

# Interfering Effects of Retrieval in Learning New Information

Bridgid Finn and Henry L. Roediger III  
Washington University in St. Louis

In 7 experiments, we explored the role of retrieval in associative updating, that is, in incorporating new information into an associative memory. We tested the hypothesis that retrieval would facilitate incorporating a new contextual detail into a learned association. Participants learned 3 pieces of information—a person's face, name, and profession (in Experiments 1–5). In the 1st phase, participants in all conditions learned faces and names. In the 2nd phase, participants either restudied the face–name pair (the restudy condition) or were given the face and asked to retrieve the name (the test condition). In the 3rd phase, professions were presented for study just after restudy or testing. Our prediction was that the new information (the profession) would be more readily learned following *retrieval* of the face–name association compared to *restudy* of the face–name association. However, we found that the act of retrieval generally undermined acquisition of new associations rather than facilitating them. This detrimental effect emerged on both immediate and delayed tests. Further, the effect was not due to selective attention to feedback because we found impairment whether or not feedback was provided after the Phase 2 test. The data are novel in showing that the act of retrieving information can inhibit the ability to learn new information shortly thereafter. The results are difficult to accommodate within current theories that mostly emphasize benefits of retrieval for learning.

*Keywords:* memory malleability, updating, memory retrieval

For over a century memory formation has been characterized as a dynamic process that extends over time (Müller & Pilzecker, 1900). According to one classic view of memory formation, after a period of initial encoding, memories become consolidated, a process that is thought to unfold over time and result in stable, long term storage (e.g., McGaugh, 1966, 2000). The view that memories remain unchanged after initial consolidation has been challenged by studies demonstrating that under some circumstances, reactivation and retrieval of memories renders the retrieved trace malleable and subject to revision. The idea that memories can be revised during retrieval has a long history in human experimental psychology (e.g., Bartlett, 1932; Bjork, 1975; Loftus & Palmer, 1974). Research in animal models also suggests that once memories are reactivated they can be modified as they are reconsolidated (D. J. Lewis, 1979; Misanin, Miller, & Lewis, 1968; Nader, Schafe, & LeDoux, 2000; for a recent review see Alberini, 2011). Work using reconsolidation paradigms has also been extended to humans (see Schiller & Phelps, 2011, for a review), and other recent studies of memory reactivation (e.g., Gisquet-Verrier & Riccio, 2012) have also examined the role of retrieval in memory modification. Key questions that result from these research tradi-

tions are aimed at understanding the functional role of memory malleability and the conditions under which such lability of memory is adaptive.

Prior research in humans has primarily been focused on changing the nature of the previously learned information directly, as in the A–B, A–D paradigm of paired-associate learning. In this task, participants first learn an association (couch–eagle) and later must learn a new response to the previous stimulus (couch–muskrat); a rich literature abounds on using this task to study processes of proactive and retroactive interference (e.g., Barnes & Underwood, 1959; see Crowder, 1976, for a partial review). In the studies presented here, we explored a different but related issue, namely, the role of the retrieval process in acquiring new information (somewhat like adding a C term to a previously learned A–B pair, another task with a rich prior history, C. H. Lewis & Anderson, 1976). We tested the hypothesis that the act of retrieving information (relative to restudying information) will facilitate the updating of recently established memories when a new contextual detail is added. In short, will the act of retrieval permit new learning to occur more readily than without such retrieval?

A large body of cognitive research has demonstrated that retrieval can modify memory, both positively and negatively (e.g., Bartlett, 1932; Gates, 1917). There is abundant evidence that retrieval on a first test can enhance retention of the retrieved information on a second test given later (see Roediger & Butler, 2011, for a review). There is also ample evidence that memories can be revised with new information introduced during the retrieval process, for example, in the classic studies of the misinformation effect (see Loftus, 2005, for a review). One interpretation of these latter findings is that participants update their memory of an event with the new (mis)information introduced during retrieval, a process not dissimilar from A–B, A–D interference. On

---

This article was published Online First April 8, 2013.

Bridgid Finn and Henry L. Roediger, III, Department of Psychology, Washington University in St. Louis.

Support for this research was provided by a James S. McDonnell Foundation 21st Century Science Initiative Grant: Bridging Brain, Mind and Behavior/Collaborative Award. We thank Larry Jacoby, David Riccio, and Chris Wahlheim for helpful comments.

Correspondence concerning this article should be addressed to Bridgid Finn, who is now at the Educational Testing Service, 660 Rosedale Road, Princeton, NJ 08541. E-mail: [bfinn@ets.org](mailto:bfinn@ets.org)

a different front, an example of adaptive (rather than maladaptive) updating can be found in studies on test-potentiated learning. These experiments show that the act of attempting retrieval, even if it fails, enhances new learning (Arnold & McDermott, 2013; Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012; Izawa, 1966; Kornell, Hays, & Bjork, 2009). For example, Kornell et al. (2009) had subjects learn weakly related cue–target pairs (frog–pond). Some of the pairs were pretested (frog–?) by having subjects generate a response before presentation of the intact cue–target pair. Because the target was a weak associate, the participants produced a different associate (e.g., toad) most of the time during the pretesting. Immediately after generating the incorrect response, they were presented with the correct cue–target pair as feedback (frog–pond). Kornell et al. found that on a final follow up test, items that were pretested were remembered better than those that were not pretested but rather studied for a longer presentation time. Together these findings suggest that when new information is introduced just after a retrieval attempt, there is a ready updating of knowledge.

Research on the phenomenon of reconsolidation can also provide insights into the functional role of memory malleability, and Hardt, Einarsson, and Nader (2010) have provided an extensive discussion of reconsolidation as a link between memory research traditions in cognitive psychology and neuroscience. The reconsolidation hypothesis proposes that when a memory is reactivated by means of a reminder cue, the trace may return to a labile state after which time it is reconsolidated or restabilized (e.g., Dudai, 2004, 2006). According to the hypothesis, during a time-limited window following reactivation, the malleable trace can potentially be updated, strengthened, or weakened.

Two prominent (but not mutually exclusive) theories have been proposed for the functional role of memory malleability via reconsolidation (see Alberini, 2011, for a review). One hypothesis suggests that the function of reconsolidation is to strengthen and maintain memories and to keep them in an accessible state (e.g., Sara, 2000), much as in the tradition of testing effect research in humans (e.g., Roediger & Karpicke, 2006b). The second idea, and one we adapt in the experiments reported here, proposes that the reconsolidation process resulting from retrieval allows updating of established memories with new information (e.g., Dudai, 2006; Forcato et al., 2007; Lee, 2009; Lewis, 1979; Morris et al., 2006; Nadel, Hupbach, Hardt, & Gómez, 2008; Schiller et al., 2010).

### The Present Experiments

Our experiments address the issue of whether the act of retrieval may permit the retrieved memories to be updated with an *additional* contextual detail that complements rather than conflicts with the original information (in contrast to the new information introduced in A–B, A–D interference paradigm or the misinformation effect paradigm). Lee (2009) wrote “Memories are retrieved often in situations presenting additional complementary information. Thus, the capacity for plastic changes in memory strength or content following memory retrieval seems potentially adaptive in terms of maintaining a memory’s relevance in guiding future behavior.” (p. 413). Indeed, in the reconsolidation literature, it has been suggested that the presence of new information at retrieval is potentially necessary for reconsolidation to occur (Rodríguez-

Ortiz, De la Cruz, Gutiérrez, & Bermudez-Rattoni, 2005; Tronson & Taylor, 2007).

Complementary information presented after retrieval of a previously learned association presents an opportunity for updating, that is, in incorporating new information into an associative memory. If retrieval serves to make new learning easier (as the test-potentiated learning literature suggests), then the act of retrieval may encourage learning of new information if that information is introduced during or just after retrieval. As an analogy to our experimental question and procedure, imagine the following commonplace scenario in which updating would be adaptive: You run into a new colleague in the hall. Last week, when you met your colleague for the first time, you learned that his name was Luca (you successfully associated his face with his name). This week when you run into your new colleague, you recall his name (“Hello, Luca!”), and immediately after your greeting he tells you that his lab will be located in the Bario building (so you now need to update your representation of Luca with a new piece of information, the location of his lab). Does having retrieved Luca’s name when seeing his face, potentially returning that face–name association to a malleable state, facilitate integration of the new detail (lab location) with your previously established association? This basic question (in its general form) motivated the present research.

In 7 experiments, we tested whether new information would be more easily learned in the context of having just retrieved an associative bond (a face–name pair) relative to a control condition in which the face–name pair was restudied rather than retrieved. We developed and tested a new paradigm used to explore updating after retrieval that followed the Luca–Bario building example presented in the previous paragraph. The general paradigm is presented in Figure 1. In each of the experiments, participants learned three pieces of information: a person’s face, his or her name, and his or her profession. In the first phase of the experiment, participants studied the faces with names. In the second phase, participants in the *restudy* condition restudied the face and name. Then, with the face and name still on the screen, a new, to-be-learned piece of information (the person’s profession) was presented for study. In the second phase of the *test* condition, participants were shown each face and asked to retrieve the name. After the retrieval attempt, participants were shown the correct name presented as feedback. Then, with the face and the corrective feedback still on the screen, the profession was presented for study. Thus, the only difference between conditions was the active retrieval process occurring (or attempted) in the test condition.

Both the restudy and test groups were given a final test in which the face was presented as a cue for recall of both the name and the profession with which it had been associated. Our primary question was whether the newly presented profession would be more readily integrated and retained following retrieval of the face–name association compared to restudy of the face–name association. If retrieval facilitates new learning via an updating process in accord with ideas about memory malleability, then final recall of professions should be enhanced in the test condition despite equated study time for the profession in each condition. However, to our surprise, the act of retrieval undermined rather than aided new learning.

One problematic issue for the experimental comparison of retrieval and restudy conditions is that in the retrieval condition

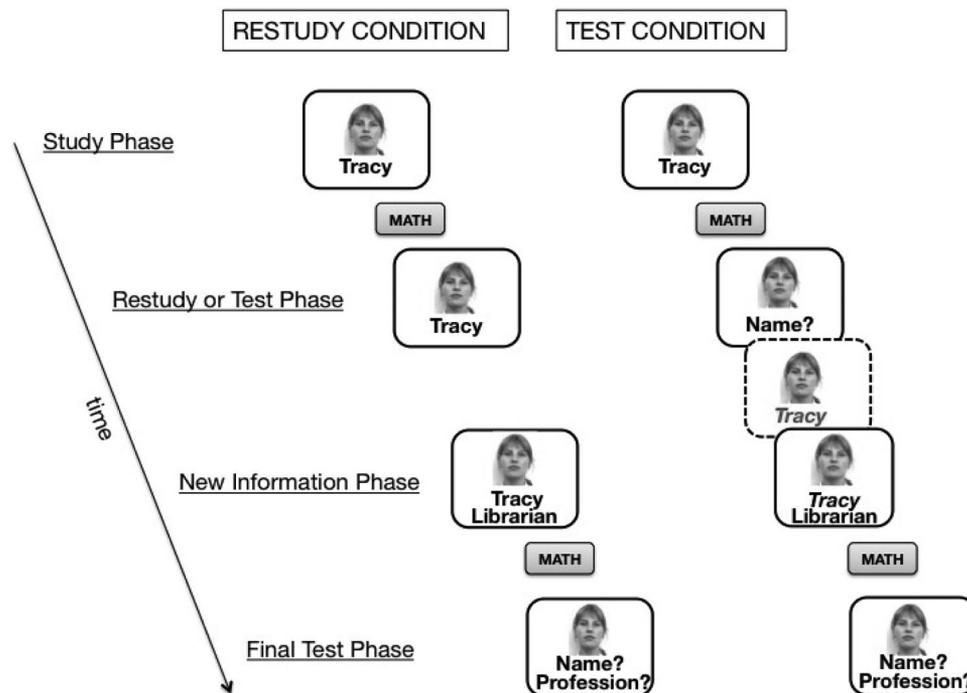


Figure 1. Schematic of the experimental procedure used for the restudy and test conditions in Experiment 1a. The box with the dashed lines represents feedback presentation in the test with feedback condition. Feedback was not presented in the test without feedback condition. Photograph from Psychological Image Collection at Stirling, n.d. (<http://pics.psych.stir.ac.uk/>).

people may selectively attend to the feedback during the time they were supposed to be learning the newly presented profession, especially on those occasions in which they failed to retrieve the name. If they continued to focus attention on the feedback while the profession was presented, a cost to learning of the newly presented profession would be expected rather than facilitation of learning. To examine this possibility, we conducted several experiments using tests without feedback in the retrieval condition. In all other respects, the retrieval without feedback condition was identical to the retrieval with feedback condition. To presage, our basic results did not change as a function of whether feedback was provided.

In sum, 7 experiments tested whether retrieval, compared to restudy, would facilitate the incorporation of new information about professions with the face–name associations established during earlier study. We predicted superior learning of professions in the testing/retrieving condition based on the memory malleability hypothesis. This outcome would demonstrate that retrieval opens a window of time in which new information can be learned better than if retrieval had not occurred. After our first few experiments, our project turned into a quest to understand why we obtained inhibition rather than facilitation of new learning following retrieval.

### Experiment 1a

#### Method

**Participants.** One hundred and eighty-five people participated in Experiment 1a. They were recruited from Amazon.com’s

Mechanical Turk (MTurk) online survey program, which allows researchers to post experiments to be completed by Amazon.com users in return for a small contribution toward a gift voucher. The platform records each participant’s IP address in order to minimize the possibility that any given individual would complete the same questionnaire or procedure more than once. (For a review of the psychological properties of samples obtained through MTurk, and the issues involving with Internet sampling more generally, see Sargis, Skitka, & McKeever, *in press*). Cheating (e.g., writing down material as it was presented) can be a problem in testing with MTurk. Our solution has been not to prohibit this form of behavior in instructions but rather to question participants after they have completed the experiment and to eliminate data from those who admit to the practice (although they were still compensated).

Ten participants were excluded for indicating on a questionnaire given at the end of the experiment that they had written down all or most of the names and professions during the experiment. Three participants were excluded for failing to answer a single item correctly during any part of the experiment, and three participants were excluded for failing to complete the experiment. This left 62 participants in the restudy condition, 59 participants in the test with feedback condition, and 48 participants in the test context without feedback condition, for a total of 169 participants. In this and in the experiments that follow, all participants were treated according to the ethical guidelines of the American Psychological Association.

**Design.** The experiment constituted a 3 (Context condition: restudy vs. test with feedback vs. test without feedback)  $\times$  2 (Item type: name vs. profession) mixed measures design. Context condition was manipulated between subjects and item type was ma-

nipulated within subject. The primary dependent variable of interest was recall on a final cued recall test for new learning (the profession) following the context manipulation. In the restudy context condition, the new information (profession) was presented for initial study just after restudy of the face and name. In the test with feedback context conditions, the profession was presented for initial study immediately after the name had been tested with feedback. In the third condition, the test for the name was given with no feedback, and then the profession was presented.

**Materials.** Study materials were composed of 20 face–name–profession triples. Faces were chosen from the *Psychological Image Collection at Stirling* (n.d.). Ten of the faces were female and were paired with typical female names (e.g., Catherine). The other 10 faces were male and were paired with typical male names (Donald). Stereotypically male (e.g., construction worker) or female (e.g., nurse) jobs were assigned to male or female faces, respectively. All other professions were randomly assigned (e.g., veterinarian). The face–name–profession stimuli sets were fixed for all participants. Order of stimulus presentation was randomized for each participant. Names and professions can be found in *Appendix A*.

**Procedure.** The general procedure is illustrated in *Figure 1*. At the start of the experiment, participants were told that they would be learning people’s faces, names, and professions. They were told that they would first learn the face and name of each person and that later in the experiment they would learn each person’s profession. Participants were told to study for a final test in which they would be shown each person’s face and be asked to remember both the name and the profession.

During the first phase of the experiment, 20 face–name pairs were presented for 5 s each. After a pair was presented, the next pair immediately followed. The profession was not presented at this time. After studying all 20 face–name pairs once, participants solved multiplication problems for 1 min to remove effects of short-term memory (Glanzer & Cunitz, 1966). In the next phase of the experiment (restudy vs. test phase) the between-participants context manipulation was introduced.

In the restudy condition, participants were presented with each face–name pair for restudy. The face and name were presented for 5 s, after which time the profession appeared underneath the name. The face–name–profession set was shown for 5 s, thus giving the profession a 5 s presentation time. In the test context conditions, participants saw a face and were prompted to enter the person’s name to the cue: Name? They typed in their response into a response field and clicked an Enter button when they were done. Participants were never forced to supply an answer (i.e., they could leave the response field blank), and they had as much time as they wanted to enter their response.

In the test with feedback condition, the name was provided immediately upon the participant entering their response (or when he or she hit the return key without recalling the name). Feedback was provided whether or not the participant’s response was correct. The face together with the name feedback was shown for 2 s, after which time the profession appeared underneath the face–name pair for 5 s. In the test without feedback condition, immediately after participants typed in their response (or hit enter), the profession appeared underneath the face, and the response that they had entered (or a blank field if they had not entered anything) stayed on the screen for 5 s. Thus, in all three conditions, the total

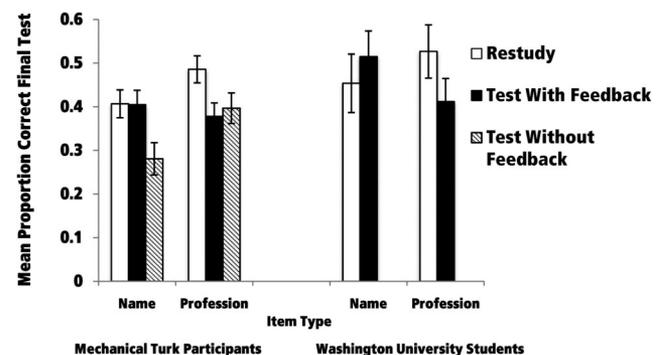
study time for professions was 5 s. The only difference among the conditions was the context immediately preceding presentation of the profession. For all groups, the face–name pairs (restudy context) or the faces (test contexts) were presented in a different random order from the one used during the initial study phase.

After all 20 professions had been presented, participants completed another minute of multiplication problems and then took the final test. During the final test, the face cues were presented in a newly randomized order. A single face was presented, and participants were prompted to provide the name (Name?) and the profession (Profession?). They typed in their response at each prompt and clicked an Enter button when they were done. They were not forced to supply answers, and they had as much time as they wanted to enter their responses. The average time for taking the final test was 4.24 min, with a range of 54 s to 18 min. In this and in the experiments that follow, after all of the sets were tested, participants were debriefed, thanked, and given MTurk credits for their time.

## Results and Discussion

**Recall performance.** On the initial test, performance was not significantly different for the two test context conditions,  $t < 1$ . Participants in the test with feedback condition recalled .27 ( $SE = 0.03$ ) names, whereas as those in the test without feedback condition recalled .28 ( $SE = 0.03$ ). Recall performance on the final test is shown in the left panel of *Figure 2*. Overall performance was best for the restudy condition ( $M = 0.45$ ,  $SE = 0.03$ ) followed by the test with feedback condition ( $M = 0.39$ ,  $SE = 0.03$ ) and the test without feedback condition ( $M = 0.34$ ,  $SE = 0.03$ ), with the main effect of condition only marginally significant,  $F(2, 166) = 2.84$ ,  $MSE = 0.11$ ,  $p = .06$ ,  $\eta_p^2 = .03$ . Across conditions, recall of professions showed a small but significant advantage ( $M = 0.42$ ,  $SE = 0.02$ ) over recall of names ( $M = 0.36$ ,  $SE = 0.02$ ),  $F(1, 166) = 23.68$ ,  $MSE = 0.01$ ,  $p < .001$ ,  $\eta_p^2 = .13$ . These main effects were qualified by a significant interaction between context condition and item type,  $F(2, 166) = 13.87$ ,  $MSE = 0.01$ ,  $p < .001$ ,  $\eta_p^2 = .14$ .

To identify the locus of the interaction, separate one-way analyses of variance (ANOVAs) were conducted for name and profession. We first consider recall of names. When participants did not get feedback following their retrieval attempt of the name on



*Figure 2.* Mean proportion correct on the final recall test for name and profession in Experiments 1a and 1b. Error bars depict standard errors of the mean.

the initial test, their recall of names on the final test was worse than groups that had been reexposed to all the stimuli through either restudy or testing with feedback. Name performance in both the restudy and test with feedback conditions was significantly better than in the test without feedback condition, smallest  $t(105) = 2.64$ ,  $p < .05$ . There was no significant difference between the restudy and test with feedback conditions ( $t < 1$ ); thus, we did not observe a testing effect (name performance: test with feedback  $>$  restudy), but testing effects often occur after longer retention intervals than that in the current experiment where the final test immediately followed the study session. (Experiment 2 provides a demonstration of the testing effect after a 24-hr delay using the same general paradigm).

Recall of professions was of particular interest because we predicted that professions should be learned better in the retrieval conditions than in the restudy condition. In fact, we found the opposite. Although the time for initial study of professions was identical in all three context conditions, final test performance did show differences by condition,  $F(2, 166) = 3.41$ ,  $MSE = 0.06$ ,  $p < .05$ ,  $\eta_p^2 = .04$ . Contrary to our prediction, recall of professions in the restudy condition ( $M = 0.49$ ,  $SE = 0.03$ ) was significantly better than in both the test with feedback condition,  $M = 0.38$ ,  $SE = 0.03$ ,  $t(117) = 2.17$ ,  $p < .05$ , and the test without feedback condition,  $M = 0.40$ ,  $SE = 0.04$ ,  $t(108) = 1.96$ ,  $p = .05$ , which did not differ from each other ( $t < 1$ ).

Thus, rather than finding enhancement of new learning following retrieval of previously studied information, we found impairment relative to the restudy condition. This impairment did not appear to be due to selective attention to feedback because no significant differences emerged between the test with feedback and the test without feedback conditions in recall of professions.

These results are the first that we know of showing that new learning is impaired when introduced after a retrieval attempt. Though the results of Experiment 1a did not show the predicted enhancement, they do suggest that new information presented in the period following retrieval is processed differently from information presented in the period following a restudy opportunity. The data converge on findings demonstrating that a period of postretrieval processing influences learning (e.g., Finn & Roediger, 2011; Finn, Roediger, & Rosenzweig, 2012).

We sought to replicate this novel finding and also ensure that the pattern of results would generalize across populations. Experiment 1b was identical to Experiment 1a, with two exceptions. The first difference was that the study was conducted with Washington University undergraduate students rather than MTurk participants. Like Experiment 1a, the experiment was conducted online. The second difference was that Experiment 1b used only the restudy and the test with feedback conditions because feedback presentation did not affect final recall performance of professions in Experiment 1a.

## Experiment 1b

### Method

**Participants, design, and materials.** Thirty Washington University undergraduate students participated for cash or course credit. Because of a condition assignment error by the experimenter, there were 13 participants in the restudy condition and 17

participants in the test condition. The design and materials were identical to Experiment 1a, except that the test without feedback condition was not included. Participants signed up for the experiment online using the Washington University subject pool sign up portal. Upon signing up, participants were given a link to follow to participate in the online study. All participants completed the experiment outside of the lab.

## Results and Discussion

**Recall performance.** On the initial test, participants in the test with feedback context condition recalled .31 ( $SE = 0.03$ ) names. Mean performance was not much different from that reported for the MTurk participants in Experiment 1a (.27–.28).

On the final test, the same pattern of results occurred with the Washington University sample as was observed in Experiment 1a. Overall, performance did not differ significantly for the restudy ( $M = 0.49$ ,  $SE = 0.06$ ) and test ( $M = 0.46$ ,  $SE = 0.05$ ) context groups,  $F < 1$ ,  $p > .05$ . The small recall advantage seen in Experiment 1a for profession over name was absent in the Washington University sample (restudy context:  $M = 0.48$ ,  $SE = 0.05$  vs. test context:  $M = 0.47$ ,  $SE = 0.04$ , respectively,  $F < 1$ ,  $p > .05$ ). This was most likely due to the absence of the test without feedback condition, which drove down performance for names in Experiment 1a.

As can be seen in the right panel of Figure 2, the same significant interaction emerged with context condition and item type,  $F(1, 28) = 5.60$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = .17$ , with participants in the test condition showing poorer recall of professions, a difference that was marginally significant,  $t(28) = 1.85$ ,  $p = .08$ . (We replicate this pattern repeatedly across experiments, so the borderline nature of the effect here does not indicate doubt about replicability of the effect).

In two experiments we showed that learning of new information was impaired when it was presented following a test of previously studied information compared to when it was introduced following restudy. In addition, the impairment was shown with both Washington University students and the MTurk sample. In Experiment 2, we sought to further delineate this new phenomenon by assessing whether the impairing effect would persist over a delay. In the literature on the testing effect, the impact of testing is often not observed until after a delay (e.g., Roediger & Karpicke, 2006a, for a review). In addition, several studies examining the reconsolidation hypothesis have found that manipulations introduced after reactivation often do not lead to an immediate alteration of memory (Hupbach, Gomez, Hardt, & Nadel, 2007; Nader et al., 2000; Walker, Brakefield, Hobson, & Stickgold, 2003; but see Finn & Roediger, 2011; Finn et al., 2012) yet do have an effect after a longer retention interval. Given these findings, it seemed possible that differences in the effect of retrieval relative to restudy conditions would emerge over a longer retention interval, with a possible advantage for learning new information in the test/retrieval condition appearing on a delayed test. Experiment 2 was identical to the previous experiments except that the final test came after a 24-hr delay. A restudy context condition and a test with feedback context condition were contrasted. Experiment 2 also permitted us to see whether a testing effect in recall of names would emerge after a delay, unlike the results on the immediate tests in Experiments 1a and 1b.

## Experiment 2

### Method

**Participants, design, and materials.** One hundred and twenty participants recruited from the MTurk website participated for Amazon.com credits. Twenty-four participants in the test with feedback context condition and 17 in the restudy context condition were excluded for not completing the second session (i.e., the delayed test portion of the experiment), and two were excluded for participating in the experiment twice, leaving a total of 77 participants. There were 43 participants in the restudy condition and 34 participants in the test condition. The design and materials were identical to Experiments 1a and 1b except that the final recall test was given approximately 24 hr later rather than immediately. Twenty hours after the initial session, participants were e-mailed with a link directing them to the site where they could complete the delayed test portion of the experiment within the next 4 hr.

### Results and Discussion

**Recall performance.** On the initial test, the participants in the test with feedback context condition recalled .28 ( $SE = 0.03$ ) names, which was the same as in earlier experiments. Final recall performance (see Figure 3) across names and professions did not differ significantly for the restudy ( $M = 0.26$ ,  $SE = 0.03$ ) and test with feedback ( $M = 0.28$ ,  $SE = 0.03$ ) groups ( $F < 1$ ). On the delayed test there was a small but significant recall advantage for names over professions,  $M = 0.29$ ,  $SE = 0.03$  vs. .24,  $SE = 0.02$ , respectively,  $F(1, 75) = 5.10$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = .06$ , driven by a testing effect with names in the test with feedback condition. This outcome was confirmed by a significant interaction between context condition (restudy or test) and item type,  $F(1, 75) = 21.48$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = .22$ .

A separate one-way ANOVA for name recall revealed a significant difference between the restudy and test with feedback conditions favoring the test with feedback condition,  $F(1, 75) = 6.08$ ,  $MSE = 0.05$ ,  $p < .05$ ,  $\eta_p^2 = .08$ . The testing effect thus emerged on the delayed test: As expected, the names that had been tested were recalled better than the names that had not been tested. An analysis of recall for professions showed the opposite pattern. Recall in the restudy condition was .28 ( $SE = 0.03$ ), whereas recall

for the test condition was .20 ( $SE = 0.03$ ). Although this difference did not reach conventional levels of statistical significance,  $F(1, 75) = 3.45$ ,  $MSE = 0.04$ ,  $p = .06$ ,  $\eta_p^2 = .04$ , the impairment was of roughly the same magnitude (8%) as in Experiments 1a and 1b (11% and 10%). Certainly, however, there was no hint for facilitation in learning new information presented after retrieval relative to after restudying, despite the fact that a positive testing effect on names occurred in this experiment. Testing enhanced retention of the tested names but still undermined acquisition of new information, professions.

## Experiment 3

In Experiment 3, we explored the possibility that impaired recall of professions was due to response competition between the name and the profession during the final test rather than to differences in the initial encoding of professions. Output interference has been shown to be a source of forgetting in recall tasks, with recall of one set of previously learned material impairing recall of other material (e.g., Roediger, 1974; Roediger & Schmidt, 1980). Accordingly, the aim of Experiment 3 was to mitigate response competition by using both the face and the name as cues to prompt recall of professions. The objective was to limit the potential interference caused by recalling both the name and profession, when the name would often be recalled before retrieval of the profession was attempted.

### Method

**Participants, design, and materials.** Sixty-six participants recruited from the MTurk website participated for Amazon.com credits. Eight participants in the test with feedback condition and 7 in the restudy context condition were excluded for writing down words during the study session, leaving a total of 51 participants. There were 25 participants in the restudy condition and 26 in the test context condition. The design and materials were identical to those in the previous experiments except that during the final test participants were provided with both the name and the face and asked to recall only the profession. Also, a test without feedback condition was not included. The final test occurred during the same experimental session.

### Results and Discussion

**Recall performance.** On the initial test, the participants in the test condition recalled .31 ( $SE = 0.04$ ) names, which was the same as in earlier experiments. As can be seen in Figure 4, final recall of professions was significantly better when they had been learned following restudy of the material ( $M = 0.52$ ,  $SE = 0.06$ ) relative to their learning following the test with feedback,  $M = 0.33$ ,  $SE = 0.05$ ,  $t(49) = 2.42$ ,  $p < .05$ . Thus even when both the faces and names were provided on the test (thus overcoming output interference), the test with feedback again led to impaired recall of professions on the final test relative to the restudy condition. In fact, the difference in this experiment was somewhat greater than in previous experiments when both names and professions were recalled. Thus, output interference in recall of both names and professions was not responsible for the inhibitory effects of learning professions after retrieval of names.

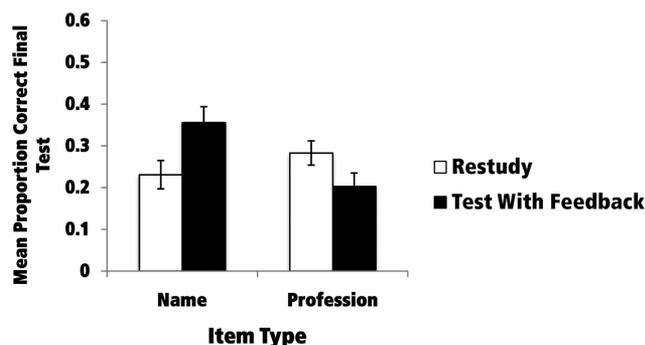


Figure 3. Mean proportion correct on the delayed final recall test for name and profession in Experiment 2. Error bars depict standard errors of the mean.

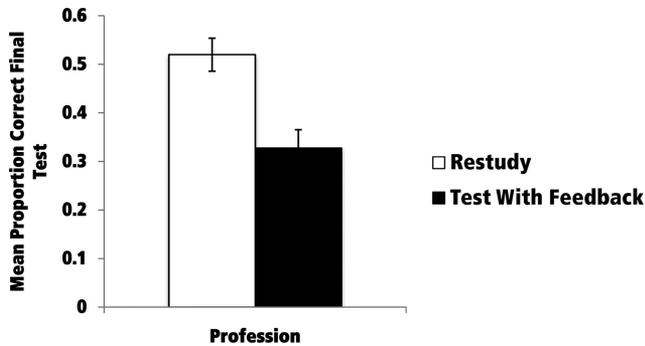


Figure 4. Mean proportion correct on the final recall test for profession in Experiment 3. Error bars depict standard errors of the mean.

The first four experiments revealed a new and puzzling phenomenon. Contrary to our expectations, we found that learning of new information just after a retrieval attempt did not enhance retention of that information but instead consistently impaired it. Because results in two of our experiments were of borderline significance, we conducted an analysis of profession recall across all four experiments comparing final recall of new information (the profession) when it followed testing versus restudying to ensure reliability of the effect across our similar experiments. Altogether, mean recall of professions across immediate and delayed tests for the 143 participants tested in the restudy conditions was .43 ( $SE = 0.02$ ), whereas recall of the 184 subjects who had learned after a test (with or without feedback) was .34 ( $SE = 0.02$ ). This difference was significant,  $t(325) = 3.18$ ,  $p < .01$ , firmly establishing the phenomenon of retrieval-induced impairment in learning of new information.

### Contrasting Explanations of Retrieval-Induced Impairment: Final Recall of Professions Conditionalized on Initial Recall Success of Names

In the following interim discussion section, we contrast two potential explanations of the impairing effect of retrieval on new learning: a blocking account and an updating account. We also present data for final recall performance for professions conditionalized on initial recall success of names in evaluating these hypotheses. Of course, such analyses introduce possibilities of results being driven in part by item selection, but they can be informative nonetheless. We therefore present conditionalized results but emphasize that they must be interpreted with caution.

The “blocking effect” shown in studies of conditioning may provide a potential parallel phenomenon to the impairing effect of retrieval on new learning.<sup>1</sup> The blocking effect refers to the seminal discovery by Kamin (1968, 1969) that a previously established association to one stimulus can block the formation of an association to an added stimulus. One explanation for why learning of an added association can be blocked in this manner is that if a first stimulus already predicts the outcome, a second stimulus adds no new information (e.g., Rescorla & Wagner, 1972). An analogous situation may be present in our experiment: Participants first learned to associate a face with a name and then later were given the task of learning an additional association to the face–name pair, a profession. If learning the first association blocks learning of an

additional association, then impairment would be expected in the tested condition because testing enhanced learning of the face–name association to a greater extent than did restudy, a pattern that was demonstrated on the delayed test used in Experiment 2. Following this logic, a blocking account would predict a greater impairment in learning a new association when retrieval of the name had been successful because successful retrieval of the name would indicate a well-learned association between the face and name.

An updating account of our results presents alternative predictions to the blocking account with respect to new learning following successful and unsuccessful retrieval. Much of the work exploring updating has also been conducted using animal learning paradigms. For example, research exploring how contextual information is updated in rodents shows that the animals can readily modify their prior representations of an environmental context when novel elements are introduced (e.g., O’Keefe & Nadel, 1978). However, a previously encountered environment must be successfully recognized as such for updating to occur. If it is not, the altered environment will be treated as a novel one rather than a variation of a familiar one (O’Keefe & Nadel, 1978; Nadel et al., 2008). In our study, successful retrieval of the name is a likely indicator that the face (and name) are recognized (but see Tulving & Thomson, 1973; Tulving & Wiseman, 1975). Presentation of novel information in this recognized context might similarly be considered an updating event and hence learned easily. In contrast, if the original association was not well established, and therefore not retrieved, the updating process may not be triggered and the new learning would not occur. If recognition facilitates integration of new learning, we might expect to find differences in learning the profession when it occurs after successful retrieval of the name relative to unsuccessful retrieval. In short, the updating and blocking accounts would predict opposite effects when learning of the profession occurs following recall of the name in our paradigm—the updating account predicts facilitation in learning of the profession when the name is recalled and the blocking account predicts interference in learning of the profession when the name is recalled.

To evaluate these alternative accounts, we compiled data across Experiments 1a, 1b, 2, and 3 and conditionalized final recall of professions on whether the preceding name had been successfully recalled at the time of the initial learning of profession. We were interested in two contrasts: first, whether final recall of professions would differ as a function of correct recall of names; and second, whether performance conditionalized on successful initial recall of names would eliminate, diminish or possibly even reverse the pattern of impairment to learning of professions that had been seen in each of the experiments. Of course, names were successfully recalled only 27% to 31% of the time across the first three experiments, so item selection problems are evident.

The first contrast showed that final recall of professions was significantly better when professions were learned following successful recall of names ( $M = 0.43$ ,  $SE = 0.03$ ) compared to unsuccessful recall of names ( $M = 0.30$ ,  $SE = 0.02$ ). This effect was significant as assessed by a 2 (Time of test: immediate vs. delayed)  $\times$  2 (successful vs. unsuccessful recall of name) mixed

<sup>1</sup> We thank David C. Riccio for a discussion of this possibility.

measures ANOVA with the tested groups only,  $F(1, 125) = 20.64$ ,  $MSE = 0.04$ ,  $p < .001$ ,  $\eta_p^2 = .14$ . Because feedback condition did not interact with any measure, we collapsed over this variable. Of course, final recall was better on the immediate test ( $M = 0.41$ ,  $SE = 0.02$ ) than on the delayed test ( $M = 0.25$ ,  $SE = 0.04$ ),  $F(1, 125) = 11.99$ ,  $MSE = 0.09$ ,  $p = .001$ ,  $\eta_p^2 = .09$ . The interaction between time of test and recall success was not significant ( $F < 1$ ). In sum, successful initial recall of the name was related to better final recall of profession (see [Wahlheim & Jacoby, 2013](#), for a similar result), though these results should be interpreted with caution due to possible item selection effects.

The second question was whether successful retrieval of the name provided an advantage in learning the profession above and beyond new learning following restudy or whether the impairment was simply eliminated when successful retrieval of the name occurred. We compared final recall of professions in the restudy condition separately to recall of professions in both the case of unsuccessful and successful conditionalized recall of professions in the test conditions. Time of test did not interact with any variable and so subsequent analyses collapsed over this variable. Results revealed a significant difference in recall of professions between the restudy ( $M = 0.40$ ,  $SE = 0.02$ ) and the unsuccessful test condition ( $M = 0.30$ ,  $SE = 0.02$ ) favoring the restudy condition,  $F(240) = 9.79$ ,  $p < .01$ . In contrast, the comparison of profession recall in the restudy condition and the successful recall of names condition (.43,  $SE = 0.03$ ) showed no difference ( $F < 1$ ). Thus, the impairing effect of retrieval in the test condition was specific to the case when retrieval was unsuccessful. When the name had been retrieved successfully, no impairment occurred in learning the professions. It is important to note, however, that we found no evidence for *enhancement* of new learning (of professions) in the case of successful retrieval of names. This null outcome occurred even with the issue of item-selection likely working toward showing such a facilitating effect; that is, only about 30% of pairs were successfully retrieved in the tested condition, and these items may have been especially “easy” in terms of their being potentially associated with new information. Although we must caution again that these item analyses are merely correlational, they still reaffirm the conclusion that no matter how we examine our results we do not find evidence for facilitation in recall of professions after retrieval.

Where do these findings leave us in terms of our competing explanations? A blocking account predicted that new learning would be poorer following successful retrieval of the name. Because we found better recall of professions following successful retrieval of names, the blocking account seems untenable. Instead, our results are more in line with the prediction of the updating account in that learning the profession was better following successful relative to unsuccessful retrieval of names (keeping in mind the correlational nature of this evidence). Still, the updating account is not fully validated because we did not find evidence of superior learning following retrieval compared to restudy, which was the original prediction we had set out to test in our experiments.

One reason that we may have failed to find superior learning of professions in the test condition was that the new information (profession) and the previously studied information (faces and names) were not particularly compatible. If the updating process involves the incorporation of new contextual detail, perhaps the

new information must be highly compatible with what has been retrieved, as is the case when memories are updated with corrective feedback or through the presentation of misinformation about an already learned detail (a stop sign rather than a yield sign). In both these cases, the new information is highly congruent with the studied information. Accordingly, updating may be more likely when new information presented after retrieval is highly compatible with the learned association that has been retrieved ([Wahlheim & Jacoby, 2013](#)). And while face–name–profession is a set of information that people often need to link, the relationship between name (or face) and profession is an arbitrary one.

In order to examine whether compatibility mattered for learning a new association in our paradigm, we manipulated the compatibility of the newly introduced information in Experiment 4. Participants studied faces, personality traits (i.e., neat, messy), and professions. Half of the professions were compatible with the previously studied traits (caring–nurse) and half were incompatible (messy–architect). The first question was whether professions that were compatible with previously studied traits would be learned more easily than those that were incompatible. We expected they would be. The second, more interesting question was whether the compatibility of the information would have a disproportionate influence on learning a new association in the test condition compared to the restudy condition. Would using materials permitting a more compatible association of new information to previously learned material finally permit us to observe enhancement in the test context relative to the restudy context?

## Experiment 4

### Method

**Participants, design, and materials.** One hundred and seventy online participants were recruited using Mechanical Turk. Ten participants were excluded for indicating on a questionnaire given at the end of the experiment that they had written down all or most of the words during the experiment. Seven participants were excluded for failing to complete the experiment, leaving a total of 153 participants. Fifty-eight participants were in the restudy condition, 42 were in the test with feedback condition, and 53 were in the test without feedback condition. The design was a 3 (Context condition: restudy vs. test with feedback vs. test without feedback)  $\times$  2 (Item type: trait vs. profession) mixed measures design.

The general procedure was identical to that in the previous experiments except that participants studied face–trait–profession sets rather than face–name–profession sets. The final test was immediate. The faces were those in used in the previous experiments. The 20 traits and professions were selected from norms that provided ratings of stereotype congruence and incongruence between traits and professions ([Barrick & Mount, 1991](#); [Weber & Crocker, 1983](#); [White & White, 2006](#)). Using these norms, we created 10 face–trait–profession sets in which the profession was compatible with the trait (caring–nurse), and 20 sets in which the profession was incompatible with the trait (messy–architect). The trait and profession stimuli can be found in [Appendix B](#).

## Results and Discussion

**Recall performance.** On the initial test, overall recall of the traits was not significantly different for the two test context conditions,  $t < 1$ . Participants in the test with feedback context condition recalled .35 ( $SE = 0.03$ ) of the traits, whereas those in the test without feedback context condition recalled .34 ( $SE = 0.03$ ) of the traits. Results for the final test are shown in Figure 5. Final test performance was analyzed using a 3 (Context condition: restudy versus test with feedback versus test without feedback)  $\times$  2 (Item type: trait vs. profession)  $\times$  2 (compatible vs. incompatible profession) mixed measures ANOVA. Item type and compatibility were within-participant variables, and context condition was a between-participants variable. Recall of the professions was significantly better ( $M = 0.56$ ,  $SE = 0.02$ ) than recall of traits ( $M = 0.48$ ,  $SE = 0.02$ ),  $F(1, 151) = 46.28$ ,  $MSE = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .24$ . Recall for compatible items ( $M = 0.56$ ,  $SE = 0.02$ ) was significantly better than recall of incompatible items ( $M = 0.48$ ,  $SE = 0.02$ ),  $F(1, 151) = 34.87$ ,  $MSE = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .19$ . A significant interaction between compatibility and item type,  $F(2, 151) = 38.86$ ,  $MSE = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .34$ , confirmed that our manipulation of compatibility was specific to learning of professions: Compatible professions were remembered more often than incompatible professions, but there was no difference in trait recall by compatibility condition. Not surprisingly, overall performance was slightly better for the restudy ( $M = 0.58$ ,  $SE = 0.03$ ) and test with feedback ( $M = 0.55$ ,  $SE = 0.03$ ) groups than for the test without feedback group ( $M = 0.43$ ,  $SE = 0.03$ ),  $F(2, 151) = 7.39$ ,  $MSE = 0.15$ ,  $p < .01$ ,  $\eta_p^2 = .09$ . These main effects were qualified by a significant three-way interaction between context condition item type and compatibility,  $F(2, 151) = 4.45$ ,  $MSE = 0.01$ ,  $p = .01$ ,  $\eta_p^2 = .06$ .

To identify the locus of the interaction, separate ANOVAs were conducted for trait and profession performance, respectively. We report the analyses of trait performance first. Of course, when participants did not get trait feedback on the initial test, trait performance on the final test was worse ( $M = 0.31$ ,  $SE = 0.03$ ) than for the groups that had been reexposed to all the stimuli either through restudy ( $M = 0.55$ ,  $SE = 0.03$ ) or testing with feedback ( $M = 0.57$ ,  $SE = 0.03$ ),  $F(2, 151) = 20.78$ ,  $MSE = 0.09$ ,  $p < .001$ ,  $\eta_p^2 = .22$ , smallest  $t(99) = 4.74$ ,  $p < .001$ , in pairwise comparisons. The test with feedback and the restudy conditions did not

differ significantly,  $t < 1$ , probably because we used an immediate test (as in Experiments 1a and 1b).

Recall of professions was of particular interest. The interaction between compatibility and context condition was significant,  $F(2, 151) = 3.45$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = .04$ , indicating differences between conditions when the profession was incompatible with the trait (conceptually replicating prior experiments), but not when it was compatible. Separate one-way ANOVAs were conducted for the compatible and incompatible conditions. As can be seen in Figure 5, when the profession was compatible with the previously studied trait, no significant differences occurred following the test and restudy conditions ( $F < 1$ ). An analysis of the incompatible trait-profession condition confirmed the usual impairment of testing on new learning observed in earlier experiments,  $F(2, 151) = 3.10$ ,  $MSE = 0.06$ ,  $p = .05$ ,  $\eta_p^2 = .04$ . Recall of incompatible professions did not differ for the tested conditions (test with feedback:  $M = 0.46$ ,  $SE = 0.03$ , test without feedback:  $M = 0.45$ ,  $SE = .04$ ,  $t < 1$ ), but both test context conditions performed significantly worse than the restudy group in recall of incompatible professions ( $M = 0.55$ ,  $SE = .03$ ), smallest  $t(109) = 1.97$ ,  $p = .05$ .

To reiterate, when new learning of a profession was incompatible with the previously studied trait, recall of profession was impaired in the test context conditions. When the profession was compatible, we found no differences across conditions. These data confirm our previous findings that retrieval differentially affects the ability to learn new information and extend them by showing that retrieval is particularly detrimental to the integration of a new incompatible detail. Still, even though we showed that the compatibility manipulation was effective, we still failed to show that the act of retrieval opens a window to especially good learning of new information (relative to restudying) even for compatible information.

In Experiment 5, we sought to further characterize the retrieval context findings by testing whether we would find the same relationship between compatibility and context condition after a 24-hr delay. Experiment 5 was identical to Experiment 4 except for two differences: the final test came after a 24-hr test delay, and we contrasted only a restudy condition and a test with feedback condition.

## Experiment 5

### Method

**Participants, design, and materials.** One hundred and twenty five participants recruited from the Mechanical Turk website participated for cash/credit. Twenty-three participants in the test with feedback context condition and 16 in the restudy context condition were excluded for not completing the second session (i.e., the delayed test portion of the experiment), leaving a total of 86 participants. There were 44 participants in the restudy context condition and 42 participants in the test with feedback context condition. The design and materials were identical to Experiment 4 except for the following two differences. First, Experiment 5 used only a restudy and a test with feedback condition. Second the final recall test was given approximately 24 hr later rather than immediately. Twenty hours after the initial session, participants were e-mailed with a link directing them to the site where they

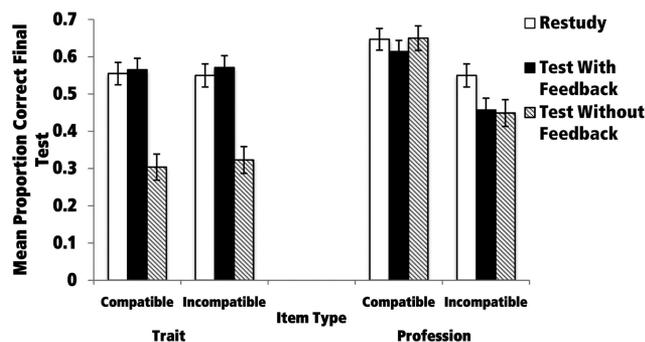


Figure 5. Mean proportion correct on the final recall test for compatible and incompatible traits and professions in Experiment 4. Error bars depict standard errors of the mean.

could complete the delayed test portion of the experiment within the next 4 hr.

## Results and Discussion

**Recall performance.** On the initial test, the participants in the test with feedback context condition recalled .35 ( $SE = 0.03$ ) of the traits. Final test performance is shown in Figure 6. The data were analyzed using a 2 (Context condition: restudy versus test with feedback)  $\times$  2 (Item type: trait versus profession)  $\times$  2 (compatible vs. incompatible profession) mixed measures ANOVA. Item type and compatibility were within-participant variables, and context condition was a between-participants variable. Overall, items in the compatible condition were recalled better than items in the incompatible condition,  $F(1, 84) = 25.93$ ,  $MSE = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .24$ . However this finding was qualified by a significant interaction between compatibility and item type,  $F(1, 84) = 55.01$ ,  $MSE = 0.01$ ,  $p < .001$ ,  $\eta_p^2 = .40$ , which indicated that while compatible professions were remembered better than incompatible professions, there was no difference in trait recall in the compatibility conditions. This finding confirmed that our manipulation of compatibility was specific to the professions.

The only other finding reaching statistical significance was the interaction between item type and condition,  $F(1, 84) = 58.85$ ,  $MSE = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .41$ . Overall recall of traits was better in the test condition than the restudy condition. As in Experiment 2, a testing effect was in evidence with trait learning after a 24-hr delay. Critically, however, new learning of professions introduced after the test was impaired whether or not they were compatible with information from the retrieval context. Recall of both compatible and incompatible professions was worse in the test context condition compared to the restudy context condition,  $F(1, 84) = 4.05$ ,  $MSE = 0.07$ ,  $p < .05$ ,  $\eta_p^2 = .05$ .

On an immediate test in Experiment 4, subjects in the restudy and test context conditions showed equivalent recall performance for compatible professions. After a 24-hr delay however, subjects in the test context condition showed impaired recall for both incompatible and compatible professions. This impairment in delayed recall suggests that the insulation of compatible new learning

from impairment in Experiment 4 on an immediate test is relatively short lived.

## Final Recall of Professions Conditionalized on Initial Recall Success of Traits

Data were compiled across Experiments 4 and 5 to investigate final recall of professions conditionalized on success in recall of traits at the time of initial learning for reasons given in our interim discussion after Experiment 3. We assessed whether final recall of professions would differ as a function of correct recall of the trait and also whether performance conditionalized on successful initial recall of the trait would yield superior performance to that in the restudy condition.

As before, recall of professions was significantly better when the profession had been learned following successful recall of the trait ( $M = 0.56$ ,  $SE = 0.02$ ) than following its unsuccessful recall ( $M = 0.35$ ,  $SE = 0.02$ ),  $F(1, 115) = 77.81$ ,  $MSE = 0.06$ ,  $p < .001$ ,  $\eta_p^2 = .40$ , as measured by a 2 (Time of test: immediate vs. delayed)  $\times$  2 (successful vs. unsuccessful recall of traits)  $\times$  2 (compatible vs. incompatible professions) ANOVA using only the tested groups. Of course, recall was better on the immediate test ( $M = 0.60$ ,  $SE = 0.02$ ) than on the delayed test ( $M = 0.31$ ,  $SE = 0.03$ ),  $F(1, 115) = 68.54$ ,  $MSE = 0.12$ ,  $p < .001$ ,  $\eta_p^2 = .37$ . A three-way interaction between time of test, recall success, and compatibility revealed that the difference between incompatible and compatible items was smallest on the delayed test in the unsuccessful condition (.15 vs. .26 in the successful condition) and smallest on the immediate test in the successful condition (.17 vs. .26 in the unsuccessful condition),  $F(1, 115) = 4.49$ ,  $MSE = 0.05$ ,  $p < .05$ ,  $\eta_p^2 = .38$ .

The second question was whether learning of professions following successful retrieval of the trait differed from learning in the restudy context. We compared final recall of professions in the restudy condition to successful conditionalized recall of professions in the tested conditions for both compatible and incompatible items. On the immediate test, when trait recall had been successful, both compatible and incompatible professions were learned more readily in the tested conditions than the restudy condition (60 vs. .45, respectively, over compatibility conditions),  $F(2, 138) = 6.73$ ,

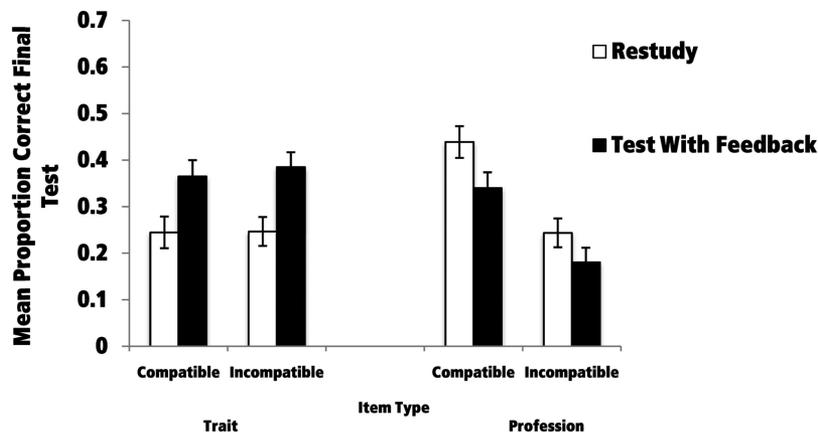


Figure 6. Mean proportion correct on the delayed final recall test for compatible and incompatible traits and professions in Experiment 5. Error bars depict standard errors of the mean.

$MSE = 0.10$ ,  $p < .01$ ,  $\eta_p^2 = .09$ . Of course, because all items are included in the restudy condition and only 35% were included in the tested condition, item selection may account for this difference. However, in this case, an item selection account appears to be unlikely because the difference only appeared on the immediate test. On the delayed test, the advantage of professions that were learned after correct retrieval was eliminated. Final recall performance was equivalent for the restudy ( $M = 0.34$ ,  $SE = 0.03$ ) and successful test groups ( $M = 0.38$ ,  $SE = 0.03$ ),  $F < 1$ . As implied by the foregoing, there was a significant interaction between time of test and recall success,  $F(1, 115) = 10.11$ ,  $MSE = 0.06$ ,  $p < .01$ ,  $\eta_p^2 = .08$ .

Experiment 4 indicated that the compatibility of new learning provided protection against the impairing effect of testing, at least on an immediate test. In Experiment 6, we sought to extend our investigation of the relationship between compatibility of new learning and retrieval to another set of materials. We used the same basic paradigm as in Experiments 4 and 5, but with semantically related and unrelated word associates. The compatible sets of words were made up of three semantically related words (ponder-wonder-think). The incompatible sets of words were made up of two related words and a third word that was semantically unrelated (dollars-money-rabbit). As in Experiments 4 and 5, our question was whether the compatibility of new learning would differentially influence learning after retrieval compared to restudy. Using a different set of materials would provide a conceptual replication of our results in Experiments 4 and 5. However, there was some chance that these materials might reveal a different pattern of results, perhaps even the enhancement from testing that we had sought when we began these experiments.

## Experiment 6

### Method

**Participants, design, and materials.** Eighty-eight participants recruited from the Mechanical Turk website participated for Amazon.com credits. Three participants in the test with feedback context condition and one participant in the restudy context condition were excluded for not completing the experiment leaving a total of 84 participants. There were 43 participants in the restudy condition and 41 in the test condition. The design was a 2 (Context condition: restudy vs. test with feedback)  $\times$  2 (compatible vs. incompatible new associate)  $\times$  2 (Item type: restudied/tested associate [ $R/T_{\text{associate}}$ ] vs. new associate [ $\text{New}_{\text{associate}}$ ]) with only the last factor manipulated between groups.

**Materials.** Study materials were composed of 20 word triples compiled from the Nelson, McEvoy, and Schreiber (1998) word association norms. Ten of the sets were compatible sets in which the first two words converged on a semantic associate, and 10 were incompatible sets in which the third word was unrelated to the first two words. The first word in each set was the cue, the second word was the restudied or tested associate ( $R/T_{\text{associate}}$ ), and the third word was the associate slated for new learning ( $\text{New}_{\text{associate}}$ ). In the compatible sets, all three words were semantically related to one another. In the incompatible sets, the cue word and the  $R/T_{\text{associate}}$  word were semantically related and the  $\text{New}_{\text{associate}}$  word was unrelated. The materials can be found in Appendix C. The FSA (Forward Strength Association, the measure of associa-

tive relatedness between a cue and target used in the Nelson et al., 1998, norms) between related words was .33. There was no difference in associative strength for related words in the compatible and incompatible sets ( $t < 1$ ). The selection of unrelated words was made by randomly selecting a word from the Nelson et al. (1998) norms. First, we ensured that the word was not listed as being related to any of the other words the set. If the word was related, it was rejected and another word was selected at random. To further confirm there was no associative relationship between the unrelated words and related words in a given set, we conducted a latent semantic analysis (LSA; Landauer & Dumais, 1997, and see <http://cwl-projects.cogsci.rpi.edu/msr/>). To examine the association strength between the unrelated word and the related items, the LSA value of the cue and the unrelated word and the LSA value of the  $R/T_{\text{associate}}$  and the unrelated word in each set were determined. The mean LSA between the unrelated words and the other items in the set was .01 and not significantly different from zero,  $t < 1$ , confirming the lack of an associative relationship between the unrelated  $\text{New}_{\text{associate}}$  and the other words in the set. All words ranged between 6 and 8 letters with no differences between sets ( $t < 1$ ).

**Procedure.** The procedure was similar to that in the previous experiments and can be seen in Figure 7. At the start of the experiment, participants were told that they would be learning 20 three words sets. They were told that they would first learn two words from the set and that later in the experiment they would learn the third word that went with the set. Participants were told to study for a final test in which they would be shown the first word from the set and be asked to remember both the second and third word.

During the first phase of the experiment, 20 cue- $R/T_{\text{associate}}$  pairs were presented for 5 s each. After studying all 20 cue- $R/T_{\text{associate}}$  pairs, participants solved multiplication problems for 1 min. In the next phase of the experiment (restudy vs. test phase) the between-participants context manipulation was introduced. During the restudy or test phase, the items were presented in a different random order from the one in which they had been presented during the initial study phase. In the restudy condition, participants were presented with each cue- $R/T_{\text{associate}}$  pair for restudy for 5 s, after which the  $\text{New}_{\text{associate}}$  word appeared underneath the  $R/T_{\text{associate}}$ . The 3-word set was shown for 5 s, thus giving the  $\text{New}_{\text{associate}}$  word a 5 s presentation time. In the test context conditions, participants saw the cue word, and participants were prompted to enter the  $R/T_{\text{associate}}$ . They typed in their response into a response field and clicked an Enter button when they were done. Participants were never forced to supply an answer (i.e., they could leave the response field blank), and they had as much time as they wanted to enter their response. (If they could not remember the associate, they pressed enter to proceed). Immediately upon entering their response, the correct answer was provided as feedback. The cue together with the correct  $R/T_{\text{associate}}$  was shown for 2 s after which time the  $\text{New}_{\text{associate}}$  appeared underneath the cue- $R/T_{\text{associate}}$  set for 5 s. There was no test without feedback condition. In the restudy and the test with feedback conditions, the total study time for the  $\text{New}_{\text{associate}}$  was 5 s. As in the other experiments, the only difference between conditions was the context immediately preceding presentation of the  $\text{New}_{\text{associate}}$  word.

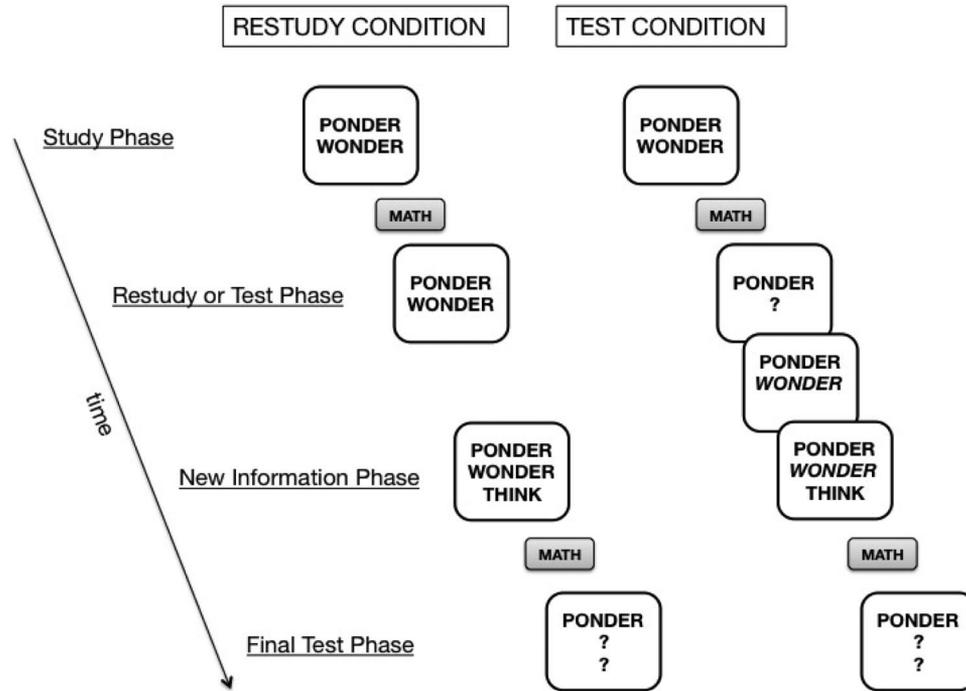


Figure 7. Schematic of the experimental procedure used in Experiment 6. A compatible set of items is presented.

After all 20  $New_{associate}$  items had been presented, participants completed another minute of multiplication problems and then they took the final test. During the final test, the cue words were presented in a newly randomized order. A single cue was presented, and participants were prompted to provide both the  $R/T_{associate}$  and the  $New_{associate}$  word. They typed in their responses at each prompt and clicked an Enter button when they were done. As in the initial test, participants were not forced to supply an answer, and they had as much time as they wanted to enter their response.

## Results and Discussion

**Recall performance.** On the initial test, average recall of the  $R/T_{associate}$  word was .74 ( $SE = 0.02$ ), which was much higher than in our previous experiments with names and faces. Final test performance was analyzed using a mixed measures ANOVA. As can be seen in Figure 8, overall final recall of the  $R/T_{associate}$  items was better than final recall of the  $New_{associate}$  items,  $F(1, 82) = 328.00$ ,  $MSE = 0.03$ ,  $p < .001$ ,  $\eta_p^2 = .80$ . Not surprisingly, compatible (or related) new associates were better remembered than incompatible (unrelated) items,  $F(1, 82) = 110.16$ ,  $MSE = 0.04$ ,  $p < .001$ ,  $\eta_p^2 = .57$ . There was also a significant interaction between item type and condition,  $F(1, 82) = 17.87$ ,  $MSE = 0.03$ ,  $p < .001$ ,  $\eta_p^2 = .18$ , such that overall recall performance for  $R/T_{associate}$  items was marginally better for the tested relative to the restudy group,  $t(82) = 1.74$ ,  $p = .08$ , whereas performance for  $New_{associate}$  items was significantly worse for the tested condition,  $t(82) = 2.16$ ,  $p < .05$ . Though the impairment to new learning was twice as large when the item was incompatible (–14%) than when

it was compatible (–7%), the three-way interaction did not reach significance,  $F(1, 36) = 1.78$ ,  $MSE = 0.09$ ,  $p > .05$ ,  $\eta_p^2 = .05$ . The interaction between item type and compatibility,  $F(1, 82) = 227.17$ ,  $MSE = 0.05$ ,  $p < .001$ ,  $\eta_p^2 = .74$ , confirmed that our manipulation was specific to the  $New_{associate}$  items only. Overall, there was no significant difference in performance for  $R/T_{associate}$  items as a function of compatibility. In contrast  $New_{associate}$  items showed a large difference in recall with performance favoring compatible over incompatible items. This pattern was as expected since the compatibility manipulation was introduced during the restudy/test phase. No other main effects or interactions were significant. Once again, we replicated the finding of impaired new learning following a retrieval attempt when the new word was both semantically related and semantically unrelated to prior words.

**Final recall of professions conditionalized on initial recall success of  $R/T_{associate}$  items.** Final recall of the  $New_{associate}$  items was conditionalized by  $R/T_{associate}$  recall success at the time of initial learning. We assessed whether final recall of the  $New_{associate}$  items would differ as a function of correct recall of  $R/T_{associate}$  items. We also examined whether performance conditionalized on  $R/T_{associate}$  recall success would yield performance greater than or equal to that of performance in the restudy condition.

Recall of compatible  $New_{associate}$  items ( $M = 0.68$ ,  $SE = 0.04$ ) was significantly better than recall of incompatible  $New_{associate}$  items ( $M = 0.23$ ,  $SE = 0.05$ ),  $F(1, 36) = 51.81$ ,  $MSE = 0.07$ ,  $p < .001$ ,  $\eta_p^2 = .59$ , as measured by a 2 (compatible vs. incompatible)  $\times$  2 (successful vs. unsuccessful recall of  $R/T_{associate}$ ) ANOVA with the tested context group only. There was no

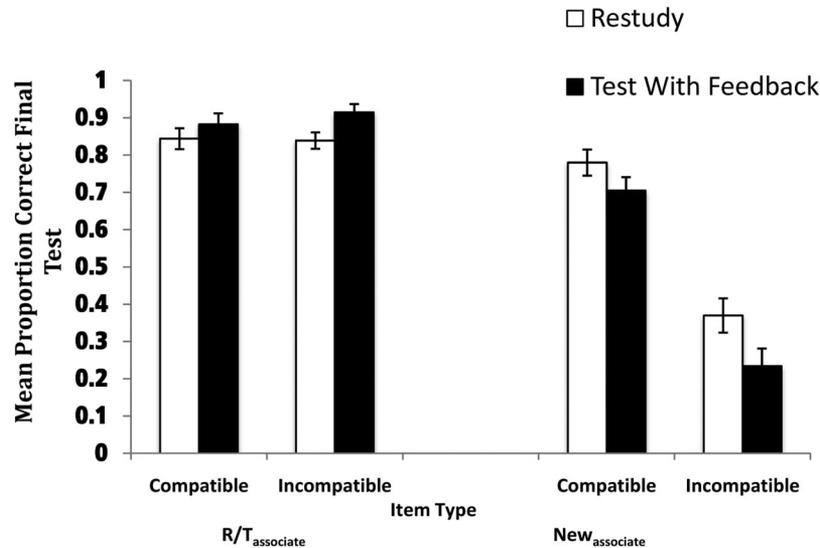


Figure 8. Mean proportion correct on the final recall test for compatible and incompatible associations in Experiment 6. Error bars depict standard errors of the mean. R/T = restudied or tested.

significant main effect of recall success,  $F(1, 36) = 1.09$ ,  $MSE = 0.07$ ,  $p > .05$ ,  $\eta_p^2 = .03$ , nor was there a significant interaction between compatibility and recall success,  $F(1, 36) = 1.78$ ,  $MSE = 0.09$ ,  $p > .05$ ,  $\eta_p^2 = .05$ . Unlike in Experiments 4 and 5, recall success did not significantly differentiate the learning of new information. This difference is likely due to our new set of materials.

We assessed whether learning of New<sub>associate</sub> items following successful retrieval of the R/T<sub>associate</sub> was equivalent to new learning in the restudy context. With compatible items, when recall had been successful, final recall performance was equivalent for the restudy ( $M = 0.78$ ,  $SE = 0.03$ ) and test ( $M = 0.74$ ,  $SE = 0.04$ ) conditions,  $t < 1$ . When recall had been unsuccessful however, compatible items were recalled significantly worse than when learned following restudy ( $M = 0.78$ ,  $SE = 0.03$  vs.  $M = 0.56$ ,  $SE = 0.07$ , for the restudy and test conditions, respectively),  $t(82) = 2.87$ ,  $p < .01$ . For incompatible items, learning the new word following restudy ( $M = 0.37$ ,  $SE = 0.05$ ) tended to be better than following either successful ( $M = 0.24$ ,  $SE = 0.05$ ) or unsuccessful ( $M = 0.23$ ,  $SE = 0.06$ ) tests, though these differences did not reach significance,  $t(80) = 1.85$ ,  $p = .06$ ,  $t(80) = 1.79$ ,  $p = .07$ , respectively.

### General Discussion

Seven experiments examined whether the act of retrieval would facilitate associative updating, the learning of new information. We tested the hypothesis that new information would be learned more readily when introduced immediately after retrieval of previously studied information than when introduced after a restudy opportunity. However, our experiments consistently showed that, rather than enhancing new learning, presenting new information just after a test generally undermined acquisition of new associations. Previous studies exploring updating in human memory have targeted how manipulations introduced during the retrieval process affect retrieval of previously learned information, hurting it (e.g.,

Loftus & Palmer, 1974) or helping it (e.g., Finn & Roediger, 2011). Our paradigm differed from previous studies targeting updating of information in that our procedure provided a new association to be learned during the period after retrieval rather than revision of previously learned information (see Morris et al., 2006, for related research with rodents). Ours are the first studies showing that a retrieval attempt can negatively affect the ability to learn new information. Rather than opening a temporal window for updating learning, the act of retrieval seems to partially close a window. We obtained inhibition in learning a new association after participants attempted retrieval.

Experiments 1–3 showed that people were impaired in their learning of a profession when it was introduced following a test of a previously studied face–name pair, compared to a restudy opportunity of the face–name pair. This impairment occurred on an immediate final test, on a 24 hr delayed final test, when feedback was provided on the initial test, and when feedback was not provided. This last result is particularly important because it rules out the idea that participants were simply focusing on the feedback about the name to the exclusion of learning the profession. We also ruled out the hypothesis that the inhibition in recall of professions was due to output interference on the final test because the impairment in recall of professions occurred even when both the face and the name were provided as cues (Experiment 3).

Subsequent experiments (Experiments 4–6) varied the compatibility of the newly introduced information to explore (a) whether learning was superior when the new information was congruent with prior study and (b) whether the compatibility of new learning with prior learning would mitigate the impairment seen in the retrieval context condition. Not surprisingly, compatible associations were learned more readily in both the restudy and test context conditions. This was true when contrasting compatible versus incompatible trait–profession associations (Experiments 4 & 5) and semantically related and unrelated word associates (Experiment 6). More interestingly, when the new learning was compat-

ible, the impairing effect of testing was reduced; still, no enhancement relative to the restudy condition was obtained in any experiment. On the immediate criterial test in Experiment 4, compatible items were learned equally well for the test and restudy conditions. However, protection against impairment for compatible items in the tested condition was relatively short lived. With the same procedure as used in Experiment 4 except for a 24-hr delayed test, recall even of compatible items was worse following a test context relative to a restudy context. A similar outcome occurred in Experiment 6.

We examined recall of professions in the test context conditions conditionalized on whether or not name recall was successful. As we noted, such analyses introduce possibilities of results being driven in part by item selection and so must be interpreted with caution. We presented them to provide a comprehensive view of our findings. In the test context conditions, retrieval success of names seemed to influence new learning. A blocking account (Kamin, 1968, 1969) predicted that new learning would be weakest when it followed successful retrieval, but we found the reverse outcome; new learning was better following successful retrieval than unsuccessful retrieval (although of course item selection could play a role—faces to which names were correctly recalled may simply be easier to associate to new information such as professions). Nevertheless, successful retrieval did not provide an advantage to new learning above and beyond new learning following restudy, a finding that would seem to run counter to an item selection account because the same level of final recall occurred in the restudy condition (that included all items) and in the successful retrieval condition (that included only a small proportion of items). In contrast to our original predictions and to those of an updating account, we did not find enhancement of new learning in the retrieval context condition in our experiments (one finding from Experiment 4 showed no interference but no facilitation, either).

Rather than finding that the act of retrieval provided a period that would permit greater opportunity for new learning, we found that retrieval (or attempted retrieval) undermined new associative learning. Why? Nadel and collaborators (Nadel et al., 2008) have suggested that familiarity or recognition of “oldness” is a critical determinant of whether updating will occur. In the animal literature on updating, research has shown that if a previously encountered context has changed substantially, the animal may not recognize it as being familiar (for a review, see Nadel et al., 2008). If the context is not recognized, updating will not occur and instead an entirely new representation will be established. Updating is predicted when the context is recognized as a familiar one containing minor variations. Wahlheim and Jacoby (2013) provided direct evidence for this prediction in an A–B, A–D interference paradigm.

In the conditionalized analyses, we found that impairment was specific to new learning following unsuccessful retrieval. In applying Nadel’s et al.’s (2008) analysis from the animal literature, we may assume that in the case of both successful retrieval and restudy, the original pair was either seen for a second time or retrieved, depending on the condition. Thus, restudy and successful cued recall may be functionally more similar to the initial cue–target study episode than unsuccessful retrieval. That is, in contrast to unsuccessful retrieval, restudy and successful retrieval may provide more familiar instantiations of the study context since in both cases all information is presented for a second time (i.e.,

both the cue and the target). Indeed, a second presentation via restudy or testing may involve spontaneous study-phase retrieval, namely, retrieval of contextual aspects of the original study episode (e.g., Hintzman & Block, 1973; Hintzman, 2011). Accordingly, new information presented in this familiar context (i.e., successful retrieval and restudy) may be akin to a small variation to a familiar context, i.e., a situation in which updating is likely to occur. In the case of unsuccessful retrieval, failure to recognize the cue, or recognition of the cue coupled with an omission or commission error, is quite contextually distinct from the initial study episode. Updating might not therefore be expected in this case, and in fact, some evidence suggests that interference is the expected outcome (Wahlheim & Jacoby, 2013). The foregoing paragraph represents speculation, of course, though speculation informed by the reconsolidation literature and recent behavioral evidence. The main purpose of our article is to describe a new phenomenon that seems surprising in the context of past research on reconsolidation.

While a clear theoretical account of our findings is not yet in hand, these initial results do support the idea that retrieval creates conditions that can alter learning and add to the growing body of work demonstrating the dynamic nature of retrieval processes in human memory. William James wrote of the dynamism of learning, “In this gradual process of interaction between the new and the old . . . the new is modified and determined by the particular sort of old which apperceives it.” (James, p. 114, 1899/2001). Our findings demonstrate that the context in which new information is introduced has a substantial impact on how well it will be learned.

## References

- Alberini, C. M. (2011). The role of reconsolidation in the dynamic process of long-term memory formation and storage. *Frontiers in Behavioral Neuroscience*, 5, 1–10. Retrieved from [http://www.frontiersin.org/behavioral\\_neuroscience/10.3389/fnbeh.2011.00012/abstract](http://www.frontiersin.org/behavioral_neuroscience/10.3389/fnbeh.2011.00012/abstract)
- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/a0029199
- Barnes, J. M., & Underwood, B. J. (1959). “Fate” of first-list associations in transfer theory. *Journal of Experimental Psychology*, 58, 97–105. doi:10.1037/h0047507
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26. doi:10.1111/j.1744-6570.1991.tb00688.x
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.
- Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Erlbaum.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology*, 55, 51–86. doi:10.1146/annurev.psych.55.090902.142050
- Dudai, Y. (2006). Reconsolidation: The advantage of being refocused. *Current Opinion in Neurobiology*, 16, 174–178. doi:10.1016/j.conb.2006.03.010
- Finn, B., & Roediger, H. L. (2011). Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later

- recall. *Psychological Science*, 22, 781–786. doi:10.1177/0956797611407932
- Finn, B., Roediger, H. L., & Rosenzweig, E. (2012). Reconsolidation from negative emotional pictures: Is successful retrieval required? *Memory & Cognition*, 40, 1031–1045. doi:10.3758/s13421-012-0203-7
- Forcato, C., Burgos, V. L., Argibay, P. F., Molina, V. A., Pedreira, M. E., & Maldonado, H. (2007). Reconsolidation of declarative memory in humans. *Learning & Memory*, 14, 295–303. doi:10.1101/lm.486107
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 6(40).
- Gisquet-Verrier, P., & Riccio, D. C. (2012). Memory reactivation effects independent of reconsolidation. *Learning & Memory*, 19, 401–409. doi:10.1101/lm.026054.112
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5, 351–360. doi:10.1016/S0022-5371(66)80044-0
- 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
- Hardt, O., Einarsson, E. O., & Nader, K. (2010). A bridge over troubled water: Reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annual Review of Psychology*, 61, 141–167. doi:10.1146/annurev.psych.093008.100455
- Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the “coordinates of truth”. *Perspectives on Psychological Science*, 6, 253–271. doi:10.1177/1745691611406924
- Hintzman, D. L., & Block, R. A. (1973). Memory for the spacing of repetitions. *Journal of Experimental Psychology*, 99, 70–74. doi:10.1037/h0034761
- Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, 40, 514–527. doi:10.3758/s13421-011-0167-z
- Hupbach, A., Gomez, L., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, 14, 47–53. doi:10.1101/lm.365707
- Izawa, C. (1966). Reinforcement test sequences in paired associate learning. *Psychological Reports*, 18, 879–919. doi:10.2466/pr0.1966.18.3.879
- James, W. (2001). *Talks to teachers on psychology and to students on some of life's ideals*. New York, NY: Dover. (Original work published 1899)
- Kamin, L. J. (1968). “Attention-like” processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior: Aversive Stimulation* (pp. 9–31). Miami, FL: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.
- 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
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240. doi:10.1037/0033-295X.104.2.211
- Lee, J. L. C. (2009). Reconsolidation: Maintaining memory relevance. *Trends in Neurosciences*, 32, 413–420. doi:10.1016/j.tins.2009.05.002
- Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology*, 8, 311–335. doi:10.1016/0010-0285(76)90010-4
- Lewis, D. J. (1979). Psychobiology of active and inactive memory. *Psychological Bulletin*, 86, 1054–1083. doi:10.1037/0033-2909.86.5.1054
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12, 361–366. doi:10.1101/lm.94705
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the inter-action between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585–589. doi:10.1016/S0022-5371(74)80011-3
- McGaugh, J. L. (1966). Time-dependent processes in memory storage. *Science*, 153, 1351–1358. doi:10.1126/science.153.3742.1351
- McGaugh, J. L. (2000). Memory: A century of consolidation. *Science*, 287, 248–251. doi:10.1126/science.287.5451.248
- Misanin, J. R., Miller, R. R., & Lewis, D. J. (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science*, 160, 554–555. doi:10.1126/science.160.3827.554
- Morris, R. G., Inglis, J., Ainge, J. A., Olverman, H. J., Tulloch, J., Dudai, Y., & Kelly, P. A. (2006). Memory reconsolidation: Sensitivity of spatial memory to inhibition of protein synthesis in dorsal hippocampus during encoding and retrieval. *Neuron*, 50, 479–489. doi:10.1016/j.neuron.2006.04.012
- Müller, G. E., & Pilzecker, A. (1900). Experimentelle beiträge zur lehre vom gedächtnis [Experimental contributions on the theory of memory]. *Zeitschrift für Psychologie Ergänzungsband*, 1, 1–300.
- Nadel, L., Hupbach, A., Hardt, O., & Gómez, R. (2008). Episodic memory: Reconsolidation. In E. Dere, J. P. Huston, A. Easton, & L. Nadel (Eds.), *Handbook of episodic memory research* (pp. 43–56). Oxford, England: Elsevier Press. doi:10.1016/S1569-7339(08)00204-X
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406, 722–726. doi:10.1038/35021052
- 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://www.usf.edu/FreeAssociation/>
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, England: The Clarendon Press.
- Psychological Image Collection at Stirling. (n.d.). [Photograph]. Retrieved from <http://pics.psych.stir.ac.uk/>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Rodriguez-Ortiz, C. J., De la Cruz, V., Gutiérrez, R., & Bermudez-Rattoni, F. (2005). Protein synthesis underlies post-retrieval memory consolidation to a restricted degree only when updated information is obtained. *Learning & Memory*, 12, 533–537. doi:10.1101/lm.94505
- Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition*, 2, 261–269. doi:10.3758/BF03208993
- Roediger, H. L., & 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., & 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., & 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
- Roediger, H. L., & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired associate lists. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 91–105. doi:10.1037/0278-7393.6.1.91
- Sara, S. J. (2000). Strengthening the shaky trace through retrieval. *Nature Reviews Neuroscience*, 1, 212–213. doi:10.1038/35044575
- Sargis, E. G., Skitka, L. J., & McKeever, W. (in press). The Internet as psychological laboratory revisited: Best practices, challenges, and solutions. In Y. Amichai-Hamburger (Ed.), *The social net: The social psychology of the Internet*. Oxford, England: Oxford University Press.

- Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, *463*, 49–53. doi:10.1038/nature08637
- Schiller, D., & Phelps, E. A. (2011). Does reconsolidation occur in humans? *Frontiers in Behavioral Neuroscience*, *5*, 1–12. Retrieved from [http://www.frontiersin.org/Behavioral\\_Neuroscience/10.3389/fnbeh.2011.00024/abstract#](http://www.frontiersin.org/Behavioral_Neuroscience/10.3389/fnbeh.2011.00024/abstract#)
- Tronson, N. C., & Taylor, J. R. (2007). Molecular mechanisms of memory reconsolidation. *Nature Reviews Neuroscience*, *8*, 262–275. doi:10.1038/nrn2090
- Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373. doi:10.1037/h0020071
- Tulving, E., & Wiseman, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, *6*, 79–82.
- Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive reminders in proactive effects of memory. *Memory & Cognition*, *41*, 1–15. doi:10.3758/s13421-012-0246-9
- Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*, 616–620. doi:10.1038/nature01930
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977. doi:10.1037/0022-3514.45.5.961
- White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles*, *55*, 259–266. doi:10.1007/s11199-006-9078-z

## Appendix A

### Names and Professions Used in Experiments 1–3

Name	Profession
Jenny	florist
Colleen	secretary
Allison	veterinarian
Paul	electrician
Amanda	receptionist
Catherine	doctor
Nathan	professor
Kevin	dentist
Adrian	architect
Jacob	cashier
Donald	geologist
Annie	teacher
Brandon	programmer
Deborah	chemist
Nicole	librarian
Karen	paralegal
Daniel	accountant
John	student
Chris	coach
Natalie	nurse

(Appendices continue)

## Appendix B

### Traits and Professions Used in Experiments 4 and 5

Trait	Compatible sets	Profession
caring		nurse
confident		lawyer
curious		reporter
efficient		accountant
friendly		teacher
neat		receptionist
responsible		doctor
quiet		librarian
sarcastic		construction worker
talkative		hairdresser
Incompatible sets		
bashful		actress
conscientious		banker
diligent		musician
impractical		economist
messy		architect
outdoorsy		computer programmer
religious		biologist
risky		nun
stubborn		counselor
well dressed		farmer

## Appendix C

### Word Stimuli Used in Experiment 6

Compatible sets (3 related words)	Profession
ponder	think
launch	rocket
faculty	staff
lettuce	tomato
slumber	sleep
wicker	basket
shatter	glass
shorts	shirt
reptile	lizard
guardian	angel
Incompatible sets (2 related words and 1 unrelated word)	
railroad	picture
solution	dinner
educate	layer
sibling	forest
bizarre	income
stream	market
citrus	career
creature	engine
minister	needle
dollars	rabbit

Received November 26, 2012  
Revision received January 29, 2013  
Accepted February 5, 2013 ■