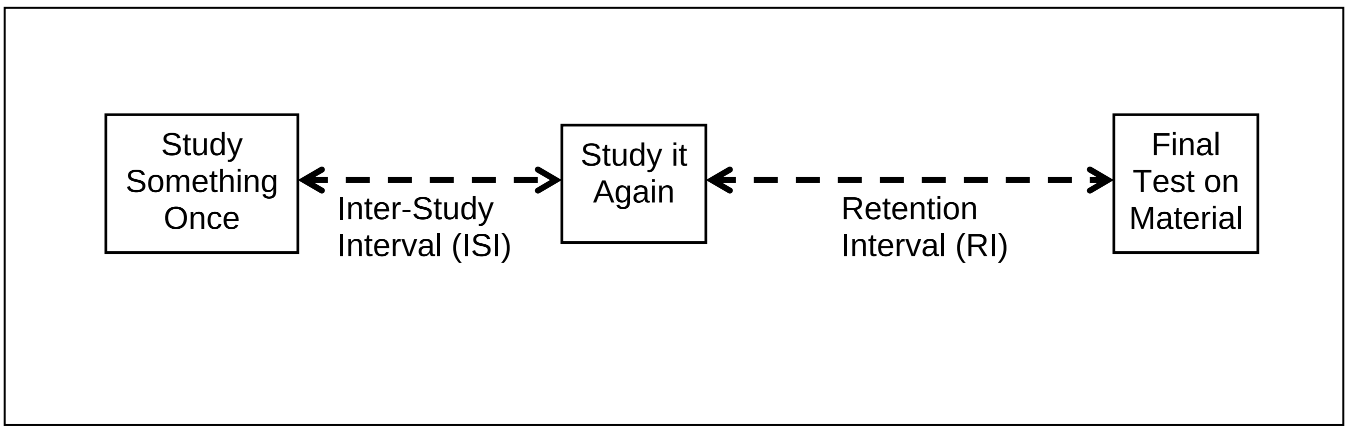

Figure 1

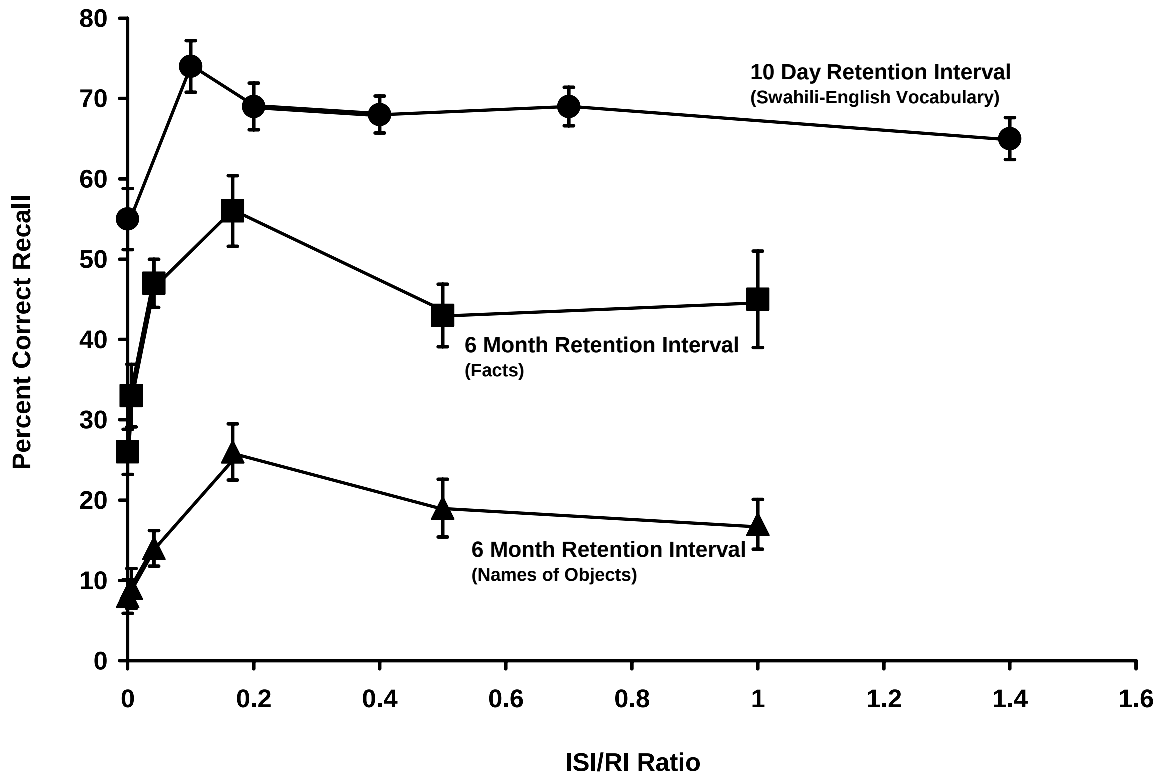
80
70
60
50
40
30
20
10
0
Percent Correct Recall
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
ISI/RI Ratio
10 Day Retention Interval
(Swahili-English Vocabulary)
6 Month Retention Interval
(Facts)
6 Month Retention Interval
(Names of Objects)

Figure 2

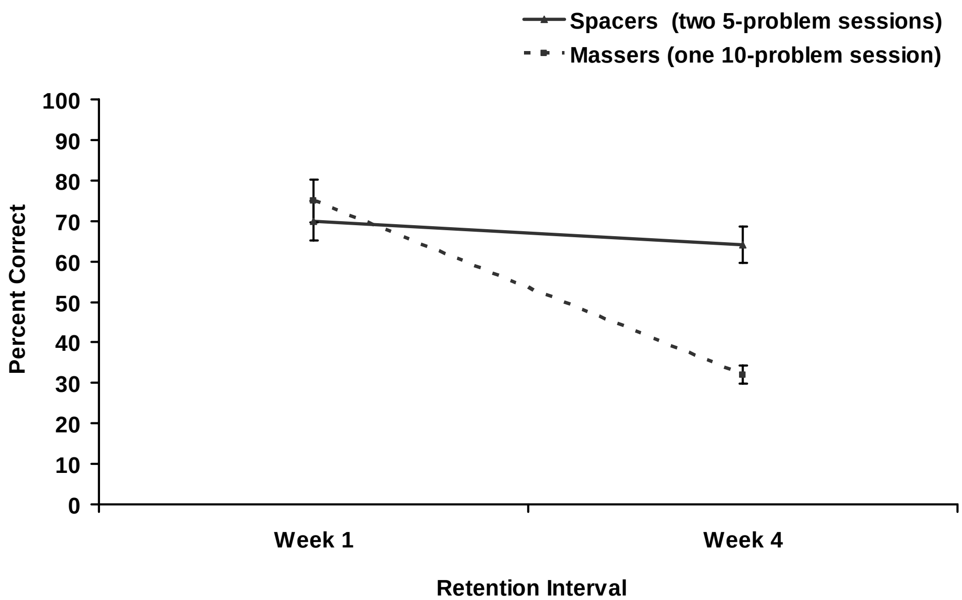
Spacers (two 5-problem sessions)
Massers (one 10-problem session)
100
90
80
70
60
50
40
30
20
10
0
Percent Correct
Week 1
Week 4
Retention Interval

Figure 3

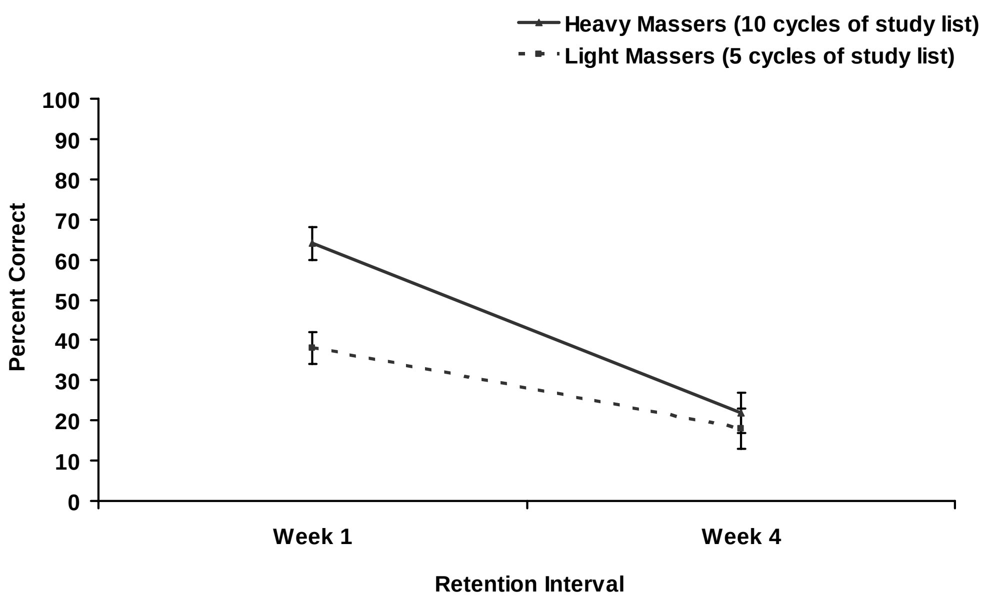
Heavy Massers (10 cycles of study list)
Light Massers (5 cycles of study list)
Percent Correct
100
90
80
70
60
50
40
30
20
10
0
Week 1
Week 4
Retention Interval

Figure 4

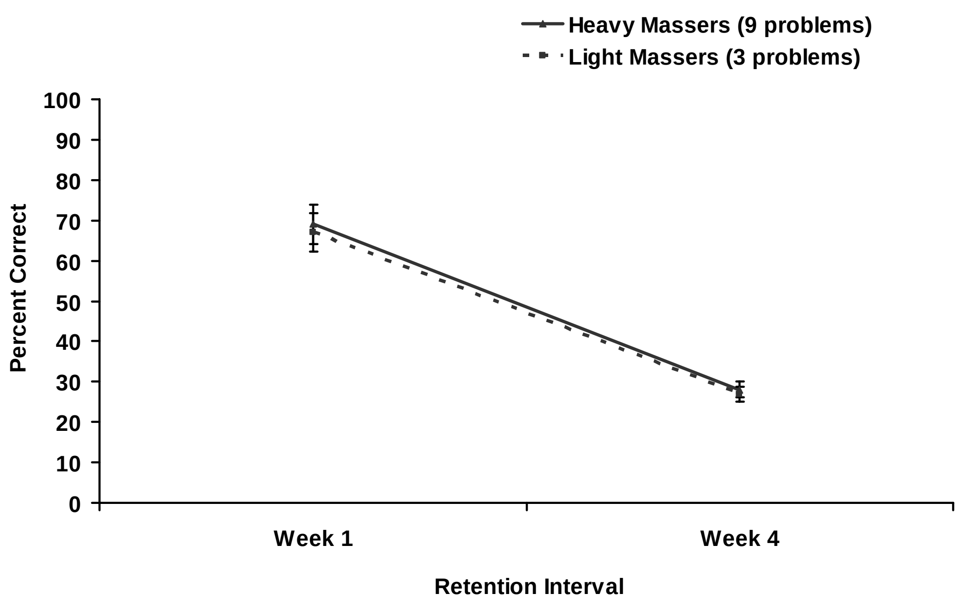
Heavy Massers (9 problems)
Light Massers (3 problems)
Percent Correct
100
90
80
70
60
50
40
30
20
10
0
Week 1
Week 4
Retention Interval

Figure 5

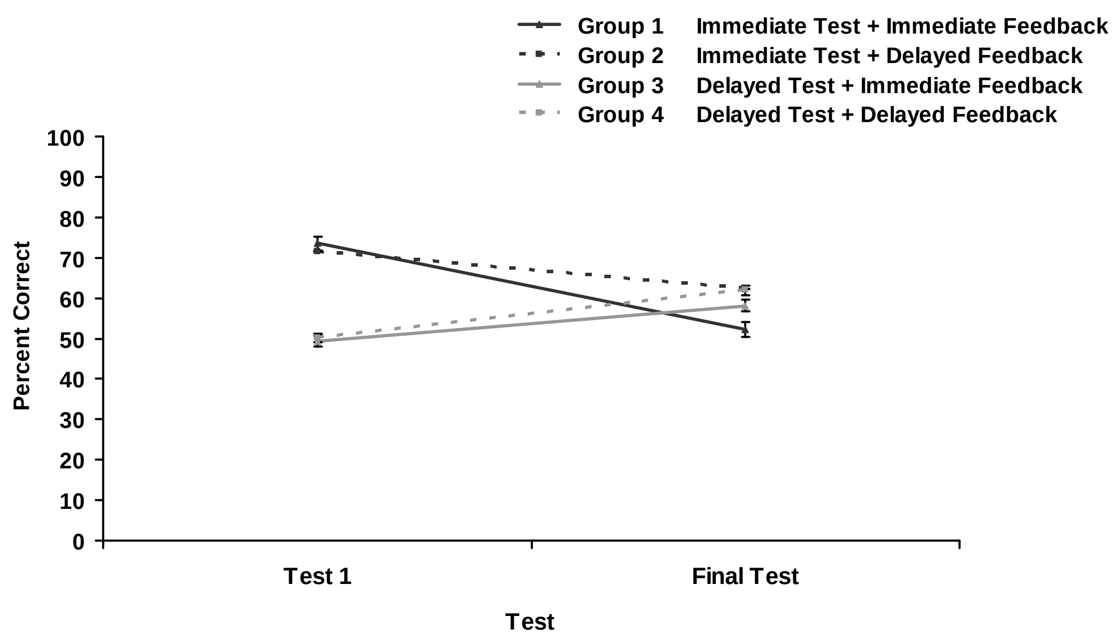
Group 1 Immediate Test + Immediate Feedback
Group 2 Immediate Test + Delayed Feedback
Group 3 Delayed Test + Immediate Feedback
Group 4 Delayed Test + Delayed Feedback
Percent Correct
100
90
80
70
60
50
40
30
20
10
0
Test 1
Final Test
Test

Figure 6

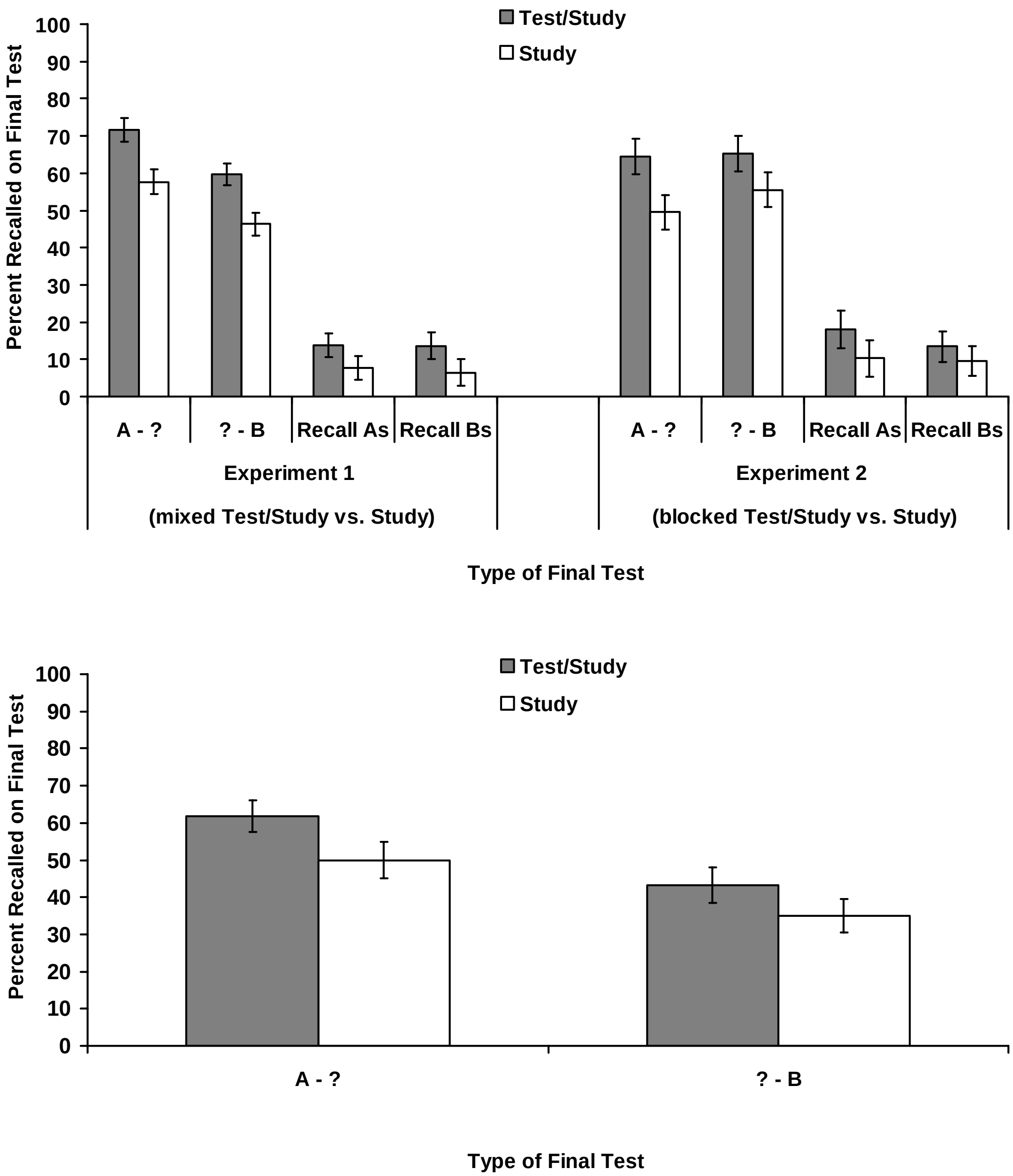
Test/Study
Study
Percent Recalled on Final Test
100
90
80
70
60
50
40
30
20
10
0
A - ?
? - B
Recall As
Recall Bs
Experiment 1
(mixed Test/Study vs. Study)
A - ?
? - B
Recall As
Recall Bs
Experiment 2
(blocked Test/Study vs. Study)
Type of Final Test
Test/Study
Study
Percent Recalled on Final Test
100
90
80
70
60
50
40
30
20
10
0
A - ?
? - B
Type of Final Test

Figure 7

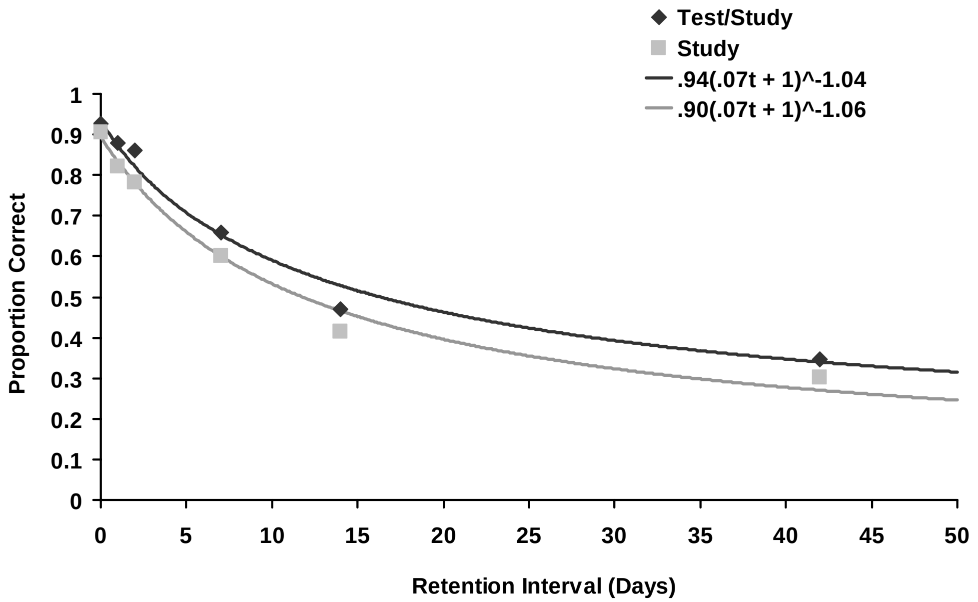

Figure 8

Submitted to Psychonomic Bulletin and Review; please do not quote without permission.

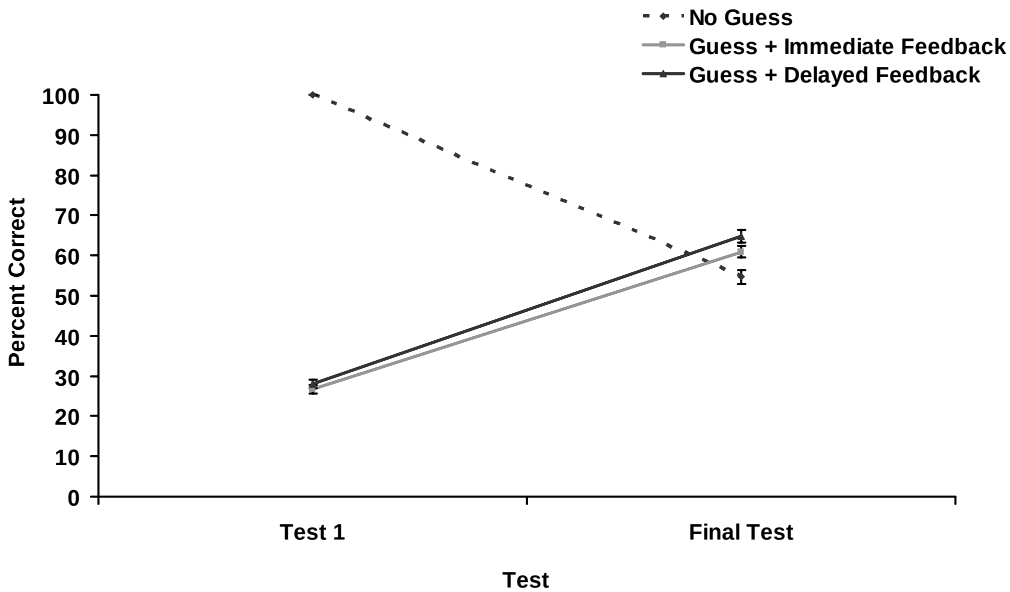

Figure 9

Submitted to Psychonomic Bulletin and Review; please do not quote without permission.