Sciencexpress Report

Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping

Jeffrey D. Karpicke* and Janell R. Blunt

Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA.

*To whom correspondence should be addressed. E-mail: karpicke@purdue.edu

Educators rely heavily on learning activities that encourage elaborative studying, while activities that require students to practice retrieving and reconstructing knowledge are used less frequently. Here, we show that practicing retrieval produces greater gains in meaningful learning than elaborative studying with concept mapping. The advantage of retrieval practice generalized across texts identical to those commonly found in science education. The advantage of retrieval practice was observed with test questions that assessed comprehension and required students to make inferences. The advantage of retrieval practice occurred even when the criterial test involved creating concept maps. Our findings support the theory that retrieval practice enhances learning by retrieval-specific mechanisms rather than by elaborative study processes. Retrieval practice is an effective tool to promote conceptual learning about science.

Most thought on human learning is guided by a few tacit assumptions. One assumption is that learning happens primarily when people encode knowledge and experiences. A related assumption is that retrieval—the active, cue-driven process of reconstructing knowledge—only measures the products of a prior learning experience but does not itself produce learning. Just as we assume that the act of measuring a physical object would not change the size, shape, or weight of the object, so too people often assume that the act of measuring memory does not change memory (*1*, *2*). Thus most educational research and practice has focused on enhancing the processing that occurs when students encode knowledge – that is, getting knowledge "in memory". Far less attention has been paid to the potential importance of retrieval to the process of learning. Indeed, recent National Research Council books about how students learn in educational settings (*3*–*5*) contain no mention of retrieval processes.

It is beyond question that activities that promote effective encoding, known as elaborative study tasks, are important for learning (*6*). However, research in cognitive science has challenged the assumption that retrieval is neutral and uninfluential in the learning process (*7*–*11*). Not only does retrieval produce learning, but a retrieval event may actually represent a more powerful learning activity than an encoding event. This research suggests a conceptualization of mind and learning that is different from one in which encoding places knowledge in memory and retrieval simply accesses that stored knowledge. Because each act of retrieval changes memory, the act of reconstructing knowledge must be considered essential to the process of learning.

Most prior research on retrieval practice has been conducted in the verbal learning tradition of memory research (*12*). The materials used have often not reflected the complex information students learn in actual educational settings (*13*). Most prior research has not used assessments thought to measure meaningful learning, which refers to students' abilities to make inferences and exhibit deep understanding of concepts (*14*, *15*). But perhaps the greatest impediment to broad application of retrieval practice is that we do not know whether retrieval activities are more effective than other active, elaborative learning activities. Retrieval practice might produce levels of learning that are essentially the same as those produced by elaborative studying. Alternatively, if there are retrieval-specific mechanisms that promote learning, then retrieval practice may represent a way to promote student learning that goes beyond elaborative study activities used in STEM education.

The present experiments put retrieval practice to a test. Elaborative learning activities hold a central place in contemporary education. We examined the effectiveness of retrieval practice relative to elaborative studying with concept mapping (*16*–*18*). In concept mapping, students construct a diagram in which nodes are used to represent concepts and links connecting the nodes represent relations among the concepts. Concept mapping is considered an active learning task, and it serves as an elaborative study activity when students construct concept maps in the presence of the materials they are learning. Under these conditions, concept mapping bears the defining characteristics of an elaborative study method: It requires students to enrich the material they are studying and encode meaningful relationships among concepts within an organized knowledge structure.

In two experiments, we compared the effectiveness of retrieval practice and elaborative studying with concept mapping for producing meaningful learning of science materials. Eighty undergraduate students participated in Experiment 1. The students first studied a science text under one of four conditions within a single initial learning session. In the study-once condition, students studied the text in a single study period. In the repeated study condition, students studied the text in four consecutive study periods (*8*). In the elaborative concept mapping condition, students studied the text in an initial study period and then created a concept map of the concepts in the text. The students were instructed about the nature of concept mapping, viewed an example of a concept map, and created their concept maps on paper while viewing the text. This is a typical way concept mapping is used as an elaborative study activity (*16–18*). Finally, in the retrieval practice condition, students studied the text in an initial study period and then practiced retrieval by recalling as much of the information as they could on a free recall test. After recalling once, the students restudied the text and recalled again. The total amount of learning time was exactly matched in the concept mapping and retrieval practice conditions (*19*).

At the end of the learning phase, we assessed students' metacognitive knowledge of the effectiveness of these learning activities by having students make judgments of learning. After completing the learning phase, students predicted the percentage of information from the text they would remember in one week (*20*).

The students then returned to the laboratory one week later for a final short-answer test. To assess meaningful learning, the test included both verbatim questions, which assessed conceptual knowledge stated directly in the text, and inference questions, which required students to connect multiple concepts from the text. Both question types are conceptual, but verbatim and inference questions are thought to assess different depths of conceptual knowledge (*14*, *15*).

The proportion of ideas produced on the initial concept maps and recalled in the retrieval practice condition was nearly identical [.78 and .81, respectively, $F(1, 38) = 0.46$, n.s.]. Therefore, the interpretation of any differences on the final test is not clouded by differences in initial learning time or differences in the initial proportion of ideas correctly produced in the concept mapping and retrieval practice conditions.

On the final test one week later, the repeated study, elaborative concept mapping, and retrieval practice conditions all outperformed the study-once condition on both verbatim and inference questions (Fig. 1, A and B). Retrieval practice produced the best learning, better than elaborative studying with concept mapping, which itself was not significantly better than spending additional time reading. Collapsed across question type (verbatim and inference), the advantage of retrieval practice ($M = .67$) over elaborative studying with concept mapping ($M = .45$) represented about a 50% improvement in long-term retention scores [$d = 1.50$, $F(1, 38) = 21.63$, $\eta_p^2 = .36$].

Students' judgments of learning, solicited in the initial learning session, reflected little metacognitive knowledge of the benefits of retrieval practice (Fig. 1C). Students predicted that repeated studying would produce the best long-term retention and that practicing retrieval would produce the worst retention, even though the opposite was true (*7*, *8*).

We carried out a second experiment to replicate the results of our first experiment and extend them in three ways. First, we sought to generalize our results to texts that represent different knowledge structures commonly found in science education, because under some circumstances the effectiveness of different learning activities can depend on the structure of the materials that students are learning (*21*). We used texts with enumeration structures, which describe a list of concepts (e.g., a text describing properties of different muscle tissues), and texts with sequence structures, which describe a continuous and ordered series of events (e.g., a text describing the sequence of events involved in the process of digestion) (*22*).

Second, to determine the robustness of our retrieval practice effects, we examined the relative effectiveness of retrieval practice and elaborative concept mapping for each individual learner. We tested a total of 120 students and used a within-subject design. Each student created a concept map of one science text and practiced retrieval of a second text. This experimental design allowed us to determine how many students showed an advantage of retrieval practice over concept mapping, how many showed the opposite result, and how many showed no difference between learning activities.

Third, we assessed long-term learning with two different final test formats. In Experiment 1, retrieval practice produced better performance than elaborative studying with concept mapping on a final short-answer test. It may be that the similarity of initial learning and final testing scenarios was important and that the final short-answer test was more similar to the initial retrieval practice task than to the initial concept mapping task. Therefore, in Experiment 2, half of the students took a final short-answer test, like the one used in Experiment 1, and half took a final test in which they created concept maps of the two texts, without viewing the texts on the final test. If retrieval practice helps students build the conceptual knowledge structures they need to retain knowledge over the long-term, then it should produce better performance than elaborative studying with concept mapping even when the final test involves creating a concept map.

Initial learning time was again exactly matched in the elaborative concept mapping and retrieval practice conditions. However, in Experiment 2, students produced a greater proportion of ideas on the initial concept maps than they did in on the initial tests in the retrieval practice condition [.74 versus .65, respectively, $F(1, 117) = 23.13$, $\eta_p^2 = .17$]. Therefore, the initial level of performance favored the concept mapping condition.

The results on the final short-answer test were similar for verbatim and inference questions (Fig. 2), as was the case in Experiment 1. Therefore, the results were collapsed across question type. Retrieval practice produced better performance than elaborative concept mapping for both types of science text (Fig. 2, A and B). Collapsed across the two text formats, the advantage of retrieval practice was again large [$d = 1.07$, $F(1, 59) = 68.54$, $\eta_p^2 = .54$].

Fig. 2, C and D, shows performance on the final concept mapping test. If the nominal similarity of initial learning and final test conditions were important, one might expect initial elaborative study with concept mapping to produce the best performance when the final test also involved creating concept maps. That was not the case. Even when the final test involved using memory to construct a concept map, practicing retrieval during original learning produced better performance than engaging in elaborative study by creating concept maps during original learning [$d = 1.01$, $F(1, 59) = 58.42$, $\eta_p^2 = .50$].

We again examined whether students exhibited metacognitive knowledge of the benefits of retrieval practice. Students' judgments of learning were solicited after students had experienced each text in the initial learning phase. In general, students erroneously predicted that elaborative concept mapping would produce better long-term learning than retrieval practice (Fig. 2, E to H).

Finally, we examined the relative effectiveness of retrieval practice and elaborative study with concept mapping for every individual learner in the experiment. Table 1 shows the number of subjects who performed better following retrieval practice than concept mapping, the number who showed the opposite result, and the number who performed equivalently in both conditions. Overall, 101 out of 120 students (84%) performed better on the final test after practicing retrieval than after elaborative studying with concept mapping. Table 1 also shows students' judgments of learning. Ninety out of 120 students (75%) believed that elaborative concept mapping would be just as effective or even more effective than practicing retrieval. Most students did not expect that retrieval practice would be more effective than elaborative concept mapping, but indeed it was.

Retrieval practice is a powerful way to promote meaningful learning of complex concepts commonly found in science education. Here we have shown that retrieval practice produces more learning than elaborative studying, and we used concept mapping as a means of inducing elaboration while students studied. We hasten to add that concept mapping itself is not inherently just an elaborative study task. When students create concept maps in the presence of materials they are learning, the activity involves elaborative studying. Students could also create concept maps in the absence of materials they are learning, and then the activity would involve practicing retrieval of knowledge. Nevertheless, both elaborative concept mapping and retrieval practice are active learning tasks, and our results make it clear that whether a task is considered "active" is not diagnostic of how much learning the task will produce. The specific nature of the activity determines the degree and quality of learning, so understanding the nature of encoding and retrieval processes is crucial for designing educational activities.

There are several theoretical reasons to expect that the processes involved in retrieving knowledge differ fundamentally from the processes involved in elaborative studying. During elaboration, subjects attain detailed representations of encoded knowledge by enriching or increasing the number of encoded features. But during retrieval, subjects use retrieval cues to reconstruct what happened in a particular place at a particular time. In free recall, subjects must establish an organizational retrieval structure (*23*) and then discriminate and recover individual concepts within that structure (*24*). Retrieval practice likely enhances the diagnostic value of retrieval cues, which refers to how well a cue specifies a particular piece of knowledge to the exclusion of other potential candidates (*25–27*). Rather than multiplying or increasing the number of encoded features, which occurs during elaboration, retrieval practice may improve cue diagnosticity by restricting the set of candidates specified by a cue to be included in the search set (*23*, *25–27*). Thus mechanisms involved in retrieving knowledge play a role in producing learning.

Research on retrieval practice suggests a view of how the human mind works that differs from everyday intuitions. Retrieval is not merely a read out of the knowledge stored in one's mind – the act of reconstructing knowledge itself enhances learning. This dynamic perspective on the human mind can pave the way for the design of new educational activities based on consideration of retrieval processes.

References and Notes

1. R. A. Bjork, in *Information Processing and Cognition: The Loyola Symposium*, R. L. Solso, Ed. (Erlbaum, Hillsdale, NJ, 1975), pp. 123–144.
2. J. D. Karpicke, H. L. Roediger, *J. Mem. Lang.* **57**, 151 (2007).
3. J. D. Bransford, A. L. Brown, R. R. Cocking, Eds., *How People Learn: Brain, Mind, Experience, and School* (National Research Council, Washington, DC, 2000).
4. M. S. Donovan, J. D. Bransford, Eds., *How Students Learn: Science in the Classroom* (National Research Council, Washington, DC, 2005).

5. R. A. Duschl, H. A. Schweingruber, A. W. Shouse, Eds., *Taking Science to School: Learning and Teaching Science in Grades K-8* (National Research Council, Washington, DC, 2005).
6. F. I. M. Craik, E. Tulving, *J. Exp. Psy. Gen*. **104**, 268 (1975).
7. J. D. Karpicke, H. L. Roediger, *Science* **319**, 966 (2008).
8. H. L. Roediger, J. D. Karpicke, *Psychol. Sci.* **17**, 249 (2006).
9. S. K. Carpenter, H. Pashler, *Psychon. Bull. Rev.* **14**, 474 (2007).
10. H. Pashler, D. Rohrer, N. J. Cepeda, S. K. Carpenter, *Psychon. Bull. Rev.* **14**, 187 (2007).
11. M. A. Pyc, K. A. Rawson, *J. Mem. Lang*. **60**, 437 (2009).
12. For a review see H. L. Roediger, J. D. Karpicke, *Perspectives Psychol. Sci*. **1**, 181 (2006).
13. But see J. D. Karpicke, H. L. Roediger, *Mem. Cog.* **38**, 116 (2010); M. A. McDaniel, D. C. Howard, G. O. Einstein, *Psychol. Sci.* **20**, 516 (2009).
14. L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich *et al*., *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* (Longman, New York, 2000).
15. R. E. Mayer, *Learning and Instruction* (Prentice Hall, Upper Saddle River, NJ, 2008).
16. J. D. Novak, D. B. Gowin, *Learning How to Learn* (Cambridge Univ. Press, NY, 1984).
17. J. D. Novak, *Res. Sci. Ed.* **35**, 23 (2005).
18. J. C. Nesbit, O. O. Adesope, *Rev. Ed. Res*. **76**, 413 (2006).
19. Materials, methods, and additional results are available as supporting material on *Science* Online.
20. J. Dunlosky, J. Metcalfe, *Metacognition* (Sage, Thousand Oaks, CA, 2009).
21. M. A. McDaniel, G. O. Einstein, *Ed. Psy. Rev*. **1**, 113 (1989).
22. L. K. Cook, R. E. Mayer, *J. Ed. Psy*. **80**, 448 (1988).
23. J. G. W. Raaijmakers, R. M. Shiffrin, *Psy. Rev*. **88**, 93 (1981).
24. R. R. Hunt, M. A. McDaniel, *J. Mem. Lang*. **32**, 421 (1993).
25. J. S. Nairne, *Mem*. **10**, 389 (2002).
26. J. S. Nairne, in *Distinctiveness and Memory*, R. R. Hunt, J. Worthen, Eds. (Oxford Univ. Press, NY, 2006), pp. 27–46.
27. J. D. Karpicke, F. M. Zaromb, *J. Mem. Lang*. **62**, 227 (2010).
28. This research was supported by a grant from the National Science Foundation (#0941170). We thank Claire Ballas, Beth Byrer, Hannah Cannon, and Brittany Etchison for help with this research.

Supporting Online Material
www.sciencemag.org/cgi/content/full/science.1199327/DC1
Materials and Methods
Fig. S1
Table S1
References

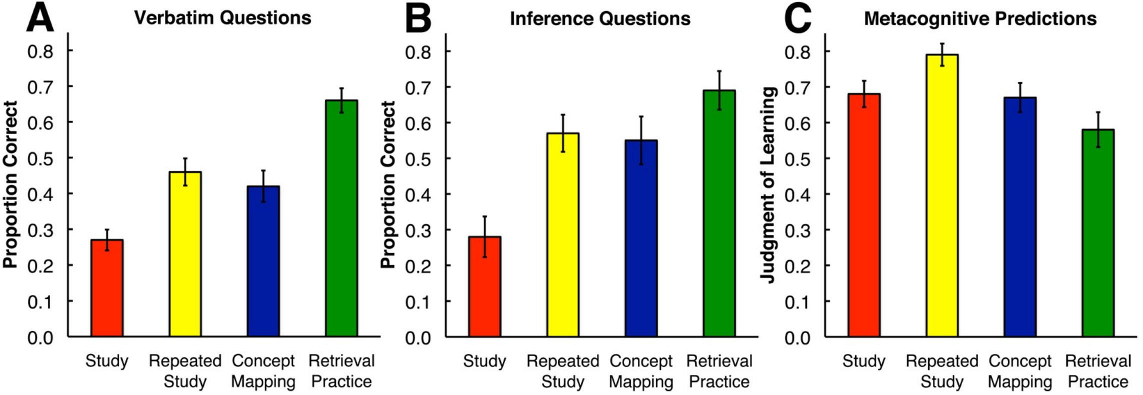
20 October 2010; accepted 10 January 2011
Published online 20 January 2011; 10.1126/science.1199327

Fig. 1. Results of Experiment 1. (**A** and **B**) show the proportions correct on verbatim and inference short-answer questions, respectively. (**C**) shows the proportion of information subjects predicted they would recall on the final test (their metacognitive judgments of learning). Error bars indicate SEMs. On the final short answer test, retrieval practice enhanced long-term learning above and beyond elaborative study with concept mapping by one and a half standard deviations (d = 1.50), yet students were largely unable to predict this benefit.

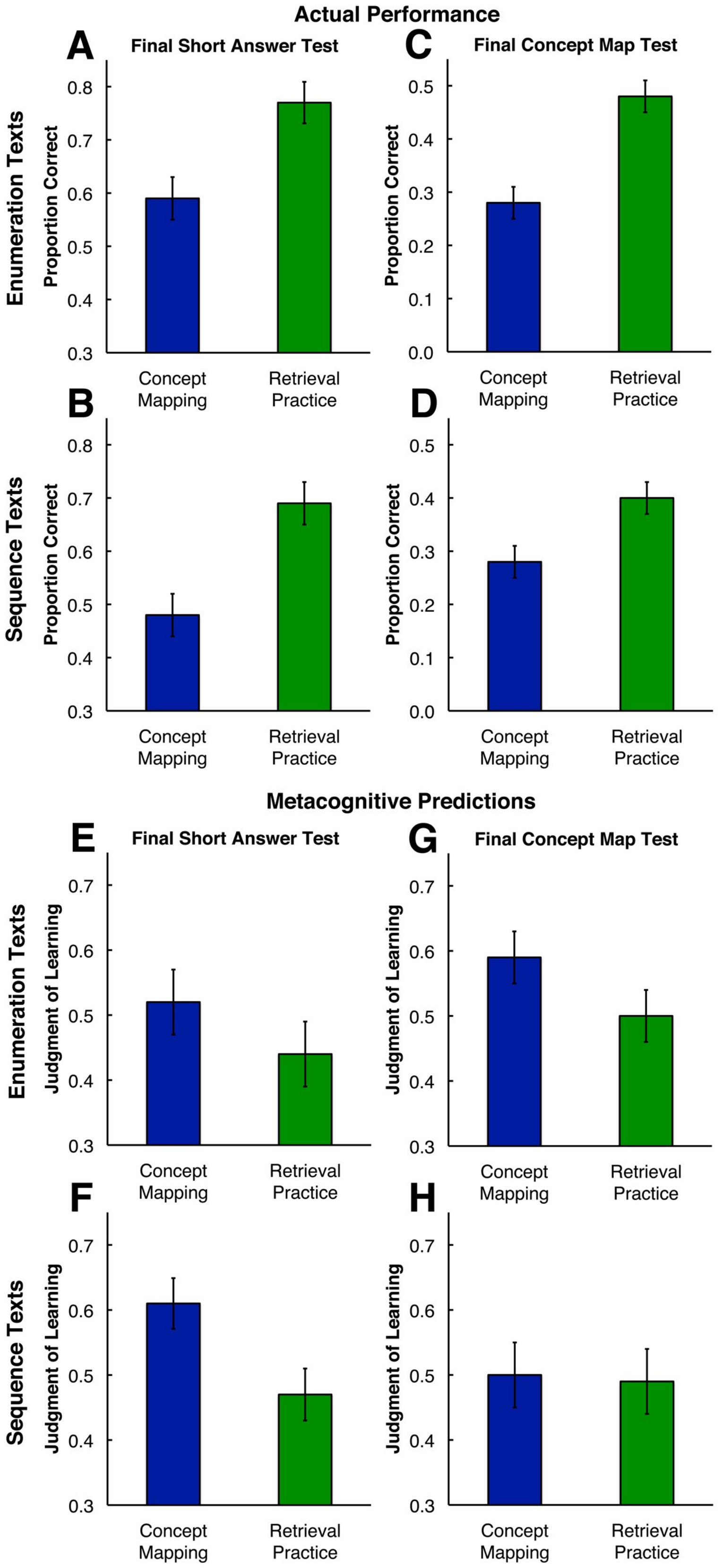
Fig. 2. Results of Experiment 2. (**A** and **B**) show the proportions correct on the final short-answer tests for enumeration and sequence texts, respectively. (**C** and **D**) show the proportions correct on the final concept mapping tests for enumeration and sequence texts, respectively. Error bars indicate SEMs. Retrieval practice enhanced long-term learning above and beyond elaborative concept mapping by more than one standard deviation on both types of final test (*d*s = 1.07 and 1.01, respectively). (**E** to **H**) show the proportion of information subjects predicted they would recall on the final test in each initial learning condition. Students tended to believe that elaborative concept mapping would produce the same or even greater learning than retrieval practice, even though the opposite was true, as shown in (A) to (D).

Table 1. Number of subjects showing different patterns of actual performance and metacognitive judgments in Experiment 2. Retrieval: the retrieval practice condition. Mapping: the elaborative concept mapping condition. Final SA Test: the final short answer test condition. Final Map Test: the final concept mapping test condition. Total: the sums across the two final test conditions.

	Actual Performance		
	Retrieval > Mapping	Retrieval = Mapping	Retrieval < Mapping
Final SA Test	52	3	5
Final Map Test	49	3	8
Total	101	6	13
	Metacognitive Predictions		
	Retrieval > Mapping	Retrieval = Mapping	Retrieval < Mapping
Final SA Test	12	14	34
Final Map Test	18	17	25
Total	30	31	59



A
Verbatim Questions
Proportion Correct
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Study
Repeated Study
Concept Mapping
Retrieval Practice
B
Inference Questions
Proportion Correct
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Study
Repeated Study
Concept Mapping
Retrieval Practice
C
Metacognitive Predictions
Judgment of Learning
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Study
Repeated Study
Concept Mapping
Retrieval Practice



Actual Performance
A
Final Short Answer Test
C
Final Concept Map Test
Enumeration Texts
Proportion Correct
0.8
0.7
0.6
0.5
0.4
0.3
0.5
0.4
0.3
0.2
0.1
0.0
Concept Mapping
Retrieval Practice
B
D
Sequence Texts
Metacognitive Predictions
E
Final Short Answer Test
G
Final Concept Map Test
Judgment of Learning
F
H