Published in final edited form as:
J Exp Psychol Learn Mem Cogn. 2013 May ; 39(3): 940–945. doi:10.1037/a0029199.

Test-Potentiated Learning: Distinguishing Between Direct and Indirect Effects of Tests

Kathleen M. Arnold and **Kathleen B. McDermott**
Washington University in St. Louis

Abstract

The facilitative effect of retrieval practice, or testing, on the probability of later retrieval has been the focus of much recent empirical research. A lesser-known benefit of retrieval practice is that it may also enhance the ability of a learner to benefit from a subsequent restudy opportunity. This facilitative effect of retrieval practice on subsequent encoding is known as *test-potentiated learning*. Thus far, however, the literature has not isolated the indirect effect of retrieval practice on subsequent memory (via enhancing the effectiveness of restudy) from the direct effects of retrieval on subsequent memory. The experiment presented here uses conditional probability to disentangle test-potentiated learning from the direct effects of retrieval practice. The results indicate that unsuccessful retrieval attempts enhance the effectiveness of subsequent restudy, demonstrating that tests do potentiate subsequent learning.

The facilitative effect of retrieval practice on memory has garnered a great deal of empirical attention in recent years (for a recent review see Roediger & Butler, 2011). Most of this research has focused on how retrieval practice (i.e., testing) enhances later retrieval. Testing slows forgetting and therefore improves the likelihood of later retrieval (relative to a comparable condition with no intervening retrieval practice). This finding, known as *the testing effect*, is a direct effect of retrieval because the act of retrieving an item directly enhances later memory.

In addition to this direct effect, testing can also have indirect effects, or effects mediated through subsequent events. One such indirect effect of testing is the mediated effect of retrieval practice through later encoding. Attempting to retrieve items may improve later encoding of those items even when the retrieval attempt fails and feedback is not given. This effect, known as *test-potentiated learning*, was first identified by Izawa (1966), who suggested that more retrieval attempts lead to greater potentiation (but see Izawa, 1970). In the following decade, several researchers explored this phenomenon (e.g., Lachman & Laughery, 1968; LaPorte & Voss, 1974; Rosner, 1970; Royer, 1973; Tulving, 1967; Young, 1971), but, with few exceptions (e.g., Karpicke & Roediger, 2007; Kornell, Hays, & Bjork, 2009; Hays, Kornell, & Bjork, in press), recent literature has ignored this effect.

To examine test-potentiated learning, in multiple experiments Izawa (e.g., 1966, 1968, 1970, 1971) had participants learn paired items (e.g., $BIY - 97$) using multi-trial cued recall paradigms. She varied the number of test trials between interspersed study trials, and found that, when considered as a function of number of preceding study trials, conditions with more interspersed test trials had steeper learning curves (i.e., items were learned at a faster rate). She assumed that neither learning nor forgetting occurred during test trials, and

Correspondence to: Kathleen Arnold, Department of Psychology, CB1125, Washington University in St Louis St Louis, MO 63130-4899, kathleen.arnold@wustl.edu.

Portions of the data were presented at the 51st annual meeting of the Psychonomic Society in November 2010.

therefore concluded that the tests must have potentiated learning during subsequent study trials.

However, research on the testing effect calls into question the assumption that learning does not occur during the tests (Izawa, 1966; 1970). As discussed previously, the act of retrieval has a direct facilitative effect on later retention (Roediger & Butler, 2011). Therefore, the effect observed by Izawa may have been a result of a direct effect of the tests themselves through improved retention rather than an indirect, potentiating effect of the tests on subsequent study trials (but see description of additional analyses Izawa conducted in the Discussion). Hence, finding an advantage at final recall for conditions with more interspersed test trials is not conclusive evidence of test-potentiated learning, although it has been the dominant approach thus far (e.g., Karpicke, 2009; Karpicke & Roediger, 2007; LaPorte & Voss, 1974; Royer, 1973; Young, 1971).

In the present experiment, a modified version of Izawa's (1966, 1971) paradigm is used with two goals in mind. The first is to replicate Izawa's findings using educationally relevant stimuli (i.e., foreign-language vocabulary words). The second (and more important) goal is to expand Izawa's findings so that the indirect, potentiating effects of tests can be disambiguated from the testing effect. That is, the results will be analyzed using conditional probability to separate the indirect effect of tests on subsequent restudy from the direct effect of taking tests. Does attempting (but failing) to retrieve target items enable better acquisition of these items during subsequent encoding?

Method

Participants

Sixty-one Washington University in St. Louis undergraduate students participated in exchange for either class credit or $10. The participants were randomly assigned to either the 1-Test or 5-Test condition. One participant was excluded from the final analysis because that participants initial test performance (96%) was more than 4 standard deviations above the mean, indicating probable prior knowledge of Russian. After this exclusion, the conditions had the following number of participants: 1 Test (n = 31) and 5 Tests (n = 29).

Design

The design is illustrated in Figure 1. After initial study, participants took one or five tests (manipulated between-participants) before restudying the material. They repeated this test-restudy cycle for a total of three (5-Test condition) or nine (1-Test condition) iterations and then took a final test. In this way, all participants studied or were tested on the material 20 times; in the 5-Test condition, participants studied the material 4 times and were tested 16 times, and in the 1-Test condition, participants studied the material 10 times and were tested 10 times.

The test immediately prior to a restudy is denoted as the pre-restudy test, and the test immediately following a restudy is denoted as the post-restudy test. In the 1-Test condition, the post-restudy test of the previous restudy is the pre-restudy test for the next restudy.

Materials

Twenty-five Russian-English word pairs (e.g., *medved – bear*) were used as stimuli. For each study and test, a new random order was determined on a participant-by-participant basis.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Procedure

Participants were tested in groups of 1 to 4, and all instructions were presented on a computer. Participants were told this experiment would test their verbal ability by having them learn foreign language word pairs. Participants first studied all 25 word pairs. Each word pair was presented on the screen for 3 s with a 250 ms interstimulus interval. Participants were told to study each word pair so that when later presented with the Russian word, they would be able to recall the English translation.

Following initial study, participants took one or five tests. On each test, each Russian word was presented alone on the screen for 3 s during which time participants attempted to type the English translation. A 250 ms interstimulus interval separated the presentation of each Russian word. After taking the test(s), participants restudied the word pairs in the same manner as in the initial study. Participants repeated this test(s)-restudy cycle three (5-Test condition) or nine (1-Test condition) times. After the final restudy, participants took a final test, which proceeded in the same manner as the prior tests.

The experiment took approximately 1 hour, after which participants were debriefed.

Results

To provide an overview of the results, the data from all tests are displayed in Figure 2. This figure shows the average proportion of items recalled on each test in both conditions. The arrows represent the increase in proportion of words recalled after each of the first three restudy trials. Performance on test 1 did not differ between groups ($t < 1$), indicating no substantial pre-experimental group differences, as the groups were treated identically until after this test. The two groups not only began at the same point on the scale but also ended up at approximately similar performance levels (1-Test: $M = .91$, 5-Test: $M = .86$, $t(58) = 1.45$, $p = .15$). What differed, of course, was the progression of learning.

Consider first the 1-Test condition, which demonstrated a standard negatively-accelerated learning curve (Ebbinghaus, 1885/1964). The key question is how taking four additional interspersed tests changes the traditional learning curve. In the 5-Test condition, learning was also negatively accelerated. However, the rate of learning during restudy trials appears to have been greater than in the 1-Test condition, at least on the first restudy trial. That is, there was a larger change in performance from the pre-restudy test to the post-restudy test in the 5-Test condition. Below, these observations will be analyzed, and the differences between conditions will be decomposed.

Although the data from all tests are displayed in Figure 2, the following analyses are limited to three restudy trials because participants in the 5-Test condition only restudied the material three times.

Rate of Learning

Overall—The results were first analyzed to determine if they replicated Izawa's (e.g., 1970) experiments. Izawa found that performance did not vary across sequential tests without intervening study trials. She interpreted this finding as evidence that no learning or forgetting occurred during test trials themselves and therefore any differences in rate of learning must be due to differences occurring during study trials. In our experiment, we replicated this finding; performance did not significantly vary across each set of five tests between restudy trials in the 5-Test condition as illustrated in Figure 2 and confirmed by three repeated-measures analyses of variance (ANOVAs) (largest $F(4,112) = 1.84$, $p = .13$, $\eta_p^2 = .06$).

Izawa (e.g., 1970) found that the rate of learning during restudy trials was faster in conditions with more prior tests. To examine this, we compared learning, defined as the proportion of words recalled on the post-restudy test, in both conditions while holding number of study opportunities constant. That is, could participants recall more items in the 5-Test condition than in the 1-Test condition after the first, second, and third restudy opportunities?

As can be seen in Table 1, after each of the restudy trials, more items were recalled in the 5-Test than in the 1-Test condition. A 3 (Restudy: 1, 2, 3) X 2 (Tests: 1, 5) mixed AN OVA on the post-restudy tests confirmed that the difference between conditions was significant; averaged across restudy trials, recall was greater in the 5-Test than in the 1-Test condition (M = .72 vs .62), $F(1, 58) = 5.49$, $p = .02$, $\eta_p^2 = .09$. There was also a significant main effect of restudy, $F(2,116) = 216.91$, $p < .001$, $\eta_p^2 = .79$. There was no significant interaction between restudy and test condition ($F < 1$), indicating that recall in the 5-Test condition was consistently greater than in the 1-Test condition across all three restudy trials. These results indicate that taking more interspersed tests increased the overall rate of learning. That is, given the same number of study trials, participants who had taken more tests could recall more items.

Rate of Learning on each Restudy Trial: Adjusted Means—Although the preceding analysis indicates a greater overall rate of learning, it cannot speak to the rate of learning on, or effectiveness of, each individual restudy trial. It does not indicate whether participants in the 5-Test condition continued to learn more effectively during each restudy trial because it does not account for differences in prior knowledge. This is especially important for the second and third restudy trials because there were substantial differences between conditions prior to restudying the material that may account for post-restudy test differences (see Table 1 and Figure 2). In both Restudy 2 and 3, the pre-restudy tests were at least marginally statistically different, $t(58) = 1.93$, $p = .06$, $d = .51$ and $t(58) = 2.65$, $p = .01$, $d = .70$, respectively. Removing variance due to these pre-restudy differences will allow for comparisons of the change in performance across restudy trials.

The change in performance for each restudy trial was analyzed by comparing post-restudy recall via three between-participants analyses of covariance (ANCOVAs) with recall on the pre-restudy test as the covariate. Because pre- and post-restudy tests in the 1-Test condition are not independent, learning during each restudy was analyzed separately. First, performance on the post-restudy test following the first restudy was examined. As seen in Figure 3, after adjusting for the pre-restudy test, performance on the post-restudy test was greater in the 5-Test condition than in the 1-Test condition (M = .54 vs. .44). A main effect of test condition, $F(1, 57) = 11.75$, $p = .001$, $\eta_p^2 = .17$, indicated that learning at this early stage in the experiment was more effective in the 5-Test condition.

Following the second restudy, there was no difference between conditions on the post-restudy test after adjusting for pre-restudy performance. Although the adjusted recall was numerically greater in the 5-Test than in the 1-Test condition (M = .72 vs. .68), this difference was not significant, $F(1, 57) = 1.80$, $p = .19$, $\eta_p^2 = .03$. Similarly, adjusted post-restudy test performance following the third restudy was not statistically different in the 5-Test and 1-Test conditions (M = .80 vs. .82), $F < 1$. These results indicate that after the first restudy, change in performance was approximately equal in both conditions after adjusting for prior knowledge.

Although the advantage of the 5-Test condition diminished as overall learning increased, it should be noted that the differences on pre-restudy tests before Restudy 2 and 3 mean that participants in the 5-Test condition were farther along on the learning curve than participants in the 1-Test condition (see Figure 2). Because learning is negatively accelerated, participants farther along on the learning curve would not be expected to learn at as fast a rate as participants at an earlier point on the curve. Although the ANCOVA adjusted for differences on the pre-restudy test, it could not adjust for the inherent differences in learning rates that occur at different points along the learning curve. These differences could mask any potential benefit of the interspersed tests.

Conditional Probability

Differences in final recall using both the adjusted and unadjusted means indicate that the rate of learning during restudy trials was faster when more initial tests had been taken, at least early in the learning curve. However, as has been previously mentioned, research on the testing effect has indicated that retrieving items provides direct benefits in the form of improved retention of retrieved items (Roediger & Butler, 2011). These tests may therefore have enhanced the retention of retrieved items (i.e., protected against forgetting), which is an effect that would not be apparent from these analyses. The increased rate of learning may have stemmed, at least in part, from this direct benefit of the tests. To determine what role test-potentiated learning may have played in increasing the rate of learning, conditional probability analyses are used below to decompose post-restudy recall into two components: items that were retained from the pre- to the post-restudy tests and items that were retrieved for the first time on the post-restudy test. Items that were retained may have benefited from both direct and indirect effects of the tests. That is, both the testing effect and the test-potentiated learning effect may have contributed to the recall of these items. Items that were newly retrieved, however, may have only benefited from the indirect, or potentiating, effect of tests. Therefore, if taking more prior tests enhanced the proportion newly retrieved items, the additional tests can be assumed to have potentiated learning during the restudy trial.

Retention—The left panel in Figure 4 shows the proportion of items retrieved on the pre-restudy test that were also retrieved on the post-restudy test (i.e., retained). In Restudy 1, a larger proportion of items were retained in the 5-Test than in the 1-Test condition (M = .96 vs .89), $t(57) = 2.13$, $p = .04$, $d = .56$. In Restudy 2, however, although there is a numerical advantage in the 5-Test condition as compared to the 1-Test condition ($M = .95$ vs .89), this difference was not significant, $t(57) = 1.64$, $p = .11$, $d = .43$. In Restudy 3, an equivalent proportion of items were retained in both conditions (both *M*s = .93), $t < 1$. The additional interspersed tests increased the proportion of items that could be retained, at least during Restudy 1. However, the degree to which this advantage came from a direct versus an indirect benefit of the tests cannot be ascertained. Further, although it is tempting to conclude that this benefit decreases as more items are learned, because an overall AN OVA cannot be conducted due to lack of independence of the tests in the 1-Testcondition, no across-restudy conclusions can be made. In addition, ceiling effects arose after Restudy 1.

Newly Retrieved—The right panel in Figure 4 shows the proportion of items not correctly retrieved on the pre-restudy test that were correctly retrieved on the post-restudy test. That is, this analysis shows the proportion of items available to be learned that were learned during the restudy trial. Therefore this analysis can be used to answer the primary aim of this study: Does failing to retrieve items enhance encoding during subsequent restudy trials?

In Restudy 1, a larger proportion of items were newly retrieved in the 5-Test than in the 1-Test condition ($M = .42$ vs .32), $t(58) = 2.41$, $p = .02$, $d = .63$. Similarly, in Restudy 2, a larger proportion of items were also newly retrieved in the 5-Test than in the 1-Test condition ($M = .58$ vs .46), $t(58) = 2.22$, $p = .03$, $d = .58$. In Restudy 3, a numerically larger proportion of items were newly retrieved in the 5-Test than in the 1-Test condition ($M = .68$ vs .59). However, this difference was not significant, $t(58) = 1.10$, $p = .27$, $d = .29$. These results indicate that the additional tests potentiated learning such that a larger proportion of items that were available to be learned were indeed learned during restudy in the 5-Testthan in the 1-Test condition, at least during Restudies 1 and 2. Again, because an overall ANOVA cannot be conducted, no across-restudy conclusions can be drawn.

Because more new items were acquired during the restudy trials when more prior tests had been taken, these results demonstrate test-potentiated learning.

Discussion

There were two main goals of this experiment. The first was to replicate Izawa's (e.g., 1966, 1971) findings that more interspersed tests increase the rate of learning on subsequent restudy trials using educationally relevant stimuli (i.e., foreign-language word pairs). As Figure 2 illustrates, this pattern was replicated. The rate of learning during restudy trials was faster in the 5-Test than in the 1-Test condition; given the same number of study trials, participants could recall more items in the 5-Test condition (see also Karpicke & Roediger, 2007). Further, the relative impact of each restudy trial was examined. The first restudy opportunity was more advantageous when more prior tests had occurred. This effect disappeared by the second and third restudy opportunities, although interpretation of these results is complicated by differences in prior knowledge that may have also influenced the rate of learning.

The second, and primary, goal of this experiment was to isolate test-potentiate learning from the testing effect. To do this, conditional probability was used to separately examine items that were retained from the pre- to the post-restudy tests and items that were newly retrieved on the post-restudy test. More items were retained in the 5-Test than in the 1-Test condition, at least in Restudy 1. Both the testing effect and the test-potentiated learning effect may have contributed to the larger proportion of retained items in the 5-Test condition. Retrieving the items on the additional tests may have directly enhanced retention (i.e., the testing effect) and/or the effect of retrieving the items may have been mediated by restudying (i.e., the test-potentiated learning effect). Further research is needed to separate these effects.

In contrast, items that were newly retrieved on the post-restudy test were likely learned on the restudy trial, and therefore any effect of the tests can be assumed to be mediated through this restudy opportunity. A larger proportion of items were newly retrieved in the 5-Test than in the 1-Test condition during Restudies 1 and 2, suggesting that the effect of the tests enhanced learning during these restudy trials. This is the strongest evidence yet that prior tests potentiate learning during subsequent restudy.

Although Izawa's (e.g., 1966, 1971) primary method for identifying test-potentiated learning was through the analysis of learning curves, she did at times supplement these results with conditional probability analyses. Collapsing across multiple study trials, she found preliminary evidence that more interspersed tests enhanced the proportion of newly retrieved items, suggesting that tests potentiate subsequent learning. The present experiment goes beyond Izawa's work by examining newly retrieved items following each restudy trial individually. This method has two advantages. First, the effect of potentiated tests on the rate of learning can be unconfounded from inherent changes in learning rate due to negative acceleration (at least during the first restudy trial). Second, changes in the effect of test-potentiated learning as learning progresses can be observed. The results provide preliminary evidence that the effect of test-potentiated learning may diminish as learning progresses.

Recently, Kornell, Hays, and Bjork (2009) introduced another paradigm that has been used to examine the effect of a retrieval attempt on a subsequent study trial (see also Slamecka &

Fevreiski, 1983). In this paradigm, participants are not given an initial study trial. Instead, they attempt to retrieve or, more accurately, guess the response to a cue before studying the cue-target pair. On a later test, participants were more likely to recall targets that they had initially attempted to generate than targets that had been studied without an initial generation attempt, suggesting that the initial generation attempt potentiated learning on the subsequent study trial. However, because there is no explicit instruction to recall (see Karpicke & Zaromb, 2010), this effect is more accurately called generation-potentiated learning rather than test-potentiated learning, and there may be important differences between these effects. For instance, unlike test-potentiated learning, generation-potentiated learning may only occur when the cue and target are related and the target is presented immediately after the generation attempt (see Grimaldi & Karpicke, 2012; Hays, Kornell, & Bjork, in press; Huesler & Metcalfe, 2012).

Two caveats in interpreting the retention and new retrieval results should be mentioned. First, because of the use of conditional probability, a larger proportion of items does not necessarily indicate a larger number of raw items. This is particularly true for Restudies 2 and 3 because pre-restudy test performance was not equated between test conditions.

Second, there is a possible confound that complicates these results. The spacing of the study trials is not equated in the test conditions. To have four additional interspersed tests in the 5-Test condition, the study trials had to be spaced farther apart (see Figure 1). Because increased spacing has been shown to benefit learning (for a review, see Cepeda, Pashler, Vul, & Wixted, 2006), the effects observed in this study could have been at least partially due to the spacing difference between conditions. Evidence from Izawa's (1968, 1971) research suggests that tests have a potentiating effect that is independent of the time it takes to take the tests, which would suggest that the results observed in this study are not solely due to spacing. However, further research is needed that explicitly controls for spacing while using analyses that can disambiguate the test-potentiated learning effect from the testing effect. Preliminary research that has been conducted in our own lab suggests that spacing alone cannot account for the observed potentiating effects.

In addition to these caveats, it should be noted that item-selection effects might hinder the observation of a test-potentiated learning effect during Restudies 2 and 3. Prior to these restudy trials, more items were already learned in the 5-Test than in the 1-Test condition, and therefore fewer items were available to be learned. This means that the items that were available to be learned in the 5-Test condition were presumably items that were more difficult for the participants to learn. Therefore, the finding that participants in this condition still learned a larger proportion of available items during Restudy 2 is all the more impressive. Additionally, this difference in prior learning means that to retain the same proportion of learned items, participants in the 5-Test condition had to retain more items, which presumably included more difficult items. This item-selection effect may have made it difficult for an advantage in the 5-Test condition to emerge.

The results presented here indicate that test-potentiated learning does occur such that prior tests enhance the encoding of not-yet-retrieved items. This study goes beyond previous research and provides the most stringent evidence yet for test-potentiated learning by disambiguating the test-potentiated learning effect from the testing effect. We propose that test-potentiated learning is most definitively identified when these effects are disambiguated such that the indirect effect of tests can be separated from the direct effect of tests, and future work on test-potentiated learning should aim to differentiate these effects. Similarly, we propose that when the focus of the research is on the direct effect of tests (i.e., the testing effect), the possible influence of indirect effects should be taken into account when the paradigm includes a restudy opportunity. For instance, experiments that show that feedback enhances the testing effect (e.g., Kang, McDermott, & Roediger, 2007) may actually be showing that the testing effect in conjunction with the test-potentiated learning effect is more powerful than the testing effect alone. Future research on the relative contribution of each effect will enhance the overall understanding of the role of tests in learning.

Acknowledgments

These experiments were supported by a grant from the James S. McDonnell Foundation and the National Institutes of Health training grant T32 GM081739-05. We thank Allison Cantor for help with data collection and members of the McDermott and Roediger laboratories for comments on an early version of this manuscript. Additionally, we thank Michael Strube who provided advice on statistical analyses.

References

Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin. 2006

Ebbinghaus, H. Memory: A contribution to experimental psychology. Ruger, HA.; Bussenius, CE., translators. New York: Dover; 1964. (Original work published 1885)

Grimaldi PJ, Karpicke JD. When and why do retrieval attempts enhance subsequent encoding? Memory & Cognition. 2012; 40:505–513. [PubMed: 22238214]

Hays MJ, Kornell N, Bjork RA. When and why a failed test potentiates the effectiveness of subsequent study. Journal of Experimental Psychology: Learning, Memory, Cognition. (in press).

Huelser BJ, Metcalfe J. Making related errors facilitates learning, but learners do not know it. Memory & Cognition. 2012; 40:514–527. [PubMed: 22161209]

Izawa C. Reinforcement-test sequences in paired-associate learning. Psychological Reports. 1966; 18:879–919.

Izawa C. Effects of reinforcement, neutral and test trials upon paired-associate acquisition and retention. Psychological Reports. 1968; 23:947–959.

Izawa C. Comparison of reinforcement and test trials in paired-associate learning. Journal of Experimental Psychology. 1969; 81:600–603.

Izawa C. Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. Journal of Experimental Psychology. 1970; 83:340–344.

Izawa C. The test trial potentiating model. Journal of Mathematical Psychology. 1971; 8:200–224.

Kang SHK, McDermott KB, Roediger HL. Test format and corrective feedback modify the effect of testing on long-term retention. European Journal of Cognitive Psychology. 2007; 19:528–558.

Karpicke JD. Metacognitive control and strategy selection: Deciding to practice retrieval during learning. Journal of Experimental Psychology: General. 2009; 138:469–486. [PubMed: 19883131]

Karpicke JD, Roediger HL. Repeated retrieval during learning is the key to long-term retention. Journal of Experimental Psychology: General. 2007; 138:469–486. [PubMed: 19883131]

Karpicke JD, Zaromb FM. Retrieval mode distinguishes the testing effect from the generation effect. Journal of Memory and Language. 2010; 62:227–239.

Kornell N, Hays MJ, Bjork RA. Unsuccessful retrieval attempts enhance subsequent learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:989–998.

Lachman R, Laughery KR. Is a test trial a training trial in free recall learning? Journal of Experimental Psychology. 1968; 76:40–50.

LaPorte R, Voss JF. Paired-associate acquisition as a function of a number of initial nontest trials. Journal of Experimental Psychology. 1974; 103:117–123.

Roediger HL, Butler AC. The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences. 2011; 15:20–27. [PubMed: 20951630]

Rosner SR. The effects of presentation and recall trials on organization in multitrial free recall. Journal of Verbal Learning and Verbal Behavior. 1970; 9:69–74.

Royer JM. Memory effects for test-like events during acquisition of foreign language vocabulary. Psychological Reports. 1973; 32:195–198.

Slamecka NJ, Fevreiski J. The generation effect when generation fails. Journal of Verbal Learning and Verbal Behavior. 1983; 22:153–163.

Tulving E. The effects of presentation and recall of material in free-recall learning. Journal of Verbal Learning and Verbal Behavior. 1967; 6:175–184.

Young JL. Reinforcement-test intervals in paired-associate learning. Journal of Mathematical Psychology. 1971; 8:58–81.

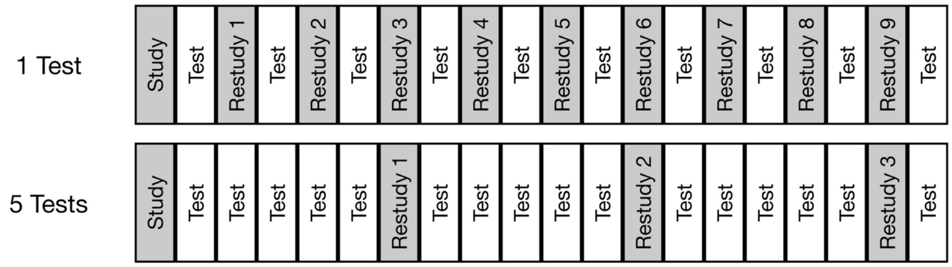

Figure 1.

Experimental design. The top row represents the 1-Test condition, and the bottom row represents the 5-Test condition. After initial study, participants took 1 or 5 tests before restudying the word pairs. This test(s)-restudy cycle repeated nine (1-Test) or three (5-Test) times. All participants then took a final test.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

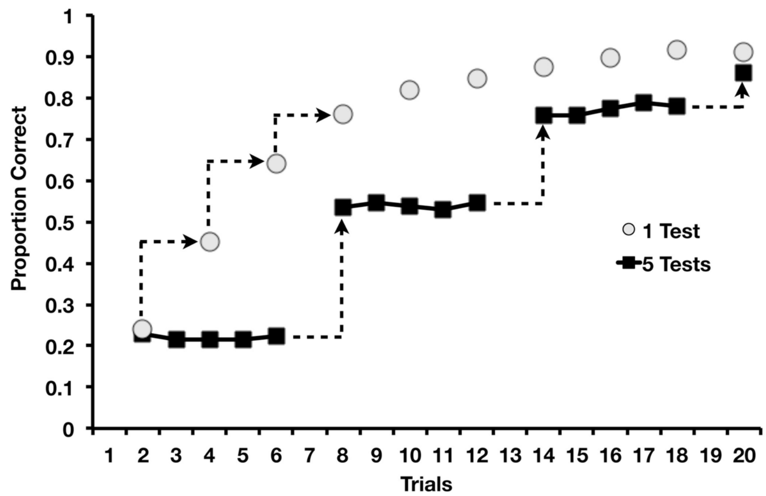

Figure 2.

Proportion of items recalled on each test in both conditions as a function of trial number. Arrows represent the change in performance from the pre- to the post-restudy test for the first three restudy trials.

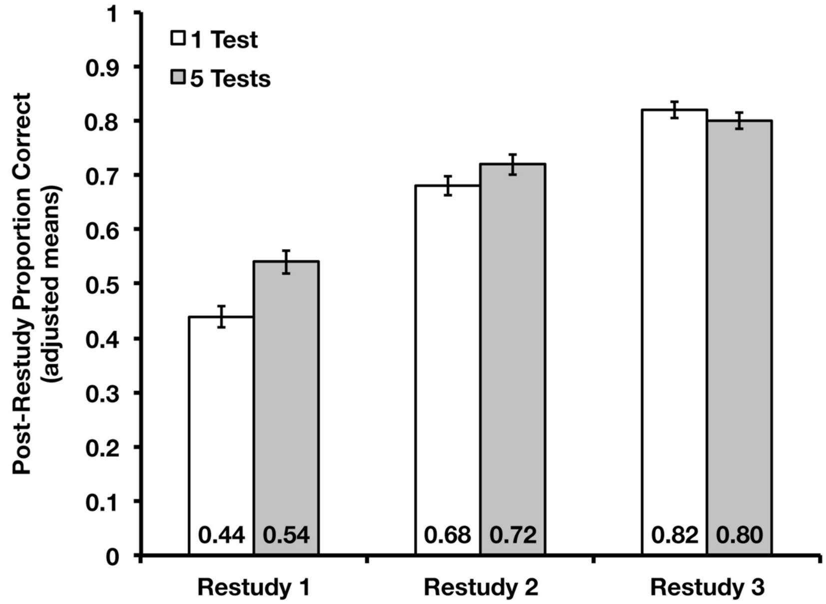

Figure 3.

Proportion of items recalled on the post-restudy tests after adjusting for recall on the pre-restudy tests for Restudies 1, 2, and 3. Means for each condition are displayed in their respective bars. Error bars represent standard errors of the mean.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

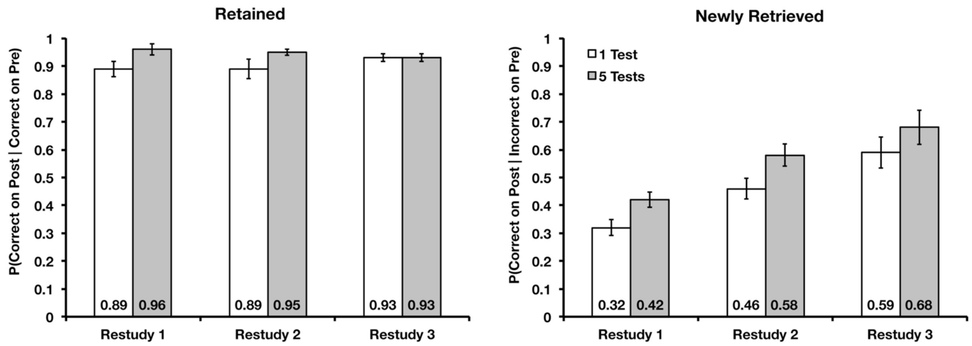

Figure 4.

Left panel: The proportion of items retrieved on the pre-restudy tests that were also retrieved on the post-restudy tests after Restudies 1, 2, and 3 as a function of test condition. Right panel: The proportion of items not retrieved on the pre-restudy tests that were retrieved on the post-restudy tests after Restudies 1, 2, and 3 as a function of test condition. Means for each condition are displayed in their respective bars. Error bars represent standard errors of the mean.

Table 1

The mean proportion of items retrieved (unadjusted) on the pre- and post-restudy tests for Restudy 1, 2, and 3. Standard errors of the mean are shown in parentheses.

		Pre-Restudy Test	Post-Restudy Test
Restudy 1	1 Test	0.22 (.02)	0.45 (.03)
	5 Tests	0.22 (.02)	0.54 (.03)
Restudy 2	1 Test	0.45 (.03)	0.64 (.04)
	5 Tests	0.55 (.04)	0.76 (.04)
Restudy 3	1 Test	0.64 (.04)	0.76 (.04)
	5 Tests	0.78 (.04)	0.86 (.03)