

## Free recall enhances subsequent learning

Kathleen M. Arnold · Kathleen B. McDermott

Published online: 8 January 2013  
© Psychonomic Society, Inc. 2012

**Abstract** Testing, or retrieval practice, has become a central topic in memory research. One potentially important effect of retrieval practice has received little attention, however: It may enhance, or potentiate, subsequent learning. We introduce a paradigm that can measure the indirect, potentiating effect of free recall tests on subsequent learning, and then test a hypothesis for why tests may have this potentiating effect. In two experiments, the benefit of a restudy trial was enhanced when prior free recall tests had been taken. The results from a third, correlational study suggested that this effect might be mediated by the effect of testing on organization. Not only do encoding conditions affect later retrievability, but also retrieval attempts affect subsequent encoding effectiveness.

**Keywords** Recall · Memory · Learning

Testing is much more than a measure of memory; taking a test, or retrieval practice, modifies memory (for a recent review, see Roediger & Butler, 2011). A well-studied demonstration of this phenomenon is seen in the *testing effect*, the finding that retrieval enhances subsequent retention (e.g., Roediger & Karpicke, 2006). However, retrieval may have another enhancing effect that has been largely overlooked: It may enhance subsequent encoding, an effect known as *test-potentiated learning* (Arnold & McDermott, 2012; Izawa, 1971).

The dearth of research on test-potentiated learning is especially notable for free recall tests, possibly in part because of a conclusion that Tulving (1967) came to decades ago. Tulving claimed that in multitrial free recall paradigms, tests and study trials have equivalent effects on learning. He

concluded that subsequent recall “depends primarily on the total amount of time spent on the task, and that it is relatively little affected, if at all, by the distribution of this time between studying and recalling the material” (p. 181). That is, practicing recalling items has the same effect on learning as does studying the material.

However, Tulving (1967) also observed that the mechanisms underlying the equivalent effects of study and test trials were not the same. In a condition with three successive tests between study trials, forgetting occurred after the first test, but this loss was counteracted by a large increase in recall after each study trial. That is, in conditions with consecutive test trials, more learning seemed to occur on subsequent study trials than in conditions without consecutive test trials. These results suggest the additional tests may have potentiated learning during study.

Later researchers demonstrated that the distribution of study and test trials does affect learning, and that taking free recall tests between study trials enhances performance (Donaldson, 1971; Karpicke & Roediger, 2007; Lachman & Laughery, 1968; Roediger & Smith, 2012; Rosner, 1970). Although this research has shown that free recall tests enhance performance, how these tests do so is still unknown. Is learning enhanced because tests directly improve retention of the retrieved items (i.e., the testing effect), as was suggested by Donaldson? Or is performance enhanced because free recall tests potentiate subsequent learning, an indirect effect of testing, as was suggested by Rosner?

By manipulating both the number of prior tests and whether or not the material was restudied, we can identify whether the benefit incurred from restudying is enhanced by a recall test preceding the restudy phase. That is, do prior recall tests boost, or potentiate, the enhancement seen from restudying the material? To answer this question, in Experiment 1 we varied the number of initial tests (zero or three) and whether or not a restudy opportunity occurred following the initial tests. Taking initial tests boosted the

K. M. Arnold (✉) · K. B. McDermott  
Department of Psychology, CB1125, Washington University in St. Louis, St. Louis, MO 63130-4899, USA  
e-mail: kathleen.arnold@wustl.edu

amount of information acquired from the restudy phase (i.e., the difference in final recall between the groups who did and did not receive restudy was greater for participants who had taken initial tests). In Experiment 2, these findings were extended and replicated in an Internet sample to ensure the replicability and generalizability of the initial findings. Finally, in Experiment 3 we explored the role that enhanced organization may play in test-potentiated learning. Previous research had shown that testing enhances organization and that this enhancement partially underlies the testing effect (Zaromb & Roediger, 2010). We asked whether enhanced organization from testing might also underlie the test-potentiated learning effect.

## Experiment 1

Do free recall tests potentiate learning during subsequent study? This question was addressed in an undergraduate population using a between-subjects design.

### Method

**Participants** One hundred and seventy-three Washington University in St. Louis undergraduate students participated in exchange for class credit or \$10. The policies of the University's Human Studies Committee were followed for all experiments.

**Design** A 2 (initial tests: zero, three)  $\times$  2 (restudy, no restudy) between-subjects factorial design was used (see Fig. 1). Each participant was randomly assigned to one of four between-subjects conditions: initial tests and restudy ( $n = 43$ ), initial tests but no restudy ( $n = 43$ ), no initial tests but restudy ( $n = 44$ ), or no initial tests and no restudy ( $n = 43$ ).

Prior to this main task, all of the participants completed a separate recall task, which was included as a baseline measure of individual memory performance.

**Materials** For the baseline memory task, three lists of 15 related words from Roediger and McDermott's (1995)

Experiment 1	Experiment 2
S <sub>40</sub> T T T S <sub>40</sub> T	S <sub>40</sub> T T S <sub>20</sub> T
S <sub>40</sub> T T T - T	S <sub>40</sub> - - S <sub>20</sub> T
S <sub>40</sub> - - - S <sub>40</sub> T	
S <sub>40</sub> - - - T	

**Fig. 1** Designs of Experiments 1 and 2. Each row represents a between-subjects condition. S, study trial; T, test trial; subscripts indicate the numbers of pictures studied

report served as the stimuli. Each list contained words (e.g., *nurse*, *sick*, and *lawyer*) related to one target item (e.g., *doctor*) that never appeared in the list. All words were unrelated to the images used in the main experiment. The lists were studied in an order randomized across participants. Within each list of related words, items were presented in a random order.

In the main task, the stimuli were 40 line drawings taken from the Snodgrass and Vanderwart (1980) norms. The images were chosen for their high name agreement (86 %–100 %) and depicted simple concrete nouns (e.g., a carrot). On study trials, the order of images was randomized for each participant.

**Procedure** The participants were tested in groups of one to four. Instructions were presented on the computer and were the same for all participants.

In Part 1 (the baseline measure), participants studied 45 words. Each word was presented individually on the screen for 3 s, with a 500-ms interstimulus interval. Participants were then given 3 min to recall as many of the items as they could in any order via the computer keyboard.

During Part 2 (the main experiment), participants first studied 40 pictures, which were presented individually on the screen for 3 s, with a 250-ms interstimulus interval. The participants then worked on math problems for 30 s to eliminate primary memory effects. Next, those in conditions with initial testing were given 3 min to recall as many of the pictures as they could, in any order, by typing the name of the object depicted in each picture. All of the responses remained onscreen for the duration of the test. This procedure was repeated twice, for a total of three tests. During this time, the participants in the no-initial-test conditions played three games of Tetris, each lasting 3 min. Next, half of the participants restudied the pictures, and the other half worked on math problems. Restudy followed the same procedure as the initial study. All of the participants then worked on math problems for an additional 30 s before taking a final test. The participants were given 5 min to recall as many pictures as possible.

### Results and discussion

**Baseline measure** A one-way analysis of variance (ANOVA) between the four conditions revealed no significant differences,  $F < 1$ , and therefore, the baseline measure was not included in the remaining analyses.

**Main experiment** The final-test data<sup>1</sup> were analyzed using a 2 (initial tests: zero, three)  $\times$  2 (restudy, no restudy) between-

<sup>1</sup> For analyses of the initial-test data from all experiments, as well as the demographic and postexperimental question data from Experiments 2 and 3, see the supplemental materials.

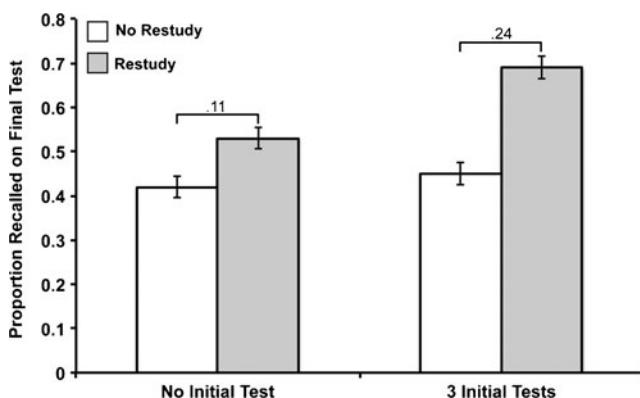
subjects ANOVA. Restudying enhanced later recall relative to not restudying ( $M = 0.61$  vs.  $0.43$ ; see Fig. 2),  $F(1, 169) = 50.87$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.23$ . Similarly, taking initial tests enhanced later recall relative to not taking initial tests ( $M = 0.57$  vs.  $0.47$ ),  $F(1, 169) = 15.07$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.08$ .

Test-potentiated learning would be substantiated by an interaction between initial tests and restudy conditions. That is, the difference between the restudy and no-restudy conditions should be larger in the condition with initial tests, indicating that the benefit of having a restudy trial was enhanced when initial tests had been taken. As can be seen in Fig. 2, this pattern emerged; the difference between the restudy and no-restudy conditions was larger when initial tests had been taken ( $M = 0.24$ ) than when they had not been taken ( $M = 0.11$ ),  $F(1, 169) = 8.21$ ,  $p = 0.005$ ,  $\eta_p^2 = 0.05$ , indicating that the initial tests potentiated learning during the restudy trial.

Another way to interpret this interaction would be to consider that a testing effect occurred only when participants restudied. There was no difference between the zero- and three-test conditions when participants had not restudied,  $t < 1$ . In contrast, when participants had restudied, recall was greater in the three-test condition ( $M = 0.69$ ) than in the zero-test condition ( $M = 0.53$ ),  $t(85) = 4.83$ ,  $p < 0.001$ ,  $d = 1.03$ . One might wonder whether this lack of a testing effect is problematic, in that a robust literature has demonstrated testing effects in the absence of a restudy condition. These testing effects, however, often do not emerge until after a delay (Roediger & Karpicke, 2006), and the present experiment was performed within a single session.

## Experiment 2

Experiment 1 demonstrated that taking free recall tests potentiated learning during a subsequent restudy trial. Experiment 2



**Fig. 2** Proportions of items recalled on the final test in Experiment 1. The mean differences between the restudy and no-restudy conditions for the two test conditions are displayed above the respective bars. Error bars represent standard errors of the means

was designed to replicate this finding and extend it to a paradigm in which restudy was manipulated within participants. Furthermore, a more diverse population was used. Rather than testing college undergraduates, participants were recruited through the Internet using the Amazon Mechanical Turk.

## Method

**Participants** One hundred and eighty participants completed the experiment on the Internet via the Amazon Mechanical Turk in exchange for \$2. Thirty-three of the participants were excluded from the analyses because, on a postexperimental question, they indicated that they had written words and/or picture names during the study phases. Additionally, seven participants were excluded because they failed to follow the instructions. After these exclusions, 140 participants remained in the final analyses.

**Design and materials** Prior to the main task, a separate recall test, identical to that used in Experiment 1, was given as a baseline measure of memory performance.

The main experiment had a 2 (initial tests: zero, two)  $\times$  2 (restudy, no restudy) mixed factorial design, with initial tests manipulated between-subjects and restudy manipulated within-subjects. The participants were randomly assigned to the initial-tests ( $n = 71$ ) or the no-initial-tests ( $n = 69$ ) condition. The stimuli were the same as in Experiment 1.

**Procedure** The procedure was similar to that of Experiment 1, with two major changes: (1) Participants were tested online and (2) all participants restudied half of the pictures. Which pictures were restudied was determined randomly for each participant.

Additional changes were made to accommodate the online testing format. Participants answered demographic questions prior to the beginning of the experiment, as well as postexperimental questions after completing the study. Furthermore, the participants studied each picture for 4 s (rather than 3 s), with a 500-ms (rather than 250-ms) interstimulus interval. Also, those in the initially tested condition took two (rather than three) initial free recall tests. Similarly, those who were not given initial tests played two (rather than three) games of Tetris. Finally, on the final test, the participants had 3 min (rather than 5 min) to recall the pictures.

## Results and discussion

**Baseline measure** A  $t$  test revealed a marginally significant difference in the proportions of words recalled in the baseline measure by participants who did ( $M = 0.38$ ) and did not ( $M = 0.43$ ) subsequently receive initial tests in the main

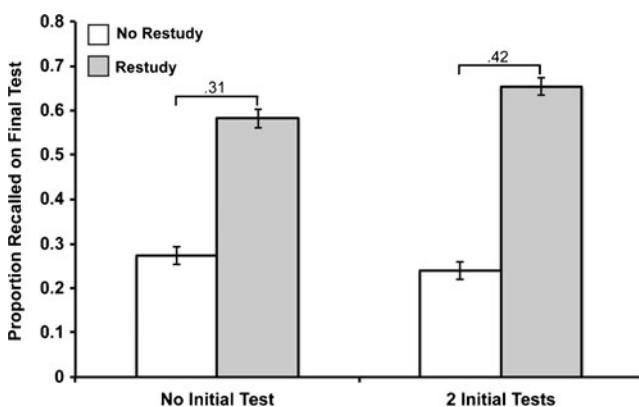
experiment,  $t(138) = 1.86$ ,  $p = 0.07$ ,  $d = 0.32$ . Because of this marginal effect, in analyses of the main experiment, baseline memory performance was used as a covariate to ensure that any differences between conditions were not due to preexperimental differences. However, including this baseline measure as a covariate did not change any conclusions (see the supplemental materials).

**Main experiment** The final-test data were analyzed using a 2 (initial tests: zero, two)  $\times$  2 (restudy, no restudy) mixed analysis of covariance (ANCOVA). The proportion of words recalled on the baseline measure was used as the covariate.

Restudied pictures were more likely to be recalled than those that were not restudied ( $M = 0.62$  vs. 0.26; see Fig. 3),  $F(1, 137) = 446.54$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.77$ . Unlike in Experiment 1, we found no main effect of testing; taking initial tests did not significantly enhance recall relative to not taking initial tests ( $M = 0.45$  vs. 0.43),  $F < 1$  (however, three tests had been used in Experiment 1, rather than two).

Test-potentiated learning would be indicated by a greater benefit of restudying pictures in the initial-tests condition than in the no-initial-tests condition. As Fig. 3 illustrates, this pattern was found; a larger difference was apparent between the proportions of restudied and not-restudied pictures recalled in the initial-tests condition than in the no-initial-tests condition ( $M = 0.42$  vs. 0.31),  $F(1, 137) = 9.32$ ,  $p < 0.003$ ,  $\eta_p^2 = 0.06$ .

As in Experiment 1, this interaction can also be interpreted as indicating a significant difference between the initial-test conditions for items that were restudied,  $F(1, 137) = 6.09$ ,  $p = 0.02$ ,  $\eta_p^2 = 0.04$ , but not for items that were not restudied,  $F(1, 137) = 1.30$ ,  $p = 0.26$ ,  $\eta_p^2 = 0.009$ . These results again show that a significant testing effect occurred only when participants were given the opportunity to restudy.



**Fig. 3** Adjusted means indicating the proportions of items recalled on the final test in Experiment 2. Adjusted mean differences between the restudy and no-restudy conditions for the two test conditions are displayed above the respective bars. Error bars represent standard errors of the means

### Experiment 3

Experiments 1 and 2 demonstrated that taking initial free recall tests can potentiate learning during a subsequent restudy trial. In Experiment 3, we explored why tests have this potentiating effect. Several previous researchers have suggested that tests enhance learning by improving the organization of already-learned material (e.g., Donaldson, 1971; Lachman & Laughery, 1968; Rosner, 1970). More recent work by Zaromb and Roediger (2010) provided evidence that free recall tests improve organization, and that this improvement partially underlies the testing effect in free recall. Does this enhancement also underlie the test-potentiated learning effect in free recall? Improving the organization of already-learned material may increase the ability to encode new items, by creating or improving a structure, or schema, that can be used to incorporate new items with already-learned items.

To test this hypothesis, in Experiment 3 we used categorized words. This change allowed organization to be measured through clustering, or recalling members of the same category together. Our measurement of clustering was adjusted for the total number of items recalled by using the adjusted ratio of clustering (ARC; Roenker, Thompson & Brown, 1971). If enhanced organization underlies test-potentiated learning, organization on an initial test should be related to the amount of information learned during a subsequent restudy trial. That is, more organized recall prior to restudying should be related to more learning during restudy. The relationship between prior organization and subsequent learning was tested in a correlational study.

### Method

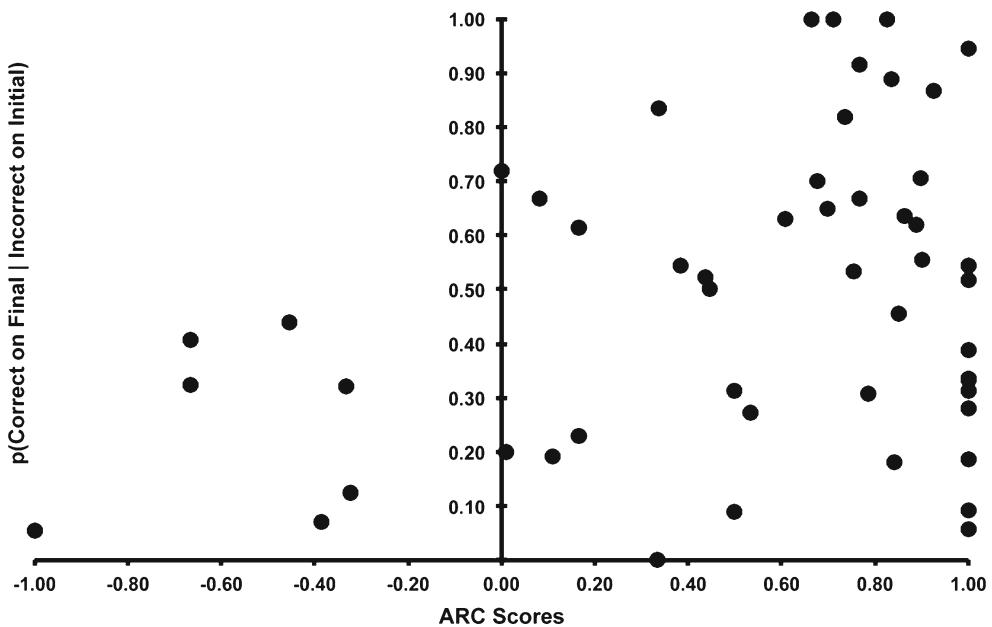
**Participants** Sixty-two participants completed the experiment via the Amazon Mechanical Turk in exchange for \$3. Seven of the participants were excluded from the final analysis because they reported writing down words and/or picture names. After these exclusions, 55 participants remained in the final analysis.

**Design** This was a correlational study. A baseline task was given prior to the main task, and in that main task, all participants took three initial tests and restudied all items.

**Materials** For the baseline task, 30 line drawings of easily identifiable nouns (all unrelated to the categorized words) were chosen from the Snodgrass and Vanderwart (1980) norms.

For the main task, five medium-frequency words from eight categories (total of 40 words) were chosen from the expanded and updated version of the Battig and Montague word norms (Van Overschelde, Rawson & Dunlosky, 2004).

**Fig. 4** Scatterplot of the proportions of words recalled on the final test given that they had not been recalled on any previous test, for each participant in Experiment 3, graphed as a function of adjusted-ratio-of-clustering (ARC) scores from the test prior to the restudy trial



**Procedure** As in Exp. 2, the participants were tested online via the Amazon Mechanical Turk and answered demographic questions before and postexperimental questions after the experiment.

The procedure was similar to that of Experiment 1—specifically, to the initial-tests-with-restudy condition—with two differences: On the baseline task, participants learned 30 (rather than 45) items, and on the final test, participants had 3 min (rather than 5 min) to recall items.

## Results and discussion

The role that organization may play in test-potentiated learning was examined by measuring the correlation between organization on the test prior to the restudy trial and learning during the restudy trial. Organization on the test prior to the restudy trial was measured using ARC scores, which can range from  $-1.0$  to  $1.0$ , with  $1.0$  indicating perfect organization,  $0$  indicating chance-level organization, and negative scores indicating below-chance-level organization.

Learning on the restudy trial was estimated using a conditional probability measure: the proportion of items recalled on the final test, given that they had not been recalled on any previous test. If organization of already-learned material improves subsequent learning, greater organization prior to restudying should be related to more learning during the restudy trial. As can be seen in Fig. 4, this pattern was found:<sup>2</sup> Higher ARC scores

on the test prior to the restudy trial were associated with a greater proportion of items learned during the restudy trial,  $r(49) = 0.31$ ,  $p = 0.03$ .

However, this relationship could be driven by a third variable. Specifically, individuals with better “memory ability” could tend to both have better organization and learn more during study trials. To test this possibility, the baseline memory measure was used as a covariate. When controlling for this estimate of memory ability, the correlation remained significant,  $r(48) = 0.34$ ,  $p = 0.02$ . That is, higher ARC scores were still associated with greater learning, suggesting that differences in memory ability did not drive this relationship.

## General discussion

The primary finding in this report is that free recall tests potentiate learning during subsequent restudy trials. The benefit of restudying the material was enhanced when initial free recall tests had been taken. This pattern was obtained when restudy was manipulated both between-subjects (Exp. 1) and within-subjects (Exp. 2), and in both an undergraduate population (Exp. 1) and a more diverse population recruited online (Exp. 2).

This potentiating effect may at least in part be due to enhanced organization. Previous research has shown that testing improves organization (Zaromb & Roediger, 2010). Experiment 3 indicated that better organization prior to restudying was associated with more learning during restudying. This relationship does not seem to be mediated by memory ability. Although this finding is only correlational, it suggests that tests may potentiate learning by enhancing organization prior to learning.

<sup>2</sup> Four participants did not recall enough items to calculate an ARC score and were not included in the correlation.

## Other explanations of test-potentiated learning

Other hypotheses have been proposed to explain test-potentiated learning. These alternative hypotheses are not mutually exclusive with the enhanced-organization hypothesis. Multiple factors may contribute to test-potentiated learning.

One such hypothesis is that test-potentiated learning may be driven by enhanced metacognitive knowledge. That is, tests may increase metacognitive accuracy (Roediger & Karpicke, 2006), which could be used to improve restudy strategies. For instance, testing may allow participants to better determine which items they cannot remember, and therefore which items they should focus on during the next restudy opportunity (Lachman & Laughery, 1968).

Recent work using functional magnetic resonance imaging has suggested another possible underlying mechanism for the enhancing effect of tests. Nelson, Arnold, Gilmore, Najjar, Finn and McDermott (2012) observed greater activation in the left posterior inferior parietal lobule during restudy of word pairs that had been tested on a cued recall test than during restudy of pairs not previously tested. That specific region has previously been associated with successful recognition memory (McDermott, Szpunar & Christ, 2009; Nelson et al., 2010), an observation that led the authors to suggest that the initial tests may have increased the tendency for study-phase retrieval, or reminders (Hintzman, 2004), during subsequent restudy. Although this work involved a different set of procedures than those used here, the results provide an intriguing possibility.

Several hypotheses have been proposed to explain the finding of enhanced encoding following a failed generation attempt (Grimaldi & Karpicke, 2012; Hays, Kornell & Bjork, 2012; Kornell, Hays & Bjork, 2009; see also Slamecka & Fevreiski, 1983), or what could be called *generate-potentiated learning*. In this paradigm, there is no initial study, and participants are asked either to guess the target of a cue word (e.g., *tide*–?) before studying the complete pair (e.g., *tide*–*beach*) or to study the pair without the initial guess. Final recall is enhanced for pairs that have an initial guess. The hypothesis favored by Grimaldi and Karpicke (2012), which they called the *search set theory*, posits that attempting to guess the target initiates a search process that activates related items. The experimentally defined target and items related to the target may become activated even if they are not given as the response, and this activation may enhance encoding when the target is subsequently presented.

However, this theory does not seem to generalize to free recall learning, especially in a paradigm in which the stimuli are unrelated to each other, as was the case in the first two experiments presented here. Furthermore, Grimaldi and Karpicke (2012) proposed that this activation process is

very short-lived and that the enhancing effect only occurs when the target is presented immediately after the retrieval attempt. In the experiments presented here, the delay between the initial tests and restudy suggests that any activation would have already dissipated.

## Conclusion

These experiments provide strong evidence that free recall tests have a potentiating effect on subsequent study and suggest that enhanced organization may underlie this potentiating effect. They introduce a new paradigm for studying test-potentiated learning and provide the first steps to a better understanding of the role of retrieval in learning. Not only does retrieval directly benefit future recall, but it also prepares the learner for future learning. In short, subsequent memory is enhanced by retrieval practice and by repeated study, and the combination of the two is an especially potent memory enhancer.

**Author note** These experiments were supported by a grant from the James S. McDonnell Foundation and by National Institutes of Health Training Grant No. 5T32GM081739. We thank Allison Cantor, Sophie Crumpacker, Laura D'Antonio, Andrew Fishell, and Fan Zou for help with the data collection and analyses.

## References

- Arnold, K. M., & McDermott, K. B. (2012). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/a0029199
- Donaldson, W. (1971). Output effects in multitrial free recall. *Journal of Verbal Learning and Verbal Behavior*, 10, 577–585.
- Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, 40, 505–513. doi:10.3758/s13421-011-0174-0
- Hays, M. J., Kornell, N., & Bjork, R. A. (2012). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/a0028468
- Hintzman, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, 32, 336–350. doi:10.3758/BF03196863
- Izawa, C. (1971). The test trial potentiating model. *Journal of Mathematical Psychology*, 8, 200–224. doi:10.1016/0022-2496(71)90012-5
- Karpicke, J. D., & Roediger, H. L., III. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57, 151–162. doi:10.1016/j.jml.2006.09.004
- 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
- Lachman, R., & Laughery, K. R. (1968). Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology*, 76, 40–50.

- McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, 47, 2290–2298.
- Nelson, S. M., Arnold, K. M., Gilmore, A. W., Najjar, L. M., Finn, B., & McDermott, K. B. (2012). *A neural signature of test-potentiated learning in parietal cortex*. Manuscript submitted for publication.
- Nelson, S. M., Cohen, A. L., Power, J. D., Wig, G. S., Miezin, F. M., Wheeler, M. E., & Petersen, S. E. (2010). A parcellation scheme for human left lateral parietal cortex. *Neuron*, 67, 156–170. doi:10.1016/j.neuron.2010.05.025
- Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15, 20–27. doi:10.1016/j.tics.2010.09.003
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255. doi:10.1111/j.1467-9280.2006.01693.x
- Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803–814. doi:10.1037/0278-7393.21.4.803
- Roediger, H. L., III, & Smith, M. A. (2012). The “pure-study” learning curve: The learning curve without cumulative testing. *Memory & Cognition*, 40, 989–1002. doi:10.3758/s13421-012-0213-5
- Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, 76, 45–48.
- Rosner, S. R. (1970). The effects of presentation and recall trials on organization in multitrial free recall. *Journal of Verbal Learning and Verbal Behavior*, 9, 69–74.
- Slamecka, N. J., & Fevreiski, J. (1983). The generation effect when generation fails. *Journal of Verbal Learning and Verbal Behavior*, 22, 153–163. doi:10.1016/S0022-5371(83)90112-3
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215. doi:10.1037/0278-7393.6.2.174
- Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 6, 175–184.
- Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, 50, 289–335. doi:10.1016/j.jml.2003.10.003
- Zaromb, F. M., & Roediger, H. L., III. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition*, 38, 995–1008. doi:10.3758/MC.38.8.995