

A ROLE FOR RECONSOLIDATION IN THE TESTING EFFECT

BY

SARAH ELIZABETH STANLEY

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Adviser:

Professor Aaron Benjamin

ABSTRACT

The act of recalling information from memory results in a higher probability of successfully remembering that information later than does an additional study opportunity. This *testing effect* is both powerful and well documented, yet there are only a few suggestions for why testing benefits memory. The three existing theoretical positions that have often been brought to bear on the testing effect are the *desirable difficulties hypothesis*, the *semantic mediator hypothesis*, and *transfer-appropriate processing hypothesis*. Recently, a new explanation for why the testing effect may occur has been put forth based on the idea of *reconsolidation*. Following results in the neuroscience of animal learning and human learning, the *reconsolidation hypothesis* assumes the act of retrieval renders the original memory for the retrieved event more malleable than a re-study opportunity. We expand on this hypothesis, proposing that this malleability results in a “generalized” contextual trace that includes elements of the original study episode and of the circumstances of testing. If the final criterion test is performed in a different context, then this more generalized context engenders a higher probability of successful recall. In the current experiments, we evaluate two predictions derived from this hypothesis. The first is that individuals may lose details of the original study context following retrieval practice (Experiment 1a, 1b, and 2). Second, there should be more interference between the original study context and a retrieval practice context than during the original study context and a restudy context (Experiment 3). The results did not support these predictions and thus do not support the reconsolidation hypothesis.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: THE INFLUENCE OF RETRIEVAL PRACTICE ON CONTEXT MEMORY ...	9
CHAPTER 3: THE INFLUENCE OF RETRIEVAL PRACTICE ON CONTEXT MEMORY GIVEN SUCCESSFUL RECALL.....	15
CHAPTER 4: THE INFLUENCE OF RETRIEVAL PRACTICE THAT INTRODUCES INTERFERING CONTEXTS.....	18
CHAPTER 5: GENERAL DISCUSSION.....	23
REFERENCES.....	25

CHAPTER 1

INTRODUCTION

Testing is frequently thought of exclusively as a means of assessing knowledge. In educational settings, instructors administer tests as a means of gauging student learning. However, testing can also be a useful device to enhance learning. Testing as a means of enhancing learning has been discussed for hundreds of years; as early as 1620, Francis Bacon wrote:

“If you read anything over twenty times you will not learn it by heart so easily as if you were to read it only ten, trying to repeat it between whiles, and when memory failed looking at the book.” (F. Bacon, 1620/2000)

This phenomenon is known as the testing effect and has been studied by psychologists sporadically over the years.

In one recent study (Roediger and Karpicke, 2006a), college students read two passages and then attempted to recall one of the passages (retrieval practice) and reread one of the passages (restudy). The condition in which participants are tested as a means of practice will be referred to as the *retrieval practice* rather than *testing* to avoid confusion with a later criterion test. After the retrieval practice and restudy events, participants were given a criterion test on the passage after a retention interval of five minutes, two days or one week. When the retention interval was only five minutes, participants in the restudy condition outperformed those in the retrieval practice condition. However, for both the two-day and one-week retention intervals, the opposite occurred, revealing a testing effect. This delay in the appearance of the testing effect is

poorly understood and is one motivation for considering new theories of why testing helps memory. Here I consider the implications of a theoretical position that has a prominent role for reconsolidation—the malleability of memories that results from retrieval.

How does testing affect memory?

Despite the fact that testing effects are among some of the most important effects being imported from cognitive psychology into education, there are only a few suggestions for how testing benefits memory. The three theories relevant to the testing effect reviewed here are the *desirable difficulties hypothesis*, *semantic mediator hypothesis*, and the *transfer-appropriate processing hypothesis*.

Desirable difficulties hypothesis

The theory of desirable difficulties argues that certain effortful events can enhance long-term memory despite appearing to impede performance during initial practice (Bjork, 1994). One such *desirable difficulty* is spaced practice. When attempting to learn a new skill or memorize information you may see more improvement during a massed practice session than over several spaced sessions. However, this may not be representative of long-term learning—on later tests, those who used spaced practice perform better (Bjork & Allen, 1970; for review, see Benjamin & Tullis, 2010; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Rowland, 2014). Similarly, when trying to learn multiple things, interleaving different tasks is helpful to long-term retention despite hindering performance during training (Rohrer & Taylor, 2007). Testing effects may reveal a similar pattern (Halamish & Bjork, 2011; for review, see Roediger & Karpicke, 2006b). The desirable difficulties hypothesis is an excellent way of describing

numerous phenomena, including the testing effect, but it does not make strong predictions, postulate any particular underlying cause, or provide an explanation for why many manipulations that render learning difficult do not result in superior retention (Yue, Castel, & Bjork, 2012).

Semantic mediator hypothesis

The semantic mediator hypothesis proposes that retrieval leads to activation of semantically related words which later serve as additional retrieval cues (Carpenter, 2011). For example, if a person had studied a word pair “*bread:jelly*,” during retrieval practice in which they are given only the cue “*bread*,” they are more likely to activate semantically related information such as “*toast*” or “*crust*” while trying to retrieve the target than they would be from a simple restudy of the word pair. This activation would be beneficial because it provides extra information that can serve as cues to recall the target at a later time. One datum in support of this hypothesis is that for difficult retrieval practice (in which the likelihood of activating additional related information is greater), the testing effect is larger. This effect is found by comparing weakly and strongly related word pairs (Carpenter, 2009), delayed and immediate retrieval practice (Karpicke & Roediger, 2007; Pyc & Rawson, 2009; Whitten & Bjork, 1977), and retrieval practice using recall versus recognition. One concern with this theory is that it fails to provide any explanation for the fact that the testing effect only appears at long retention intervals. Additionally, it cannot explain why retrieval practice is superior to elaborative study (Karpicke & Smith, 2012).

Transfer-appropriate processing hypothesis

The transfer-appropriate processing hypothesis suggests that retrieval practice is superior to restudy due to a greater similarity in the processes utilized during retrieval practice and the final test (Morris, Bransford, & Franks, 1977). Some support for this hypothesis was found by McDaniel and Fisher (1991). They had participants study trivia facts followed by either retrieval practice or restudy opportunity of the fact. The final criterion test questions were phrased in either the same or different format. For example, participants may have been given the fact, “In Moslem countries white is the mourning color.” During retrieval practice, participants were asked, “What is the mourning color in Moslem countries?” and on the final test, they were asked “What does the color white symbolize in Moslem countries?” Items that had been successfully recalled during retrieval practice were more likely to be recalled on the final criterion test if the questions were phrased in a similar fashion than if they were phrased differently; however, question format was not relevant for items answered incorrectly. These findings indicate that retrieval practice is more beneficial when the similarity between retrieval practice and the final test is greater. At a theoretical level, it suggests that specific retrieval cues given during retrieval practice may contribute to the testing effect. While this hypothesis seems likely to be a factor in the testing effect this theory also fails to address the delay in the testing effect.

A reconsolidation view of the testing effect

Recently, Finn and Roediger (2011) proposed a new explanation for the testing effect drawing on the idea of *reconsolidation*, a concept commonly discussed in neuroscience but rarely addressed in the cognitive literature. Immediately after a new memory is formed it is in a malleable form and is susceptible to numerous manipulations (Burnham, 1903; Duncan, 1949;

Hebb, 1949; Dudai, 1996). The process through which a new memory is converted into a stable memory is called *consolidation*. It is thought that when a stable memory is retrieved from long-term memory, it re-enters this malleable state and must undergo reconsolidation before becoming stable once again (Nader, Scafe & LeDoux, 2000; Finn & Roediger, 2011).

Nader et al. demonstrated reconsolidation with rats by first training them to fear a tone by exposing them to the tone and giving them a small shock. One day later, after the memory should have been fully consolidated, the rats were exposed to the tone (in the absence of the shock) causing them to recall the fearful shocking experience. Immediately after this “retrieval practice,” the rats were given the protein synthesis inhibitor anisomycin, which is known to interfere with consolidation. The rats who had been given this drug showed less of a fear response to a later exposure than those who had not (Nader, Scafe & LeDoux, 2000), suggesting that the fearful memory had reentered a malleable state in which anisomycin had an effect. That is, anisomycin interferes with not only consolidation but also reconsolidation.

In human learning, Finn and Roediger (2011) provided some support to the idea that reconsolidation plays a role in the testing effect through a pair of experiments. In one experiment, participants studied Swahili-English word pairs and were then given a cued recall test. Immediately following successful recall participants were shown a blank screen, a neutral picture or a negatively arousing picture. Prior work had shown that negative arousal during consolidation can enhance memory (Anderson, Wais, & Gabrieli, 2006); thus, it would be reasonable to expect a similar effect from negative arousal during reconsolidation. Word pairs that were followed by a negatively arousing picture during retrieval practice were more likely to be recalled on the final criterion test, which can be taken to imply that the retrieval practice created a period of reconsolidation. A similar experiment using restudy instead of retrieval

practice found no benefit to negatively arousing pictures. This disparity lends support to the reconsolidation explanation of the testing effect, since reconsolidation appears to happen after retrieval practice but not after restudy.

Reconsolidation and context-dependent memory

Finn and Roediger (2011) proposed that reconsolidation may play a role in the testing effect but did not offer a specific mechanism through which reconsolidation may actually benefit memory. We propose a more detailed explanation using mechanisms of context dependence in memory—reconsolidation enhances long-term memory by influencing the set of contextual cues available during later recall.

Every act of encoding is tied to a complex context that includes bits of information about a learner's physical (Smith & Vela 2001) and psychological (Bower, 1981) state, as well as information about other recently encoded information (Kahana & Howard, 2005). Because the likelihood of successful retrieval is strongly influenced by the overlap of encoding and retrieval contexts (Tulving, 1983), conditions that promote encoding of a diverse set of contextual aspects are more likely to lead to retrieval success than conditions that tie encoding to a more selective set of characteristics. For example, you may have experienced the minor panic associated with a forgotten locker combination. You find yourself unable to remember the series of numbers until you are standing in front of the locker with the knob in your hand, and suddenly, the numbers come back to you. Any number of contextual clues may contribute to this effect: the feel of the lock in your hand, the look of the numbers, the color of the room, the amount of light, the smell, the time of day, etc. Any one of these things may not be necessary to remember, but the more cues that are present, the easier recall is.

The malleability created by retrieval practice may result in a single contextual trace that includes elements of the original study episode and of the circumstances of retrieval practice. Imagine, for example, that you are given a new locker and learn the combination at 8 o'clock Wednesday morning. You know it then, but will you remember it the next time you need it? This will depend on a vast number of factors, but all else being equal you are more likely to remember at 8 AM since that is when you first learned it (the time itself may be a contextual clue or any number of other factors correlated with time). Now let's say the next day you return and decide to check that you still remember the combination on Thursday at 10 AM (thus engaging in retrieval practice) and fortunately you do still remember the code. Now what time of day are you most likely to remember the combination? You might expect that either 8 or 10 would be equally good, as you now have two memories of using the combination at two times which you can draw on. But perhaps it is more complicated. We have established that after the retrieval practice the initial memory becomes malleable and subject to change. What happens to the contextual elements tied to these memories? Do these also change? If so, how?

One possibility would be to replace the old trace with the new trace. The locker combination was associated with 8 but after retrieval practice it is now associated with 10. The problem with this idea is that it fails to explain the testing effect; it provides no basis for generalization or enhancement of memory.

A second possibility is that you connect and generalize—in other words, you now associate your locker combination with “morning” instead of either 8 or 10 AM. This could be a very powerful way to enhance a memory because it would associate the memory with a much broader context. This explanation is similar to that invoked by advocates of encoding variability theory in the context of spacing effects: a greater diversity of contexts enhances memory because

of the greater likelihood that the testing effect will overlap with some portion of those prior contexts (Bower ref; Estes, 1955; cf. Benjamin & Tullis, 2010). So you might be just as good at remembering your combination at 9 AM as you would at 8 or 10 AM.

Another possibility is that you may connect the two times without generalizing, i.e. then you may associate your locker combination with both 8 and 10 AM. But how would this differ from restudy? Restudy can be thought of as creating two distinct memories, one associated with 8 and one associated with 10 AM. Retrieval practice, on the other hand, would be one memory associated with both 8 and 10 AM. Of course, in the real world there are *many* contextual details contributing to this effect. So, let's expand our example and consider also the day of the week. In the restudy condition, you have two memory traces, one on Wednesday at 8 AM, and one on Thursday at 10 AM. Alternatively, after retrieval practice you have one memory associated with Wednesday, Thursday, 8 and 10 AM. So if later you try to remember your combination on a Thursday at 8 AM, then each of the restudy memories has only one matching element while the retrieval practice memory is associated with both elements.

If indeed you connect the two contexts without generalizing, specific details of the original may remain intact. However, it is reasonable to think that the two context may be easier to confuse with one another, which is investigated in Experiment 3.

CHAPTER 2

THE INFLUENCE OF RETRIEVAL PRACTICE ON CONTEXT MEMORY

Experiments 1a and 1b

If reconsolidation causes context memory to become more general, then remembering contextual details should be more difficult after retrieval practice. Experiments 1a and 1b sought to test this hypothesis by having participants study word pairs in different contexts (in this case, the context was locations on the screen). For the study session, words were presented in one of eight possible locations, and the review (either restudy or retrieval practice) occurred in the center of the screen. If participants create more general contextual traces (i.e. remembering “morning” instead of 8 or 10), then it may be difficult to retrieve the original screen location. For example, during study participants may encode that the item was in the upper right of the screen, but when it is later retrieved in the middle of the screen, they may update this contextual trace with something more general, like it was on the right half of the screen.

The stimuli in these experiments were low-association word pairs, presented in one of eight possible locations on a computer screen. The only difference between Experiment 1a and Experiment 1b is the number of word pairs used, with 48 and 96 respectively. Participants were asked to study the word pairs and were informed that there would be a later test in which they would be given the first word of the pair (the cue) and asked to provide the second word in the pair (the target). The critical contrast is whether memory for location on the screen is hurt by retrieval practice of the word pairs.

Participants. Twenty-nine undergraduate students from the University of Illinois at Urbana-Champaign participated in Experiments 1a and 1b, each, in partial fulfillment of a course requirement. Three participants in Experiment 1a had incomplete data and were excluded from

analysis (two due to a failure to attend day two of the experiment and one due to computer difficulties). Four participants in Experiment 1b had incomplete data and were excluded from analysis (all due to a failure to attend day two of the experiment).

Design. The experiment was a 3 x 2 within-subject design. The first variable was the review type in which word pairs could be restudied, undergo retrieval practice or receive no additional review time (control). The second variable was the final test, in which word pairs were either tested with cued recall or tested for context memory. For each participant, all conditions had an equal number of word pairs.

Materials. Ninety-six word pairs (cue to target association of 0.028 – 0.03) were selected from the University of South Florida Free Association Norms database (Nelson, McEvoy & Schreiber, 1998). For Experiment 2 only 48 of the original 96 word pairs were used. The word pairs were presented on a computer screen in one of eight possible boxes, each equidistant from the center (see Figure 1). For each subject, a quasi-random order was generated such that each box was used an equal number of times for each condition and no box was used twice a row. For the review phase of the experiment, the boxes were no longer on the display, and the word pair (for restudy) or prompt and blinking cursor (for retrieval practice) was shown in the center of the screen. In the final phase, cued recall tests were presented in the middle of the screen; context memory tests displayed the word pair in the center of the screen surrounded by eight clickable boxes.

Procedure. Participants were told that they were going to be shown a series of word pairs and that they were to study them for a later test in which they would be prompted with the first word of a pair and asked to provide the second word from the pair. Participants were not told that they would later be tested on the location of the word pairs. The study phase presented

the word pairs, for 5 s each with a 1-s inter-stimulus interval. After all the word pairs were presented, participants were then given a 1-minute distractor task in which they performed simple 1-2 digit addition.

For the review phase, participants restudied one-third of the word pairs and performed retrieval practice on one-third of the word pairs. For each participant, the word pairs were randomly assigned to conditions. The restudied word pairs were presented again for 5 s, this time in the center of the screen. For the word pairs that were assigned to the retrieval practice condition, a cued recall test was given in the center of the screen. Participants were instructed to guess if they did not know the answer, and they could not continue until at least three characters were typed. The retrieval practice was self-paced. The restudy and retrieval practice were interleaved with no more than three of one condition in a row. The remaining third of the words comprised the control/single presentation condition and were not revisited in this phase of the experiment. At the completion of this phase, participants left. They returned two days later to complete the experiment.

For the final phase, two days later, participants were given a cued-recall test in the center of the screen on one half of the word pairs (an equal amount from each condition). Participants were instructed to guess if they did not know the answer, and they could not continue until at least three characters were typed. For the other half of the word pairs, participants were tested for context memory; a word pair was presented on the screen along with the eight boxes from the study phase and participants were asked to click on the box in which the word-pair had originally been presented. The test types were blocked and the order of the test blocks was counterbalanced.

Results. The results of Experiments 1a and 1b are shown in Figure 2. The two experiments exhibited the same pattern and thus were combined for all analysis. Cued recall was significantly better in the retrieval practice condition than in the restudy condition, $t(50) = 4.48$, $p < .001$ and significantly better in the restudy condition than in the control condition, $t(50) = 7.24$, $p < .001$. The critical prediction is that context memory should be lower in the retrieval practice condition than in either the restudy or the control condition. This is not the case; in fact, performance on the context memory test was significantly better in the retrieval practice condition than in either the restudy or the control condition $t(50)=2.09$, $p < .05$. Context memory test performance was not significantly different between the restudy and control conditions.

Discussion. The critical prediction of the reconsolidation hypothesis was not confirmed in Experiment 1. Context memory was superior following retrieval practice than restudy, in direct opposition to the prediction made by the reconsolidation view. However, there is the potential for a measurement confound in this experiment that could work against the reconsolidation hypothesis. Specifically, because context memory may be dependent on item memory is possible that the expected difference in memory for the word pairs is biasing our measure of the context memory. Perhaps, in order for people to successfully retrieve the context memory of an item, they must first successfully recall the word itself. If this were true, context memory test performance may be higher in the retrieval practice condition not because the context memory is better, but because the memory for the word pair is better. What is needed is a way to conditionalize our measures of context memory on item memory, but the between-item nature of the testing procedure in Experiment 1 precludes such an analysis.

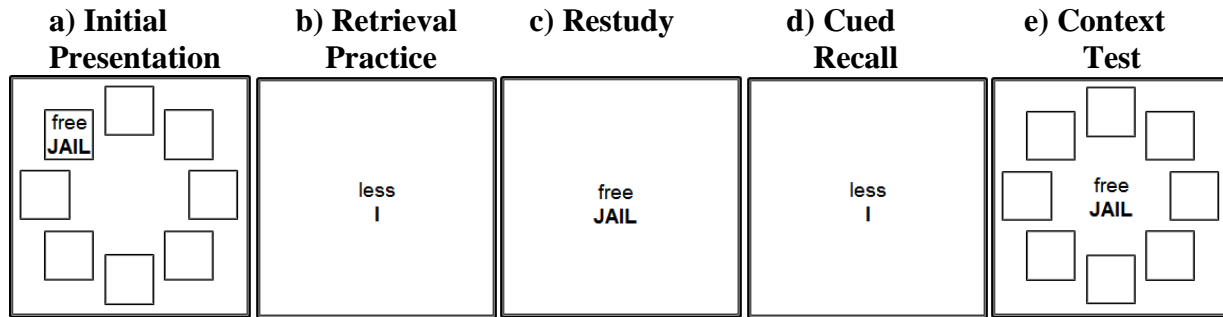


Figure 1. Sample display designs. (a) Eight boxes equidistant from the center of the screen with a word pair in one of the boxes used during the initial presentation of word pairs. (b) Cue from word pair with a blinking cursor below presented during retrieval practice trials. (c) Word pair is shown in the center of the screen for restudy trials. (d) Cue from word pair with a blinking cursor below presented for cued recall. (e) Word pair in the center of the screen surrounded by eight boxes which participants can click on presented during context test.

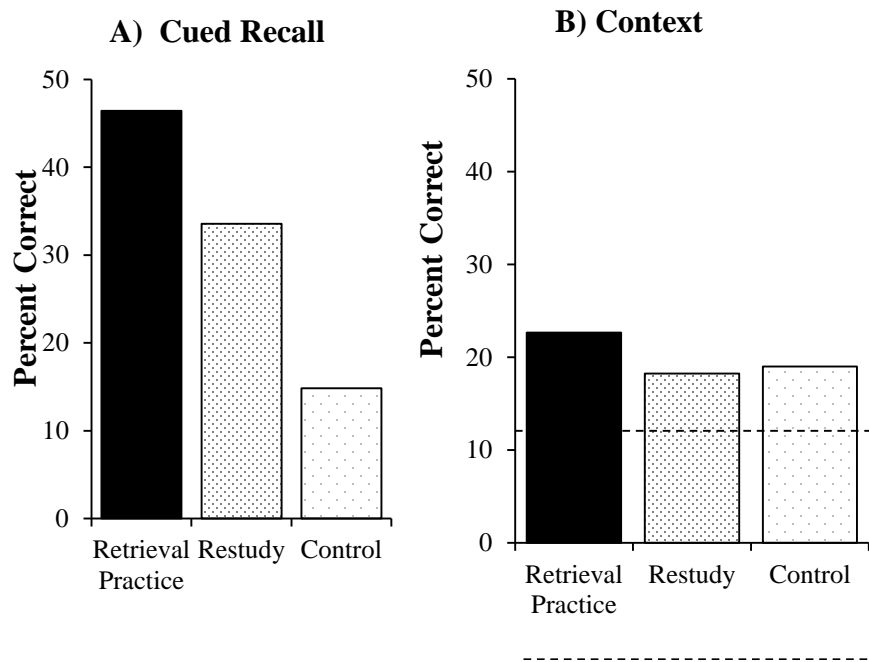


Figure 2. Combined results of Experiments 1a and 1b. (A) Percent correct on cued recall test in the retrieval practice, restudy and control conditions. (B) Percent correct on the context test in the retrieval practice, restudy and control conditions. Dashed line indicates chance performance.

CHAPTER 3

The Influence of Retrieval Practice on Context Memory Given Successful Recall

Experiment 2

This experiment was similar to Experiment 1b, but with the modification that each word pair was tested using both cued recall and context memory. This modification allows us to look at context memory selectively for the words that were successfully recalled, thus removing some of the confound from the Experiment 1 (though at the expense of introducing a potential for a carryover confound across conditions).

Participants. Fifty-nine students from the University of Illinois at Urbana-Champaign participated in the experiment in partial fulfillment of a course requirement. Ten participants had incomplete data and were excluded from the analysis (two due to a failure to attend day two of the experiment and eight due to computer problems).

Materials. The same materials were used as in Experiment 1b.

Procedure. The study phase and review phase were identical to Experiment 1b. In the final phase, each word pair was first tested using cued recall and then forced-choice context memory. The second test was administered using the complete word pair as a prompt, regardless of the response on the first (cued recall) test. Everything else remained the same.

Results. The results of Experiment 2 are shown in Figure 3. As expected, cued recall was significantly better in the retrieval practice condition than in the restudy condition, $t(48) = 3.01$, $p < .05$ which in turn was significantly better than the control condition $t(48) = 8.07$, $p < .001$. Replicating Experiment 1, we again found that context memory was superior in the retrieval practice condition, $t(48) = 2.07$, $p < .05$. Context memory test performance was not

significantly different in the restudy and control conditions, $t(48) = 0.83$. The critical prediction is that context memory of only those items successfully recalled on the cued recall task should be lower in the retrieval practice condition than in either the restudy or the control condition. This hypothesis was not confirmed; performance on the context memory test given successful recall was better in the retrieval practice condition than in the restudy condition (though this opposite effect was not significant, $t(48) = 1.58$, $p > .05$). Context memory test performance given successful recall was also not significantly different between the restudy and control conditions, $t(48) = 0.11$, $p > .05$.

Discussion. Experiment 2 again revealed a strong testing effect on memory for the word pairs. The context memory for all items (regardless of successful recall) replicated the results of Experiment 1. The same pattern was evident for context memory conditional upon successful cued recall, thus indicating that the prior results were not simply an artifact of confounding based on memory for the word pairs. From these experiments, it seems that retrieval practice does not hinder context memory. This set of results caused to revisit our interpretation of the reconsolidation hypothesis and seek alternative sources of evidence for the operation of reconsolidation.

In our earlier example in which a person took learned a lock combination at 8 AM and retrieved the combination at 10 AM, we suggested that they may replace the two traces with a single more general memory trace (such as “morning”), or that they may retain two memory traces and connect them, i.e. “8am or 10am”. We have found no evidence thus far for the former but have yet to address the later possibility. We revisit another prediction of the reconsolidation hypothesis based on this alternative interpretation in Experiment 3.

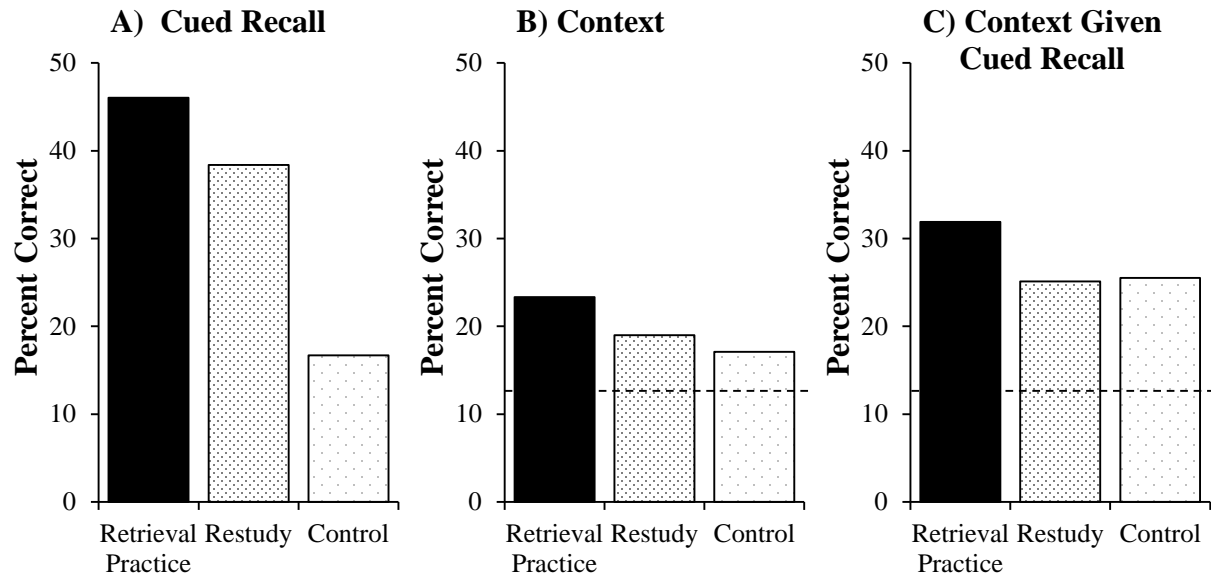


Figure 3. Results of Experiments 2. (A) Percent correct on cued recall test in the retrieval practice, restudy and control conditions. (B) Percent correct on the context test in the retrieval practice, restudy and control conditions. (C) Percent correct on the context memory test for only those word pairs that were successful recalled on the cued recall.

CHAPTER 4

THE INFLUENCE OF RETRIEVAL PRACTICE THAT INTRODUCES INTERFERING CONTEXTS

Experiment 3

If during reconsolidation a person connects contextual traces without generalizing, you would still expect a benefit of retrieval practice, because there would still be more contexts associated with the memory, i.e. both 8 AM and 10 AM in our prior example. If this is the case, then the two connected traces may be more easily recalled, but their respective contextual details may be more easily confused. Because Experiments 1, 2, and 3 always used the same “neutral” center position for the retrieval practice and never used that position for the initial presentation, there is little opportunity for confusion; participants knew that any recollection of a center position was from a restudy or retrieval practice and that any other location was from the original presentation. Experiment 4 seeks to produce interference during retrieval practice by having both restudy and retrieval practice occur in one of the boxes around the circle.

Experiment 3 is similar to Experiment 2, with the small modification that during the review phase, items are presented in one of the boxes around the circle (as in the study phase), rather than in the center of the screen. No individual word pair is ever presented in the same location for both the initial presentation and study retrieval practice phase. The purpose of this modification is to create retroactive interference. If the reconsolidation hypothesis is correct and memories are connected in the retrieval practice condition, it is reasonable to predict that the retrieval practice condition would suffer more from such interference, and thus the ability to remember the context of the original presentation will be lower.

Participants. Fifty-four students from the University of Illinois at Urbana-Champaign participated in the experiment in partial fulfillment of a course requirement. Four students were omitted from analysis due to a failure to attend day two.

Materials. The same materials were used as in Experiment 2 for all but the review phase. In the restudy phase the word pairs were no longer in the center of the screen, as in prior experiments, but rather shown in one of the eight boxes used in the initial presentation (see Figure 4a). In the retrieval practice condition, the prompt and blinking cursor was shown in one of the boxes (see Figure 4b). For both conditions, an individual word pair was never shown or tested in the same box as it was originally presented.

Procedure. With the exception of the changes described above, the procedure was identical to Experiment 2.

Results. Results of Experiment 3 are shown in Figure 5. Cued recall was significantly higher in the retrieval practice condition than in the restudy condition, $t(49)=4.62$, $p < .05$ and significantly higher in the restudy condition than in the control condition, $t(49)=8.67$, $p < .001$. We again found that context memory was superior in the retrieval practice condition relative to restudy, $t(49)=2.97$, $p < .05$ and in the restudy relative to the control, $t(49)=2.58$, $p < .05$. The critical prediction is that context memory for those items successfully recalled on the test of paired associates should be lower in the retrieval practice condition than in either the restudy or the control condition. The reconsolidation hypothesis was again not supported; in fact, performance on the context memory test given successful recall was significantly higher in the retrieval practice condition than the restudy, $t(47) = 2.68$, $p < .05$, which was also higher than in the control, though not significantly so $t(41)=0.89$.

Discussion. Experiment 3 again successfully demonstrated a strong testing effect. Memory for contexts was also highest in the retrieval practice condition, a finding that is concordant with prior results here but inconsistent with the reconsolidation hypothesis.

a) Restudy



b) Retrieval Practice

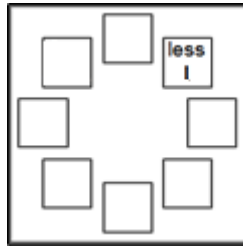


Figure 4. Sample display designs for Experiment 3. (a) Eight boxes equidistant from the center of the screen with a word pair in one of the boxes used during restudy. (b) Eight boxes again with a cue from word pair and a blinking cursor below presented in one of the boxes during retrieval practice trials.

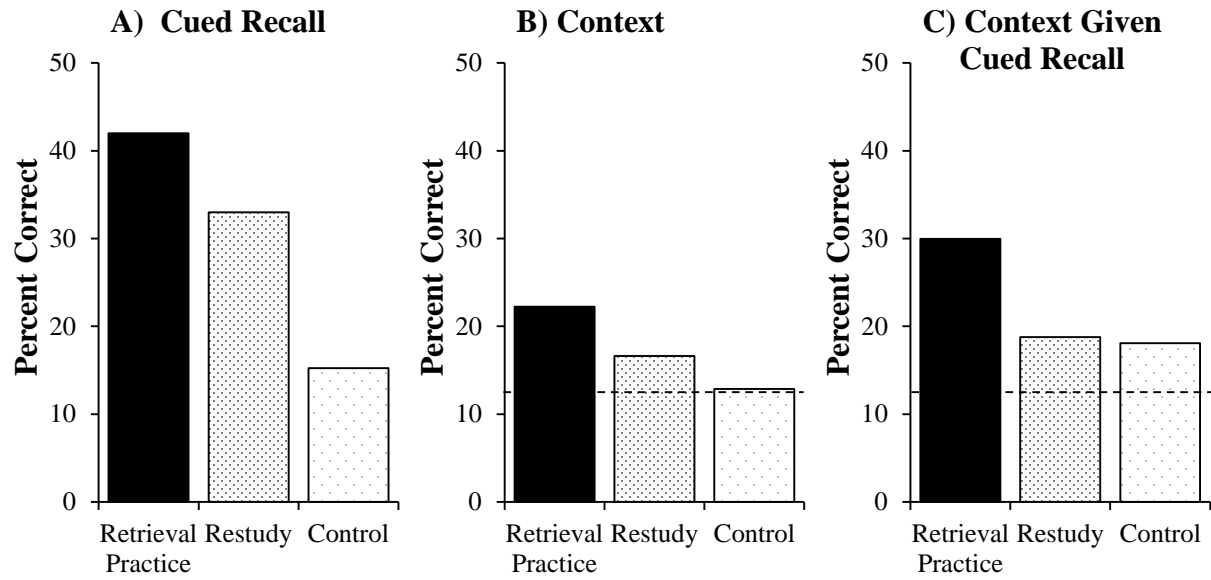


Figure 5. Results of Experiment 3. (A) Percent correct on cued recall test in the retrieval practice, restudy and control conditions. (B) Percent correct on the context. (C) Percent correct on the context memory test for only those word pairs that were successful recalled. Dashed line indicates chance performance.

CHAPTER 5

GENERAL DISCUSSION

Retrieval practice is an excellent way to enhance long-term memory; this broadly accepted claim has been supported by these experiments. Our work so far has found no support for our version of the reconsolidation hypothesis, but it is possible that we have been looking for signs in the wrong places. It is possible that reconsolidation may allow to events to be connected in memory without losing any individual detail. Future studies will seek to examine the reconsolidation hypothesis without the assumption that contextual details will be lost, as this does not appear to be the case.

One ongoing line of research is focused on how changes in context between the original study, review and final test influence the magnitude of the testing effect. If our hypothesis is correct, that testing increases the likelihood of recall by creating more general contextual cues, than one might anticipate that the greater the disparity in original context and retrieval practice context, the more likely there will be overlapping context with a final test, and, therefore, the more powerful the testing effect will be. Similarly, a greater difference in context between learning and the final test should engender a larger testing effect. This should also hold with increasing the difference in context between retrieval practice and the final test.

Despite the fact that our experiments did not reveal the decreased contextual memory we anticipated, there is still a lot of work to be done to fully examine our theory. The evidence for reconsolidation seems compelling and the notion that reconsolidation benefits memory through contextual cues cannot yet be dismissed. However, there may be other ways, as of yet unexamined, through which reconsolidation may cause the testing effect. Further,

reconsolidation may only play a small part in the testing effect with semantic mediators and the transfer appropriate processing also playing a role.

REFERENCES

- Anderson, A., Wais, P., & Gabrieli, J. D. E. (2006). Emotion enhances remembrance of neutral events past. *Proceedings of the National Academy of Sciences, USA*, 103, 1599–1604.
- Bacon, F. (2000). *Novum organum* (L. Jardine & M. Silverthorne, Trans.). Cambridge, England: Cambridge University Press. (Original work published 1620).
- Bjork, R. A. & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior*, 9, 567-572.
- Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. Solso (Ed.), *Information processing and cognition: the Loyola symposium*(pp. 123-144). Wiley, NY: Wiley, John & Sons, Incorporated.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalf & J. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-205). Cambridge, MA: MIT Press.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36, 129-148.
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1563–1569.
- Carpenter, S.K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1547-1552.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354-380.

- Dudai, Y. (1996). Consolidation: Fragility on the road to the engram. *Neuron*, 17, 367-370.
- Duncan, C. P. (1949). The retroactive effect of electroconvulsive shock. *Journal of comparative and Physiological Psychology*, 42, 32-44.
- Estes W.K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, 62, 145-154.
- Finn, B., & Roediger, H. L. (2011). Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later recall. *Psychological Science*, 22, 781-786.
- Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 801-812
- Hebb, D. O. (1949). *The organization of behavior*. Wiley, NY:
- Kahana, M. J. & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, 12, 159-164.
- Karpicke, J. D., & Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 704-719.
- Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, 67, 17-29
- McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology*, 16, 192-201.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, 16, 519-533.

- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). The labile nature of consolidation theory. *Nature Reviews Neuroscience*, 1, 216–219.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>.
- Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, 60, 437–447.
- Roediger, H.L. & Karpicke, J.D. (2006a). Test-enhance learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249-255.
- Roediger, H.L. & Karpicke, J.D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspective on Psychological Science*, 1, 181-210.
- Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.
- Smith, S. M. and Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods*, 42(1), 292-301.
- Smith, S. M., & Vela, E. (2001). Environmental context dependent memory: A review and meta-analysis. *Psychonomic Bulletin and Review*, 8, 203-220.
- Whitten, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior*, 16, 465–478.
- Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*.