Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/11/$12.00 DOI: 10.1037/a0024140

RESEARCH REPORT

Semantic Information Activated During Retrieval Contributes to Later Retention: Support for the Mediator Effectiveness Hypothesis of the Testing Effect

Shana K. Carpenter
Iowa State University

Previous research has proposed that tests enhance retention more than do restudy opportunities because they promote the effectiveness of mediating information—that is, a word or concept that links a cue to a target (Pyc & Rawson, 2010). Although testing has been shown to promote retention of mediating information that participants were asked to generate, it is unknown what type of mediators are spontaneously activated during testing and how these contribute to later retention. In the current study, participants learned cue–target pairs through testing (e.g., *Mother:* _____) or restudying (e.g., *Mother: Child*) and were later tested on these items in addition to a never-before-presented item that was strongly associated with the cue (e.g., *Father*)—that is, the *semantic mediator*. Compared with participants who learned the items through restudying, those who learned the items through testing exhibited higher false alarm rates to semantic mediators on a final recognition test (Experiment 1) and were also more likely to recall the correct target from the semantic mediator on a final cued recall test (Experiment 2). These results support the mediator effectiveness hypothesis and demonstrate that semantically related information may be 1 type of natural mediator that is activated during testing.

Keywords: testing effect, retrieval practice, cued recall, mediator effectiveness

One of the more widely studied and robust phenomena in current memory research is the *testing effect*—the finding that taking a memory test on some information enhances later memory for that information (see e.g., Roediger & Karpicke, 2006a). This effect emerges even when testing is compared with an equivalent number of opportunities to engage in further encoding of the material (see e.g., Butler & Roediger, 2007; Carrier & Pashler, 1992) and when corrective feedback is not provided on the initial test (see e.g., Carpenter & DeLosh, 2005, 2006). These findings have led to the notion that there is something beneficial about the retrieval process itself that contributes to later retention.

We currently know little about what this process involves, however, and how it benefits retention. Toward developing a theoretical explanation for the testing effect, Carpenter (2009) recently proposed an *elaborative retrieval hypothesis*. According to this view, recalling a target from a cue (e.g., *Basket:* _____) is more likely to involve the activation of information that is related to that cue (e.g., *Eggs, Wicker*) than is simply restudying the cue–target pair (*Basket: Bread*). Activating these concepts is beneficial because it provides extra information than can facilitate recall of the target at a later time.

In support of this notion, a number of studies have shown that the testing effect is greater under conditions in which the initial test is made more difficult, for example by administering tests of recall instead of recognition (see e.g., Kang, McDermott, & Roediger, 2007), administering the initial test after longer as opposed to shorter time intervals (see e.g., Karpicke & Roediger, 2007; Pyc & Rawson, 2009; Whitten & Bjork, 1977), and by testing a target item (e.g., *Bread*) with a weakly related cue (e.g., *Basket:* _____) as opposed to a strongly related cue (e.g., *Toast:* _____; Carpenter, 2009). Each of these manipulations renders the target information less accessible at the time of an initial test, which would increase the likelihood of activating this "extra" information that is beneficial for retention.

One shortcoming of this hypothesis, however, is that it fails to specify exactly what this extra information is. What type of information is activated during initial retrieval that is helpful for later retention? A recent article by Pyc and Rawson (2010) has begun to address this important question. The authors propose a *mediator effectiveness hypothesis*, in which tests are beneficial because they are more likely than restudy opportunities to enhance the link between a cue and target via *mediating information* (i.e., a word or concept that links a cue to a target). Participants in this study learned Swahili–English word pairs (e.g., *Wingu: Cloud*) either through repeated restudying of the intact word pair or through an equivalent number of attempts to recall the English word (*Wingu:* _____) followed by restudy of the intact word pair. During re-

The author thanks Natalie Clark, Michael Clary, Riebana Sachs, Matthew Erdman, Jessica Kloeppel, Kellie Olson, Jill Feipel, Sam White, and Hana Ibrik for their assistance with data collection and scoring.

Correspondence concerning this article should be addressed to Shana K. Carpenter, Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, IA 50011-3180. E-mail: shacarp@iastate.edu

study, all participants were asked to generate a mediator that could help them remember the target from the cue (e.g., *Wingu* → "Wing" → *Cloud*). On a later cued recall test, participants who engaged in prior recall of the English word were not only better at recalling the English word from the Swahili cue (i.e., the classic testing effect) but were also better at recalling the mediators they had previously generated, along with the association between the mediator and the English word.

These results nicely demonstrate that tests with feedback promote memory of mediating information that participants were asked to generate during learning. In usual studies of the testing effect, however, participants are not specifically asked to generate mediators. Instead, they are given a cue and simply asked to recall the target. An important question, therefore, is how mediating information might account for the testing effect that is observed under normal circumstances.

The current study addressed this question by exploring what type of information is spontaneously activated during an initial test and how this contributes to later retention. Carpenter (2009) proposed that one type of information activated during retrieval might be semantic in nature. For example, the cue *Basket:* _____ may activate concepts such as "Eggs," "Flour," or "Wicker." If the correct target is eventually retrieved (*Bread*), there is now a structure of mediating semantic information that links *Basket* to *Bread*. Although Carpenter's results support this idea by demonstrating a stronger testing effect for targets recalled from weakly related cues (*Basket:* _____) compared with strongly related cues (*Toast:* _____), memory for the supposed mediating information (e.g., "Eggs," "Flour") was never tested.

The purpose of the current study was to directly measure the information that is activated during retrieval and how this contributes to later retention. In two experiments, participants learned weakly related cue–target pairs (e.g., *Mother: Child*) through either testing (*Mother:* _____) or restudying (*Mother: Child*). In Experiment 1 participants completed a final test of single-item recognition that contained cues (e.g., *Mother*) and targets (e.g., *Child*), as well as new items that had never appeared before. Each new item was either unrelated to any items that had appeared before (e.g., *Rabbit*) or strongly related to one of the cues (e.g., *Father*). Because of its strong preexisting association with the cue, this latter item is referred to as a *semantic mediator*. If semantic mediators are more likely to be activated during an initial test compared with during restudy, then participants who learn the cue–target pairs through testing should exhibit higher false alarm rates to the semantic mediators relative to participants who learned the cue–target pairs through restudying.

In Experiment 2, final retention of the same target items was measured through cued recall. This time, participants were asked to recall the targets from the same cues as before (e.g., *Mother:* _____), from a new cue that was weakly related to the target (e.g., *Birth:* _____), or from the semantic mediator (*Father:* _____). If an initial cued recall test (e.g., *Mother:* _____) involves activation of a semantic mediator (e.g., *Father*) during the process of retrieving the correct target (*Child*), then the link between the mediator and target (*Father* → *Child*) should be stronger for word pairs that were learned through testing compared with through restudying. Furthermore, this link should be stronger than the link between the target and a new related cue (*Birth* → *Child*), because the latter was not as likely to have been activated during initial retrieval.

Experiment 1

Method

Participants. Fifty-nine undergraduate students volunteered to participate in order to fulfill partial requirements for an introductory psychology course at Iowa State University. Thirty participants were randomly assigned to learn the word pairs through testing, and 29 through restudying. These groups are hereafter referred to as the test and study conditions, respectively.

Materials. Participants learned 16 cue–target pairs that had an average associative strength of 3% according to the norms of Nelson, McEvoy, and Schreiber (1998). These norms were also used to obtain a semantic mediator (i.e., a word with a high preexisting association with the cue). For example, when given the cue *Mother:* _____, the word *Father* (semantic mediator) is generated 60% of the time. Each semantic mediator was associated with its respective cue at 57%, on average, but was not associated with the target member of the cue–target pair (e.g., *Father* bears no preexisting semantic association with *Child*) or with any of the other 15 targets. A complete list of the materials can be found in the Appendix.

In addition, 16 new items were obtained that were not related to any of the cues, targets, or semantic mediators (see the unrelated items in the Appendix). The recognition test thus contained 32 old items (16 cues and 16 targets) and 32 new items (16 semantic mediators and 16 unrelated items).

Design and procedure. Participants first underwent an initial encoding phase in which they saw each of the 16 cue–target pairs one at a time and were asked to rate the relatedness between the two words on a scale of 1 (*not at all related*) to 5 (*highly related*). The cue and target appeared side-by-side in the center of the computer screen. The cue appeared in a box on the left, and the target appeared underlined in a box on the right. Participants were asked to try to remember these words for an upcoming memory test.

Immediately after entering a rating between 1 and 5 for each cue–target pair, participants completed a 20-s distractor task in which they were asked to add together a series of single-digit numbers appearing rapidly on the screen. Immediately after adding the numbers together and entering a response to indicate the sum total, participants in the test condition completed a test on the word pairs, and participants in the study condition were shown the same word pairs again.

Participants in the test condition saw the cue member of the pair presented by itself on the left side of the screen, and they were asked to type in the underlined word that had previously been paired with it. Participants' responses appeared on the right side of the screen, and after pressing the *Enter* key, the program advanced to the next item without providing feedback. Participants in the study condition were shown each of the 16 word pairs again, along with the same 1–5 rating scale, and were asked to rate the relatedness between the two words just as before. Each of the 16 word pairs was presented once for both test and study conditions, in an order that was randomized and different for each participant.

After all items had been presented as test or study, all participants engaged in a 5-min distractor task in which they were asked to list as many of the U.S. states as they could. Participants were shown a list numbered 1–50 on the computer screen, and they were instructed to enter a state name next to each number.

At the end of 5 min, all participants were given the same recognition test, which involved the 32 old items (16 cues and 16 targets) and the 32 new items (16 semantic mediators and 16 unrelated items) in a randomly presented order. Each of the 64 items was presented one at a time, in the center of the computer screen, with a scale labeled 1 (*I definitely did NOT see this word*) to 6 (*I definitely saw this word*). Participants responded by pressing the 1–6 key for each item and were not given feedback about the correctness of their responses.

Results and Discussion

Recognition test performance was calculated for both old and new items as a function of whether words were learned through test or study. First, the 6-point recognition scale was divided in half such that a response of 4 or higher was considered a "yes." All analyses were then repeated, this time for those items that received only a 5 or 6 (i.e., "yes" responses that were given with higher confidence). Finally, all analyses were again repeated for those items that received a 6 (i.e., "yes" responses given with the highest confidence).

Table 1 contains the hit rates for old items as a function of learning condition (test vs. study) and confidence of the response. Significant testing effects emerged regardless of confidence level. Cues that were learned through test produced significantly higher hit rates than did cues learned through study, and this was true whether the hit rate was based on all "yes" responses, $t(57) = 2.65$, $p < .011$; medium-confidence "yes" responses, $t(57) = 2.98$, $p < .005$; or high-confidence "yes" responses, $t(57) = 2.51$, $p < .016$. Hit rates for targets also revealed a significant testing effect whether they were based on all "yes" responses, $t(57) = 1.97$, $p = .05$ (two-tailed; $p = .02$ one-tailed); medium-confidence "yes" responses, $t(57) = 2.77$, $p < .008$; or high-confidence "yes" responses, $t(57) = 2.61$, $p < .013$.

Table 2 contains false alarm rates to new items as a function of learning condition (test vs. study) and the type of new item (semantic mediator vs. unrelated). False alarm rates for semantic mediators were higher for test than for study, whereas the opposite pattern was observed for the unrelated items. A 2 × 2 (Learning Condition × Item Type) mixed analysis of variance (ANOVA) revealed that this interaction was significant for all "yes" responses, $F(1, 57) = 7.46$, $p < .009$, $MSE = .007$; medium-confidence "yes" responses, $F(1, 57) = 6.12$, $p < .017$, $MSE = .006$; and high-confidence "yes" responses, $F(1, 57) = 4.61$, $p < .037$, $MSE = .005$. All three analyses also revealed a main effect for item type (all Fs > 54, ps < .001), such that the false alarm rate was greater overall for semantic mediators than for unrelated items. No main effect emerged for learning condition (all Fs < 1.16).

These results confirm the prediction that semantic mediators (e.g., *Father*) are more likely to be activated as a result of recalling a target from a cue (*Mother:* _____) compared with simply restudying the cue and target together (*Mother: Child*). Experiment 2 was conducted to explore whether semantic mediators are more likely to be linked with targets (e.g., *Father* → *Child*) as a result of testing compared with as a result of restudying.

Table 1
Hit Rates to Cues and Targets in Experiment 1 as a Function of Learning Condition and Level of Confidence

Confidence level	Test	Study
All "yes" responses (4s, 5s, 6s)		
Cues	.973 (.048)	.927 (.082)
Targets	.958 (.058)	.925 (.073)
Medium-confidence "yes" responses (5s, 6s)		
Cues	.971 (.048)	.918 (.084)
Targets	.954 (.057)	.905 (.078)
High-confidence "yes" responses (6s only)		
Cues	.954 (.075)	.830 (.261)
Targets	.948 (.062)	.825 (.250)

Note. Standard deviations are given in parentheses.

Table 2
False Alarm Rates to Semantic Mediators and Unrelated Items in Experiment 1 as a Function of Learning Condition and Level of Confidence

Level of confidence	Test	Study
All "yes" responses (4s, 5s, 6s)		
Semantic mediators	.181 (.158)	.135 (.133)
Unrelated items	.023 (.056)	.060 (.135)
Medium-confidence "yes" responses (5s, 6s)		
Semantic mediators	.169 (.141)	.127 (.122)
Unrelated items	.019 (.047)	.050 (.112)
High-confidence "yes" responses (6s only)		
Semantic mediators	.148 (.137)	.097 (.106)
Unrelated items	.019 (.047)	.026 (.054)

Note. Standard deviations are given in parentheses.

Experiment 2

Experiment 2 was identical to Experiment 1, in which half of the participants learned the 16 cue–target pairs through test, and the other half through study. This time, however, final retention was tested via cued recall in which participants were asked to recall the target (e.g., *Child*) from the same cue as before (e.g., *Mother:* _____) or from a new cue that they had not seen before. The new cues were either items that were weakly related to the target (e.g., *Birth:* _____) or semantic mediators that were not related to the target (e.g., *Father:* _____).

If the process of recalling a target from a cue establishes a link between a semantic mediator and a target, then recall of a target from a semantic mediator cue should be higher for participants who learned the cue–target pairs through test compared with through study. Furthermore, participants who learned the cue–target pairs through test should have an easier time recalling a target from a semantic mediator cue (e.g., *Father:* _____) compared with from a new related cue (e.g., *Birth:* _____), because the latter was presumably less likely to have been activated and linked with the target during the process of initial retrieval.

Method

Participants. Sixty-one participants were recruited from the same participant pool as in Experiment 1. Of these participants, 31 learned the cue–target pairs through test, and 30 through study. Eleven participants in the test condition were given the final test with same cues, 10 with new related cues, and 10 with semantic mediator cues. Ten participants in the study condition were given the final test with same cues, 10 with new related cues, and 10 with semantic mediator cues.

Materials, design, and procedure. In addition to the 16 cue–target pairs from Experiment 1, 16 new cues were obtained that had similar preexisting associative strength to the targets as did the original cues (see the related items in the Appendix). The new related cues were associated with only their respective targets and not with any of the other 15 targets.

The learning phases for test and study were identical to those in Experiment 1. In Experiment 2, however, participants were given the final cued recall test after a 30-min retention interval during which they worked on answering a variety of unrelated trivia questions.

At the beginning of the final test in Experiment 2, participants in all conditions were informed that the cues might or might not be the same ones that they originally saw. Participants were instructed to read each cue and to report which one of the previously learned underlined words came to mind when they saw that cue. Participants entered their responses by typing them onto the right-hand side of the screen next to the cue that was presented. Each of the 16 cues appeared one at a time, in a different random order for each participant. As in Experiment 1, the final test was self-paced and feedback was not provided.

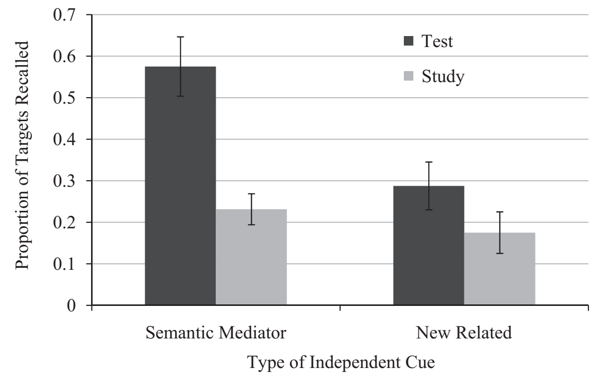

Figure 1. Final cued recall accuracy in Experiment 2, according to learning condition and type of independent cue. Participants learned cue–target pairs (e.g., *Mother: Child*) through either test (*Mother:* _____) or study (*Mother: Child*) and later were asked to recall the target from an independent cue that had some preexisting association to the target—that is, a new related cue (e.g., *Birth:* _____) or from an independent cue that had no preexisting association with the target but was strongly associated with the original cue—that is, a semantic mediator cue (e.g., *Father:* _____). Relative to participants who learned the items through study, those who learned the items through test showed a substantial advantage in recalling targets from semantic mediator cues compared with from new related cues.

Results and Discussion

The proportion of targets correctly recalled on the final test was calculated for both test and study conditions according to the three types of cues that participants received (same cues, new related cues, or semantic mediator cues). Results for the same cue condition revealed a classic testing effect, in that participants who learned the word pairs through test recalled more targets (M = .960, SD = .042) than did those who learned the word pairs through study (M = .900, SD = .084), $t(19) = 2.101$, $p < .05$.

Figure 1 displays the average proportion of targets recalled on the final test from independent cues (i.e., semantic mediator cues vs. new related cues) for both test and study conditions. For participants in the test condition, recall of targets from semantic mediator cues was greater than from new related cues. The advantage of semantic mediator cues over new related cues was not as strong for participants in the study condition, however. A 2 × 2 (Learning Condition × Cue Type) between-subjects ANOVA revealed that this interaction was significant, $F(1, 36) = 4.34$, $p < .045$, $MSE = .031$.

A main effect also emerged for learning condition, $F(1, 36) = 16.91$, $p < .001$, $MSE = .031$, indicating that targets were better recalled from independent cues by participants in the test condition (M = .431, SD = .248) compared with the study condition (M = .203, SD = .139). Finally, a main effect emerged for type of cue, $F(1, 36) = 9.60$, $p < .005$, indicating that targets were better recalled overall from semantic mediator cues (M = .403, SD = .249) than from new related cues (M = .231, SD = .176).

These results demonstrate that testing enhances performance on a cued recall test, even when the final test cues are not the same ones with which the targets were originally learned. Not all independent cues are created equal, however. Despite the fact that new related cues had greater preexisting associative links to the targets than did semantic mediator cues, the latter produced better final recall of targets. The advantage of semantic mediator cues over new related cues occurred to a much greater degree when targets were learned through test compared with study. It appears, therefore, that semantic mediators are more likely to be linked with targets as a result of initial retrieval compared with as a result of restudy.

General Discussion

In two experiments, final retention of target items was better when these items were learned through testing compared with through restudying. This replicates a number of studies reporting beneficial effects of testing (see e.g., Bjork, 1988; Carpenter, Pashler, & Vul, 2006; Chan, McDermott, & Roediger, 2006; Cull, 2000; Izawa, 1992; Karpicke & Roediger, 2008; Kuo & Hirshman, 1996; Mcdaniel, Roediger, & McDermott, 2007). The current article helps advance greater understanding of the causal mechanisms underlying this phenomenon by proposing a theoretical explanation for why tests are beneficial.

According to the mediator effectiveness hypothesis (Pyc & Rawson, 2010), tests are beneficial because they are more likely than restudy opportunities to promote the effectiveness of mediators (i.e., words or concepts that link a cue to a target). Pyc and Rawson found that, supporting this view, participants who learned Swahili–English word pairs (e.g., *Wingu: Cloud*) through testing (*Wingu:* _____) compared with through restudying (*Wingu: Cloud*) were more likely to remember a mediator that they had generated to link the cue to the target (e.g., “Wing”) and were also more likely to remember the target when cued with the mediator.

The current results address an important question that the mediator effectiveness hypothesis does not, however. That is, how does mediating information play a role in the typical testing effect, in which participants are not specifically asked to generate mediators? Participants in the current study were never instructed to generate mediators during learning but were simply asked to learn cue–target pairs through either cued recall or restudy. In Experiment 1, participants who learned these items through cued recall (e.g., *Mother:* _____) were more likely to make false alarms to an item that was never presented but closely associated with the cue—that is, a semantic mediator (e.g., *Father*)—than were participants who learned the same items through restudy (*Mother: Child*).

In Experiment 2, participants who learned the items through testing (e.g., *Mother:* _____) were more likely to produce the correct target (*Child*) on the final test in response to a semantic mediator cue (e.g., *Father:* _____) compared with a new cue that had some preexisting association with the target (e.g., *Birth:* _____). Participants who learned the items through restudy, on the other hand, did not demonstrate as large an advantage for semantic mediator cues over new related cues. The link between a semantic mediator and a target (e.g., *Father* → *Child*) is therefore more likely to be strengthened through testing compared with through restudying.

The current results support the notion that retention is facilitated by the presence of mediating information that links a cue to a target. They also expand upon this notion by demonstrating direct evidence of the type of mediating information that is spontaneously activated during retrieval—that is, when participants are not specifically asked to generate mediators. It appears that retrieving a target from a cue can activate a word or concept that is semantically related to the cue, which then becomes linked with the target and serves to facilitate its retrieval at a later time. Activation of this mediator is less likely to occur during restudy, because the availability of the answer reduces the need to activate additional information in an attempt to retrieve the target.

This semantic mediator hypothesis provides a more specific theoretical explanation for the testing effect than do previous views based on elaborative retrieval (see e.g., Carpenter, 2009; Carpenter & DeLosh, 2006; Glover, 1989). The semantic mediator hypothesis proposes that, beyond the notion that tests are beneficial because of extra information that is activated during retrieval, this extra information can be a semantic word or concept that is activated by the cue and becomes linked with the target during the process of retrieving the target from the cue.

This hypothesis accounts for the testing effect itself, as well as the tendency for the effect to be greater under conditions in which initial retrieval is made more difficult (see e.g., Carpenter & DeLosh, 2006; Kornell, Hays, & Bjork, 2009; Pyc & Rawson, 2009; Whitten & Bjork, 1977). Rendering target information less accessible at the time of initial retrieval would presumably increase the likelihood of activating semantically related information that can mediate the link between a cue and target.

The hypothesis also accounts for the tendency of testing effects to be greater when the final test is given after a longer, as opposed to a shorter, time delay (see e.g., Roediger & Karpicke, 2006b; Runquist, 1983). The organization of long-term memory is believed to be semantic in nature (see e.g., Bartlett, 1932), as contrasted with the more perceptual nature of short-term memory (see e.g., Baddeley, 1976). If an initial test is beneficial because it activates and links semantically related information from the cue to the target, then the advantages of testing would be more likely to show up on a final test that is delayed rather than immediate, because the former is more sensitive to semantic processing. The finding that testing reduces the rate of forgetting over time, therefore (see e.g., Carpenter, Pashler, Wixted, & Vul, 2008; Roediger & Karpicke, 2006b; see also Kornell, Bjork, & Garcia, in press), could be due to the tendency for tests to enhance semantic content that is more likely to remain in long-term memory.

Finally, it is important to note that the semantic mediator hypothesis provides an explanation for testing effects that have been observed in verbal cued recall tasks. It is likely that this hypothesis may not account for the effect that has been observed with a variety of different types of tasks, such as those that involve spatial processing (see e.g., see Carpenter & Pashler, 2007; Kang, 2010; Rohrer, Taylor, & Sholar, 2010). It is hoped that future research will explore potential boundary conditions of this and other theoretical accounts of the testing effect, as well as how best to use tests to maximize learning and retention.

References

Baddeley, A. D. (1976). *The psychology of memory.* New York, NY: Basic Books.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 397–401). New York, NY: Academic Press.

Butler, A. C., & Roediger, H. L., III. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19,* 514–527. doi:10.1080/09541440701326097

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1563–1569. doi:10.1037/a0017021

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology, 19,* 619–636. doi:10.1002/acp.1101

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. doi:10.3758/BF03193405

Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review, 14,* 474–478. doi:10.3758/BF03194092

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. doi:10.3758/BF03194004

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36,* 438–448. doi:10.3758/MC.36.2.438

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20,* 633–642. doi:10.3758/BF03202713

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. III. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135,* 553–571. doi:10.1037/0096-3445.135.4.553

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology, 14,* 215–235. doi:10.1002/(SICI)1099-0720(200005/06)14:3<215::AID-ACP640>3.0.CO;2-1

Glover, J. A. (1989). The "testing" phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology, 81,* 392–399. doi:10.1037/0022-0663.81.3.392

Izawa, C. (1992). Test trial contributions to optimization of learning processes: Study/test trials interactions. In A. F. Healy & S. M. Kosslyn

(Eds.), *Essays in honor of William K. Estes: From learning theory to connectionist theory* (Vol. 1, pp. 1–33). Hillsdale, NJ: Erlbaum.

Kang, S. H. K. (2010). Enhancing visuo-spatial learning: The benefit of retrieval practice. *Memory & Cognition, 38,* 1009–1017. doi:10.3758/MC.38.8.1009

Kang, S. H. K., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19,* 528–558. doi:10.1080/09541440601056620

Karpicke, J. D., & Roediger, H. L., III. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 704–719. doi:10.1037/0278-7393.33.4.704

Karpicke, J. D., & Roediger, H. L., III. (2008, February 15). The critical importance of retrieval for learning. *Science, 319,* 966–968. doi:10.1126/science.1152408

Kornell, N., Bjork, R. A., & Garcia, M. A. (in press). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language.*

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 989–998. doi:10.1037/a0015729

Kuo, T., & Hirshman, E. (1996). Investigations of the testing effect. *American Journal of Psychology, 109,* 451–464. doi:10.2307/1423016

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. doi:10.3758/BF03194052

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida Word Association, Rhyme, and Word Fragment Norms.* Retrieved from http://w3.usf.edu/FreeAssociation/

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60,* 437–447. doi:10.1016/j.jml.2009.01.004

Pyc, M. A., & Rawson, K. A. (2010, October 15). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330,* 335. doi:10.1126/science.1191465

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. doi:10.1111/j.1745-6916.2006.00012.x

Roediger, H. L., III., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17,* 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 233–239. doi:10.1037/a0017678

Runquist, W. N. (1983). Some effects of remembering on forgetting. *Memory & Cognition, 11,* 641–650. doi:10.3758/BF03198289

Whitten, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior, 16,* 465–478. doi:10.1016/S0022-5371(77)80040-6

Appendix

Items Used in the Current Study

Old items		New items			Associative strength		
Cue	Target	Mediator	Related	Unrelated	C–T	C–M	R–T
Mother	Child	Father	Birth	Bread	.010	.597	.015
Prescription	Doctor	Drug	Hospital	Horse	.034	.477	.027
Soil	Earth	Dirt	Continent	Teeth	.040	.717	.041
Dusk	Evening	Dawn	Morning	Train	.042	.609	.047
Donor	Heart	Blood	Liver	Circle	.042	.524	.041
Weapon	Knife	Gun	Ax	Game	.075	.592	.046
Sonnet	Music	Poem	Dancer	Floor	.059	.471	.052
Employment	Office	Job	Government	Clock	.020	.605	.024
Trash	Paper	Garbage	Ink	Forest	.013	.526	.013
Sea	River	Ocean	Flood	Bottle	.017	.456	.020
Vocabulary	School	Words	Text	Fence	.013	.507	.013
Jacket	Shirt	Coat	Hanger	Ladder	.013	.564	.014
Pedestrian	Street	Walk	Neighborhood	Noodle	.032	.597	.034
Breeze	Summer	Wind	Mosquito	Rabbit	.012	.606	.014
Coffee	Table	Tea	Banquet	Key	.020	.442	.020
Frame	Window	Picture	Shingle	Lace	.014	.811	.014

Note. Associative strength between mediators and targets was 0. C–T = cue-to-target strength; C–M = cue-to-mediator strength; R–T = related-to-target strength.

Received December 17, 2010
Revision received April 25, 2011
Accepted April 26, 2011 ■