Journal of Memory and Language 94 (2017) 103–118

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Evaluating mechanisms of proactive facilitation in cued recall

William R. Aue *, Amy H. Criss, Matthew D. Novak

Department of Psychology, Syracuse University, Syracuse, NY 13244, USA

ARTICLE INFO

Article history:
Received 18 April 2016
revision received 8 October 2016

Keywords:
Cued recall
Episodic memory
Memory models
List discrimination
Associative memory
Proactive facilitation

ABSTRACT

Confusion of older information with newer, similar information is a potent source of memory errors. The current project focused on understanding how memories for recent experiences interact, or interfere, with other related information. In the experiments, participants study multiple lists of pairs of items. Items from an initial study list (e.g., A-B) reappear on a second study list paired with new, other items (e.g., A-Br). Memory performance for A-Br pairs is contrasted with control pairs exclusive to the second study list (e.g., C-D). We observed that the correct recall of the second presentation of a target (Br) is better when cued by its partner (A) despite being studied with a different partner during the initial presentation; a phenomena called proactive facilitation. We examined multiple possible explanations for proactive facilitation, including whether proactive facilitation was driven by changes in response threshold, whether participants encoded the pairs with repeated items and associations better during the second study list, or whether participants spent more time searching memory for A-Br pairs. In general, the data appear to be most consistent with the idea that some items, when encountered a second time, are encoded more completely while others are not. Implications for models of memory are discussed.

Memory represents a connection to a person's past. Whether it is memories for a lifetime of experiences (e.g., birthdays, vacations) or a vast amount of acquired knowledge, without access to such information we would be lost in the world. Accordingly, much research has focused on understanding the circumstances under which memories are acquired, lost, or updated. For example, memory updating that occurs during testing has been the source of much interest. Factors that help memory, such as the testing effect (Karpicke, Lehman, & Aue, 2014; Roediger & Karpicke, 2006a, 2006b), and that harm memory for other items such as output interference (Criss, Malmberg, & Shiffrin, 2011; Murdock & Anderson, 1975; Ratcliff & Hockley, 1980; Roediger, 1974; Roediger & Schmidt, 1980) have been of particular interest. In the current manuscript, we focus on the memory updating that occurs during study. In particular, we examine updating that occurs when encountering information that contradicts what has been recently studied.

It is well understood that memories of similar or related information can interfere with one another (e.g., Barnes & Underwood, 1959; Burns, 1989; Postman & Gray, 1977; Postman & Keppel, 1977; Postman & Stark, 1964; Postman, Stark, & Burns, 1974; Underwood, 1949, 1957, but see Anderson & Neely, 1996; Postman & Underwood, 1973 for reviews), making them difficult to retrieve. For example, placing your keys in a different spot each day when you arrive home, may make it difficult to recall where your keys were last laid. The problem is that the item used to cue the memory search (i.e., keys) is associated with multiple target locations (e.g., key dish, entryway, kitchen table, coat pocket). The difficulty lies in isolating the most recent location of the keys in the midst of multiple associated locations. The observation that

* Corresponding author at: Purdue University, Department of Psychological Sciences, West Lafayette, IN 47907, USA.

E-mail address: william.aue@gmail.com (W.R. Aue).

http://dx.doi.org/10.1016/j.jml.2016.10.004
0749-596X/

memories for previous events interfere with newer memories has been established experimentally and is called proactive interference.

Cued recall is a task commonly used to measure the impact of such interference on memory. During cued recall participants must retrieve a target memory when given a cue with which to search. For example, retrieving a person's name when given a face, or retrieving a significant event (e.g., an anniversary) when cued with the date. The task is particularly useful for understanding memory because memory is cue driven and cued recall provides a straightforward way to manipulate the cue.

In the lab, a participant in a typical cued recall task studies a list of pairs of items (e.g., word-name pairs). Afterward, the participants are asked to retrieve one member of the pair (i.e., the target) when given the other member of the pair, such as a face (i.e., the cue). We induce interference by manipulating the number of times and the number of pairings in which individual items appear during study. A participant might study one list of word-face pairs and then a little later study a second list of word-face pairs where some of the pairs have been rearranged such that a particular face is now presented with a different word. As one might expect, performance is worse for items that are associated with many other items such as a face associated with multiple names (e.g., Anderson, 1974; Weinstein, McDermott, & Szpunar, 2011; see Table 1 for examples of different pair types employed in these designs). More recently, studies have reported an improvement in performance in typical interference designs. For example, Aue, Criss, and Fischetti (2012) observed that associating a cue (e.g., face) to multiple targets (e.g., words) facilitated recall for some items. This effect, termed proactive facilitation, has also been observed elsewhere in a variety of paradigms (Jacoby, Wahlheim, & Yonelinas, 2015; Putnam, Wahlheim, & Jacoby, 2014; Wahlheim & Jacoby, 2013).

The specific purpose of the current project was to test whether mechanisms in a model of cued could explain proactive facilitation. We accomplish this by identifying and testing potential mechanisms underlying proactive facilitation in cued recall. We begin with an overview of the research in proactive interference and proactive facilitation, followed by an evaluation of three potential mechanisms for proactive facilitation that were proposed in Aue et al. (2012). We then present the data from five experiments designed to tease apart the explanations of proactive facilitation.

Proactive interference & facilitation

Aue et al. (2012) observed both proactive interference and proactive facilitation during cued recall using the study design shown in Fig. 1. In their experiment, participants studied a list of word-face pairs during an initial incidental study list. After a short break participants studied a second list wherein half of the pairs were comprised of

Table 1
Different types of pairs employed in interference designs.

Prior list study		Potential critical list pairs			
Cue	Target	Pair type	Cue	Target	Description
Absence	Hollow	A-B, A-B	Absence	Hollow	Both cue and target are repeated together
Pupil	River	A-B, A-B′	Absence	Empty	Repeated cue, semantically similar target
		A-B, A-Br	Absence	River	Both cue and target are repeated but in different pairings
		A-B, A-D	Absence	Tissue	Repeated cue, new target
		A-B, C-B	Pillar	Hollow	New cue, repeated target
		A-B, C-D	Pillar	Tissue	New cue, new target

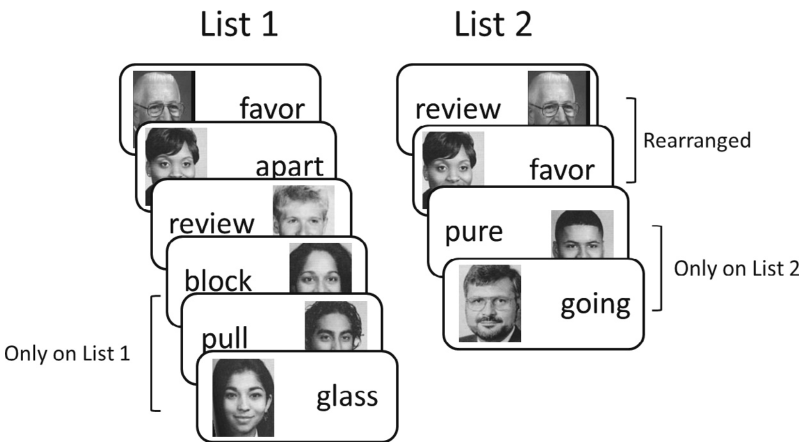

Fig. 1. The study design from Aue et al. (2012; Experiment 1). Rearranged pairs correspond to A-Br pairs, List 2 only pairs correspond to C-D pairs. Figure taken from Aue et al. (2012).

items from the first list that had been rearranged into new pairs (i.e., A-B, A-Br, herein referred to as A-Br pairs; see Table 1; Twedt & Underwood, 1959). The other pairs on the list were items that were only on List 2 (i.e., A-B, C-D, herein referred to as C-D pairs). Participants were then tested for the most recent list using cued recall. For each test cue, a response could be correct, incorrect, or the participant could indicate not remembering the target (called 'no recall' here). Aue et al. observed significantly more correct responses for A-Br cues relative to C-D cues, and more incorrect responses for A-Br cues relative to C-D cues. Simply put, participants demonstrated *both* proactive interference and proactive facilitation, similar to a pattern of data observed by Criss and Shiffrin (2005) for associative recognition.

Wahlheim and Jacoby (2013) observed proactive facilitation using an A-B, A-D design, also with cued recall. Wahlheim and Jacoby (2013) had participants study two lists where a cue is associated with different target responses on each list (e.g., A-B, A-D) and target responses were unique to their respective lists. They also employed a C-D control condition manipulated within list. At test, Wahlheim and Jacoby (2013) asked participants to recall the D item given A as a cue. Participants were also asked to self-report if another item came to mind first and, if that happened, to report what they remembered. Wahlheim and Jacoby found that the recall of D in an A-D pair was greater when they conditionalized responses based on whether participants reported B (from the initial A-B pair) coming to mind *prior* to the recall of D. When participants reported no item coming to mind before D, A-D pairs were recalled worse relative to the control C-D pair. Their results reflect a mixture of proactive interference and facilitation. Likewise, Burton, Lek, and Caplan (2017), employing a similar design, observed a positive correlation between recall of B and D given A as a cue. Postman and Gray (1977) also observed that when given A while having studied A-B, A-D, participants were most successful at retrieving D if they first recalled B. These results, in different experimental designs, provide corroborating evidence of the persistence of proactive facilitation across a variety of paradigms.

Possible explanations of proactive facilitation

Wahlheim and Jacoby (2013) proposed that proactive facilitation in the A-B, A-D design is driven by participants detecting that the A-B pairing switched to A-D during study of List 2. Wahlheim and Jacoby contend that the detection of the change results in the recursive storage of the List 1 representation in the memory for List 2. At test, proactive facilitation is purported to occur if the participant reports another response coming to mind *prior* to the List 2 response. In theory, the List 1 response coming to mind first reflects the recollection of the change during List 2. How this theory could be integrated into extant computational models of memory is unclear. How memory for List 2 would be represented is also unclear. For example, if the memory for List 1 is embedded in the memory for List 2, then one interpretation is that participants would have to access the List 2 memory in order to access the associated List 1 response; the opposite of Wahlheim and Jacoby's observations. Additionally, while it may be the case that those responses reflecting proactive facilitation are the result of change recollection, such recollection is not an explanation for how or why a memory search leads to the recollection of some changes but not others.

In the current manuscript, we sought to further advance theorizing of proactive facilitation by focusing on basic processes that may drive the phenomenon. Aue et al. (2012) proposed possible explanations for proactive facilitation from the perspective of basic memory mechanisms; ideas that we flesh out here. The ideas were couched in the conceptual framework of the retrieving effectively from memory model (REM) model of cued recall (Diller, Nobel, & Shiffrin, 2001; Nobel & Shiffrin, 2001), with the goal of developing the broader theoretical framework of REM to encompass proactive facilitation. We considered three potential explanations for the data in Aue et al. (2012) based, in part, on mechanisms in REM: (1) a change in response threshold for the target item, (2) A-Br pairs are better encoded during List 2, or (3) participants are spending longer searching memory in response to familiar A-Br test cues. Importantly the REM framework provides a comprehensive account of memory across multiple memory tasks and manipulations. Directly linking explanations for proactive facilitation to a formal model of memory is essential for developing an integrative model of human memory.

Differences in response threshold

Given the pattern of data observed in Aue et al. (2012) wherein more responses (both correct and incorrect) were provided for A-Br pairs relative to C-D pairs, one possibility is that participants are changing the threshold of what responses to output based on the familiarity of the cue. The cues for A-Br pairs have been seen twice across lists and as a result, should be more familiar relative to the cues for C-D pairs that have been only seen once. Some models of memory (including REM, e.g., Malmberg & Shiffrin, 2005) assume that participants monitor the quality of a response before outputting it. If the quality of the response exceeds the response threshold, the response is output otherwise it is withheld and no response is provided (e.g., Koriat & Goldsmith, 1994, 1996). Cue strength may inform the decision about the required threshold for outputting a response. A familiar cue provides a feeling that the participant should know the target. This feeling of familiarity leads to a reduction in the threshold for outputting a response and a lower quality response. Cue quality drives other aspects of metacognition such as feeling of knowing (e.g., Eakin, 2005; Metcalfe, Schwartz, & Joaquim, 1993; Schwartz & Metcalfe, 1992), retrieval strategy (e.g., Reder, 1987), and willingness to provide a response (Hanczakowski, Pasek, Zawadzka, & Mazzoni, 2013), and is consistent with assumptions in models of cued recall (e.g., Diller et al., 2001).

A response threshold explanation could lead to participants being more or less willing to respond depending on the familiarity of the cue. For example, participants may be more willing to output a target in response to a more familiar cue, or less willing to output a target given a less familiar cue. The result would be an overall change in the number of responses for the cue type, impacting both incorrect and correct responses and contributing to proactive interference and facilitation. In Experiment 1 and 2 we employ forced recall, requiring a response on each trial (Koriat & Goldsmith, 1994, 1996) to circumvent response thresholds. We also examine changes in the quality of the responses by measuring memory confidence.

Better encoding for A-Br pairs during List 2

Another possible explanation for proactive facilitation discussed in Aue et al. (2012) is that pairs containing repeated, but rearranged, items were better encoded the second time they were encountered. Currently, REM (Criss & Shiffrin, 2005) assumes that the A-Br pairs studied in List 2 would be stored as a new trace, with the same fidelity as List 1 pairs (i.e., A-B). However, perhaps it is the case that when the A-Br items appear on List 2, participants update the existing representation of the *item* in episodic memory (i.e., A-B) to create the A-Br pairs. This assumes that participants are able to isolate the constituent *items* that were studied in the context of a pair during List 1. This updating, of sorts, could benefit both item information (i.e., memories for the individual items) and associative information (i.e., information indicating that two items were studied *together*).

Others have posited updating mechanisms for a type of proactive facilitation observed for free recall with an A-B, A-D study design (e.g., Burns, 1989; Hirshman, Burns, & Kuo, 1993). They suggest that when participants study A-B followed later by A-D, they update the existing memory for A and associate it to the new trace for D. As a result, the pair is represented by three elements in memory (i.e., A, B, & D). The facilitation for free recall of D is the result of A-D pairs having a shorter "list" relative to an A-B, C-D study design, where four elements are stored in memory (i.e., one each for A, B, C, D) or more retrieval routes (both A and B lead to retrieval of D). Hintzman's (2004) recursive reminding is a similar idea that has been applied to proactive facilitation (Jacoby, Wahlheim, & Kelly, 2015; Putnam et al., 2014; Wahlheim & Jacoby, 2013) by suggesting that when A-B, A-D are studied and a change is detected, the List 1 memory is embedded in a recursive representation of the List 2 memory. At test, the recall of B given A facilitates the correct recall of D. One caveat is that better encoding might explain proactive facilitation, but cannot account for proactive interference. If we find evidence for enhanced encoding of List 2 pairs, then we must also assume another process or processes to explain the full pattern of data.

In Experiment 3, we examine whether proactive facilitation results from better encoding *during* List 2 by examining memory for List 1 items. The design is the same as shown in Fig. 1 except at test participants are asked to recall the associations from List 1 (i.e, A-B). If the benefit for A-Br pairs is the result of better encoding of List 2 pairings, then memory for List 1 should not benefit from the subsequent encoding of List 2.

In Experiment 4 we added a longer retention interval. If participants are encoding A-Br pairs *better* than C-D pairs then the advantage for correct responses should persist over the delay.

Longer memory searches for A-Br pairs

The third possible explanation that we consider is that participants spend more time searching memory for A-Br pairs than C-D pairs. The idea is that search duration could change for a motivated searcher. Searchers are generally thought to be motivated based on evidence of longer response times to errors than correct responses (e.g., MacLeod & Nelson, 1984; Millward, 1964). We suggest that motivation to search is a cost-benefit analysis. For example, a person would likely spend more time searching for money that had been dropped if the value was higher (e.g., $50) than lower (e.g., $0.01). Likewise, the familiarity of the cue could influence search time with participants spending additional time interrogating memory for a familiar cue that is likely to 'pay out.' The extended search time would provide additional opportunities to output any response, either correct or incorrect. The idea that familiarity could drive search duration is supported by findings of Nelson, Gerler, and Narens (1984) who noted a positive correlation between a participant's rated feeling of knowing and the latency at which they decided to terminate unsuccessful searches. Searches tended to go on longer when participants felt they knew the answer. Further, the idea that search time is a function of cue familiarity is explicitly assumed in a model of cued recall (Diller et al., 2001). In Experiment 5, we look at differences in response times to A-Br and C-D pairs as a measure of search duration.

Examining response thresholds

The first idea that we considered is that proactive facilitation is caused by a change in the response threshold. Participants could change the response threshold based on cue familiarity. A-Br cues are more familiar than C-D cues by virtue of having appeared on both List 1 and List 2. One possibility is that participants were outputting responses they otherwise would have withheld for A-Br pairs, driving up both correct and incorrect responses. Alternatively, participants could have been withholding responses to C-D pairs, attenuating responding for both correct and incorrect responses. In the following two experiments, we tested these ideas using two approaches. First, we employed instructional manipulations to encourage participants to eliminate the influence of response threshold (Experiment 1) or to set a higher response threshold (Experiment 2). If participants are withholding C-D responses that would otherwise be provided, then forcing a response (Experiment 1) should boost performance for C-D pairs. If participants were providing A-Br responses that would have otherwise been withheld, then asking participants to adopt a more strict response threshold (Experiment 2) should have attenuated the effect.

As a second metric for evaluating the quality of response being provided for Experiments 1 and 2, we also collected retrospective confidence ratings. After providing a response, participants rated their confidence in the veracity of the response. If participants were providing lower quality responses for A-Br pairs, then they would be associated with a lower confidence rating relative to the C-D pairs.

General method

Materials

All of the experiments involve the study of word-face pairs. The face stimuli used are the same as those described in Aue et al. (2012) and Criss and Shiffrin (2005). A total of 210 faces were used. Faces were standardized for orientation position of facial landmarks. The words used in the experiments were 800 high-frequency words ($M = 130.66$ from Kučera and Francis (1967) or, alternatively, log frequency $M = 10.46$ in the Hyperspace Analog to Language corpus of Balota et al. (2007) and Lund and Burgess (1996)). These materials were used in all experiments.

Design

All of the experiments followed the same general design and procedure that is described in this section, with the critical changes described in each experiment's respective section. Participants completed the experiment individually on a Windows-based computer running Authorware (v. 7.0). The experiment was a within-subject design with two conditions (i.e., A-Br vs. C-D pairs). Participants studied the two lists of word-face pairs followed by a cued recall test. The experiment lasted approximately 30 min.

List 1 and List 2 each contained 24 word-face pairs. List 2 was comprised of 12 unique pairs (C-D pairs) plus 12 pairs of items rearranged from List 1 (A-Br pairs). The pairing of A-B and C-D pairs and the re-pairing for A-Br pairs was randomly chosen for each participant. Pairs were presented with the items side-by-side for 3 s, with the left/right position of the face and word randomly chosen on each trial for each participant. After 3 s, the pair disappeared and participants were prompted to answer a question about the pair. Encoding was incidental. An encoding task was used wherein, following the presentation of a pair, participants answered a question about the pair. Different encoding tasks were used for List 1 and List 2. The tasks are the same as those described in Aue et al. (2012) and Criss and Shiffrin (2005). No differences have been observed for cued recall performance when encoding tasks were switched (Criss, Aue, & Smith, 2011). For the List 1 encoding task (A-B pairs), participants were asked to rate the degree of association between the items on a 9-point scale (1 – *not at all associated* to 9 – *highly associated*). Upon completion of List 1, participants read a comic for 60s. For the List 2 encoding task, participants were asked to generate a sentence about each pair and then rate how difficult it was to do so using a 9-point scale (1 – *very easy* to 9 – *very difficult*). After the second study list, participants completed a running summation task for 60s.

The cued recall test immediately followed the distractor task. During the test, faces from List 2 appeared one at a time. Participants were asked to type the word that the presented face was paired with on the most recent list (i.e., List 2). The order of the faces was randomized anew for each participant, and the test was self-paced. Participants' responses were coded as correct, incorrect, or no response. A correct response consisted of a response that was either the target word or contained minor errors such as misspellings (e.g., braec instead of brace) or added suffixes (e.g., walked instead of walk). All other responses were coded as intrusions. With the exception of Experiment 1, participants were permitted to indicate they did not recall the target for a given cue and these were coded as No Recall responses.

Analysis plan

Unless otherwise specified, data were analyzed using the R programming language (R Core Team, 2013) and the "EZ" package (Lawrence, 2013) for ANOVAs, the "MBESS" package (Kelly & Lai, 2015) for effect sizes and confidence intervals, and the "BayesFactor" package (Morey, Rouder, & Jamil, 2015) for calculating the Bayes Factor (BF_{10}). The Bayes Factor represents the level of evidence supporting an alternative model with a non-zero effect size, relative to a null model with a zero effect size. Values larger than 1 represent evidence in favor of the alternative model, whereas values less than 1 represent evidence in favor of the null model. Sample size was determined by approximating that of Aue et al. (2012).

Experiment 1: Forced-report cued recall

In Experiment 1, we employed a forced report procedure. The testing scenario was the same as in Aue et al. (2012); however, participants were required to provide a response to each cue before proceeding. This effectively eliminated a response threshold from affecting performance (Koriat & Goldsmith, 1994, 1996) by eliciting sub-threshold responses that would have otherwise been withheld. If participants were setting a more conservative response threshold for C-D pairs, reducing both correct and incorrect responses, then performance for the two conditions should have been equated under forced report.

Experiment 1 method

Participants

In total, 40 participants from the Syracuse University subject pool participated in the experiment for course credit. All participants were included in the analyses. All efforts were made to ensure that samples for each experiment were independent. Our sample contained both native and non-native English speakers. The Syracuse University Institutional Review Board approved all study protocols, and all participants provided informed consent. This is true of all subsequent experiments.

Design

Participants were instructed to provide a response to each face cue, even if doing so required guessing. Occasionally a participant would disregard the instructions and enter a response that indicated they did not recall the cue for a given target (e.g., "no", "idk"). These were coded as No Recall responses and were infrequent.[1] After each recall attempt, participants were asked to report their confidence that the word they just recalled was accurate. Reporting was done on a 6-point scale (*1 – I am sure it is wrong* to *6 – I am sure it is correct*).

Results & discussion

Memory performance

We performed a two-sided paired *t*-test on the proportion of correct responses provided for the A-Br versus C-D conditions. As can be seen in Fig. 2, participants provided significantly more correct responses to A-Br pairs ($M = .285$, $SE = .036$) than C-D pairs ($M = .185$, $SE = .02$; $t(39) = 3.57$, $p = .001$, $d = .798$, 95% CI = [.34, 1.25], $BF_{10} = 27.11$), demonstrating the same pattern as the Aue et al. (2012) data. Forcing participants to respond did not improve performance for C-D pairs relative to A-Br pairs. The persistence of the A-Br advantage, is consistent with experiments investigating memory performance under forced versus free recall (e.g., Roediger & Payne, 1985) where forcing participants to respond only tends to increase the number of incorrect responses. In this case, forcing participants to respond did not improve performance for C-D pairs to the level of A-Br pairs, indicating that participants are well calibrated (e.g., Koriat & Goldsmith, 1996).

Memory confidence

Next we examined confidence ratings across conditions by averaging responses within subject separately for correct and incorrect responses. If participants were withholding responses for C-D pairs by setting a higher response threshold, one might expect higher levels of confidence in the responses provided. A portion of participants only provided correct responses, and thus confidence ratings, for A-Br pairs, but not C-D pairs, or vice versa. As a result, we performed an independent samples *t*-test for correct responses.[2] There was no difference in confidence for correct responses of A-Br pairs ($M = 4.61$, $SE = .186$) and C-D pairs ($M = 4.33$, $SE = .190$; $t(67.97) = 1.02$, $p = .313$, $d = .243$, 95% CI = [−.228, .712], $BF_{10} = .291$). However, the results of a paired *t*-test revealed that participants tended to be more confident for intrusions for A-Br pairs ($M = 2.23$, $SE = .130$) than for C-D pairs ($M = 1.83$, $SE = .092$, $t(39) = 3.36$, $p = .002$, $d = .751$, 95% CI = [.295, 1.20], $BF_{10} = 15.68$), which is expected given the interference design.

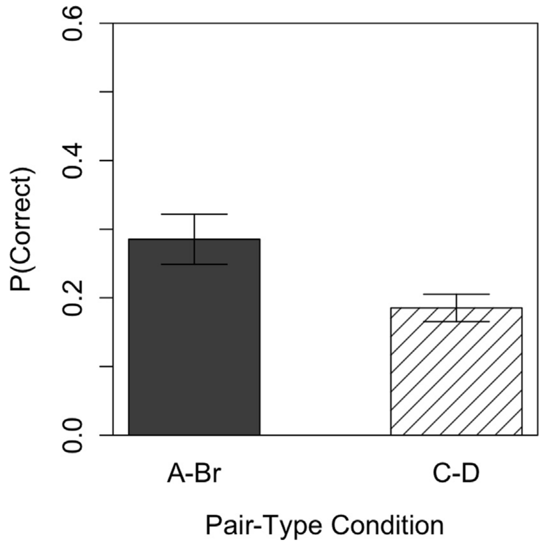

Fig. 2. Correct responses as a function of pair condition for Experiment 1. Participants had more correct responses for A-Br pairs relative to C-D pairs. Thus forcing participants to respond did not improve performance for C-D pairs indicating that a conservative response bias is not at play. Error bars represent ± 1 standard error.

The advantage of A-Br over C-D pairs for correct responses did not change when participants were forced to respond, nor were there differences in mean memory confidence for correct responses. Based on these data, we suggest that participants were not withholding responses for C-D pairs and provided similar quality responses for both types of pairs. A conservative response threshold for C-D pairs is not a likely explanation for proactive facilitation.

Experiment 2: High confidence responses

Whereas in Experiment 1 we attempted to induce the most liberal response threshold by requiring a response on every trial, here we do the opposite. The aim of the current experiment was to induce participants to adopt a conservative response threshold by asking participants to restrict responses to high confidence responses. If participants set liberal thresholds for responding to a familiar A-Br cue, then we should be able to eliminate the difference between A-Br and C-D cues with these instructions.

Method

Participants

40 Syracuse University undergraduates participated in the experiment. Participants received course credit for their participation.

Materials

The materials were identical to those used in Experiment 1.

[1] No recall responses occurred on $M = .002$ of trials for A-Br pairs and $M = .004$ of trials for C-D pairs.

[2] We chose the independent samples *t*-test to avoid dropping subjects with incomplete data. We repeated the analysis with a paired *t*-test by dropping participants who did not provide at least one response type (i.e., correct, incorrect) for each pair type condition (i.e., A-Br, C-D). The conclusions do not change. Participants were equally confident in their correct responses for A-Br and C-D pairs ($t(31) = 0.567$, $p = .575$) and participants were more confident in their incorrect responses for A-Br pairs relative to the C-D pairs ($t(31) = 3.97$, $p < .001$).

Design

The details were the same as those described in Experiment 1 with a minor modification to the testing procedure. At test, participants were provided with a cue and asked to recall the target it was studied with on the most recent list (i.e., List 2). Additionally, participants were instructed only to respond if they felt highly confident that they were correct, and if they did not feel a high level of confidence they should indicate that they did not know the answer (i.e., withhold their response). Afterward, participants rated confidence in the same manner as Experiment 1.

Results & discussion

As a manipulation check, we compared average confidence ratings for correct responses collapsed across conditions in Experiment 1 and Experiment 2. Participants in Experiment 2 tended to be more confident in their correct responses ($M = 5.04$, $SE = .12$) than participants in Experiment 1 ($M = 4.55$, $SE = .14$; $t(70.89) = 5.45$, $p < .001$, $d = 1.22$, 95% CI = [.737, 1.69], $BF_{10} = 25401.4$) indicating that they attempted to follow the instructions. Additionally, we examined the difference in correct responses in Experiment 1 to Experiment 2 using a 2 (Pair-type) × 2 (Experiment) mixed ANOVA. We observed a main effect of experimental condition where A-Br pairs ($M = .257$) were recalled better on average relative to C-D pairs ($M = .18$; $F(1,78) = 5.53$, $p = .021$, $\eta_p^2 = .07$), but no main effect of experiment ($F(1,78) = 1.14$, $p = .289$, $\eta_p^2 = .01$), nor a condition by experiment interaction ($F(1,78) = 1.37$, $p = .246$, $\eta_p^2 = .02$). Thus despite attempting to follow instructions (at least as judged by confidence ratings), accuracy in responding did not change compared to Experiment 1. We now turn to an analysis of Experiment 2 data.

Memory performance

A paired t-test revealed that participants provided more correct responses to A-Br pairs ($M = .229$, $SE = .024$) relative to C-D pairs ($M = .175$, $SE = .02$; $t(39) = 1.97$, $p = .055$, $d = .441$, 95% CI = [−.004, .884], $BF_{10} = 0.76$).[3] In spite of the marginal significance, we suggest the difference is meaningful given the fact that it replicates the pattern of several experiments in Aue et al. (2012) and several experiments in this paper. The same pattern was observed for incorrect responses to A-Br pairs ($M = .298$, $SE = .026$) relative to C-D pairs ($M = .229$, $SE = .028$; $t(39) = 2.53$, $p = .016$, $d = .566$, 95% CI = [.117, 1.01], $BF_{10} = 2.25$), replicating the pattern observed in Aue et al. (2012).

Memory confidence

The confidence data for Experiment 2 was similar to the data from Experiment 1. We performed an independent samples t-test on confidence ratings for correct responses.[4] The test revealed that participants did not differ in their confidence of correct responses for A-Br pairs ($M = 4.94$, $SE = .144$) relative to C-D pairs ($M = 5.10$, $SE = .177$; $t(64.51) = -.685$, $p = .496$, $d = -.164$, 95% CI = [−.633, .306], $BF_{10} = 0.23$) but were more confident in their intrusions to A-Br pairs ($M = 3.91$, $SE = .169$) relative to C-D pairs ($M = 3.26$, $SE = .204$; $t(68.04) = 2.46$, $p = .017$, $d = .572$, 95% CI = [.105, 1.04], $BF_{10} = 2.74$). Having participants set a stricter threshold for responding did not influence proactive facilitation. Thus, there is little evidence that participants are setting a more liberal criterion for A-Br pairs relative to C-D pairs.

Response threshold discussion

The data from the first two experiments indicate that differences in response threshold were not a viable explanation for the proactive facilitation. For Experiment 1, correct A-Br responses were higher than C-D responses even under forced report. We suggest this indicates that participants were not adopting a conservative response threshold for C-D pairs, nor were there any differences in the retrospective confidence ratings for the correct response. Thus, C-D pairs were simply less well remembered than A-Br pairs and were not being attenuated by a conservative response threshold. With respect to Experiment 2, proactive facilitation was observed despite asking participants to respond only with higher confidence responses. If participants were adopting a more liberal response threshold and, as a result, outputting more, lower quality responses, then asking participants to withhold low-quality responses should have selectively attenuated A-Br performance. This was not the case, despite the fact that participants appeared to follow instructions. That is, they tended to be more confident in their retrospective confidence ratings relative to Experiment 1. Taken together, we suggest there is no support for a response threshold explanation in either the proportion of responses provided or in the retrospective quality of the response.

Examining a better encoding explanation

Next, we examined the idea that participants were drawing on recent experience with individual items to store more accurate versions of the A-Br pairs studied on List 2. The general idea is that participants were better able to draw associations, make connections, and encode pairs for more familiar items. Thus, when seeing the second presentation of an individual item, it was more familiar and the resulting storage was more complete. In other words, the representation for A during the study A-B would have less information stored, relative to the representation of

[3] Given the marginal nature of the results, we performed a randomized permutation test to examine whether our observed data are reliably different from what one might expect from random noise (Nichols & Holmes, 2002). We simulated an experiment with null results by randomly reassigning the data within participant to a pair-type condition. We then repeated this process 50,000 times to create a null distribution. We find that the observed difference in A-Br and C-D performance for Experiment 2 is unlikely when the null hypothesis is true and falls beyond the 95% Highest Density Interval for the distribution: ($M = .054$, two-tailed $p = .029$, 95% HDI [−.054, .046]).

[4] As with Experiment 1, the pattern of data did not change when the data were analyzed by dropping participants using a paired t-test. There was no difference in correct responses ($t(29) = -0.543$, $p = .591$), and the difference in incorrect responses was significant ($t(29) = 2.67$, $p = .012$).

A during the study of A-Br which integrates and updates the representation for A from the initial presentation. Here we looked for indirect effects of better encoding of the List 2 pairs using multiple approaches.

Experiment 3: Test of List 1 associations

If better encoding were occurring *during* List 2, when the individual items are presented a second time in a different pairing and the new pair was stored as a separate memory trace from the original presentation, then the proactive facilitation benefit should not extend to the memory for the pairs studied during List 1. In this retroactive design, list being tested shifts from the second list to the first list. This design allowed us to examine the influence of subsequent learning on previously learned information. The study scenario is identical to Aue et al. (2012; Experiment 1) and of Experiments 1 and 2 in the current paper. However, at test participants were provided with a cue and asked to recall the target with which the item was first studied. The primary question of interest was whether or not proactive facilitation, or in this case retroactive facilitation, was present for first list associations.[5] With respect to the proposed hypotheses, both response bias and longer searches predict more responses overall for A-Br pairs because the cue provided at test is more familiar and would induce either longer searches or a shift in the response threshold. However, if the pairs were being encoded more completely during the *second* presentation and stored in a separate trace from the List 1 pairs, then asking participants to recall the List 1 associate should have eliminated the effect for correct responses.

Method

Participants

A total of 57[6] Syracuse University subject pool members participated in the experiment. Participants received course credit for their participation.

Materials

The materials were identical to previous experiments.

Design

The details were the same as those described in Experiment 2. The pairs were created in a manner identical to previous experiment. The only changes were at test. No instructional manipulation was performed, no confidence ratings were collected, and participants were tested for their memory for the first list. At test, participants were shown a face and asked to recall the word that it was studied with during the *very first* list.

Results and discussion

A paired *t*-test was performed to compare A-Br and C-D pairs for correct and incorrect responses. As can be seen in Fig. 3, participants tended to respond correctly more often to A-Br pairs ($M = .146$, $SE = .017$) relative to C-D pairs ($M = .105$, $SE = .014$; $t(56) = 2.38$, $p = .021$, $d = .445$, 95% CI = [.073, .816], $BF_{10} = 1.47$). Participants also responded incorrectly more often to A-Br pairs ($M = .377$, $SE = .029$) relative to C-D pairs ($M = .209$, $SE = .024$; $t(56) = 6.51$, $p < .001$, $d = 1.22$, 95% CI = [.816, 1.62], $BF_{10} = 565{,}803$), a pattern similar to proactive facilitation. Moreover, this corroborates findings by Jacoby, Wahlheim, and Kelly (2015) who also observed evidence of both retroactive and proactive facilitation.

If A-B and A-Br are stored as separate events in memory, then based on these data we suggest that proactive facilitation is not explained by storing a more complete representation for items during List 2. If the proactive facilitation observed in Aue et al. and other studies was the result of betting encoding of List 2 items, then it should not have persisted for List 1 associations given that any additional encoding would occur during the List 2 presentation. However, the data could potentially be explained by longer memory searches given that the cue was more familiar, as we will discuss later.

Experiment 4: Expanding the retention interval

In this experiment, we attempted to put the familiarity of A-Br cues and C-D cues on closer to equal footing by adding a 16 min retention interval between List 2 and final test. In this scenario, if cue familiarity were the *only* feature driving proactive facilitation then the difference between A-Br and C-D pairs should be attenuated. However, if proactive facilitation were the result of better encoding of the pairs during List 2 then the effect should have persisted.

Method

Participants

A total of 53 participants took part in the “No Delay” experiment, while 56 took part in the 16 min delay experiment. Recruitment and compensation was the same as previous experiments.

Materials

The stimulus materials were identical to previous experiments. In addition, a Harry Potter puzzle was used for the condition with an extended retention interval.

Design

This experiment includes a between-subject manipulation of retention interval. In the 1 min retention interval, after studying List 2 participants completed a 60s arithmetic task and then went immediately into the cued recall task as in Aue et al. (2012) In the 16 min delay condition participants completed the same 60s arithmetic task. Afterwards, participants were prompted that there would be a brief delay before proceeding and to please see the

---

[5] RF has also been observed elsewhere in slightly different designs (e.g., Bruce & Weaver, 1973; Postman & Stark, 1969; Robbins & Bray, 1974; Tulving & Watkins, 1974).

[6] The sample size for Experiment 3 and 4 is larger than previous experiments due to ambitious research assistants.

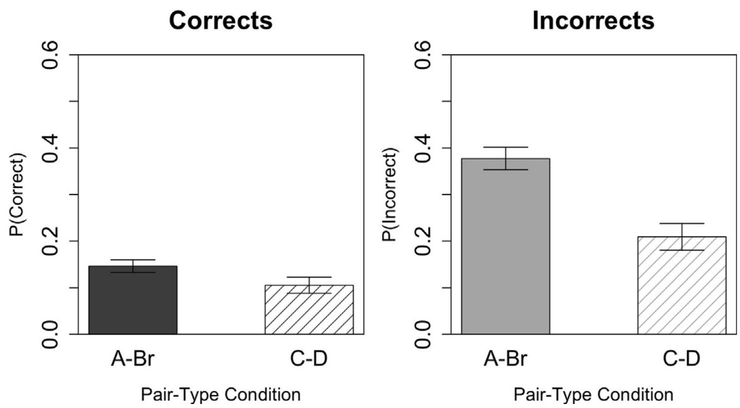

Fig. 3. Experiment 3 correct (left panel) and incorrect (right panel) responses. When participants were tested on their memory for the initial study list in a retroactive interference design, retroactive facilitation was observed relative to pairs exclusive to the initial study list. Error bars represent ± 1 standard error.

experimenter. Participants were then seated at a table with anyone else currently taking the experiment and asked to work together on the puzzle. After 15 min had passed, participants completed the cued recall task as described earlier.

We anticipated that overall performance would be worse than Experiments 1–3 following the delay. To avoid floor effects, we shortened the test and study list to 16 items. Everything else about the encoding phase and test phase was as described in the General method section.

Results and discussion

We employed a two-way mixed ANOVA with delay (between subjects; 1 min vs. 16 min) and pair condition (within subject; A-Br vs. C-D) for the correct and incorrect responses separately. For correct responses, we observed main effects of both delay and pair condition and no interaction ($F(1,107) = 2.01$, $p = .693$, $\eta_p^2 < .01$). Performance was better for the short delay condition ($M = .314$, $SE = .027$) relative to the long delay condition ($M = .164$, $SE = .019$; $F(1,107) = 20.76$, $p < .001$, $d = .866$, 95% CI = [.471, 1.26], $BF_{10} = 1267$). A-Br pairs ($M = .266$, $SE = .028$) were better recalled than C-D pairs ($M = .208$, $SE = .019$; $F(1,107) = 10.01$, $p = .002$, $d = .429$, 95% CI = [.160, .697], $BF_{10} = 8.77$), consistent with existing data.

For incorrect responses, we observed only a main effect of pair condition. Across the delay, the number of incorrect responses for the short delay condition ($M = .258$, $SE = .020$) was not different than the long delay condition ($M = .267$, $SE = .024$; $F(1,107) = .07$, $p = .790$, $d = -.051$, 95% CI = [−.427, .324], $BF_{10} = .153$). We observed more intrusions for A-Br pairs ($M = .297$, $SE = .019$) than C-D pairs ($M = .228$, $SE = .019$; $F(1,107) = 9.35$, $p = .003$, $d = .418$, 95% CI = [.149, .686], $BF_{10} = 6.95$), replicating Aue et al. (2012) and Experiments 1–3. The data are plotted as a function of delay and pair type in Fig. 4.

In summary, the persistence of proactive facilitation provides support to the better encoding explanation. The purpose of the delay was to make the cue less familiar. The cue is critical for cued recall because it is the primary information participants use to search memory, in addition to context. If List 2 pairs were better encoded during the second study list, then proactive facilitation effect should have persisted across the long retention interval, as we observed.

Discussion of the better encoding results

The current experiments present a complicated picture with regard to a better encoding explanation. The results of Experiment 4 are consistent with the better encoding idea, given effect persistence of proactive facilitation across the retention interval. However, the data from Experiment 3 contradict the narrative that better encoding is occurring during List 2, given that the proactive facilitation benefit extended retroactively to memory for List 1 associations. A possibility mentioned earlier is that differential encoding, as opposed to "better" encoding, may be able to explain these data. Aue et al. (2012), also suggested that transient associations (e.g., Provyn, Sliwinski, & Howard, 2007) between studied items might be a mechanism for proactive facilitation. For example, a transient association could be associating the List 2 response for a given word-face pairing to the initial word-face pairing from List 1. Thus after studying A-B in List 1, when A-D is studied in List 2 memory could contain information about both items. This idea is similar in spirit to the recursive representation proposed by Wahlheim and Jacoby's (2013). If so, this would account for the retroactive facilitation observed in Experiment 3 and perhaps the proactive interference, especially considering that Aue et al. (2012) reported that many intrusions for A-Br pairs come from the List 1 partner.

We fleshed out this idea more in the General Discussion. In Experiment 5, we examined the third explanation that we proposed, whether participants are searching memory longer for A-Br pairs.

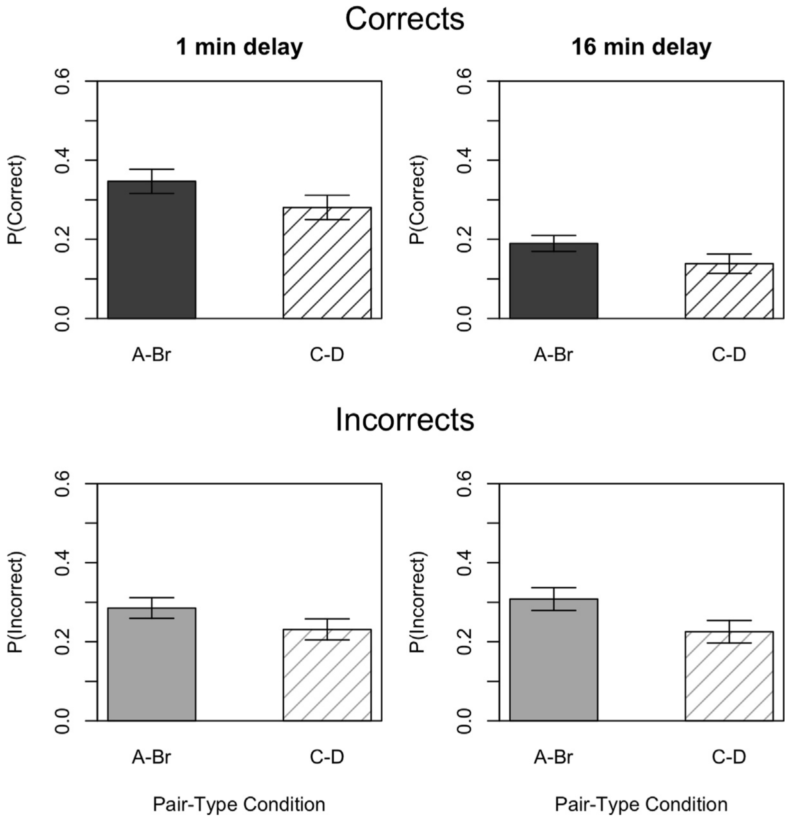

Fig. 4. Memory performance in Experiment 4 following a 1 min delay (left panels) or a 16 min delay (right panels). The A-Br pairs have more correct responses relative to C-D pairs across retention intervals. Error bars represent ± 1 standard error.

Testing longer memory searches

Experiment 5

The primary aim of Experiment 5 was to measure whether participants spent more time searching memory for A-Br relative to C-D pairs. The familiar A-Br cues could have motivated participants to spend more time searching memory for the target. It is reasonable to assume that A-Br cues were more familiar given that the individual items in A-Br pairs tend to be better recognized relative to C-D pairs (Aue et al., 2012; Criss & Shiffrin, 2005). In some memory models such as REM (Diller et al., 2001) the time spent searching memory is represented as an index of the number of times the model attempts to retrieve information from memory and either fails or rejects the retrieved information. Thus, each additional attempt at searching memory affords an additional opportunity to retrieve either the correct memory trace, retrieve an incorrect memory trace, or to fail to retrieve the memory. Having additional opportunities to retrieve something from memory increases the likelihood of retrieving either the correct or incorrect trace for A-Br pairs. As a result, longer searches could have explained both proactive facilitation and proactive interference.

If it were the case that participants were spending more time searching memory for A-Br pairs relative to the equivalent of C-D pairs, this should be reflected as longer reaction times (RT) for both correct and incorrect responses. Across experiments the overall number of responses is relatively low, common in cued recall and other difficult recall tasks. This makes it difficult to get an accurate RT measure because RT is notoriously variable. For Experiment 5, we altered the experimental design to increase the number of responses per participant. Whereas previous studies had two consecutive study lists followed by a test list, here we used a study-test-study-test design with nine study-test blocks. After the first study list, each of the subsequent study lists contained pairs composed of items that appeared on the immediately preceding study list (A-Br pairs) and pairs unique to the current study list (C-D pairs). The C-D pairs of the list were then rearranged to create the A-Br pairs for the next study-test cycle.

Method

Participants

A total of 40 participants took part in the experiment. Participants received course credit for their participation. One participant was dropped due to not providing any correct responses during the experiment, leaving 39 total participants.

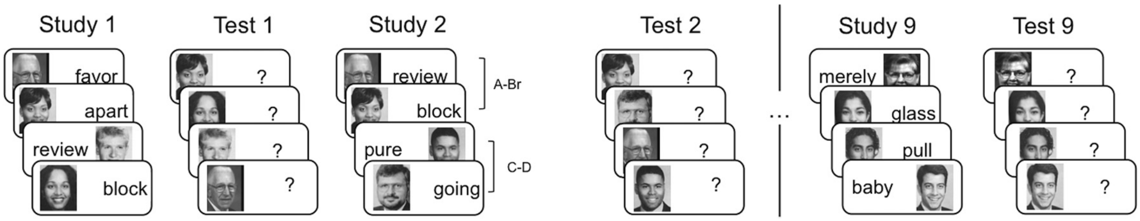

Fig. 5. The design for Experiment 5. Participants went through nine study-test cycles. For the second through ninth cycle participants studied half A-Br and half C-D pairs where A-Br were comprised of rearranged C-D pairs from the immediately preceding list.

Materials

The stimuli used in the experiment were same as those employed in previous experiments. MATLAB (R2011a, MathWorks) was used for stimulus presentation and response timing.

Design

The general design is presented in Fig. 5. During the experiment, participants completed nine study-test cycles with lists of 16 word-face pairs and 16 test trials. With the exception of the first study-test cycle, each list contained half A-Br pairs and half C-D pairs. The A-Br pairs for each list were composed of items studied on the immediately preceding list as C-D pairs, but arranged in different pairs. Pairs were studied for 3 s with a word and face appearing side by side. The cue-target presentation location (i.e., left, right) was randomized. Participants were asked to engage in one of the two previously described encoding tasks for each list. To reduce prior list intrusions, the encoding tasks alternated for each studied list such that a participant never completed the same encoding task for two consecutive lists.

The first study cycle contained all A-B pairs comprised of words and faces. Participants were then tested on the first list using cued recall. The cued recall task differed somewhat from previous experiments. On each test trial, participants were shown a face and asked *think* of the word that it appeared with on the most-recent list. Participants were instructed that once they thought of the target word or decided that they do not know it, to indicate as much by pressing the space bar.[7] The RT analyzed and reported below are time to press the space bar. Participants did not receive any instructions about how quickly to respond. Next the cue disappeared and participants were prompted to type out a response or to type 'no' if they did not recall the target. The program then advanced to the next trial. After each test list participants had a 60s arithmetic task distractor before proceeding to the next study-test cycle. There were a total of nine study-test cycles. The entire session lasted approximately one hour.

RT data were analyzed using "retimes" (Massidda, 2013) package for R. Memory performance was analyzed using the techniques described previously.

Results and discussion

Memory performance

Only the last eight test blocks were analyzed because the first list contained only A-B pairs. To maximize number of observations, we collapsed across test blocks in the analyses, after confirming that a two-way repeated measures ANOVA of test block and pair type condition revealed no interaction with test block for correct ($F(1,38) = 2.96$, $p = .09$, $\eta_p^2 = .07$) or incorrect responses ($F(1,38) = .01$, $p = .94$, $\eta_p^2 < .01$).[8] A paired *t*-test for correct responses showed proactive facilitation. Participants correctly recalled A-Br targets ($M = .436$, $SE = .034$) more often than C-D targets ($M = .377$, $SE = .033$; $t(38) = 4.65$, $p < .001$, $d = 1.05$, 95% CI = [.576, 1.53], $BF_{10} = 540.58$), replicating our previous studies. The proactive interference results that we had previously observed for incorrect responses were conspicuously absent. Indeed, there was no difference between incorrect responses to A-Br cues ($M = .210$, $SE = .021$) and C-D cues ($M = .220$, $SE = .02$; $t(38) = -1.10$, $p = .28$, $d = -.248$, 95% CI = [−.693, .198], $BF_{10} = .223$). The absence of proactive interference following testing was unexpected, but it is not without precedence. Szpunar, McDermott, and Roediger (2008) observed that testing provided a release from the build-up of proactive interference, in the same manner as switch categories provides a release from interference (e.g., Malmberg, Criss, Gangwani, & Shiffrin, 2012; Wickens, 1970). Likewise, some have theorized the release to be driven by a test-driven context shift (Jang & Huber, 2008; Klein, Shiffrin, & Criss, 2007; Lehman & Malmberg, 2013). Based on the dissociation, we suggest that proactive facilitation and proactive interference rely on different mechanisms.

Response time

Next we examined response times for the correct, incorrect, and no recall responses. As stated earlier, if participants were spending more time searching memory for A-Br pairs relative to C-D pairs then it should have been reflected in the time participants took to output a response. First, we examine the RT data using a traditional approach, followed by a more robust curve-fitting analysis.

[7] We also measured the time to first keystroke once participants entered a response, the time to response submission, and total response duration. The pattern of data did not differ from the RT data described here, a full complete analysis of these data can be found in Aue (2014).

[8] There was a main effect of test block for correct performance ($F(1,38) = 4.13$, $p = .05$, $\eta_p^2 = .1$). However the effect was not systematic or theoretically meaningful, so we do not explore it further. The effect for incorrect performance was not significant ($F(1,38) = .89$, $p = .35$, $\eta_p^2 = .02$).

Using a pair *t*-test, we observed no differences between the average median response times for correct responses for A-Br pairs ($M = 2.01$, $SE = .196$) relative to C-D pairs ($M = 1.98$, $SE = .153$; $t(38) = .352$, $p = .73$, $d = .08$, 95% CI = [−.365, .523], $BF_{10} = .133$). Likewise, there were no differences between the median response times for incorrect responses for A-Br pairs ($M = 2.65$, $SE = .318$) relative to C-D pairs ($M = 2.76$, $SE = .238$; $t(38) = -.425$, $p = .67$, $d = -.096$, 95% CI = [−.54, .348], $BF_{10} = .136$). However, there were differences for the No Recall responses. No Recall responses could be thought of as the time it took participants to terminate a memory search. Participants tended to take longer to decide they did not know the answer for A-Br pairs ($M = 3.02$, $SE = .386$) than C-D pairs ($M = 2.53$, $SE = .294$; $t(38) = 3.17$, $p = .003$, $d = .718$, 95% CI = [.258, 1.18], $BF_{10} = 9.85$). Based on these data, participants were searching memory longer for A-Br pairs, but they were not searches that resulted in recall of a word.

Applying traditional statistical approaches to RT data can misrepresent the results given the need to aggregate the data and violations of statistical assumptions, amongst other issues (e.g., Balota & Yap, 2011; Heathcote, Popiel, & Mewhort, 1991). We wanted to buttress the above analysis by fitting an ex-Gaussian distribution using maximum likelihood estimation (e.g., Heathcote et al., 1991; Rouder & Speckman, 2004)[9] to the data. The ex-Gaussian distribution is a convolution of a Gaussian distribution and an exponential distribution and tends to provide excellent fit to response time data (Heathcote et al., 1991; Hohel, 1965; Matzke & Wagenmakers, 2009). There are three parameters that influence the shape of the distribution: the mean (μ) and standard deviation (σ) of the Gaussian component which shift the distribution along the *x*-axis and change the spread of the distribution, respectively, and the mean of the exponential component (τ) which influences the thickness of the tail of the distribution.

There was not enough data per participant, response type, and condition to fit individual participant responses. As a result, we fit aggregated data across participants for each response time and condition. The descriptive statistics for the three response types are provided in Table 2. The only difference in parameter estimates are for No Recall responses. As reflected in the τ parameter, participants tend to have more long responses that resulted in a retrieval failure for A-Br pairs relative to C-D pairs. The observation is consistent with the analysis of median RT, but reflects the change in median RT for no recall trials containing a greater proportion of particularly long responses rather than a shift in the entire distribution. While there are clearly other differences in parameter estimates, the other estimates are somewhat noisy and have overlapping distributions.

To summarize, we theorized that proactive facilitation could have been driven by an extended memory search resulting in increased correct and incorrect responses. We observed significant proactive facilitation in the absence of proactive interference. This is a feature of the experimental design that we did not anticipate, but is consistent with the retrieval practice literature (Szpunar et al., 2008). The fact that proactive facilitation persisted in the absence of proactive interference is novel and we suggest it is evidence that the phenomena are driven by different mechanism. There was no evidence of longer searches leading to more correct responses for A-Br relative to C-D pairs. We observed evidence that participants took longer to terminate their memory search for A-Br pairs relative to C-D pairs, but these data alone do not explain proactive facilitation.

[9] The SIMPLEX method was used during MLE. Confidence intervals on the parameter values were generated through 1000 bootstrapped samples.

Table 2
The best-fitting ex-Gaussian parameters for reaction times by response and pair types.

	μ	σ	τ
Correct			
A-Br	.819 (.77, .866)	.252 (.204, .297)	1.45 (1.34, 1.57)
C-D	.807 (.765, .853)	.223 (.180, .265)	1.64 (1.47, 1.80)
Incorrect			
A-Br	.36 (.232, .495)	.167 (.055, .278)	2.84 (2.48, 3.21)
C-D	.279 (.154, .401)	.109 (0, .212)	3.46 (2.68, 4.52)
No recall			
A-Br	.127 (.088, .155)	.032 (0, .052)	3.18 (2.92, 3.45)
C-D	.161 (.132, .191)	.052 (.033, .074)	2.64 (2.46, 2.83)

Note. 95% Highest Density Interval of the distributions generated based on 10,000 bootstrapped samples.

General discussion

The aim of the current project was to further develop our understanding of proactive facilitation by examining basic memory processes that could explain the phenomenon. In trying to integrate proactive facilitation into a common theoretical framework of an existing model of cued recall (Diller et al., 2001), we tested three potential explanations. We examined differences in response thresholds for the target item, differences in encoding A-Br pairs during List 2 presentation, and differences in search duration. The data tended to be most consistent with the idea that an encoding advantage for some A-Br pairs is driving proactive facilitation. Next, we discuss the limitations and potential modifications to the explanation that could more completely explain the constellation of observed data.

Explanations for proactive facilitation

At the outset, the most parsimonious explanation for the pattern of data for A-Br pairs (an increase in both correct and incorrect responses) seemed to be some form of bias moderated by the familiarity of the cue. However, Experiments 1 and 2 revealed that responding was resistant to manipulations of response threshold, and produced no evidence of differences in the quality of correct responses provided. We also considered whether participants were biased to spend more time interrogating memory for A-Br pairs. However, while participants did take longer to terminate searches, the longer search was not driving the increase in responding to A-Br pairs.

Our manipulations in Experiments 1 and 2 were aimed at a 'sophisticated' guessing mechanism and we failed to change the pattern of observed data. Related to this idea that cue familiarity is critical is the idea that participants simply guess more in response to a familiar cue. Guessing, in this case from the two possible targets, might lead to more correct and to more incorrect responses. While this is a reasonable idea, it is inconsistent with Experiment 5 where facilitation but not interference was observed. That proactive facilitation is observed in the absence of interference suggests different mechanisms and cannot be accounted for by simple guessing.

The data support the idea that proactive facilitation is at least partially due to an encoding advantage that occurs during List 2 for some A-Br pairs. An encoding advantage is evidenced by the fact that the proactive facilitation persists across a long retention interval (Experiment 4). We argue that would not be the case if participants were simply relying on cue familiarity to change a response threshold or to terminate searches. However, it is clearly not the case that A-Br pairs always enjoy an encoding advantage given the presence of proactive interference. Instead, we suggest the data reflect two separate phenomena. Perhaps it is the case that a subset of trials benefit from the encoding advantage when noticing the change in pairings from the initial study list.

Wahlheim and Jacoby (2013) observed a form of proactive facilitation that they suggest is driven by participants recollecting the fact that the pairings changed across lists. In Wahlheim and Jacoby, participants studied a list of weakly associated word pairs during intentional learning of an initial list. During a second list participants are presented with another list of pairs where the cue item is repeated but the target is replaced with a new, weakly associated target (i.e., A-B, A-D). During List 2, participants are told that some of the items from List 1 may be repeated, and they should try to detect the change and recall the List 1 target if a change is detected. At test, participants are given a cue and asked to recall the word it was studied with on the most-recent list. Afterward, participants are asked about their phenomenological experience during recall. As discussed earlier, Wahlheim and Jacoby (2013, Experiment 1) find that participants were able to recall the List 2 target more often, relative to a C-D control, if the List 1 target came to mind first. Wahlheim and Jacoby have cited Hintzman's (2004, 2010) theory of a recursive representation as an explanation of proactive facilitation.

Hintzman (2004, 2010) proposed that when an item is encountered a second (or *n*th) time, spontaneous recall of the earlier events associated with the item in the experimental context could occur. The conditions of the earlier presentation are integrated into the memory for the current presentation in such a way that preserves item and order information, this is the recursive representation. Wahlheim and Jacoby (2013), apply this idea by suggesting that participants store a recursive representation of A-B embedded in A-D when a participant detects that A switched partners from List 1 to List 2. Additionally, Burton et al. (2017) observed evidence of proactive facilitation—they call it associative facilitation—and found that it was most robust when participants were instructed to form mediators between the A-B and A-D pairs. Benjamin and Tullis (2010) have also proposed that recursive reminding may also provide insight into spacing effects. Benjamin and Tullis suggest when reminding takes place during study, encoding of the second presentation is potentiated.

We propose that better encoding of the A-Br pairs takes place during List 2 as a consequence of study-phase retrieval. When a pair of items is presented for study, a participant evaluates the familiarity of the presented item. If the items are sufficiently familiar then they are recognized as having been studied before, and the item's information from the existing (e.g., A-B) trace is updated and used in the List 2 trace, thereby associating it with the new target (e.g., A-Br). The same process would take place for the target. The detection and updating of existing memory traces during study has been implemented in modeling the strength based mirror effect (Criss, 2006; Shiffrin & Steyvers, 1997) and has been demonstrated during testing (Criss et al., 2011), as a component of an explanation of output interference (Criss & Koop, 2015). However, updating has always involved existing traces. Here we are suggesting that components of existing traces are being drawn upon and updated during the storage of new associations. This idea of updating and integrating, would also account for the fact that the current data appear to reflect a combination of proactive facilitation and interference. Interference from other studied items would likely make the detection of A-Br pairs difficult and would not occur on every trial. Thus, memory performance for A-Br pairs could be better or worse depending on whether it was updated correctly during List 2.

Blurring the lines between encoding and retrieval

Beyond the specific explanations for proactive facilitation, the observation of proactive facilitation has implications for memory research more generally. In the current paper we repeatedly observed proactive facilitation in experimental designs that generate interference, but we also observed proactive facilitation when interference was not observed (Experiment 5). We suggest that the persistence of proactive facilitation in Experiment 5 indicates that proactive facilitation is a phenomenon independent from proactive interference and provides a novel perspective into the nature of memory.

Indeed, the current data reinforce the idea that memory is far more complicated than the outdated notion that learning only occurs during study and recall only occurs during test. There is ample evidence indicating that the encoding of information occurs while testing both to our detriment (e.g., output interference [Annis, Malmberg, Criss, & Shiffrin, 2013; Aue, Criss, & Prince, 2015; Criss et al., 2011; Koop, Criss, & Malmberg, 2015; Malmberg et al., 2012]) and to our benefit (e.g., retrieval-based learning [see Karpicke et al., 2014 for a recent review]). Additionally, our explanation for proactive facilitation put forth in the previous section suggests that memory retrieval is taking place during the encoding of information. Although this is not something that is measured in the current data, Wahlheim and Jacoby (2013) specifically asked participants whether they detected that an item had been studied previously. As discussed above, they found that when a participant noticed a change and the previous association (i.e., A-B) was accessible at test, they tended to recall the most recent association (i.e., A-D) more often. In other words, participants are learning while being tested, and appear to be spontaneously testing themselves while studying information. Moreover, the fact that all of the participants in the current experiments (except Experiment 5) and those of Aue et al. (2012) were unaware that their memory was going to be tested indicates that this is something that occurs spontaneously and is not the result of a particular encoding strategy adopted by a subset of participants.

The benefit extends beyond the confines of memory experimentation. Indeed, the idea that participants are building on existing memories to aid in learning new information is similar to a well-known learning strategy encouraged by educators called self-explanation. Berry (1983) found that children were more accurate at certain types of problem solving if they discussed the problem and related to their existing knowledge while they were attempting to solve it. Similarly, participants may be drawing existing representation of the studied information to provide a starting point for the encoding of the new pairs in List 2.

Future directions

The current data raise new questions about proactive facilitation. For example, what is the nature of the memory representation for A-Br pairs in List 2? One possibility is a recursive representation where the memory for List 1 is embedded in List 2. Another possibility is that the representation for an A-B pair following List 2 could be updated with the List 2 partner (e.g., A-D-B), where both associative and order information are represented. However, such a representation would get complicated given the current A-B, A-Br design where *both* the cue and the target are repeated across lists. Under the recursive representation logic for A-Br pairs the representation for List 2 studied pairs would contain the List 1 partners for both items of the A-Br pair if the repetition were noticed. For example, as is depicted in Table 1 if participants study the pairs *Absence-Hollow* and *Pupil-River* during List 1 that are then rearranged for List 2 so that they study *Absence-River*, then the implication of a representation that embeds List 1 partners into List 2 is that study of *Absence-River* (assuming noticing of the changes) should contain information about the List 1 partner for both items (e.g., *Hollow-Absence-Pupil-River*).

While it is impossible to examine the contents of the memory trace directly, one could test a participant's memory for a remote association between the items that were not studied together but were studied *with* items that were studied together. An interesting test of this idea may be to give a participant *Pupil* and measure how often the response is *Absence*. The two items were never studied together but they would be associated in a recursive representation. In this design, *Pupil* is an independent cue because the two words were never studied together, but would be directly associated if stored in a single representation. This is type of response would be similar to the concept of remote or transitive associations (e.g., Provyn et al., 2007).

It is not clear from the current data whether facilitation depends on repetition of the cue, the target, or both (as in our experiments). Differentiating the influence of the cue and target would help expand our understanding of the phenomena. Additionally, it is unclear whether overall familiarity with the individual item could drive proactive facilitation, or whether the cue and target need to be experienced in List 1 in the context of a pair. Participants are able to strategically access subsets of information (e.g., items studied in one pair type but not another) within memory (Aue et al., 2012; Criss & Shiffrin, 2005). It is conceivable that familiarizing the participants with the cue or target as individual items would not have the same facilitating effect as studying the items in the context of a pair.

The variability in recognizing repeated items across lists may also explain why proactive facilitation has gone largely unnoticed in the literature for so very long. As discussed, Burton et al. (2017) have suggested that the differences in paired-associate techniques (e.g., learning to a criterion) may have contributed to masking the proactive facilitation effects.

Conclusion

The current project focused on understanding how memories for recent experiences interacted with recent memories for related information. We examined this by testing memory for a recent event that conflicted with a prior event by having a different pairing of the material (e.g., a new word with an old face). We observed robust proactive facilitation across a variety of experimental scenarios. During these experiments, we were able to rule out decisional factors such as an altered response threshold or longer memory searches. The current data are consistent with the idea that when some items are encohttps://osf.io/6ph2z/untered a second time, pairs receive additional encoding that facilitates their later recall relative to a novel control pair.

Author notes

The order of authorship reflects the relative contribution of the authors to the project. The presented data were part of WRA's doctoral dissertation. WRA is currently at the Department of Psychological Sciences at Purdue University. This material is based upon work supported by the National Science Foundation under Grant Number (#0951612) awarded to AHC. MDN is now at the Department of Applied Behavioral Science at the University of Kansas. Data and materials can be accessed through the Open Science Framework: https://osf.io/6ph2z/.

References

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. Bjork, R. A. Bjork, E. Bjork, & R. A. Bjork (Eds.), *Memory* (pp. 237–313). San Diego, CA US: Academic Press. http://dx.doi.org/10.1016/B978-012102570-0/50010-0.

Annis, J., Malmberg, K. J., Criss, A. H., & Shiffrin, R. M. (2013). Sources of interference in recognition testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1365.

Aue, W. R. (2014). *Understanding proactive facilitation in cued recall* (Unpublished doctoral dissertation). New York: Syracuse University.

Aue, W. R., Criss, A. H., & Fischetti, N. W. (2012). Associative information in memory: Evidence from cued recall. *Journal of Memory and Language, 66*(1), 109–122. http://dx.doi.org/10.1016/B978-012102570-0/50010-0.

Aue, W. R., Criss, A. H., & Prince, M. A. (2015). Dynamic memory searches: Selective output interference for the memory of facts. *Psychonomic Bulletin & Review, 22*(6), 1798–1806.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry the power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459. http://dx.doi.org/10.1016/B978-012102570-0/50010-0.

Barnes, J. M., & Underwood, B. J. (1959). 'Fate' of first-list associations in transfer theory. *Journal of Experimental Psychology, 58*(2), 97–105. http://dx.doi.org/10.1016/B978-012102570-0/50010-0.

Benjamin, A. S., & Tullis, J. G. (2010). What makes distributed practice effective? *Cognitive Psychology, 61*(3), 228–247.

Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *Quarterly Journal of Experimental Psychology, 35*, 39–49.

Bruce, D., & Weaver, G. E. (1973). Retroactive facilitation in short-term retention of minimally learned paired associates. *Journal of Experimental Psychology, 100*(1), 9–17. http://dx.doi.org/10.1037/h0035488.

Burns, D. J. (1989). Proactive interference: An individual-item versus relational processing account. *Journal of Memory and Language, 28*(3), 345–359. http://dx.doi.org/10.1016/0749-596X(89)90038-7.

Burton, R. L., Lek, I., & Caplan, J. B. (2017). Associative independence revisited: Competition between conflicting associations can be resolved, or even reversed, in one trial. *Quarterly Journal of Experimental Psychology, 70*(4).

Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language, 55*(4), 461–478.

Criss, A. H., Aue, W. R., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language, 64*(2), 119–132. http://dx.doi.org/10.1016/j.jml.2010.10.001.

Criss, A. H., Malmberg, K. J., & Shiffrin, R. M. (2011). Output interference in recognition memory. *Journal of Memory and Language, 64*(4), 316–326.

Criss, A. H., & Koop, G. J. (2015). *7 Differentiation in Episodic Memory*. Cognitive Modeling in Perception and Memory: A Festschrift for Richard M. Shiffrin. 112.

Criss, A. H., & Shiffrin, R. M. (2005). List discrimination in associative recognition and implications for representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1199–1212. http://dx.doi.org/10.1037/0278-7393.31.6.1199.

Diller, D. E., Nobel, P. A., & Shiffrin, R. M. (2001). An ARC–REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 414–435. http://dx.doi.org/10.1037/0278-7393.31.6.1199.

Eakin, D. K. (2005). Illusions of knowing: Metamemory and memory under conditions of retroactive interference. *Journal of Memory and Language, 52*, 526–534.

Hanczakowski, M., Pasek, T., Zawadzka, K., & Mazzoni, G. (2013). Cue familiarity and 'don't know' responding in episodic memory tasks. *Journal of Memory and Language, 69*, 368–383. http://dx.doi.org/10.1016/j.jml.2013.04.005.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109*(2), 340–347.

Hintzman, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition, 32*(2), 336–350.

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition, 38*(1), 102–115.

Hirshman, E., Burns, D. J., & Kuo, T. (1993). Examining a processing tradeoff explanation of proactive interference. *Memory & Cognition, 21*(1), 5–10. http://dx.doi.org/10.3758/BF03211159.

Hohel, R. H. (1965). Inferred components of reaction time as a function of foreperiod duration. *Journal of Experimental Psychology, 69*, 382–386.

Jacoby, L. L., Wahlheim, C. N., & Kelly, C. K. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1282–1297.

Jacoby, L. L., Wahlheim, C. N., & Yonelinas, A. P. (2013). The role of detection and recollection of change in list discrimination. *Memory and Cognition, 41*, 638–649.

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 112.

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

Kelly, K. & Lai, K. (2015). MBESS. R package version 3.3.3. <https://cran.r-project.org/web/packages/MBESS/index.html>.

Koop, G. J., Criss, A. H., & Malmberg, K. J. (2015). The role of mnemonic processes in pure-target and pure-foil recognition memory. *Psychonomic Bulletin & Review, 22*(2), 509–516.

Klein, K. A., Shiffrin, R. M., & Criss, A. H. (2007). Putting context in context. In J. S. Narine (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 171–189). New York, NY: Psychology Press.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*(3), 297–315. http://dx.doi.org/10.1037/0096-3445.123.3.297.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*(3), 490–517. http://dx.doi.org/10.1037/0033-295X.103.3.490.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Lawrence, M. A. (2013). ez: Easy analysis and visualization of factorial experiments. R package version 4.3. <http://CRAN.R-project.org/package=ez>.

Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review, 120*(1), 155–189. http://dx.doi.org/10.1037/a0030851.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28*(2), 203–208. http://dx.doi.org/10.3758/BF03204766.

MacLeod, C. M., & Nelson, T. O. (1984). Response latency and response accuracy as measures of memory. *Acta Psychologica, 57*, 215–235.

Malmberg, K. J., Criss, A. H., Gangwani, T., & Shiffrin, R. M. (2012). Overcoming the negative consequences of interference that results from recognition memory testing. *Psychological Science, 23*, 115–119.

Malmberg, K. J., & Shiffrin, R. M. (2005). The 'one-shot' hypothesis for context storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 322–336. http://dx.doi.org/10.1037/0278-7393.31.2.322.

Massidda, D. (2013). retimes: Reaction Time Analysis. R package version 0.1-2. <http://CRAN.R-project.org/package=retimes>.

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798–817.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(4), 851–861.

Millward, R. (1964). Latency in a modified paired associate learning experiment. *Journal of Verbal Learning and Verbal Behavior, 3*, 309–316.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-2. <https://cran.r-project.org/web/packages/BayesFactor/index.html>.

Murdock, B. B., & Anderson, R. E. (1975). Encoding, storage, and retrieval of item information. In *Information processing and cognition: The Loyola symposium* (pp. 145–194). Chicago.

Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General, 113*, 282–300.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15*(1), 1–25.

Nobel, P. A., & Shiffrin, R. M. (2001). Retrieval processes in recognition and cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 384–413. http://dx.doi.org/10.1037/0278-7393.27.2.384.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6*(4), 451–474.

Postman, L., & Gray, W. (1977). Maintenance of prior associations and proactive inhibition. *Journal of Experimental Psychology: Human Learning and Memory, 3*(3), 255–263. http://dx.doi.org/10.1037/0278-7393.3.3.255.

Postman, L., & Keppel, G. (1977). Conditions of cumulative proactive inhibition. *Journal of Experimental Psychology: General, 106*(4), 376–403.

Postman, L., & Stark, K. (1964). Proactive inhibition as a function of the conditions of transfer. *Journal of Verbal Learning and Verbal Behavior, 3*(3), 249–259. http://dx.doi.org/10.1016/S0022-5371(64)80050-5.

Postman, L., & Stark, K. (1969). Role of response availability in transfer and interference. *Journal of Experimental Psychology, 79*(1, Pt.1), 168–177. http://dx.doi.org/10.1037/h0026932.

Postman, L., Stark, K., & Burns, S. (1974). Proactive inhibition on unpaced tests of retention. *The American Journal of Psychology, 87*(1/2), 33–56. http://dx.doi.org/10.2307/1421998.

Postman, L., & Underwood, B. (1973). Critical issues in interference theory. *Memory & Cognition, 1*, 19–40. http://dx.doi.org/10.3758/BF03198064.

Provyn, J. P., Sliwinski, M. J., & Howard, M. W. (2007). Effect of age on contextually mediate associations in paired associate learning. *Psychology & Aging, 22*(4), 846–857.

Putnam, A. L., Wahlheim, C. N., & Jacoby, L. L. (2014). Memory for flip-flopping: Detection and recollection of political contradictions. *Memory & Cognition, 42*, 1198–1210.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL <http://www.R-project.org/>.

Ratcliff, R., & Hockley, W. E. (1980). Repeated negatives in item recognition: Nonmonotonic lag functions. In Roediger, H. L. (1973). *Inhibition in recall from cueing with recall targets. Journal of Verbal Learning and Verbal Behavior*, 12(6), 644–657. doi:http://dx.doi.org/10.1016/S0022-5371(73)80044-1.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*(1), 90–138.

Robbins, D., & Bray, J. F. (1974). The spacing effect and the AB, AC paradigm: Evidence for retroactive facilitation. *Journal of Experimental Psychology, 103*(3), 420–425.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition, 2*(2), 261–269. http://dx.doi.org/10.3758/BF03208993.

Roediger, H., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x.

Roediger, H., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181–210. http://dx.doi.org/10.1111/j.1745-6916.2006.00012.x.

Roediger, H. L., & Payne, D. G. (1985). Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognize theories. *Memory & Cognition, 13*(1), 1–7. http://dx.doi.org/10.3758/BF03198437.

Roediger, H. L., & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired-associate lists. *Journal of Experimental Psychology: Human Learning and Memory, 6*(1), 91–105. http://dx.doi.org/10.1037/0278-7393.6.1.91.

Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin & Review, 11*(3), 419–427.

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievabiliby enhances feeling of knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1074–1083.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM – retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*(2), 145–166. http://dx.doi.org/10.3758/BF03209391.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. III, (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1392–1399.

Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior, 13*(2), 181–193. http://dx.doi.org/10.1016/S0022-5371(74)80043-5.

Twedt, H. M., & Underwood, B. J. (1959). Mixed vs. unmixed lists in transfer studies. *Journal of Experimental Psychology, 58*(2), 111–116. http://dx.doi.org/10.1037/h0043795.

Underwood, B. J. (1949). Proactive inhibition as a function of time and degree of prior learning. *Journal of Experimental Psychology, 39*(1), 24–34. http://dx.doi.org/10.1037/h0059550.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64*(1), 49–60.

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory & Cognition, 41*(1), 1–15. http://dx.doi.org/10.3758/s13421-012-0246-9.

Weinstein, Y., McDermott, K. D., & Szpunar, K. K. (2011). Testing protects against proactive interference in face-name learning. *Psychonomic Bulletin & Review, 18*, 518–523.

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review, 77*(1), 1–15.