Journal of Experimental Psychology: Learning, Memory, and Cognition

Manuscript version of

Elaborative Retrieval: Do Semantic Mediators Improve Memory?

Melissa Lehman, Jeffrey D. Karpicke

Funded by:
- National Science Foundation
- U. S. Department of Education, Institute of Education Sciences

Elaborative Retrieval: Do Semantic Mediators Improve Memory?

Melissa Lehman[1] and Jeffrey D. Karpicke[2]

[1]Lynn University and [2]Purdue University

Author Note

Melissa Lehman, Department of Psychology, College of Arts and Sciences, Lynn University and Jeffrey D. Karpicke, Department of Psychological Sciences, Purdue University.

This research and the writing of this paper were supported in part by grants from the National Science Foundation (DUE-0941170 and DRL-1149363) and the Institute of Education Sciences in the U. S. Department of Education (R305A110903). The opinions expressed are those of the authors and do not represent views of the Institute or the U. S. Department of Education. We thank Andrea Grovak for help with data collection.

Correspondence concerning this article should be addressed to Melissa Lehman, College of Arts and Sciences, Lynn University, 3601 N Military Trail, Boca Raton, FL 33431. Email: mlehman@lynn.edu and karpicke@purdue.edu.

Abstract

The elaborative retrieval account of retrieval-based learning proposes that retrieval enhances retention because the retrieval process produces the generation of semantic mediators that link cues to target information. We tested two assumptions that form the basis of this account: that semantic mediators are more likely to be generated during retrieval than during restudy and that the generation of mediators facilitates later recall of targets. Although these assumptions are often discussed in the context of retrieval processes, we noted that there was little prior empirical evidence to support either assumption. We conducted a series of experiments to measure the generation of mediators during retrieval and restudy and to examine the effect of the generation of mediators on later target recall. Across 7 experiments, we found that the generation of mediators was not more likely during retrieval (and may be more likely during restudy), and that the activation of mediators was unrelated to subsequent free recall of targets and was negatively related to cued recall of targets. The results pose challenges for both assumptions of the elaborative retrieval account.

Keywords: memory, retrieval practice, testing effect, elaboration, mediators

During recall of events that occurred in one's past, retrieval processes are enacted which provide access to information stored in memory. However, retrieval does more than simply provide a report of the information stored in memory; it also changes the information in a way that often makes it more accessible in the future. Memory tests enhance retention more than restudying (Roediger & Karpicke, 2006a), a phenomenon often referred to as the *testing effect.* While the testing effect is often discussed as a single effect, it actually represents a variety of direct and indirect benefits of taking memory tests. Indirect benefits of tests are those that occur not due to the retrieval processes used during testing, but due to additional processes invoked by testing that promote learning (Roediger & Karpicke, 2006b). For example, testing may provide opportunities to assess the proportion of material that has not yet been learned, motivating more study time, or tests may include feedback to correct misconceptions in knowledge. These indirect benefits occur due to mechanisms outside of the retrieval process. In contrast, taking a test can also produce direct effects on learning; learning is enhanced by the act of retrieval itself, even without subsequent study opportunities (Roediger & Karpicke, 2006b). While there are a variety of indirect effects of testing that may enhance retention, little is known about the mechanism underlying the direct benefits of retrieval. We refer to the memorial advantage produced by these direct benefits as the *retrieval practice effect*.

One explanation for the retrieval practice effect that has recently gained steam is the *elaborative retrieval hypothesis*, according to which retrieval promotes elaboration on encoded information that aids in later retrieval. Specifically, Carpenter (2009) proposed that retrieval of target information activates a network of semantically related information, which helps to provide access to the target information on a later test. The reasoning behind the elaborative retrieval hypothesis follows from the principle that semantic elaboration creates paths from cue information to target information. According to this hypothesis, activation of the cue in memory (in search of the target) produces activation of words that are semantically associated with the cue word. These semantic "mediator" words become associated with the cue and the target, providing a link or retrieval route from the cue to the target on a later test. When restudying the word pair, the generation of additional mediators does not necessarily occur because there is no search for the target word during restudy trials.

In support of the elaborative retrieval hypothesis, Carpenter (2009) showed that subjects were more likely to recall information that was initially retrieved under conditions that presumably increase the likelihood that such elaboration will occur. Subjects studied either strongly or weakly related cue-target pairs, after which they either restudied the pairs or completed a cued recall test requiring them to recall the target when presented with the cue. According to the elaborative retrieval hypothesis, the target is easily recallable when it is strongly associated to the cue, for example in the pair *toast-bread*. However, when the target is weakly related to the cue, such as in the pair *basket-bread*, the search for the cue activates other items that are semantically related to the cue, such as *eggs* and *flour*, and these items become mediators that presumably serve as additional retrieval routes from cue to target. It is assumed that as more mediators are produced for weakly related pairs, these pairs will be better remembered on a later test. Consistent with this prediction, subjects recalled more items that were part of weakly related pairs than strongly related pairs on a final free recall test (Carpenter, 2009). These findings are consistent with prior work suggesting that items are better recalled after they are subject to more "difficult" retrieval tasks (Carpenter & DeLosh, 2006).

The elaborative retrieval hypothesis is made up of two key assumptions. The first assumption, unique to this account, is that the retrieval process activates more mediators than restudying, and the second assumption is that the generation of more mediators enhances later retrieval. Mediators can refer to a variety of things, such as mental images formed to link cue and target words together; here we refer to semantic mediators, which are words that are semantically related to the cue word. This second assumption is related to ideas about the role of semantic elaborative processes in encoding and retrieval that have been around for decades (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975). As discussed in detail below, there are both theoretical and empirical reasons to question these assumptions, yet surprisingly, neither idea has been tested directly. The following sections review the support and potential problems for each assumption and then describe a set of experiments aimed at testing each assumption independently.

Assumption 1: Retrieval produces more mediators than restudy

The first assumption, that retrieval produces more semantic mediators than restudy, is consistent with Carpenter's (2009) data because conditions in which more mediators could potentially be generated produced greater recall. More direct support for this assumption comes from two experiments by Carpenter (2011) in which subjects were presented with lists of cue-target pairs (e.g., *mother-child*), after which they either restudied the pairs or completed a retrieval task in which they were presented with the cue and asked to recall the target. On a final criterial test, subjects were tested on the targets they had studied; however, their sensitivity to mediator words was also measured.

Experiment 1 used a final recognition test that included both cues and targets as correct items and unrelated words (e.g., *banquet*) and unstudied semantic mediators (e.g., *father*) as foils. The mediators were words considered to be semantically associated to the cue but not the target, based on the Nelson, McEvoy, and Shreiber (1998) word association norms. Experiment 2 used a cued recall final test, where the cues provided on the final test were either the same as the original cues, unstudied semantic mediator cues, or other non-studied items that were semantically related to targets but were not semantically related to the cues (e.g. *birth*). Carpenter (2011) proposed that if mediators are more likely to be activated during retrieval than during restudy, then on a later test, sensitivity to these mediators (as reflected by increased false alarms to semantic mediators and increased recall when semantic mediator cues are provided) should be higher in the retrieval condition than in the restudy condition. Indeed, false alarm rates for semantic mediators were higher in the retrieval condition than in the restudy condition (but this was not true for unrelated items), and the retrieval practice effect was larger on the final recall test when cues were semantic mediators (related to the cue but not the target) than when they were items that were semantically related only to the targets.

Although these data appear to be consistent with the elaborative retrieval hypothesis, the conclusion that mediators are more likely to be activated during initial retrieval is drawn from data that measures the activation of mediators only on a later criterial test (increased recognition of mediators and an increased benefit of mediator cues), not from any direct measurement of mediator activation during the initial retrieval when the activation is thought to occur. The nature of this measurement makes it difficult to determine when those mediators were activated and, as a result, whether mediator activation plays a causal role in the memory improvement produced by the retrieval process. For example, it is possible that mediators that are generated during initial study of a cue-target pair are retrieved along with the target during initial retrieval (Pyc & Rawson, 2010), which would both increase (erroneous) memory of the mediator and strengthen the episodic association between the mediator and the target. This would result in the effects observed by Carpenter (2011); however, this would mean that mediators were incidental to the retrieval process. In other words, mediators might become activated during study and be “strengthened” during initial retrieval of targets (e.g., Nelson & Goodmon, 2002), but this does not mean that the activation of mediators produces the strengthening of targets (Lehman et al., 2014).

A more direct test of mediator activation during initial retrieval would be a better way to determine whether retrieval produces more mediators than restudy. Decades of research on priming effects have given us tools to measure the activation of semantic information (e.g., McNamara & Healy, 1988; Meyer & Schvaneveldt, 1971; Tulving & Schacter, 1990), but the hypothesis that mediators are activated during the process of retrieval has not been tested by directly measuring the presence of mediators on the initial test using these tools. Experiments 1a, 1b, and 2 were aimed at testing Assumption 1 by measuring the activation of mediators during initial restudy and retrieval, when this activation is thought to occur.

Assumption 2: Generating more mediators during learning produces better subsequent recall

The second assumption, that the generation of more mediators produces better recall of targets, is consistent with broader ideas about the role of semantic elaboration in memory. Originally influenced by the finding that retrieval is enhanced when encoding occurs in the context of semantic processing rather than processing that focuses on non-semantic details, such as phonemic processing or mere repetition (Craik & Lockhart, 1972), the argument that semantic elaboration produces successful retention has been the subject of debate for many years (e.g., Baddeley, 1978; Craik & Tulving; 1975; Morris, Bransford, & Franks, 1977; Nelson, 1977; Tulving & Thompson, 1973). According to the elaborative retrieval account, mediators (like *father* and *love*) are activated in response to a cue (*mother*) during the search for a target (*child*) and become linked to both the cue and the target so that on a later test, when the cue *mother* elicits these mediators, they serve as various paths to access the target (Anderson, 1983; but see also Anderson, 1974). Accordingly, greater mediator activation means that more retrieval paths will be available, increasing performance on a later test. Interestingly, although this assumption enjoys strong anecdotal support among memory researchers, few studies have examined whether it is true, and those that have do not provide strong support (e.g., Montague & Kiess, 1968).

While there have been few experiments directly testing Assumption 2, theoretical arguments cast doubt on this assumption. As discussed recently by Lehman, Smith, and Karpicke (2014) and Karpicke, Lehman, and Aue (2014), the idea that associating more information with a cue will enhance recall of target information contradicts various models of memory that assume that retrieval of a target item is a function of a cue's ability to uniquely specify a target to the exclusion of extraneous information (Nairne, 2002; Raaijmakers & Shiffrin, 1981; Surprenant & Neath, 2009; Watkins & Watkins, 1975). According to such models, when more information is associated with the same cue, the probability of accessing target information is decreased (a situation referred to as *cue* *overload*). According to such models, which are often referred to as *cue distinctiveness* (or *cue diagnosticity*) accounts, if information that is semantically related to a cue is activated during the search for the target, later recall of that target will be decreased, rather than increased. Furthermore, cue distinctiveness accounts make the counter prediction that if retrieval invokes a process by which the activation of information that is semantically related to the cue is decreased, later recall of targets will be increased.

Consistent with such accounts is the phenomenon of *retrieval-induced forgetting*, the finding that recall of unpracticed members of a category is decreased below baseline after other members of that category have been retrieved in response to the category cue (Anderson & Spellman, 1995), suggesting that, contrary to Assumption 1, initial recall of targets produces *decreased* activation of other information that is semantically related to the cue. Similarly, the number of words that are implicitly semantically associated with a cue word is negatively associated with target recall, referred to as the *cue set size effect* (Nelson & McEvoy, 1979). Presumably if the set size were effectively increased via the activation of mediators, this would produce lower recall of targets according to cue distinctiveness models (see Karpicke et al., 2014, for a more detailed discussion of these issues).

The assumption that activating additional semantically-related information increases recall of targets seems inconsistent with some theoretical models, and also with extant data in other related paradigms, but it is a critical component of the elaborative retrieval account. Carpenter's (2009) finding that weak cues produced more retrieval-based learning of targets than strong cues is cited in support of this assumption. However, the design of Carpenter's experiments illustrates a critical limitation in many studies of semantic elaboration processes in memory: the conclusion that elaboration occurred comes from experiments that implement conditions in which elaboration might occur, but only final recall performance is measured, and elaboration is not measured directly (Baddeley, 1978). To illustrate this problem, imagine a researcher who hypothesizes that sugar produces weight gain in mice. An effective test of this hypothesis would involve manipulating the amount of sugar fed to the mice and examining its effect on weight gain, or at least allowing mice to eat freely and examining the correlation between the measured amount of sugar eaten and weight gain. An experiment in which the researcher puts mice into an empty donut box which he thinks could potentially have some sugar in it and then measures weight gain, without ever manipulating or observing the consumption of sugar, would be a tenuous way to test the hypothesis, but it is analogous to previous research on the link between semantic mediators and recall. The hypothesis that the generation of more mediators during study produces better recall has not been tested in experiments that manipulated or measured of the generation of mediators.

To examine the effect of generating mediators on later recall performance, we conducted several experiments involving both the measurement and manipulation of mediator generation. Experiments 3a and 3b were aimed at measuring the generation of mediators and examining the correlational relationship between the number of mediators generated and later recall of targets. Experiments 4 and 5 were aimed at manipulating the generation of mediators and examining the effect on later target recall.

Testing Assumption 1

The first set of experiments aimed to test the assumption that mediators are more likely to be activated during retrieval than during restudy by measuring the activation of mediator words immediately after retrieval and restudy trials. Lexical decision tasks are often used to measure implicit activation of semantic information (Meyer & Schvaneveldt, 1971) and provide an ideal method for testing mediator activation in this task. In a lexical decision task, subjects are presented with strings of letters and required to make decisions about whether the items are words or nonwords (meaningless items that look and sound like words). Semantic priming effects occur when lexical decision response times are faster for words that are related to primes. For example, if subjects are faster to make a lexical decision response for the item *father* than the item *bread* after seeing the prime *mother*, it is argued that the word *mother* has semantically primed the word *father* (Meyer & Schvaneveldt, 1971).

If mediators are activated at all by the exposure of cues during restudy or retrieval tasks, we expect that response times on a lexical decision task will be faster for mediators that are related to cue words than for unrelated words. Critical to Assumption 1, if mediators are more likely to be generated during retrieval than restudy, we expect to see faster response times to mediators on the lexical decision task in the retrieval condition. Of course, baseline responding may occur at different rates in the two conditions because the occurrence of the retrieval or restudy task immediately before the lexical decision task may produce differential task-switching costs. Thus, in Experiments 1a and 1b, we measured the difference in response time between lexical decisions for mediators and baseline response time to blank trials, in order to provide an additional index of mediator activation. Experiment 1a used a standard initial overt retrieval practice task, in which subjects retrieved an item and typed it on a computer, and compared this to a restudy task which also required an overt response (clicking a button to continue). In order to examine the effects of restudy/retrieve trials on lexical decision responses without the physical task-switching required in Experiment 1a, Experiment 1b required subjects to either restudy the items for 5 seconds or mentally recall the targets in covert retrieval trials (Smith, Roediger, & Karpicke, 2013) within 5 seconds, after which the program moved on automatically.

Experiments 1a and 1b

Method

Subjects. Subjects were 85 Purdue University undergraduates who participated in exchange for course credit. For all experiments reported in this manuscript, subjects were tested in groups of 1-4 people on individual computers. For Experiment 1a and 1b, groups were randomly assigned to experiment, resulting in 42 subjects in Experiment 1a and 43 subjects in Experiment 1b.[1] All manipulations were within-subject.

Materials and Design. A list of 32 word sets was created. Each set contained a cue word (e.g. *mother*) and a weakly associated target word (e.g. *child*) that would be studied, along with a strongly associated "mediator" word (e.g. *father*), an unrelated word (e.g. *banquet*), and a pronounceable nonword (e.g. *clett*) that were not studied. For each set, the cue shared similar associations to the target and mediator words to the materials described in Carpenter (2011). Specifically, cue and target words were weak forward associates (mean forward associative strength = .05; i.e., the target is produced as an associate to the cue during free association 5% of the time), mediator words were strong forward associates to the cue word (mean forward associative strength = .67,), and mediators and

[1] Each experiment in this paper was completed in one or two weeks; the number of subjects differs across experiments only because subjects signed up for participation more in some weeks than in others. No subjects participated in more than one of the experiments reported in this paper.

targets were weakly associated (mean forward associate strength from mediator to target = .05[2]), according to the Nelson, McEvoy, and Shreiber (1998) word association norms. In other words, during free association tasks, the target word *child* is not often produced as an associate to the cue word *mother*, but the mediator word *father* is often produced as an associate to the cue word *mother*. Word sets were unrelated to other word sets in the list. Word sets were randomly selected for each participant from a larger group of 42 possible sets. Unrelated words were matched for length and concreteness, and nonwords were matched for length. Word lists used for all experiments are provided in the Appendix.

For the 32 word sets used in the experiment, half of the cue-target pairs appeared on restudy and half on retrieve trials. For the lexical decision trial that followed each restudy or retrieve trial, 8 trials contained mediator words, 8 contained unrelated words, 8 contained nonwords, and 8 were blank trials. The order of the list and trial type was randomized for each phase of the experiment. Response times were measured for each word type (mediator, unrelated, nonword, and blank) after each trial type (restudy and retrieve), resulting in a 2 (trial type) x 4 (word type) within-subjects design.

Procedure. Experiments 1a and 1b consisted of three phases. The experiments were identical, except that in the second phase, Experiment 1a required overt response and Experiment 1b required covert responses. Subjects were instructed on each task and completed practice trials before beginning each phase of the experiment. In the first phase of the experiment, subjects were informed that they would study pairs containing a cue word (*mother*) and a target word (*child*), after which they were given an example pair and asked to identify the cue word and the target word to ensure that they understood the

[2] The mediator-target forward strength differs slightly from Carpenter (2011), who used mediators and targets that had 0 forward strength.

terminology before they began studying the cue-target pairs (e.g., *mother – child*). Each pair appeared on the screen, one at a time, for 5 seconds. After studying all pairs, subjects began the second phase of the experiment, in which they completed a series of randomly intermixed restudy or retrieve trials, followed by a lexical decision trial. During restudy trials, subjects restudied the pair, and during retrieve trials, they retrieved the target in response to the cue. After each restudy or retrieve trial, they completed a lexical decision task for one of the items from the word set. On restudy trials in Experiment 1a, subjects were required to restudy the pair for 3 seconds, after which a button labeled "Submit" became enabled and they could continue at their own pace. On retrieve trials in Experiment 1a, subjects were shown the cue word along with a two-letter word stem from the target word (e.g. *mother – ch*______) and were asked to type in the target word that they had studied in the previous phase. On restudy trials in Experiment 1b, the pair appeared on the screen for 5 seconds, and on retrieve trials, the cue and target word-stem appeared on the screen for 5 seconds, after which the program automatically advanced. During the retrieve trials, subjects asked were to mentally recall the targets.

Immediately after each restudy or retrieve trial, a "+" symbol appeared in the center of the screen for 500ms, followed by a lexical decision trial. During the lexical decision task, the mediator word that was associated to the cue (e.g., *father*), an unrelated word (*banquet*), or a nonword (*clett*) appeared in the center of the screen, or the center of the screen remained blank. If an item appeared on the screen, subjects were instructed to determine whether the item was a word or nonword, as quickly as possible, and to click the corresponding button (either "Word" or "Nonword"), which appeared under the item. On blank trials, a blank button appeared below the empty space in the middle of the screen, and subjects were instructed to click the blank button as quickly as possible. Subjects were informed that they would be tested on the cue-target pairs again later, and they were not informed that some of the items from the lexical decision task were related to the cue-target pairs. After completing two practice trials, they completed 32 trials in which the restudy or retrieve task was followed by the lexical decision task. Subjects who performed fewer than 75% of the lexical decision trials correctly were excluded, resulting in 40 subjects in Experiment 1a and 39 in Experiment 1b. These subjects performed an average of 96% of the lexical decision trials correctly.

In the third phase of the experiment, subjects completed a 60 second math distractor task (2-digit addition problems), followed by a 3-minute cued recall task, in which original cues were shown and subjects were asked to recall the targets. The cues were presented in a list on the left side of the screen along with an empty text box on the right side, and subjects were asked to type the corresponding target word next to each cue[3]. Subjects were instructed to go in order, and words remained in the textbox until the recall period was over.

Results

Experiment 1a. In the second phase of the experiment, subjects spent an average of 5 seconds on restudy trials and an average of 6 seconds on retrieve trials, and they correctly recalled an average of 86% of targets on retrieve trials. There were no differences between restudy and retrieve conditions in the correct identification of words in the lexical decision task ($t < 1$). The critical results concern the speed of lexical decision responses for

[3] The simultaneous presentation was used in all experiments rather than presenting words one at a time so that the cued recall condition would be similar to the free recall conditions present in Experiments 3-4.

each word type. Response times were submitted to a 2 (trial type) x 4 (word type) repeated measures ANOVA.

The data are shown in Figure 1. The elaborative retrieval hypothesis predicts that mediators are more likely to be generated during retrieval trials than during restudy trials. If so, then these mediators should be primed following retrieval trials, and responses should be faster for mediators in the retrieval condition relative to the restudy condition. Thus, the elaborative retrieval hypothesis predicts a trial type x word type interaction. As shown in the top left panel of Figure 1, although response times were faster for mediators than for unrelated items, $t(39) = 3.09$, $d = 0.49$ [0.16, 0.81], results of the two-way ANOVA revealed no trial type x word type interaction, $F(3,39) = 0.75$, $MSE = 52865.90$, $p = .52$.

Because the retrieval and restudy tasks could have produced overall differences in response times, we also estimated priming effects within each condition by comparing mediator response times to those for "control" trials (in Experiments 1a and 1b, we used both unrelated words and blank trials as control trials). Priming of mediators is indicated by faster response times for mediators compared to control trials. According to the elaborative retrieval hypothesis, we should observe greater priming of mediators following retrieval trials compared to restudy trials. Thus, we compared the mediator priming effect in the retrieval and restudy conditions by computing difference scores for these priming effects across the two conditions.

Priming effect data are shown in Table 1. A priming effect for retrieval versus restudy conditions would be indicated by a positive number in the priming effect columns. For Experiment 1a, there was little difference in mediator priming effects in the retrieval versus restudy conditions (regardless of whether the effect was measured by comparing response times for mediators to those for unrelated words or blank trials). However priming was relatively greater in the restudy condition, the opposite pattern of that predicted by the elaborative retrieval hypothesis.

Analyses of the final cued recall data revealed the typical retrieval practice effect. As shown in the top right panel of Figure 1, more retrieved targets were recalled than restudied targets, $t(39) = 8.80$, $d = 1.39$ [0.95, 1.82]. In addition to correct recall, we also examined erroneous recall of mediators on the final cued recall test. Although there were no differences in mediator activation across conditions according to measurements of lexical decision times, we may see support for Assumption 1 if mediators are more likely to be recalled on the final cued recall test in the retrieve condition. Contrary to this prediction, however, recall of mediators was higher in the restudy condition than in the retrieve condition, $t(39) = 2.85$, $d = 0.45$ [0.12, 0.77]. As shown in the bottom panel of Figure 1, this was true even when the analyses were restricted to targets in the unrelated and nonword conditions, in which the mediator words were never previously encountered in the experiment (non-mediator trials), $t(39) = 3.73$, $d = 0.59$ [0.25, 0.93], and also when analyses were conditionalized on whether an error was made, $t(36) = 1.53$, $d = 0.24$ [-0.07, 0.55] (the conditionalized data include only 37 subjects because 3 had no errors in the retrieve condition). That is, when an incorrect item was produced on a cued recall trial, it was more likely to be a mediator in the restudy condition than in the retrieve condition.

Experiment 1b. Whereas Experiment 1a used overt responding on restudy/retrieve trials, Experiment 1b used covert restudy/retrieval; however, all other tasks were the same, and thus the same analyses were conducted for Experiment 1b. There were no differences between restudy and retrieve conditions in the correct identification of words in the lexical decision task (t < 1). Again, as shown in the top left panel of Figure 2, response times were significantly faster for mediators than for unrelated items, $t(38) = 6.42$, $d = 1.03$ [0.63, 1.41], but the two-way ANOVA revealed no trial type x word type interaction, $F(3,38) = 0.22$, $MSE = 87997.517$, $p = 0.88$. Additionally, consistent with Experiment 1a, there was very little difference in mediator priming between retrieval and restudy conditions, as shown in Table 1.

The final cued recall test revealed the standard retrieval practice effect, $t(38) = 3.13$, $d = 0.50$ [0.16, 0.83], as shown in the top right panel of Figure 2. There were very small differences in erroneous recall of mediators on the final test, in the same direction as in Experiment 1a. As shown in the bottom panel of Figure 2, recall of mediators was slightly higher in the restudy condition than in the retrieve condition, $t(38) = 1.54$, $d = 0.24$ [-0.07, 0.56]. Again this was true in the non-mediator trials, $t(38) = 0.93$, $d = 0.15$ [-0.17, 0.47], and also when analyses were conditionalized on whether an error was made, $t(35) = 0.85$, $d = 0.13$ [-0.18, 0.44] (the conditionalized data include only 36 subjects because 3 had no errors in the retrieve condition).

Discussion

The results of Experiments 1a and 1b do not support the assumption that mediators become more activated during retrieval than during restudy. Despite a clear advantage of the retrieval task on subsequent recall, there is no evidence that mediators are more active in the retrieval condition (in fact, the evidence suggests that mediators are more active in the restudy conditions). One possibility is that although mediators are naturally more active during retrieval, the procedure used in the experiment caused mediator activation during restudy. Because subjects knew that a related word might appear during the next lexical decision trial, they may have used restudy trials to generate a related mediator. If so, the artificial activation of mediators in the restudy condition may have washed out any differences between conditions. If this were the case, however, then the elaborative retrieval hypothesis would predict that later recall of targets in the restudy condition should benefit from the generation of these mediators, yet performance on the final cued recall test showed a significant advantage for the retrieval condition. Thus, this alternative explanation seems unlikely. Although it is difficult for this alternative explanation to account for the finding that mediator generation appears to be similar in these two conditions but recall is higher in the retrieval condition, we conducted an additional experiment in which the lexical decision task occurred for all items after the restudy/retrieval phase was completed, to rule out the possibility that the lexical decision task produced the activation of mediators in the restudy condition.

Experiment 2 replicated the procedure used in Experiments 1a and 1b except that the restudy/retrieve phase contained no lexical decision trials. Instead, following the restudy/retrieve phase, an additional phase was added in which subjects completed the lexical decision task for all items. Again, the final phase consisted of a cued recall task. Because the lexical decision task occurred after the restudy/retrieve phase (rather than immediately after each trial), subjects could respond using keypresses, avoiding task switching. Additionally, because the lexical decision task occurred only after all restudy/retrieve trials were complete, there is no way that expectations on the lexical decision task could have influenced subjects' behavior during the restudy/retrieve trials. Thus any differences in priming of mediators or in final cued recall of targets would reflect pure influences of the restudy/retrieve trials.

Experiment 2

Method

Subjects, Materials, and Design. Subjects were 32 Purdue University undergraduates, who participated in exchange for course credit. All manipulations were within-subject. Materials were the same as those used in Experiments 1a and 1b. Lexical decision response times were measured for each word type (mediator, unrelated, and nonword) for each trial type (restudy and retrieve), resulting in a 2 (trial type) x 3 (word type) within-subjects design.

Procedure. The experiment consisted of four phases. In the first phase, subjects studied the 32 cue-target pairs in the same fashion as in Experiments 1a and 1b. After studying all pairs, subjects began the second phase of the experiment, in which they completed a series of randomly intermixed restudy and retrieve trials. On restudy trials, each pair appeared on the screen for 5 seconds, after which the program automatically moved on to the next pair. On retrieve trials, subjects were shown the cue word along with a two-letter stem of the target word and were asked to type in the target word that they had studied in the previous phase. They were informed that they would be tested on all cue-target pairs again later.

Once all restudy/retrieve trials were completed, subjects moved on to the third phase, the lexical decision task. In this task, a "+" symbol appeared in the center of the screen for .5 seconds, followed by a lexical decision trial. During the lexical decision task, either the mediator word that was associated to the cue, an unrelated word, or a nonword appeared on the screen. Subjects were instructed to determine whether the item was a word or nonword, as quickly as possible, and to press the corresponding button. For word trials, subjects were to push the "Z" button, and for nonword trials, subjects were to push the "M" button. They were instructed to keep their fingers on the respective buttons throughout the task. Subjects were given instructions and completed two practice trials before beginning the task. During the task, 10 trials contained mediator words, 10 contained unrelated words, and 12 contained nonwords. The order of the list and trial type was randomized for each phase of the experiment. Subjects performed an average of 95% of the lexical decision trials correctly; no subjects performed fewer than 85% of the lexical decision trials correctly, and thus no subjects were excluded. In the last phase, subjects completed a 60 s math distractor task followed by a 3-minute cued recall task, in which original cues were shown and subjects were asked to recall the targets. The procedure for these tasks was identical to that in Experiment 1a and 1b.

Results

The same analyses were conducted as in Experiments 1a and 1b. In the second phase of the experiment, subjects spent an average of 5.5 seconds on each retrieve trial and correctly recalled an average of 82% of targets. There were no differences between restudy and retrieve conditions in the correct identification of words in the lexical decision task ($t < 1$). Lexical decision response times were submitted to a 2 (trial type) x 3 (word type) repeated measures ANOVA.

The data are shown in Figure 3. As in Experiments 1a and 1b, results revealed that although response times were faster for mediators than for unrelated items, $t(31) = 7.85$, $d = 1.39$ [0.89, 1.87], according to the two-way ANOVA, there was no trial type x word type interaction, $F(3,31) = .11$, $MSE = 11356.02$, $p = .89$. As shown in the top left panel of Figure 3, there is no evidence that mediators were activated more in the retrieval condition than in the restudy condition. Table 1 shows that, as in Experiments 1a and 1b, there was little difference in priming effects for mediators in the retrieval and restudy conditions; as with Experiment 1a, the priming effect was slightly greater for the restudy condition. The results from Experiment 2 converge with the results of Experiments 1a and 1b, providing no evidence that semantically related mediators become more active during retrieval relative to restudy.

Analyses of the final cued recall data revealed the typical retrieval practice effect. As shown in the top right panel of Figure 3, more retrieved targets were recalled than restudied targets, $t(31) = 3.95$, $d = 0.70$ [0.31, 1.08]. As in Experiment 1a and 1b, we examined erroneous recall of mediators on the final cued recall test. Again, contrary to the prediction that more mediators are activated in the retrieve condition, we found that recall of mediators was higher in the restudy condition than in the retrieve condition, $t(31) = 4.44$, $d = 0.78$ [0.38, 1.18]. As shown in the bottom panel of Figure 3, this was true even when the analyses were restricted to targets in the unrelated and nonword conditions, in which the mediator words were never previously encountered in the experiments (non-mediator trials), $t(31) = 3.74$, $d = 0.66$ [0.27, 1.04], and also when analyses were conditionalized on whether an error was made, $t(28) = 3.91$, $d = 0.69$ [0.30, 1.07] (the conditionalized data include only 29 subjects because 3 had no errors in the retrieve condition). Thus, when an incorrect item was produced on a cued recall trial, it was much more likely to be a mediator in the restudy condition than in the retrieve condition. Because the lexical decision task occurred only after all restudy and retrieve trials were completed, this effect could not be due to a change in study strategy that may have been produced by the lexical decision task in Experiments 1a and 1b.

Discussion

The results of Experiments 1a, 1b, and 2 fail to show support for Assumption 1 of the elaborative retrieval account. There is no evidence from these experiments that semantic mediators are more likely to become activated during retrieval. Additionally, the measurement of mediator recall on the final test suggests that mediators may become more activated in the restudy condition than in the retrieval condition, the opposite pattern of that predicted by Assumption 1. Importantly, we observed retrieval practice effects in all three experiments.

In order to get an overall estimate of the critical effects, we combined the mediator priming data from the three experiments to provide an overall priming effect across experiments. As shown in the bottom row of Table 1, there was very little difference in mediator priming between retrieval and restudy conditions, and the data indicated slightly greater mediator priming after restudy. We also combined the retrieval practice effect data across the three experiments. Despite very little difference in mediator priming, there were robust and consistent retrieval practice effects. It is difficult for the elaborative retrieval hypothesis to explain why, if the activation of mediators produces the retrieval practice effect, we observed such effects without any activation of mediators in retrieval conditions.

One possibility is that mediators were being activated during retrieval, but for some reason, we missed the effect. For example, in Experiment 2, it is possible that because semantic priming is time-sensitive, differential activation of mediators occurred during the restudy/retrieve phase but it was not observed because it had decreased by the time the lexical decision task had begun. If that were the case, however, it remains to be explained why we found the opposite pattern: the trend across the three experiments suggested that mediators were more likely to be activated during restudy trials vs. retrieval practice trials. In the General Discussion, we consider the evidence that mediators are activated during retrieval practice in more detail. For now, we move on to an examination of Assumption 2 of the elaborative retrieval account.

Testing Assumption 2

The second critical assumption of the elaborative retrieval theory is that the act of generating mediators directly enhances subsequent retention of target words. In order to address whether the generation of mediators enhances recall, in Experiments 3a, 3b, 4, and 5, we induced mediator generation/elaboration and examined the effects of elaboration on later recall. Critically, in these experiments, we measured or manipulated elaboration during initial learning in order to examine the relationship between elaboration and subsequent recall. According to the elaborative retrieval hypothesis, we should see greater recall when more mediators are activated, because more mediators are assumed to serve as retrieval routes from the cue to the target. Thus, the elaborative retrieval hypothesis predicts a positive relationship between the number of mediators generated and later target recall. Additionally, if mediators serve to provide retrieval routes from cues to targets, we may expect the benefits of this type of elaboration to be more pronounced in cued recall than in free recall. This is because when original cues are provided, the retrieval routes would provide paths from these cues to the targets, whereas in free recall, the original cues are not provided, and thus those retrieval routes may not be activated. In contrast to these predictions, a cue distinctiveness perspective predicts that target recall will be higher when fewer nontargets are associated with a given cue (e.g., Nairne, 2002). Thus, the cue distinctiveness perspective predicts a negative relationship between the number of mediators generated and recall, particularly in cued recall – the opposite prediction of the elaborative retrieval account.

In Experiments 3a and 3b, subjects studied a list of target words and then freely generated associates to cue words (which had strong forward associations to the previously studied target words). We examined the relationship between the number of associates generated and recall of the targets on a criterial test. During the associate generation phase, subjects were not asked to recall the target items, nor were they made aware that the targets were associated with the cues, because the goal was to induce the generation of mediators while searching for the target without eliciting episodic retrieval processes, in order to disentangle the semantic generation of mediators while generating a target from the episodic retrieval of a target.[4] Thus, these experiments were intended to induce the type of semantic elaboration thought to occur during retrieval of targets, according to the elaborative retrieval hypothesis, without engaging episodic retrieval processes. Although Experiments 3a and 3b are correlational in nature, they serve as a first step in testing these predictions, as a negative relationship between the number of mediators generated and target recall would contradict the elaborative retrieval hypothesis. Experiments 3a and 3b were identical, aside from the type of mediator associate generated. Experiment 3a required the generation of semantic associates (following Carpenter, 2009) and 3b required word-stem completions (following Carpenter & DeLosh, 2006).

[4] See Karpicke et al. (2014) for a more detailed discussion of the evidence that it is the episodic nature of retrieval, and not a semantic process, that produces retrieval-based learning.

Experiment 3a and 3b

Method

Subjects. Subjects were 78 Purdue University undergraduates, who participated in exchange for course credit. Groups of 1-4 subjects were randomly assigned to each experiment, resulting in 35 subjects Experiment 3a and 43 subjects in Experiment 3b.

Materials. For each experiment, two lists of 16 target words were created. For Experiment 3a, each target word (e.g., *earth*) had a semantically associated cue word (e.g. *globe*), which was used as the cue in the association task. Experiment 3b used word stems of the targets as the cues (e.g., *ear*___) in the association task. For Experiment 3a, words were selected such that the cues had either large or small associative sets according to the Nelson et al. (1998) word association norms, and the target was either the 2nd or 3rd strongest associate to the cue. Half of the items had small associative sets (fewer than 12 associates) and half had large associative sets (greater than 15 associates). In Experiment 3b, words were selected such that, within a list, each word had a unique first letter. The word stems used in the association task were either two- or three-letter stems (randomly assigned), where there were at least 2 words that could complete the stems. Both of these set-size manipulations were intended to increase variability in the number of associates that could be generated during the association phase.

Procedure. All subjects completed one list of words under each of two final testing conditions, cued recall and free recall, which were counterbalanced across subjects. There were three phases for each condition. In the first phase, subjects were informed that they would study a list of target words (e.g., *earth*) that they would need to remember on a later test. Targets appeared on the screen one at a time for 5 seconds each. In the second phase, subjects generated associates to cues. In Experiment 3a, they were told that they would see a series of cue words (e.g., *globe*), and that for each cue word, they were to enter as many associated words as they could in 16 seconds. In Experiment 3b, subjects were told that they would see a series of word-stems (e.g., *ear*___) and for each stem they were to generate as many words as possible that would complete the stem. At the beginning of the task, subjects were informed that they would not need to remember the words from the word association task. Each cue word appeared on the screen one at a time with a response box below it, and subjects typed their responses in the box. After 16 seconds, the program automatically advanced to the next cue. After completing the last cue, subjects completed a 60 s math distractor task (two-digit addition problems).

In the third phase, subjects took a free or cued recall test, after which the entire procedure was repeated with the other type of test. For the cued recall test, subjects were asked to recall each of the target words from the first phase of the experiment. Subjects in Experiment 3a were told that they would be provided with a clue word (e.g., *globe*) to help them remember the target, and subjects in Experiment 3b were told that they would be provided with the first few letters of the target (e.g., ear___) to help them recall it. The cues shown were the same cues that appeared in the second phase of the experiment. Cues were randomly ordered, and next to each was a blank box in which subjects could enter the corresponding target. Subjects were asked to recall the targets in the order in which the clues were presented. For the final free recall test, subjects were asked to type all of the target words from the experiment into a text box on the screen. They could enter the words in any order, and the words that they typed remained in the text box until the end of the free recall period.

Results

Analyses were restricted to trials in which the targets were correctly produced among the associates generated during the second phase (the associate generation phase) of the experiment (for example, the item *earth* was only included in the recall analyses if it was correctly generated as an associate to the cue *globe* in the second phase).[5] This occurred on 33% of trials in Experiment 3a and 73% of trials in Experiment 3b. The number of items generated ranged from 1 to 11 in Experiment 3a; however, on only one trial were 1, 10, or 11 items generated, and thus, these trials were excluded from analyses. The number of items generated ranged from 1 to 8 in Experiment 3b. Because the trials analyzed were only those in which targets were correctly produced, when 1 item was generated, it was the target item and there were no "mediator" items; when 2 items were generated, one was the target item and one was a mediator; 3 items meant the target and two mediators; etc.

In each experiment, subjects were allowed to enter as many associates as possible for each cue word; thus, different subjects had different numbers of data points. For example, one subject may have produced 2 items for some cues, 4 items for other cues, and 7 items for other cues, but never generated 1, 3, 5, 6, or 8 items for any cues, whereas another subject may have produced a completely different pattern of results. These data were analyzed in two different ways: using correlations computed for each subject, and using recall probabilities averaged across subjects to compute overall correlations.

[5] Only trials in which the targets were correctly produced were included because the goal was to examine the effect of generating associates while generating the target, and this is impossible when the target was not generated.

In the first analysis, correlations between the number of mediators generated and recall of corresponding targets were computed for each subject. Correlation coefficients were transformed to z-scores using Fisher's r-to-z transformation. The z-scores were then submitted to a one-sample t-test to determine whether the average correlation was significantly different from 0. In Experiment 3a, for cued recall, this resulted in a total of 9 positive correlations and 15 negative correlations (the remaining subjects in each experiment either showed no correlation or had too few data points to compute a correlation coefficient). When averaged across subjects, the correlation coefficient (r = -.11) was not significantly different from 0 (t = 1.05). For free recall, this resulted in a total of 13 positive correlations and 11 negative correlations. When averaged across subjects, the correlation coefficient (r = .09) was not significantly different from 0 (t < 1.) In Experiment 3b, for cued recall, this resulted in a total of 12 positive correlations and 20 negative correlations. When averaged across subjects, the correlation coefficient (r = -.17) was significantly less than 0, $t(26) = -2.48$, p = .02, suggesting a significant negative correlation between number of mediators generated and target recall. For free recall, this resulted in a total of 19 positive correlations and 22 negative correlations. When averaged across subjects, the correlation coefficient (r = .03) was not significantly different from 0 (t < 1.)

The first analysis indicated no correlation for free recall, and a negative correlation in cued recall. Because this analysis excluded subjects with too few data points to calculate a correlation, we conducted an additional analysis in which we averaged recall probabilities across participants for each number of associates generated. We then examined the correlation between the number of associates generated and correct recall of targets. The recall results are shown in Figure 4. As shown in the top panel of Figure 4, Experiment 3a revealed a negative correlation between number of semantic associates generated and cued recall performance, $r = -.77$, $p = .026$, and weak relationship between number of semantic associates generated and free recall performance, $r = .30$, $p = .47$, which was not significantly different from zero. As shown in the bottom panel of Figure 4, Experiment 3b revealed the same pattern of results: a negative correlation between number of word-stem associates and cued recall performance, $r = -.89$, $p = .003$, and a weak relationship between number of word-stem associates and free recall performance, $r = -.24$, $p = .57$, which again was not significantly different from zero. These patterns are generally consistent with the previous analysis; the relationship between the generation of mediators and target recall was negative under cued recall conditions and was weak or nonexistent in free recall.

Additional analyses showed no effect of associative set size on final recall performance for either cued or free recall in Experiment 3a (both *t*s < 1). In Experiment 3b, targets associated with 3-letter cues (which produced fewer associates) were recalled better than targets associated with 2-letter cues in cued recall, $t(41) = 4.90$, $d = .76$ [.41, 1.10], but there was no difference in free recall ($t < 1$).

Discussion

Experiments 3a and 3b provide no support for the hypothesis that generating more mediators produces higher target recall. In both experiments, the number of mediators generated during the initial task was not correlated with final free recall, and the number of mediators generated during the initial task was negatively related to cued recall, which is the opposite pattern from what would be predicted according to the elaborative retrieval hypothesis. Instead, the results are consistent with a cue distinctiveness hypothesis. When subjects generated more associates to a cue, final cued recall of targets decreased. Also consistent with a cue distinctiveness account, the number of associates generated in response to a cue was not related to free recall performance, when the original cues were not provided during criterial retrieval. Thus, when the amount of elaboration occurring during the initial task is measured, we failed to find support for Assumption 2.

However, the conclusions that we can draw from Experiments 3a and 3b are limited for a number of reasons. First, these data are correlational, so no causal conclusions about the effects of generating mediators on later recall of targets can be made. For example, while we propose that the negative correlation between the number of associates generated and recall is observed because generating more associates produces a situation of cue overload, it is possible that items for which fewer associates were generated are simply more memorable items for some reason. In addition, subjects were free to generate as many associates as possible, and thus the number of associates generated was not controlled. Finally, many trials were discarded in Experiment 3a because the target word was not generated during the first phase. Experiment 4 was conducted to address these issues by experimentally manipulating mediator generation.

Experiment 4

In Experiment 4, the number of mediators that subjects were asked to generate was manipulated. Similar to Experiments 3a and 3b, subjects studied a list of targets in the first phase and then generated associates to cues in the second phase. However, subjects were instructed to generate 2, 4, or 6 associates to the cues. Because one of the associates was the target, the conditions are referred to as the 1-mediator, 3-mediator, and 5-mediator conditions, respectively. Additionally, word stems were provided to encourage generation of a specific set of associates, which included the target word along with other strong associates. For example, if subjects had studied the target word *earth* in the first phase, they might see the cue *globe* in the second phase, along with the word stems *wo*_____, *ro*_____, *ea*_____, and *at*_____. This procedure allowed us to control the number of associates generated for each item, increased the likelihood of successful generation of the target during the second phase, and produced equivalent successful target generation regardless of the number of associates generated. As in Experiments 3a and 3b, subjects then completed a final free or cued recall test for the targets.

Method

Subjects. Subjects were 50 Purdue University undergraduates who participated in exchange for course credit.

Materials. Two lists of 15 target words were created, and each target word had a semantically associated cue word that was used in the association task. In addition to the target word, the other 5 most closely associated target words with unique 2-letter cue stems were selected for each cue word from to the Nelson et al. (1998) word association norms. Again, the target was one of the top 3 strongest associates to each cue. For example, for the cue word *globe*, the 6 most closely related associates were *world, round, earth* (the target), *atlas, ball,* and *circle*. For each list, 5 targets were assigned to the 1-mediator condition, 5 to the 3-mediator condition and 5 to the 5-mediator condition. For example, for the target word *earth*, in the 1-mediator condition, subjects would see the cue *globe* along with the word stems *wo*____ and *ea*_____; in the 3-mediator condition, subjects would see the cue *globe* along with the word stems *wo*_____, *ro*_____, *ea*_____, and *at*_____; and in the

5-mediator condition, subjects would see the cue *globe* along with the word stems *wo*_____, *ro*_____, *ea*_____, *at*_____, *ba*_____, and *ci*_____. List order was randomized for all phases, as was order of word stems in the second phase.

Procedure. All subjects completed one list of words under each of two final testing conditions, free recall and cued recall. In the first phase of the experiment, subjects were informed that they would study a list of target words, and that they would need to memorize target words for a later test. Targets appeared on the screen one at a time for 5 seconds each. In the second phase, subjects generated associates to cues. They were told that they would see a series of cue words and that they would have to generate 2, 4, or 6 associated words for each cue word. They were not informed that the some of the associates would be targets from the previous phase. They were informed that they would be shown word stems to help them generate the associated words and that they must enter a response for each word stem before moving on to the next cue. Finally, they were told that they would not need to remember the cues or associates for a later task. The task was self-paced. A blank box appeared next to each word stem, with a "Submit" button at the bottom of the page. After entering an associate into each box, they were allowed to push the "Submit" button to move on to the next cue. Subjects watched an example trial in which the computer completed word stems, and then they completed one practice trial, after which they completed this task for all cue words. Response time was measured for each trial, from the appearance of the cue word until the "Submit" button was clicked.

In the third phase, subjects completed a 60 second math distractor task then completed a free recall test, after which the entire procedure was repeated with a final cued recall test. Free recall always came before cued recall in order to avoid subjects noticing that the cues were the same as those that appeared during the second phase and then intentionally studying the cues during the second phase of the next cycle. The procedure for the cued and free recall tests was identical to that in Experiment 3a.

Results

Analyses were restricted to trials in which the targets were correctly produced among the associates generated during the second phase, which was 67% of trials. There was no difference in the probability of correctly producing the target across conditions ($F < 1$). Response time was positively associated with the number of associates generated, $F(1,49) = 262.17$, $MSE = 44.37$, $p < .001$ ($M = 8.7$s to generate 2 items, $M = 20.2$s to generate 4 items, and $M = 32.5$s to generate 6 items).

The recall data are shown in Figure 5. For cued recall, there were small differences between conditions, $F(1,49) = 2.92$, $MSE = .071$, $p = .059$. Recall was better for targets that were generated along with 1 mediator than for targets generated with 3 mediators, $t(48) = 2.43$, $d = .35$ [.06, .64], or targets generated with 5 mediators, $t(48) = 1.87$, $d = .27$ [-.02, .56]. There was no difference between recall of targets generated with 3 mediators and recall of targets generated with 5 mediators, $t < 1$. For free recall, there were no differences between any of the conditions, $F(1,49) = .099$, $MSE = .096$, $p = .91$ (all ts < 0.50).

Discussion

The findings of Experiment 4 are consistent with the findings from Experiments 3a and 3b. While there was no relationship between the number of mediators generated and free recall performance, the number of mediators generated in addition to the target was negatively associated with cued recall performance; more targets were recalled when only one mediator was generated than when three or five mediators were generated in addition to the target. Due to the limited number of data points, it is not clear whether there is a linear negative relationship between the number of mediators generated and target recall in cued recall, as suggested by a cue distinctiveness account, or whether the effect was driven by something unique to the 1-mediator condition. Experiment 5 was conducted in order to further explore this pattern in cued recall with a greater number of conditions. The procedure was similar to that of Experiment 4, except that only a cued recall final test was used, and the second phase of the experiment involved the generation of 2, 3, 4, 5, and 6 associates.

Experiment 5

Method

Subjects, Materials, and Procedure. Subjects were 33 Purdue University undergraduates, who participated in exchange for course credit. Materials were the same as those used in Experiment 4, but a single list of 20 targets was randomly selected from the larger set used in Experiment 4. Four targets each were assigned to the 1-mediator (2 associate), 2-mediator (3 associate), 3-mediator (4 associate), 4-mediator (5 associate) and 5-mediator (6 associate) conditions. The procedure was the same as that in Experiment 4 except that only one list and only final cued recall was used. Additionally, after completing the final cued recall task, subjects were asked to indicate whether they noticed during the word-association phase that some of the associates were targets that appeared in the previous phase.

Results and Discussion

Analyses were restricted to trials in which the targets were correctly produced among the associates generated during the second phase, which was 62% of trials. There was no difference in the probability of correctly producing the target across conditions ($F < 1$). Subjects spent more time generating more associates, $F(1,32) = 242.03$, $MSE = 76.27$, $p < .001$ (M = 10.1s to generate 2 items, M = 18.4s to generate 3 items, M = 29.1s to generate 4 items, M = 34.0s to generate 5 items, and M = 40.2s to generate 6 items).

The recall data are shown in Figure 6. To examine the effect of the number of mediators generated on recall, the cued recall data were submitted to a repeated measures ANOVA. There was a significant negative linear trend, $F(1,32) = 4.70$, $MSE = .12$, $p = .04$, shown in the left panel of Figure 6. Additionally, these data, collapsed across subjects for each number of mediators (as in Experiments 3a and 3b), were submitted to a correlational analysis, revealing a significant negative correlation between the number of mediators generated and target cued recall,$r = -.95$, $p = .02$. Analyses of responses to the final question revealed that 14 subjects (42% of subjects) noticed at some point during the association phase that some of the associates were target words from the previous phase. To ensure a pure measure of the effect of mediator generation on target recall, we excluded those subjects who noticed and may have intentionally retrieved targets during the association phase, and we repeated the analysis, revealing a nearly identical negative linear trend, $F(1,17) = 9.13$, $MSE = .11$, $p = .008$, shown in the right panel of Figure 6.

One possibility is that target items that were generated alongside fewer mediators enjoy the benefit of a more closely associated semantic mediator network, and the cohesion of this network enhances recall of targets (e.g., Anderson, 1976). That is, because the mediators chosen were those most closely associated to the cue, and because the target is related to the cue, it is possible that the semantic relatedness among the target and the mediators was higher when there were fewer mediators generated. To examine this possibility, we calculated the semantic relatedness between each target and the mediators generated in the set using Latent Semantic Analysis (Landauer & Dumais, 1997) to determine whether the averaged semantic relatedness was correlated with recall. There were no significant correlations among averaged semantic relatedness of the mediators and recall of targets for any number of mediators generated, nor was there a significant correlation between overall recall and overall semantic relatedness, or between the number of mediators generated and semantic relatedness when collapsed across subjects. Although the semantic relatedness was numerically higher when the target was generated with only one mediator (M = .28) than for any other number of mediators (M = .22 for 2-mediator sets, M = .19 for 3-mediator sets, M = .19 for 4-mediator sets, and M = .21 for 5-mediator sets), semantic relatedness does not appear to explain the relationship between the generation of semantic mediators and recall.

As in Experiments 3a, 3b, and 4, in Experiment 5, the number of mediators generated along with the target was negatively associated with subsequent target recall. Additionally, Experiment 5 suggests that, consistent with a cue overload prediction, the trend is linear; as more items become associated with a given retrieval cue, it becomes more difficult to recall the target item.

General Discussion

Across 7 experiments, we have failed to find any support for either of the two assumptions critical to the elaborative retrieval hypothesis. The results of Experiments 1a, 1b, and 2 suggest that mediators are not more likely to be generated during retrieval than during restudy. In fact, our data suggest that if there are any differences in mediator activation across conditions, mediators may be activated more often in the restudy condition than in the retrieve condition. The results of Experiments 3a, 3b, 4, and 5 suggest that the generation of more mediators has no effect on final free recall, and more importantly, has a negative effect on final cued recall, the exact opposite pattern of that predicted by the elaborative retrieval hypothesis. These findings, however, are consistent with a cue distinctiveness hypothesis – the generation of more mediators produced cue overload, rendering cues less effective at eliciting targets (see Karpicke et al., 2014, for a description of a cue distinctiveness account that is consistent with the current findings).

There are a handful of possible limitations of the present experiments that are worth considering. It is possible that semantic elaboration normally occurs during initial retrieval but that our procedure did not induce elaboration or mediator generation in the best ways. For example, it may be that our failure to find any evidence of mediator activation in Experiments 1a, 1b, and 2 occurred because we gave subjects stems of the target words. We used this method because in pilot work, when we gave subjects only the cue words, we were unable to obtain results like Carpenter's (2011); initial recall was extremely low, so there was no advantage of initial recall over restudy on the criterial test. In the present experiments, giving subjects stems of the target words may have restricted memory search and prevented the generation of mediators that do not start with the same letters as the targets. For example, providing the cue *mother-ch____* may restrict the mediators that people generate to words that start with *ch*. This could explain the findings that the selected mediators (such as *father*) were more likely to show up in the restudy conditions, when there was no such "limitation" on what could be generated. However, it seems unlikely that presenting this cue in the retrieval condition would initiate an automatic spreading activation process for mediators related to the word *mother* (as suggested by Carpenter, 2009) but simultaneously initiate some other controlled process where only words beginning with *ch* can become mediators.

Further, if it were true that the present procedure reduced or eliminated elaboration during initial retrieval, then the present results would pose a different but nonetheless substantial challenge to the elaborative retrieval account. All of the present experiments showed robust and consistent retrieval practice effects. If the conditions used in the present experiments somehow reduced or eliminated elaboration or mediator generation, then the retrieval practice effects observed here occurred in the absence of elaborative retrieval (see too Karpicke & Smith, 2012). That outcome would be problematic for the theory that elaborative retrieval is the mechanism that causes retrieval practice effects.

Of similar concern, in the critical stage in Experiments 3a, 3b, 4, and 5, subjects were provided with the cue and asked to generate associates (one of which was the target), but they were not directly asked to generate associates that connected the cue to the target. Perhaps the intention of generating mediators that link cues to targets is critical for the elaborative retrieval account of retrieval practice. However, the idea that intentional generation of mediators matters seems unlikely for several reasons. First, it is difficult to see how a person could intentionally generate a word that links a cue to a target when the target is not known (i.e., while the person is searching for but has not yet recalled the target). Although it is possible that linking words could be generated after the target is successfully recalled, it is not clear what purpose that would serve if the target could be recalled without linking words. More importantly, Karpicke and Smith (2012) had subjects intentionally generate mediators to link cues to targets and found that generating linking words after successful retrieval produced no benefits over additional study, whereas repeatedly retrieving targets after successful retrieval produces large benefits.

Along the same lines, one possible limitation of Experiment 3b is that, unlike Experiment 3a, it does not allow for the activation of information that is semantically related to the target, because only word-stems were provided, and thus any word that fits the word stem could be generated (even if it is not semantically related to the target). However, this is a similar type of retrieval task to that used by Carpenter and DeLosh (2006) in their Experiments 2 and 3, the results of which they explained by "elaborative retrieval processing" (p. 274). Moreover, the results of Experiment 3b were similar to those of Experiment 3a, suggesting that the semantic nature of the task was irrelevant.

Despite these contradictions to the elaborative retrieval hypothesis, it still may be that we did not elicit the type of elaboration that the elaborative retrieval hypothesis assumes should occur during retrieval. If that is the case, it is difficult to see how one might test the elaborative retrieval account unequivocally. To our knowledge, the present experiments are the most direct test of both the assumption that mediators are more likely to be generated during retrieval and the assumption that the generation of mediators aids later recall, and we failed to find support for either assumption. The present work helps clarify the assumptions of the elaborative retrieval account, and support for the account rests on results consistent with these two assumptions.

The present investigation highlights the challenges inherent in examining the role of elaboration in remembering. It is well established that elaborative or deep encoding tasks improve memory. The effect of elaborative study is usually examined by manipulating the task performed by subjects during encoding and measuring performance on a later memory test. However, when one encoding condition produces better memory performance than another, a researcher may simply attribute the different to elaboration, without independent evidence that elaboration occurred during study. Specifically, the presence of mediators during encoding has often been inferred from performance on a criterial test, instead of being examined by manipulating or measuring mediator generation during the original learning activity. Craik and Tulving (1975) pointed out that "there are obvious dangers of circularity present in that any well-remembered event can too easily be labeled *deeply processed*" (p. 271), and they suggested that researchers should use independent indices of depth of processing during encoding. Yet 40 years after Craik and Tulving's landmark paper, there is still no standard index of elaboration during encoding.

A related challenge is that to measure elaboration during learning, one must provide a precise definition of elaboration. The term "elaboration" has been used to refer to a wide variety of activities in previous research, but a strength of the elaborative retrieval account is that it offers a precise definition of elaboration: In that account, elaboration is defined as the activation of mediator words that are semantically related to cue words. To determine whether elaboration occurs in a particular learning task, it is necessary to measure elaboration in some way, and to determine whether elaboration enhances subsequent retention, it is necessary to manipulate elaboration during initial learning. Otherwise, it is impossible to know whether elaboration plays a causal role in memory or is merely epiphenomenal to other processes (Underwood, 1972). In the present experiments, based on the definition of elaboration from the elaborative retrieval account, we measured the occurrence of elaboration during retrieval and study trials, and we induced elaboration and assessed its effects on a subsequent test, but the results did not support the idea that elaboration occurred during retrieval or was responsible for retrieval practice effects.

It is worth considering how the current findings might relate to previous research on mediator effectiveness more broadly. A wealth of research has shown that the production of mediators improves retention in verbal learning tasks (see Richardson, 1998, for a review). However, much of this work has been focused on imagery or other strategies rather than semantic elaboration. In studies measuring semantic elaboration, mediator generation has frequently been confounded with item type: pairs for which mediators were successfully generated were more likely to consist of related items relative to pairs for which mediators were not generated (e.g., Richardson, 1998), making it impossible to draw causal conclusions about the effects of mediator generation, per se. Further, prior work on mediator generation has been focused on the generation of mediators during study trials, when both cue and target were present (e.g., Bellezza & Poplawsky, 1974; Pyc & Rawson, 2010). The issue of interest in the present work is whether mediators are produced during the process of retrieval, and we are aware no studies examining or manipulating elaborative processing during retrieval.

The goal of the present investigation was not to suggest that elaboration does not improve memory. Accessing semantic meaning, forming mental images, thinking about relationships among items, and many other elaborative study tasks have been shown to improve memory when people are instructed to use these strategies during encoding (e.g., Craik & Tulving, 1975; Richardson, 1998). Instead, the purpose of the present studies was to examine the role of elaboration in retrieval practice effects by testing two assumptions of the elaborative retrieval account: first, that mediators (elaborations) are more likely to be activated during retrieval than they are during restudy, and second, that the generation of more mediators during retrieval enhances subsequent retention. While the results pose challenges for the elaborative retrieval account, they also highlight broader challenges inherent in examining the role of elaboration in remembering.

References

Anderson, M.C., & Spellman, B.A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102*, 68-100.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 5,* 451-474.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261-295.

Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review, 85*(3), 139-152.

Bellezza, F. S., & Poplawsky, A. J. (1974). The function of one-word mediators in the recall of word pairs. *Memory & Cognition, 2,* 447-452.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, & 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, & Cognition, 37*, 1547-1552.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition, 34*, 268-276.

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6) 407-428.

Collins, A. M., & Quillian, M. R. (1972). Experiments on semantic memory and language comprehension. In L. Gregg (Ed.), *Cognition and Learning* (pp. 117–138). New York, NY: Wiley.

Coppens, L.C., Verkoeiken, P.P.J.L., & Rikers, R.M. (2012). *Semantic mediators and the testing effect in cue-target pairs: Investigating the role of mediator-cue association strength.* Poster presented at the 53rd Annual Meeting of the Psychonomic Society, Mineapolis, MN.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal behavior,* 11, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation, Vol. 61* (pp. 237-284). San Diego, CA: Elsevier Academic Press.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding*. Journal of Memory and Language, 67*, 17-29.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semanctic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-140

Lehman, M. & Malmberg, K. J. (2013). A buffer model of encoding and temporal correlations in retrieval. *Psychological Review, 120*(1), 155-189*.*

Lehman, M., Smith, M. A., & Karpicke, J. D. (in press). Toward an episodic context account of retrieval-based learning: Dissociating retrieval practice and elaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

McNamara, T.P., & Healey, A.F. (1988). Semantic, phonological, and mediated priming in reading and lexical decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 398-409.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology: General, 90*, 227-234.

Montague, W. E., & Kiess, H. O. (1968). The associability of CVC pairs. *Journal of Experimental Psychology, 78*(2), 1–23.

Morris, C. D., Bransford, J., & Franks, J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519-533.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10*, 389-395.

Nelson, D. L., & Goodmon, L.B. (2002). Experiencing a word can prime its accessibility and its associative connections to related words. *Memory & Cognition, 30*, 380-398.

Nelson, D. L., & McEvoy, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 292–314.

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/.

Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 151-171.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis*. Science, 330*(6002), 335.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychnomic Bulletin & Review, 5,* 597-614.

Roediger, H. L., & Karpicke, J. D. (2006a). Test-enhanced 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. *Perspectives on Psychological Science, 1,* 181-210.

Smith, M. A., Roediger, H. L., & Karpicke, J. D. (2013). Covert retrieval practice benefits retention as much as overt retrieval practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1712-1725.

Surprenant, A. M., & Neath, I. (2009). *Principles of memory*. New York: Psychology Press.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373.

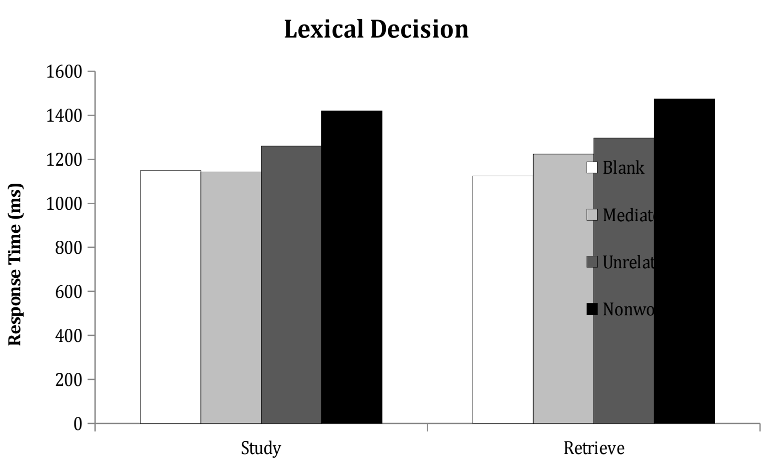
Underwood, B.J. (1972) Are we overloading memory? In A.W. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington, D.C.: Winston & Sons.

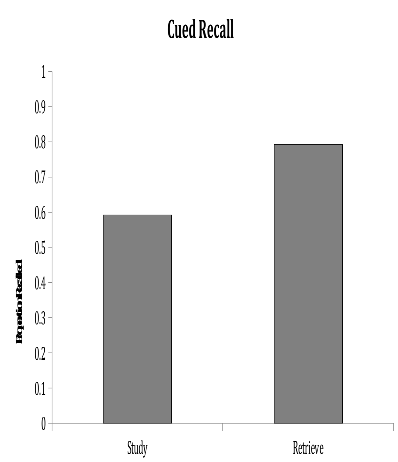
Watkins, O.C. & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1*(4), 442-452.

Table 1: Priming effect for mediators during retrieval compared to restudy, and retrieval practice effects for Experiments 1a, 1b, and 2.

	Priming Effect for Mediators vs. Unrelated Words	Priming Effect for Mediators vs. Blank Trials	Retrieval Practice Effect
Experiment 1a	-46 [-190, 98]	-106 [-248, 37]	1.39 [0.95, 1.82]
Experiment 1b	26 [-126, 180]	-48 [-236, 139]	0.50 [0.16, 0.83]
Experiment 2	-14 [-76, 46]	-	0.70 [0.31, 1.08]
Overall	**-11 [-86, 63]**	**-77 [-192, 37]**	**0.88 [0.49, 1.26]**

Note: Priming effects were calculated by computing difference scores in lexical decision response times (in milliseconds) for mediators and comparison conditions (unrelated words and blank trials), and then calculating the priming differences between retrieval and restudy conditions. A priming effect of zero in this table indicates no difference in priming between restudy and retrieval conditions, a positive number indicates greater priming in retrieval conditions, and a negative number indicates greater priming in restudy conditions. Experiment 2 did not include blank trials. The right column shows retrieval practice effects.





Mediator Intrusions in Cued Recall

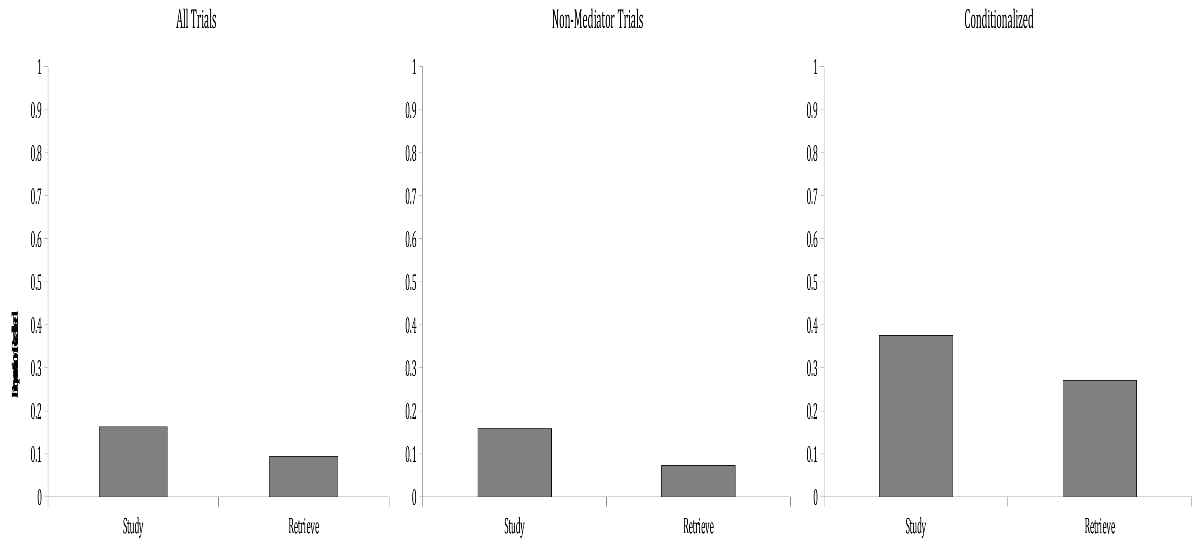
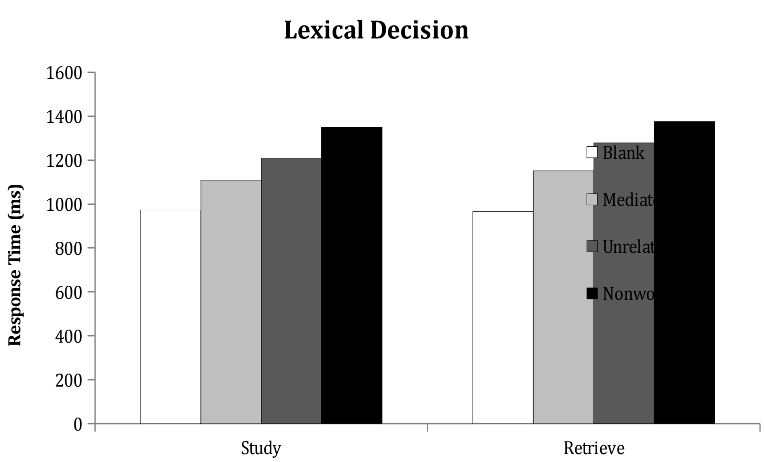
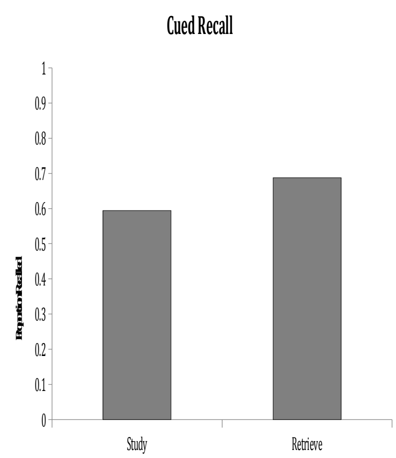


Figure 1: Data from Experiment 1a. The top left panel shows the lexical decision response times for each item type. The top right panel shows the proportion of targets recalled on the final cued recall test. The bottom panel shows the proportion of mediators incorrectly recalled on the final cued recall test for all trials, non-mediator trials only (i.e. recall of mediators that did not previously appear in the lexical decision task), and the proportion of mediators recalled conditionalized on whether an error was produced during the cued recall test. Error bars represent standard error.





Mediator Intrusions in Cued Recall

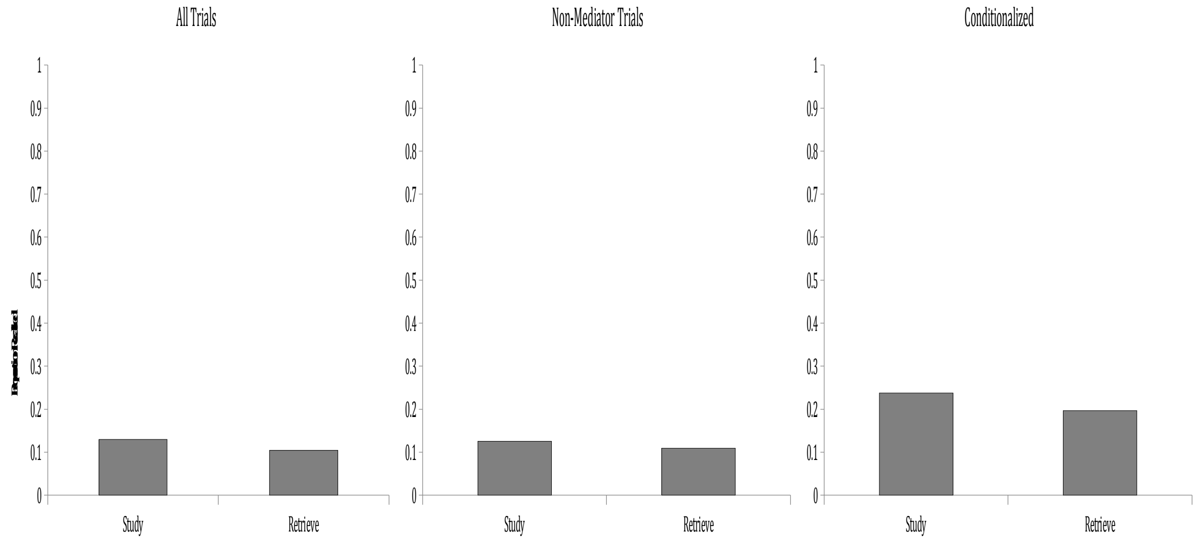
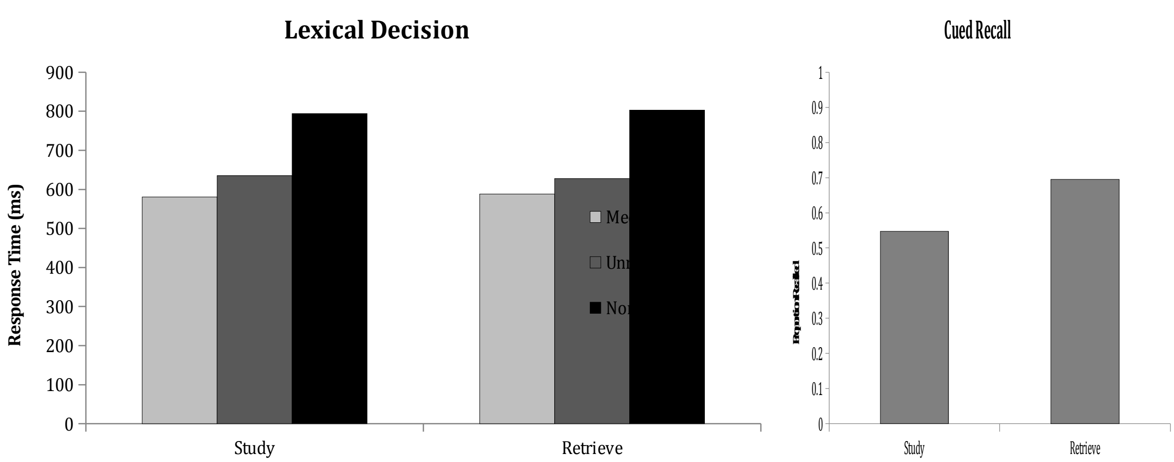


Figure 2: Data from Experiment 1b. The top left panel shows the lexical decision response times for each item type. The top right panel shows the proportion of targets recalled on the final cued recall test. The bottom panel shows the proportion of mediators incorrectly recalled on the final cued recall test for all trials, non-mediator trials only (i.e. recall of mediators that did not previously appear in the lexical decision task), and the proportion of mediators recalled conditionalized on whether an error was produced during the cued recall test. Error bars represent standard error.



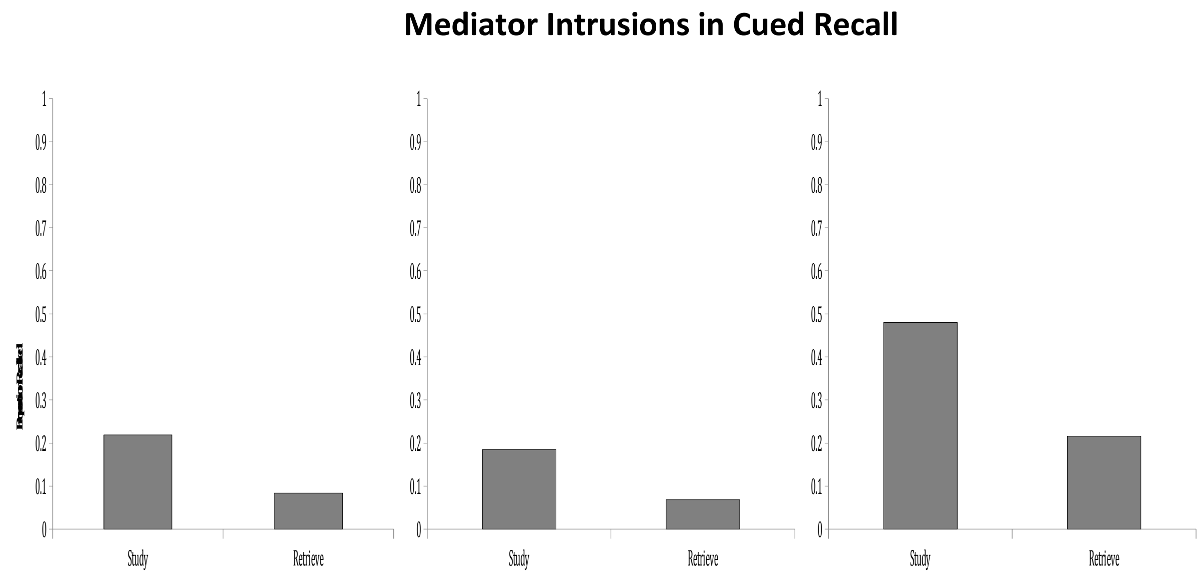
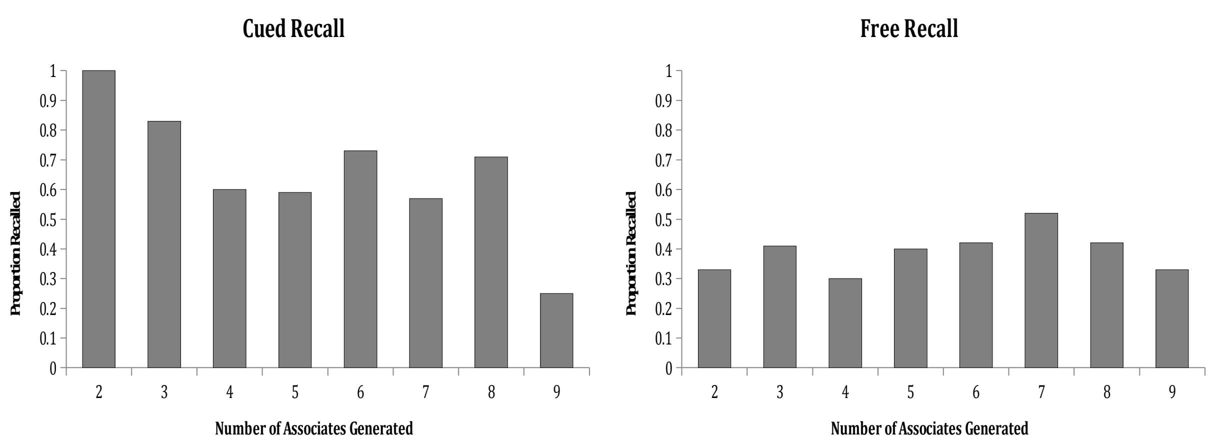


Figure 3: Data from Experiment 2. The top left panel shows the lexical decision response times for each item type. The top right panel shows the proportion of targets recalled on the final cued recall test. The bottom panel shows the proportion of mediators incorrectly recalled on the final cued recall test for all trials, non-mediator trials only (i.e. recall of mediators that did not previously appear in the lexical decision task), and the proportion of mediators recalled conditionalized on whether an error was produced during the cued recall test. Error bars represent standard error.



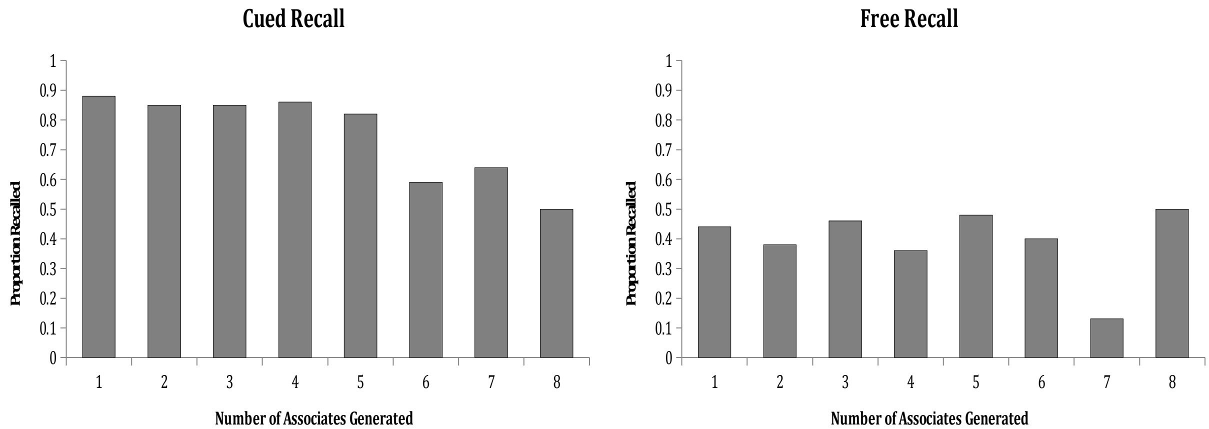


Figure 4*:* Proportion of targets recalled in cued recall and free recall for each number of associates generated in Experiment 3a (top panels) and 3b (bottom panels). The number of associates generated ranged from 2-9 in Experiment 3a and from 1-8 in Experiment 3b.

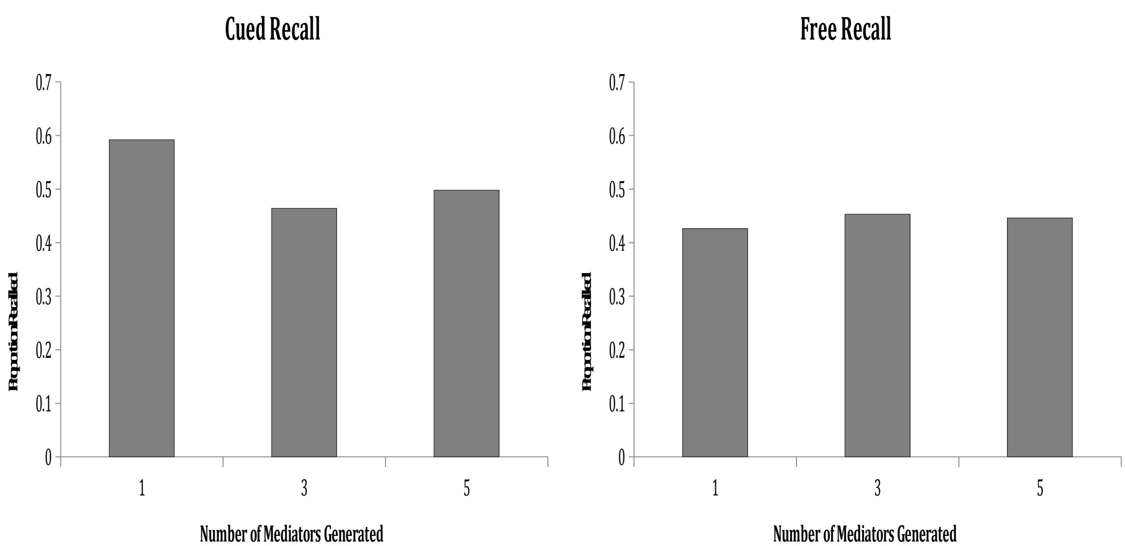


Figure 5*:* Proportion of targets recalled in cued recall and free recall for condition in Experiment 4. The number of mediators generated refers to the number of associates generated in response to a cue, in addition to the target item (i.e. for 1-mediator trials, 2 associates to the cue were generated: 1 target and 1 mediator; for the 3-mediator trials, 4 associates were generated: 1 target and 3 mediators; etc.). Error bars represent standard error.

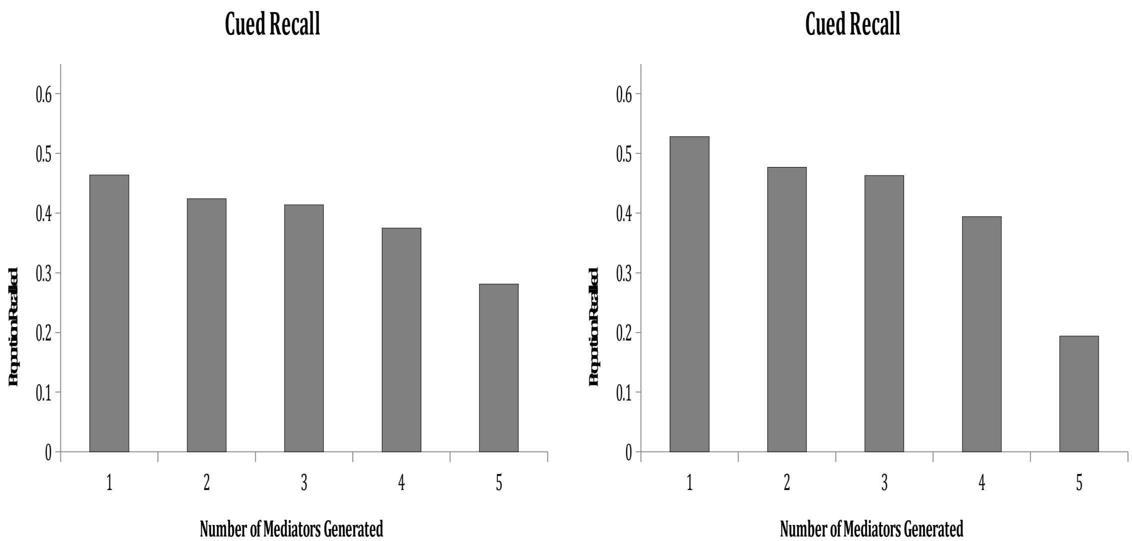


Figure 6: Proportion of targets recalled in cued recall Experiment 5. The number of mediators generated refers to the number of associates generated in response to a cue, in addition to the target item (i.e. for 1-mediator trials, 2 associates to the cue were generated: 1 target and 1 mediator; for the 3-mediator trials, 4 associates were generated: 1 target and 3 mediators; etc.). The left panel shows data from all subjects. The right panel shows data only from subjects who did not notice that some words in the associate-generation phase were targets from the previous phase. Error bars represent standard error.

Appendix

Words used in Experiments 1a, 1b, and 2

Cue	**Target**	**Mediator**	**Unrelated**	**Experiment**
antler	fawn	deer	garden	E1a, E1b, E2
arm	thigh	leg	mosquito	E1a, E1b, E2
bulb	lamp	light	morning	E1a, E1b, E2
calculus	equation	math	liver	E1a, E1b, E2
calf	bull	cow	game	E1a, E1b, E2
cash	bank	money	noodle	E1a, E1b, E2
cathedral	steeple	church	circle	E1a, E1b, E2
chalk	bulletin	board	floor	E1a, E1b, E2
cob	husk	corn	frog	E1a, E1b, E2
cod	trout	fish	journal	E1a, E1b, E2
crib	diaper	baby	banquet	E1a, E1b, E2
donor	plasma	blood	winter	E1a, E1b, E2
exam	quiz	test	walk	E1a, E1b, E2
film	cinema	movie	poetry	E1a, E1b, E2
frame	portrait	picture	shingle	E1a, E1b, E2
gums	braces	teeth	volcano	E1a, E1b, E2
handbag	pocketbook	purse	teeth	E1a, E1b, E2
hanger	wardrobe	clothes	foam	E1a, E1b, E2
hive	buzz	bee	birth	E1a, E1b, E2
hog	pork	pig	tank	E1a, E1b, E2
icing	frosting	cake	dancer	E1a, E1b, E2
instructor	professor	teacher	train	E1a, E1b, E2
jacket	mink	coat	forest	E1a, E1b, E2
juice	tangerine	orange	rabbit	E1a, E1b, E2
knob	hinge	door	hospital	E1a, E1b, E2
marrow	skeleton	bone	continent	E1a, E1b, E2
nest	canary	bird	bottle	E1a, E1b, E2
nurse	physician	doctor	hall	E1a, E1b, E2
occupation	career	job	letter	E1a, E1b, E2
pal	buddy	friend	knife	E1a, E1b, E2
pane	sill	window	wine	E1a, E1b, E2
pen	eraser	pencil	salad	E1a, E1b, E2
petals	tulip	flower	key	E1a, E1b, E2
pistol	trigger	gun	lace	E1a, E1b, E2
pony	saddle	horse	ladder	E1a, E1b, E2
sail	yacht	boat	clock	E1a, E1b, E2
slither	serpent	snake	tower	E1a, E1b, E2
sock	sneaker	shoe	tobacco	E1a, E1b, E2
stone	boulder	rock	text	E1a, E1b, E2
table	seat	chair	flood	E1a, E1b, E2
tin	opener	can	fence	E1a, E1b, E2
yolk	omelet	egg	ink	E1a, E1b, E2

Note: average forward strength for cues to targets was .05, for cues to mediators was .67, and for mediators to targets was .05.

<u>Nonwords used in Experiment 1a, 1b, and 2</u>
shrumped
shround
rop
grourn
clett
snurfs
steaves
twarked
phrinsed
swaught
shroons
phleague
blowns
stroobs
wofts
whols
flince
spirped
stilch
shourned
clulched
fusk
phrup
vapse
fenth
trebe
tarb
crolt
croiced
gwoints
streeved
soys
brepth
flane
plail
skoal
micked
ghumped
phlands
swoists
yusks
wumps

Words used in Experiment 3a and 3b

Experiment 3a

List 1

Cue	**Target**
bee	hive
bus	driver
fight	box
fish	sea
fork	lift
gem	jewel
job	money
map	direction
meal	lunch
paste	tooth
salt	water
skull	skeleton
string	kite
team	sport
test	grade
yarn	thread

List 2

Cue	**Target**
book	school
face	nose
germ	bacteria
globe	earth
king	ruler
leg	foot
lemon	orange
lens	eye
petal	rose
police	law
river	stream
room	house
sand	castle
seal	envelope
skunk	stink
wrench	hammer

Experiment 3b

List1

Target
absent
blank
elephant
note
faint
matter
tangle
wood
dragon
client
lock
kill
purple
queen
trouble
year

List 2

Target
antenna
east
list
opera
pity
waste
coach
diamond
gasp
castle
house
interest
probe
knife
last
united

Note: average forward strength for cues to targets in Experiment 3a was .10.

Words used in Experiments 4 and 5

Cue	**Associates**
book	read, worm, school*, study, learn, novel
bus	car, driver*, school, stop, ride, people
cloud	sky, rain*, white, nine, soft, fluffy
face	eyes, nose*, smile, person, head, pretty
fight	fist, box*, hurt, war, argue, win
fish	water, swim, sea*, scales, smell, cod
fork	spoon, knife, eat, lift*, food, stab
gem	diamond, jewel*, stone, ruby, pearl, ring
germ	disease, bacteria*, sick, virus, wheat, infection
globe	world, round, earth*, atlas, ball, circle
insect	bug, bite*, fly, mosquito, ant, bee
job	work, money*, occupation, career, employment, pay
king	queen, monarch, ruler*, kong, crown, powerful
lagoon	water, blue, lake*, swamp, pond, island
lava	volcano, hot, lamp*, rock, molten, soap
ledge	cliff, window*, edge, fall, jump, building
leg	arm, foot*, walk, body, knee, long
lemon	lime, sour, orange*, tree, fruit, yellow
map	road, direction*, travel, world, globe, lost
passenger	plane, car*, driver, train, bus, ride
police	cop, help, law*, car, pig, man
reptile	snake, lizard*, alligator, scales, animal, frog
river	lake, stream*, water, flow, boat, canoe
salt	pepper, water*, sugar, bitter, food, ocean
sand	beach, castle*, dirt, dune, ocean, pebble
seal	animal, envelope*, close, stamp, water, ball
string	rope, kite*, thread, guitar, ball, cord
team	football, sport*, group, together, baseball, effort
throat	neck, sore, mouth*, swallow, tonsil, voice
wrench	tool, hammer*, screwdriver, pliers, pull, fix

*target words

Note: average forward strength for cues to associates was .11.