Published in final edited form as:

Behav Processes. 2014 March ; 103: 43–51. doi:10.1016/j.beproc.2013.11.008.

Failure to Observe Renewal following Retrieval-Induced Forgetting

Gonzalo Miguez, Lisa E. Mash, Cody W. Polack, and Ralph R. Miller
State University of New York at Binghamton

Abstract

Recent studies have pursued the nature of inhibition observed in retrieval-induced forgetting (RIF) tasks. In a RIF paradigm, participants are trained on category-exemplar pairs in Phase 1. Then, some exemplars from select categories (Rp+ items) receive further practice in Phase 2. At test, impaired recall for non-practiced exemplars of the practiced categories (Rp– items) is observed relative to exemplars from non-practiced categories (Nrp items). This difference constitutes RIF. Prior reports of spontaneous recovery from RIF indicate that RIF represents a lapse rather than a loss of memory. Empirical analogs and theoretical considerations suggest that RIF should also be reversible through a change of context between Phase 2 and testing (i.e., renewal). We conducted two experiments using human participants to evaluate the context dependency of RIF. In both experiments, Phases 1 and 2 occurred in distinctly different contexts with subsequent testing occurring in either the Phase 1 context or the Phase 2 context. RIF was observed in both experiments. Experiment 1 additionally found that the magnitude of RIF was not reduced by testing in the Phase 1 context relative to testing in the Phase 2 context. Experiment 2 further tested context dependency of RIF by 1) increasing the dissimilarity between the two contexts and 2) inserting a retention interval between Phase 2 and test for half of the participants in each test context condition. The data again indicated no effect of the context manipulation. Thus, no renewal from RIF was observed in either experiment; moreover, these null findings were supported by Bayesian analyses. These results are compared with analogous inhibitory processes in the animal memory literature that typically show both physical and temporal context dependency.

Keywords

Renewal; spontaneous recovery; retrieval induced forgetting; inhibition

1. Introduction

Researchers in the field of memory have long studied the nature of forgetting as well as its benefits (e.g., Ebbinghaus, 1885/1964; Jenkins and Dallenbach, 1924; McGeoch, 1932; Tulving and Psotka, 1971). Intuitively, one may appreciate the adaptive value of forgetting. In order to effectively cope in a dynamic environment, memories must be constantly revised and re-prioritized, with retrieval of momentarily less appropriate memories being

© 2013 Elsevier B.V. All rights reserved.

Mailing Address: Ralph R. Miller, Department of Psychology, SUNY – Binghamton, Binghamton, NY 13902-6000, USA, TEL: (607) 777-2291, FAX: (607) 777-4890, rmiller@binghamton.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

suppressed. This flexibility is essential for everyday functioning. For example, it would be maladaptive if every time a person went shopping, he remembered every item he had ever bought at that store. Research has supported the popular speculation that forgetting is an active, purposeful mechanism that prevents unessential information from interfering with the retrieval and processing of immediately critical information (e.g., Anderson, Bjork, and Bjork; 1994; Wixted, 2005).

One phenomenon that provides evidence for this view is known as retrieval-induced forgetting (RIF). In the standard design for RIF developed by Anderson et al. (1994), participants study multiple category-exemplar pairs (e.g. metal-mercury) from multiple categories (e.g. metals, colors, animals, etc.) each containing a unique set of exemplars in Phase 1 (the study phase). Then half of the exemplars from half of the categories are practiced using a category-stem cued recall task (e.g. metal-me___) in Phase 2 (the retrieval practice phase). These exemplars belong to the retrieval-practice positive condition (Rp+). Non-practiced exemplars from practiced categories constitute the retrieval-practice negative condition (Rp-), while exemplars from non-practiced categories serve as a baseline, constituting the no retrieval-practice condition (Nrp). During the test phase, which usually occurs after a brief retention interval, participants are prompted to complete a recall task in which they are provided with category names, and must recall as many exemplars as possible from each category. Within this paradigm, Rp+ words are recalled more frequently than both Rp- and Nrp words, which is unsurprising given they were subject to retrieval practice during Phase 2. However, the less intuitive but more interesting finding is that Rp- words are recalled less frequently relative to the baseline Nrp words. This provides evidence that the retrieval practice of Rp+ items not only facilitates later recall of Rp+ exemplars, but results in the decreased recall of Rp- items (i.e., RIF; e.g., Anderson et al.; for recent reviews, see Anderson, Bjork, and Bjork, 2000; Storm, 2010; Storm and Levy, 2012)

Several accounts of RIF have been proposed. Anderson and colleagues (1994, 1995) have favored an active inhibitory mechanism which suppresses retrieval of non-practiced exemplars belonging to the same category as the practiced exemplars. That is, they suggest that during Phase 2 the retrieval of an incorrect item (that was cued by the immediate practice category), followed by recognition that the item is incorrect, results in active suppression of subsequent retrieval of that item. This inhibits future recall of the item (i.e., Rp-), thereby reducing interference with future retrieval of the practiced (i.e., Rp+) items. Rp- items presumably are not recognized as incorrect until it is noticed that they do not fit the exemplar stem provided on that retrieval practice trial, whereas Nrp items are rarely recalled during retrieval practice because they are unrelated to the categories undergoing retrieval practice. This active inhibition account is widely accepted (for reviews, see Anderson, 2003; Anderson and Levy, 2010), but there is also some research implicating mechanisms that differ from the specific inhibitory mechanism postulated by Anderson and colleagues (for reviews, see Raaijmakers and Jakab. 2012, 2013; Verde, 2012, 2013). However, an in-depth discussion of the competing accounts of RIF is beyond the scope of this paper.

Retrieval-induced forgetting presents an interesting analogy to other, more extensively studied phenomena that are reported in both the human and animal learning literature. Specifically, the RIF preparation parallels those used to study retroactive associative outcome interference (for reviews of these parallels, see Ortega-Castro and Vadillo, 2013; Vadillo et al., 2013). Retroactive outcome interference is often observed when pairings of X-O1 in Phase 1 are followed by pairings of X-O2 in Phase 2 (where X is a cue and O1 and O2 are two distinctly different outcomes). At test, participants will ordinarily respond to X in a manner consistent with the more recent pairings (i.e., X-O2). Pavlovian extinction serves as a well-known example of this type of interference, in which the conditioned

stimulus (CS) serves as X, the unconditioned stimulus (US) serves as O1, and non-reinforcement serves as O2. After undergoing the phases described above, subjects will initially fail to respond to CS X as a result of recent nonreinforced presentations. This can be viewed as similar to RIF in which Phase 2 (i.e., retrieval practice) enhances subsequent retrieval of a potentially interfering association that at test inhibits retrieval of associations acquired in Phase 1 which were not practiced in Phase 2, but share a common element (i.e., category) with a practiced association.

Research in the animal literature has shown that retroactive associative outcome interference is often attenuated by shifts between Phase 2 and testing of the physical or temporal context (for a review, see Bouton, 2010). For example, Bouton and Peck (1992) demonstrated the effects of varying the time between Phase 2 and testing using a counterconditioning paradigm. In a traditional counterconditioning preparation, a CS is paired with an appetitive or aversive US (O1) in Phase 1, and is then paired with a US of the opposite valence (O2) in Phase 2. Although the rat's response to the CS at test following counterconditioning ordinarily reflects the second-learned CS-O2 association, these researchers found that if a long retention interval was inserted between Phase 2 and testing, subjects' responding was more consistent with the CS-O1 training experience. This observation indicates that retroactive outcome interference is attenuated by a long retention interval (i.e., spontaneous recovery). Turning to changes in physical context, Peck and Bouton (1990) demonstrated sensitivity of retroactive outcome interference to shifts in physical context between Phase 2 and testing. Again using a counterconditioning design, Phases 1 and 2 were delivered in distinctly different contexts (A and B, respectively), followed soon thereafter by a test in Context A or B. Testing in the Phase 1 context (A) caused greater recovery of the behavior reflecting the CS-O1 pairings relative to testing in the Phase 2 context (B); this is known as ABA renewal.

The effects of spontaneous recovery and renewal are not confined to counterconditioning preparations. Researchers as early as Pavlov (1927) found that when responding to a CS is extinguished, responding to that CS recovers over long intervals between extinction (Phase 2) and testing. Moreover, recovery from extinction in the form of ABA renewal of extinguished responses has also been demonstrated (e.g., Bouton and Bolles, 1979; Bouton and King 1983). That is, if an association is acquired in Context A during Phase 1 and extinguished in Context B during Phase 2, testing in Context A tends to result in stronger responding (i.e., ABA renewal) than is observed with testing in Context B (i.e., ABB baseline control condition).

To date, the published research best supports the account of outcome interference and recovery therefrom that was proposed by Bouton (1993). This model proposes that the CS acquires an excitatory association with the US during Phase 1 and an inhibitory association with the US during Phase 2. At test, the test context acts as a discriminative stimulus that determines which of these two competing associations will be expressed. Miller and Laborda (2011) updated Bouton's model by adding a role for the relative strengths of the acquisition memory and extinction memory. This model provides an elegant explanation of interference between associations that share a common element, as well as attenuation of this interference through shifts in physical and temporal context between training and testing. Research within this framework has established that extinction appears to represent inhibitory processes (Pavlov, 1927), rather than an erasure of acquired associations as proposed by Rescorla and Wagner (1972). For a more general discussion of this account that includes other types of associative interference, see Miller and Escobar (2002). Although there is a general consensus regarding the role of inhibitory mechanisms in outcome interference, there is some debate concerning the details of this account. For example, Bouton's model specifically posits the establishment of inhibitory cue-outcome associations,

whereas Anderson and Spellman (1995) argue that it is the activation of the exemplar (or outcome) representation that is inhibited independent of any particular association.

To summarize the most critical points with respect to the current study, it has been established that retroactive associative outcome interference is context-specific (e.g., Peck and Bouton, 1990). Likewise, it has been demonstrated that long retention intervals can reduce the effects of outcome interference (Bouton and Peck, 1992; Pineño and Matute, 2000). Recent research has examined the validity of analogies between RIF and retroactive outcome interference. These studies have suggested that the same or at least similar mechanisms that are responsible for outcome interference may also be responsible for RIF (Ortega-Castro and Vadillo, 2013; Vadillo et al. 2013). Vadillo et al. used an associative outcome interference design in which experimental participants (undergraduates) received X-O1 presentations in Phase 1, X-O2 presentations in Phase 2, and were then tested on the rate of learning of a new Y-O1 association in Phase 3, relative to participants who received control training for outcome interference in which there was no common element across Phases 1 and 2. They found that associative interference between outcomes O1 and O2, based on their mutual associations to X, impaired acquisition (or at least expression) of a new association involving O1 (i.e., Y-O1), presumably because of attenuated activation in Phase 3 of the representation of the target outcome (O1). Their observations suggest similarities between the outcome interference observed in their contingency learning task and RIF in that both appear to arise from active inhibition of the representation of O1 as hypothesized by Anderson et al. (1994). Similarly, Ortega-Castro and Vadillo (2013) observed that RIF-like impairment of recall occurred regardless of whether the stimuli tested (in Phase 3) were cues or outcomes from Phase 1. In their design, Phase 2 consisted of either retrieval practice with cue-outcome pairs from Phase 1 that shared cues with other outcomes from Phase 1 that were not practiced during Phase 2 (a design for conventional RIF of outcomes), or retrieval practice with cue-outcome pairs from Phase 1 that shared outcomes with other cues from Phase 1 that were not practiced during Phase 2 (a design for RIF of cues). Notably, Anderson's et al.'s exemplar-inhibition account cannot explain RIF of cues because in Phase 2 there was no basis for the nonpracticed cues to have been retrieved during the retrieval practice phase, and retrieval is presumably necessary for subsequent active inhibition according to the model (see Miller and Escobar, 2002, for an account that addresses both outcome and cue interference). The findings of Ortega-Castro and Vadillo provide a further parallel between RIF-like forgetting and prior reports of retroactive outcome and cue interference. Moreover, there are now many demonstrations that RIF, like retroactive associative interference, is subject to spontaneous recovery given a long (e.g., 24 h) retention interval between Phase 2 (i.e., retrieval practice) and testing (e.g., Carroll, Campbell-Ratcliffe, Murnane, and Perfect, 2007; Chan, 2009; MacLeod and Macrae, 2001; Saunders and MacLeod, 2002; Saunders, Fernandes, and Kosnes, 2009). Some researchers have failed to see this spontaneous recovery from RIF (Garcia-Bajos, Migueles, and Anderson, 2009; Migueles and Garcia-Bajos, 2007; Saunders et al., 2009; Storm, Bjork, Bjork, and Nestojko, 2006); however, these failures to observe spontaneous recovery from RIF often arise from subjects being given multiple tests for the same exemplars with different retention intervals (i.e., within-subject designs) which seems to mitigate the effect of spontaneous recovery (e.g., Storm, Bjork, and Bjork, 2012). Thus, spontaneous recovery appears to be a well-established property of RIF, suggesting that the forgetting of RIF is a retrieval failure rather than an irreversible loss of the Rp- memories from Phase 1. However, to our knowledge, there has been no investigation of whether an analog of renewal occurs in RIF to date.

Here we conducted two experiments to assess the context specificity of RIF. If similar processes underlie retroactive outcome interference and RIF, given appropriate parameters one should be able to observe renewal as well as spontaneous recovery following RIF.

Experiment 1 used an ABA renewal design to determine whether RIF was sensitive to context shift effects known to influence associative outcome interference. ABA renewal was sought because ABA renewal has been seen to be one of the more robust types of renewal in associative outcome interference, hence a type of renewal likely to be observed following RIF. Experiment 2 built upon Experiment 1 through the exaggeration of the differences between physical contexts, as well as the insertion of a somewhat longer retention interval between Phase 2 and testing than was used in Experiment 1. However, given that spontaneous recovery following RIF has been reported repeatedly in the past, we were not pursuing spontaneous recovery here, but simply using it as a potential means of increasing sensitivity to renewal. Thus, our ‘long’ retention interval in Experiment 2 was intentionally kept short with the intent of avoiding sufficient spontaneous recovery to eliminate RIF, which would have left us insensitive to the renewal that we were seeking. Recent associative interference preparations have demonstrated that the effects of renewal and spontaneous recovery can summate (Laborda and Miller, 2013; Rosas and Bouton, 1998; Rosas, Vila, Lugo, and Lopez, 2001). Thus, stronger attenuation of RIF was expected in Experiment 2, in which both the temporal and physical contexts were appreciably changed between Phase 2 and testing, possibly increasing sensitivity to renewal. Our central question was whether renewal alone would make an appreciable contribution to any observed recovery from RIF.

2. Experiment 1

The objective of Experiment 1 was to determine whether changing the physical context between Phase 2 and testing would attenuate RIF. The impetus for this study was a growing literature of recent findings suggesting that similar mechanisms may underlie RIF and retroactive associative outcome interference (Ortega-Castro and Vadillo, 2013; Vadillo et al., 2013) and research showing that outcome interference is susceptible to ABA renewal (e.g., Peck and Bouton, 1990; Bouton and Bolles, 1979; Bouton and King 1983). We expected such an ABA context manipulation applied to the study, retrieval practice, and recall phases of a RIF design, respectively, would induce renewal of Rp– words, thereby attenuating the RIF effect.

2.1. Materials and Methods

2.1.1. Participants—Participants in this study were 35 undergraduate students (13 males and 22 females) at the State University of New York at Binghamton, who participated for partial fulfillment of a course requirement. Participants were assigned to one of two treatment groups (ABA with $n = 17$ or ABB with $n = 18$). All participants gave written informed consent and were debriefed when their session was completed.

2.1.2. Design—One between-subjects factor (context of testing) and one within-subjects factor (retrieval practice status) were manipulated (see Table 1). There were two levels of testing context (Context A [ABA] and Context B [ABB]). Both groups received study and retrieval practice in Contexts A and B, respectively. Group ABA was then tested in Context A, while Group ABB was tested in Context B. Retrieval practice status levels were Rp+, Rp–, and Nrp. Rp+ words were each practiced during Phase 2 three times in a cued-recall test (e.g., metals-me___) using a retrieval practice schedule that expanded the time between successive practice trials with the same Rp+ words (Bjork and Bjork, 1992). Rp– words were not presented during Phase 2, but they belonged to the same categories as the Rp+ words. Nrp words (from different categories than the Rp+ and Rp– words) were not presented during Phase 2, and provided a baseline against which we could compare the effects of Rp+ and Rp–. The proportional frequency of recall of each type of word during the Test Phase (which consisted of a recall test cued by the name of the category) served as the dependent variable.

2.1.3. Materials—Two distinctly different contexts were used. Context 1 was a small office, containing a desk and computer. Audio instructions were read aloud in a female voice, and delivered through headphones connected to the computer. Booklets in Context 1 were printed on white, 21.5 × 27.9 cm paper. Participants were provided with #2 pencils, which were used to write in the booklets during Phase 2 and Testing (Phase 3). Context 2 was set in a spacious classroom, and participants were seated at school desks. Audio instructions were read in a male voice, and delivered through external speakers connected to the computer. Participants were provided with a jumbo novelty pen to write in the Context 2 booklets, which were printed on orange, 21.5 × 27.9 cm paper. Within each condition (ABA and ABB), Context 1 served as Context A for half the participants in each group and as Context B for the remaining participants.

Forty-eight exemplar words evenly divided among 8 categories (6 exemplars/category) were borrowed from Anderson et al. (1994) to be used as Rp+, Rp−, and Nrp items. Additionally, 2 filler categories each containing 6 filler words were used. Word trials and instructions were presented in the printed booklets. Booklets were constructed following the methods of Anderson et al. In the Phase 1 booklets, 48 experimental and 12 filler word pairs were presented in 6 blocks of 10 words (one from each category). Blocks were presented in one of two orders. The two orders differed in the position of the categories within each block as well as the position of the block. Phase 2 booklets contained a category label across the top of each page, and an exemplar stem in the middle of the page. Stems consisted of the first two letters of an exemplar followed by an underscore (the length of which was constant for all exemplars), which could be completed by writing in the booklet (e.g., metals-me____). Four alternative sets of words to serve as Rp+ were created following the methods of Anderson et al. (1994). Each target item (i.e., Rp+, Rp−, or NRP) was presented three times in this phase on an expanding schedule (from 3.5 to 6.5 trial-average spacing, accomplished by the insertion of an increasing number of retrieval trials with filler items, as in Anderson et al.). Test booklets presented the 12 category words alone as cues for recall) in one of three possible orders, one category to a page. Specific content of Phase 1 booklets, Phase 2 booklets, and Test booklets was counterbalanced within each group.

2.1.4. Procedure

2.1.4.1. Phase 1: In Phase 1, each participant was seated at a computer station (Context A) and asked to read instructions from the provided booklet. The instructions were also delivered via a recorded audio clip. The instructions were as follows:

Welcome. You are participating in a memory study. In this phase, you will have five seconds to study category-exemplar pairs from the booklet provided. You should spend all of this time relating the exemplar to its category. After each five seconds, a voice will ask you to turn the page. Please turn the page only when you are told to do so. This sequence will be repeated until all word pairs in the booklet have been presented.

The audio clip instructed each participant to turn to the next page (with one pair of words/page) after 5 s of study, with a 1-s intertrial interval provided for turning the page.

2.1.4.2. Phase 2: After Phase 1 was completed, an experimenter led participants to Context B to complete Phase 2. This took approximately 2 min. A new instruction booklet was provided and new instructions were delivered through an audio clip. These instructions asked participants to fill in incomplete exemplars by writing in the booklet. The instructions were as follows:

In this phase of the experiment, each page in the new booklet will contain the name of one of the categories that you studied in the previous phase along with a hint

concerning the exemplar word you should retrieve. The hint will consist of the first two letters of the appropriate exemplar word. Please use the pen we provided to complete the blank space by writing the missing part of the appropriate word. You will have ten seconds to complete the word. After each ten seconds, a voice will ask you to turn the page. This sequence will be repeated until all word pairs in the booklet have been presented.

The incomplete exemplar was presented on the same page as the category word. The first two letters of each exemplar were provided. Each retrieval-practice trial was 10 s in duration, with a 1-s intertrial interval. In Phase 2, participants from both ABA and ABB groups were randomly assigned to one of four practice conditions, counterbalanced within groups.

2.1.4.3. Test Phase: After the completion of Phase 2, the participants were instructed by the audio clip to immediately contact the experimenter and were led out of Context B. For the Test phase, participants in the ABA group returned to Context A, while participants in the ABB group returned to Context B. Both destinations took approximately 2 minutes to reach. Participants in each group were randomly assigned to one of three orders of category test presentation. Participants both read and listened to audio clip instructions which asked them to write all of the exemplars they could remember for each of the given categories, with one category per page. The instructions were:

In this phase of the experiment, at the top of each page there will be a name of one of the categories that you previously studied. Please recall ALL of the exemplars of that category that you had been shown at any time in the experiment. You will have 30 seconds to write the exemplars. After each 30 seconds, a voice will ask you to turn the page. This sequence will be repeated until all categories have been presented.

For each category, participants were given 30 s to recall as many exemplars as possible. After the Test phase was completed, participants were debriefed and dismissed.

2.1.5. Data Analysis—The proportion of each type of word recalled was analyzed with a 2 (Context: ABA vs. ABB) $\times$ 3 (Retrieval Practice Status: Rp+ vs. Rp- vs. Nrp) mixed analysis of variance (ANOVA), with Retrieval Practice Status as a within-subjects variable and Context as a between-subjects variable. Planned comparisons using the error term from the overall ANOVA were conducted to test specific hypotheses. Effect sizes were computed for significant effects as were 95% confidence intervals (CIs) for these effect sizes.

2.2. Results and Discussion

Results are depicted in Figure 1. The ANOVA yielded a significant effect of Retrieval Practice Status, $F(2, 66) = 34.46$, $p < .001$, $MSE = .026$, Cohen's $f = 1.38$ (95% CI = 0.97; 1.80). There was no significant effect of Group Assignment or an interaction between Group and Retrieval Practice Status, $F(1, 33) = 0.04$, $p = .83$, $MSE = .045$, and $F(2, 66) = 0.02$, $p = .90$, $MSE = .026$, respectively. The F values here are appreciably below 1.0, which is ground for pause. Inspection of the data suggests that these unusually low F values arose from large individual differences inflating within-group variance, but small variance across items (i.e., Rp+, Rp-, and Nrp) because the Retrieval Practice Status variable was manipulated within subjects. Overall, Rp+ words were recalled more frequently than baseline Nrp words, $F(1, 33) = 49.57$, $p < .001$, Cohen's $d = 0.42$ (95% CI = 0.32; 0.48) indicating a benefit of retrieval practice in Phase 2. More importantly, we observed impaired recall of Rp- words relative to the Nrp baseline, $F(1, 33) = 4.48$, $p = .04$, Cohen's $d = 0.22$ (95% CI = 0.03; 0.31). Therefore, we successfully replicated the basic RIF effect (e.g., Anderson et al., 1994). The relatively large 95% CIs for estimated effect sizes throughout this experiment

likely reflect the relatively small number of participants that were used. Experiment 2 included near replications of all of the conditions in Experiment 1 using over three times the number of participants and obtained the same relationships with smaller CIs around the effect sizes. Thus, Experiment 2 largely corrects for the degree to which Experiment 1 may have been underpowered.

In order to determine whether the context shift affected RIF, a planned 2 [Rp– vs. Nrp] $\times$ 2 [Group: ABA vs. ABB] ANOVA was conducted. It yielded no significant interaction, suggesting that testing outside the context of retrieval practice (i.e., Context B), as opposed to inside the context of retrieval practice (i.e., Context A) did not significantly affect the recall of Rp– words relative to Nrp baseline words, $F(1, 33) = 0.03$, $p = .86$. Moreover, a planned comparison confirmed that Rp– recall at test did not significantly differ between groups, $F(1, 33) < 0.01$, $p = .93$. Thus, no recovery from RIF in the form of renewal was observed. Given that this null result is our principle finding in this experiment, it seemed appropriate to assess the likelihood that these two null hypotheses were correct relative to the alternative hypotheses that there were actual differences in RIF as a function of test context. This was done using Bayes factors to compute odds ratios (Rouder et al., 2009). The similarity between Groups ABA and ABB in differences in recall of Rp– and Nrp items was supported with odds of 4.01 in favor of the null hypothesis, and the similarity between groups in recall of Rp– items alone was supported with odds of 4.05 in favor of the null hypothesis, both of which constitute “substantial” support for their respective null hypotheses (Rouder et al.).

3. Experiment 2

Given the null results of Experiment 1 with respect to renewal following RIF, the purpose of Experiment 2 was to increase sensitivity to potential renewal by making the physical contexts more dissimilar. Additionally, a slightly longer retention interval (20 min) between Phase 2 and testing, as well as the retention interval of Experiment 1 (2 min), was used, with the thought that the 20-min retention interval might weaken RIF, thereby increasing sensitivity to renewal. Based on prior research demonstrating that spontaneous recovery and renewal often summate to produce greater recovery from retroactive outcome interference than either treatment alone (e.g., Laborda and Miller, 2013; Rosas and Bouton, 1998; Rosas et al., 2001), we expected that each treatment, renewal and spontaneous recovery, would independently work toward attenuating RIF. However, as spontaneous recovery following RIF has been reported previously (e.g., Carroll, Campbell-Ratcliffe, Murnane, and Perfect, 2007; Chan, 2009; MacLeod and Macrae, 2001; Saunders and MacLeod, 2002; Saunders, Fernandes, and Kosnes, 2009), our goal in Experiment 2 was to demonstrate renewal either with or without the increased retention interval.

3.1. Materials and Methods

3.1.1. Participants—Participants were 252 undergraduate students (90 male and 162 female) at the State University of New York at Binghamton, who participated for partial fulfillment of a course requirement. Participants were randomly assigned to one of four treatment groups (ABA-Imm, $n = 63$; ABA-Delay, $n = 54$; ABB-Imm, $n = 66$; and ABB-Delay, $n = 69$). All participants gave written informed consent, and were debriefed when their session was completed.

3.1.2. Design—Two between-subjects factors (context of testing and interval between Phase 2 and testing) and one within-subjects factor (retrieval practice status) were manipulated (see Table 2). There were two levels of test context (Context A [ABA] vs. Context B [ABB]). Participants in Condition ABA were tested in Context A, while those in Condition ABB were tested in Context B. There were two levels of interval between

retrieval and test (Immediately after Phase 2 [Imm] vs. Delay after Phase 2 [Delay]). Like Experiment 1, retrieval practice status levels were Rp–, Rp+, and Nrp, as in Experiment 1. The proportion of each type of word recalled during the Test Phase served as the dependent variable.

3.1.3. Materials—The booklets were constructed as specified in Experiment 1. In addition to the features described in Experiment 1 for Contexts 1 and 2, we added distinct tactile and olfactory features to Context 2 in order to further enhance the distinctiveness of the two contexts. Specifically, a wooden block scented with two drops of methyl salicylate was placed on each desk, and participants were asked to place their non-dominant hands in a container filled with six inches of room-temperature water, which was placed on an adjacent desk. Like Experiment 1, Contexts 1 and 2 served as Contexts A and B, counterbalanced within each of the four groups.

3.1.4. Procedure—Phase 1 and Test were identical to Experiment 1. Phase 2 differed from Experiment 1 in that participants were instructed to mentally complete each cued exemplar, rather than actually writing in the booklet. Due to this difference, the instructions were:

In this phase of the experiment, each page in the new booklet will contain the name of one of the categories that you studied in the previous phase along with a hint concerning the exemplar word you should retrieve. The hint will consist of the first two letters of the appropriate exemplar word. You will have ten seconds to complete the word mentally. You should not write on the paper. After ten seconds, a voice will ask you to turn the page. This sequence will be repeated until all word pairs in the booklet have been presented.

An important difference between the two experiments was the addition of the delay between Phase 2 and Test for the Delay condition. After Phase 2, participants in the Imm condition were tested immediately (after contacting the experimenter, as in Experiment 1), whereas participants in the Delay condition were exposed to a 20-min retention interval before the Test Phase, during which they completed a nonverbal task (Raven Progressive Matrices Test).

3.1.5. Data Analysis—The proportions of Rp+, Rp–, and Nrp words recalled were analyzed using a 2 (Context: ABA vs. ABB) $\times$ 2 (Interval Before Test: Imm vs. Delay) $\times$ 3 (Retrieval Practice Status: Rp+ vs. Rp– vs. Nrp) mixed ANOVA, with Retrieval Practice Status being a within-subjects variable and Context and Interval Before Test being between-subjects variables. Planned comparisons were conducted to test specific hypotheses using the error term of the overall ANOVA.

3.2. Results and Discussion

Results are depicted in Figure 2. The $2 \times 2 \times 3$ ANOVA yielded an effect of Retrieval Practice Status, $F(2, 496) = 131.50, p < .001, MSE = .022$, Cohen's $f = 1.02$ (95% CI = .88; 1.16). Significance was not found for the effects of Context, Interval Before Test, or any interaction among the three factors, all $F_s < 1$. Investigating the main effect of Retrieval Practice Status, Rp+ words were recalled more frequently than baseline Nrp words, $F(1, 248) = 192.88, p < .001$, Cohen's $d = 0.23$ (95% CI = 0.21; 0.25), indicating a benefit in recall of items that were retrieval practiced during Phase 2. More importantly, impaired recall of Rp– words was observed relative to that of baseline Nrp words, $F(1, 248) = 13.08, p < .001$, Cohen's $d = .09$ (95% CI = 0.08; 0.14), replicating the basic effect of RIF observed in Experiment 1.

A planned 2 (Retrieval Practice Status: Rp– vs. Nrp) $\times$ 2 (Context: ABA vs. ABB) $\times$ 2 (Interval Before Test: Imm vs. Delay) contrast found no significant three-way interaction, $F(1, 248) = 0.11, p = .73$, suggesting that testing outside the physical and temporal contexts of retrieval practice had no appreciable effect on the recall of Rp– words at test relative to their respective Nrp baselines in this preparation. The context shift failed to interact with retrieval of Rp– words relative to Nrp words, $F(1, 248) = 0.26, p = .60$. However, a nonsignificant tendency toward an interaction between Interval Before Test and difference in recall of Rp– and Nrp words, $F(1, 248) = 2.73, p = .10$, suggests a weak spontaneous recovery effect which would be congruent with the larger spontaneous recovery from RIF effects with longer retention periods that were described previously (e.g., Saunders et al., 2009). In the Immediate condition Rp– items were recalled less often than Nrp items, $F(1, 248) = 14.33, p < .05$; however, in the Delay condition Rp– items were not reliably recalled less than Nrp items, $F(1, 248) = 1.86, p > .17$. Thus, the present results suggest that RIF is subject to spontaneous recovery, even with our relatively short 20-min retention interval relative to the Immediate condition (i.e., 2 min).

Focusing first on the Immediate condition, there was no interaction between Context and the difference in recall of Rp– and Nrp words, $F(1, 248) = 0.01, p > .90$. However, significant RIF (i.e., difference in recall of Rp– and Nrp items) was present in both Group ABA-Imm, $F(1, 248) = 6.56, p < .05$, and Group ABB-Imm, $F(1, 248) = 7.80, p < .05$. Additionally, a planned contrast between Groups ABA-Imm and ABB-Imm confirmed that recall of Rp– items at test did not significantly differ, $F(1, 248) = 0.05, p = .82$. Thus, as in Experiment 1, significant RIF was observed in both the ABB and ABA conditions, but no recovery from RIF was observed as a result of testing in the Phase 1 context (ABA) as opposed to the Phase 2 context (ABB). The similarity of Groups ABA-Imm and ABB-Imm on recall of Rp– items as well as on the difference in recall of between Rp– and Nrp items were confirmed by Bayesian odds ratios of 7.14 and 7.27, respectively, in favor of the null hypothesis (Rouder et al., 2009).

Now consider the Delay condition data. Recall that the purpose of the Delay retention interval was an attempt to partially attenuate RIF in order to increase sensitivity to the effect of a context shift (i.e., renewal). As intended, the Delay treatment seemingly reduced RIF relative to the Immediate condition, as no significant RIF effect was observed in the Delay condition when Groups ABA-Delay and ABB-Delay were pooled. Moreover, a planned contrast indicated that Group ABB-Delay failed to exhibit reduced recall of Rp– items relative to Nrp items, $F(1, 248) = 2.19, p = .13$. That is, in this renewal control group, there was a tendency toward RIF, but it was not significant, which one would expect to reduce sensitivity to any potential renewal effect that might have occurred in Group ABA-Delay. Critically, the interaction between Context and Item (Rp– vs. Nrp) in the Delay condition was qualitatively in the expected direction (i.e., attenuated RIF in the ABA condition), but it was not statistically significant, $F(1, 248) = 2.22, p = .13$. A Bayes factor analysis found an odds ratio of 2.56 in favor of the null hypothesis for this interaction, which is weak but positive support according to Rouger et al. (2009). Collectively, the data suggest that the 20-min delay between Phase 2 and testing (relative to the 2-min interval of the Immediate condition) both decreased RIF, thereby reducing the range over which renewal might be expected, and increased sensitivity to renewal by making more labile the remaining tendency toward RIF. In summary, the Delay treatment failed to produce a significant renewal from RIF, at least in part because there was so little RIF to recover from.

4. General Discussion

Experiment 1 was successful in replicating the RIF effect first demonstrated by Anderson et al. (1994). Not surprisingly, words that received retrieval practice in Phase 2 (Rp+) were

consistently recalled more frequently than either non-practiced words from practiced categories (Rp–) or baseline words from non-practiced categories (Nrp). Critically, Rp– words were recalled less frequently than Nrp words; that is, RIF was observed. However, although our procedure yielded RIF, the context change between retrieval practice and testing in Experiment 1 failed to attenuate this RIF effect. That is, no renewal was observed following RIF. These findings are contrary to what one would expect based on prior research using analogous preparations to study retroactive outcome interference (Bouton and Bolles, 1979; Bouton and King, 1983; Peck and Bouton, 1990).

Experiment 2 also obtained the basic RIF effect. To improve sensitivity to potential recovery from RIF in Experiment 2, we took measures to further increase the dissimilarity between Contexts A and B (see Methods section of Experiment 2 for details). Additionally we exposed half of the participants in each context condition (ABA and ABB) to a 20-min retention interval between Phase 2 and testing with the thought that this might slightly weaken RIF, in an attempt to make RIF more labile with respect to the context shift. Based on findings from earlier research in both retroactive outcome interference and RIF (see Introduction), we had expected that recovery from RIF due to a long retention interval and to a shift in physical context between Phase 2 and testing would summate to attenuate RIF to a greater degree than either the retention interval or change in physical contexts alone, similar to the way in which they summate to reduce outcome interference (Laborda and Miller, 2013; Rosas and Bouton, 1998; Rosas et al., 2001). However, the longer retention interval attenuated RIF to the point that there was little left on which any potential renewal might have acted in the Delay condition of Experiment 2 (see Figure 2). More centrally, in neither the Imm or Delay conditions did we observe renewal following RIF. Moreover, these null results were supported by Bayesian analysis which produced odds ratios favoring the null hypotheses.

The failure of this series of experiments to detect ABA renewal following RIF leads us to suggest that the mechanisms which underlie RIF may not fully parallel those responsible for retroactive associative outcome interference. A recent study by Vadillo et al. (2013) offers an interesting perspective on our findings. Their experiment sought to assess the mechanisms underlying RIF, and it found compelling evidence for mechanisms that parallel those of retroactive outcome interference. However, their design differed markedly from ours in that they used a Pavlovian task and found RIF-like phenomena, whereas we used a RIF task and sought Pavlovian phenomena. Their design, while exceptionally clever, diverges from traditional RIF preparations in several major ways: 1) Phase 2 involved learning a new pair of words (the Phase 1 cues with new outcomes) instead of receiving retrieval practice with an already learned pairs of words and 2) participants were tested for retarded acquisition of an association between a novel cue and the inhibited outcome from Phase 1, whereas the traditional RIF test phase prompts participants to recall previously presented exemplars (i.e., outcomes). Although their study yielded important new information about potential RIF mechanisms, we recommend that this approach be applied within a conventional RIF procedure.

Our data also speak to a widely studied phenomenon in human memory known as encoding-retrieval match (e.g., Tulving, 1983). This hypothesis posits that items are better remembered when testing takes place in the context of learning. This could be extended to support our [empirically refuted] hypothesis in the present study. In agreement with this theory, we expected that when the test context matched the context of encoding (Context A), recall at test would be more indicative of the learning that took place in Phase 1 than the interference learning that presumably took place during Phase 2. According to this logic, individuals in Group ABA of Experiment 1 should have shown a reduced RIF effect. However, as was seen, this reduction in RIF was not observed in Experiment 1, nor was it

seen in Experiment 2 after we implemented changes to make the two contexts more distinct. Therefore, the encoding-retrieval hypothesis is not supported by the present data, although it has been widely supported by prior research in other preparations (e.g., Morris, Bransford, and Franks, 1977; Tulving, 1983). Extrapolating from Nairne (2002), it may be that our manipulation of context simply did not provide sufficient discriminative cues to allow the presumably inhibitory processes that suppressed recall of Rp– items to be specific to a given physical context (also see, Leon, Abad, and Rosas, 2010).

Some limitations to our study preclude our drawing definitive conclusions without further research. One such limitation concerns our differentiation of contexts. Although we took measures to increase the dissimilarity of Contexts A and B in Experiment 2, our failure to attenuate RIF allows the possibility that our physical contexts still may not have been perceived by the participants as sufficiently different to engender renewal. Due to time constraints, the contexts were necessarily placed in two rooms within the same building, which may have reduced the perceived difference between the two contexts. Prior studies using rats have demonstrated acute sensitivity to even the most subtle changes in context, such as light, orientation, and background noise in other outcome interference designs (e.g., Polack, Laborda, and Miller, in press). Although humans have generally shown less sensitivity to such variables, there is ample evidence from our own laboratory and others to suggest that outcome interference in humans can be affected by subtle contextual changes as well, either overtly or implicitly (e.g., Rosas and Callejas-Aguilera, 2006; Gonzalez-Martin, Cobos, Moris, and Lopez, 2012). Furthermore, with a more extreme view of ‘context,’ we may consider that in most RIF experiments the somewhat different tasks presented to the subjects in Phase 1, Phase 2, and Test may be considered contextual changes as well. This perspective has received support from studies that have failed to observe RIF when the retrieval practice context is not activated at test (e.g., Jonker, Seli, and MacLeod, in press). However, our focus in the present research was on determining the impact of a physical (i.e., spatial) context change on the mechanism underlying RIF phenomena, as such a context shift is analogous to that which results in renewal (i.e., contextual control) in retrospective associative outcome interference preparations which are also assumed to arise from inhibitory processes created during Phase 2 (e.g., Bouton and Bolles, 1979).

An additional potential shortcoming of this research concerns the length of our retention interval. As previously mentioned, there is some disagreement in the literature regarding the durability of the RIF effect. A number of studies have found RIF to be far stronger at 0–20 min following Phase 2, but greatly reduced at 24 h, which is why in Experiment 2 we increased the retention interval between Phase 2 and testing to only 20 min. That is, we feared that a longer retention interval would have eliminated RIF, leaving no deficit on which renewal might act. In retrospect, even 20 min (compared with the 2-min retention interval of the Immediate condition) seems to have largely eliminated RIF in our preparation. Perhaps a retention interval of more than 2 min but less than 20 min would have revealed renewal. Of course with the design used we have no way of knowing whether the tendency of reduced RIF with increased retention interval was due to the passage of time during the retention interval per se or the distractor task that filled most of the 20-min retention interval. Different experiments in the literature have used diverse retention intervals, with most of them being between 5–30 min (e.g., Carroll et al., 2007; Chan, 2009; Garcia-Bajos et al., 2009; MacLeod and Macrae, 2001; Migueles and Garcia-Bajos, 2007; Saunders and MacLeod, 2002; Saunders et al., 2009; Storm et al., 2006). Most of these studies obtained RIF, although they did not provide comparisons across diverse retention intervals. In contrast, with longer retention intervals (e.g., 24 h), recovery from RIF has often been observed (e.g., Carroll, Campbell-Ratcliffe, Murnane, and Perfect, 2007; Chan, 2009; MacLeod and Macrae, 2001; Saunders and MacLeod, 2002; Saunders, Fernandes, and Kosnes, 2009).

A final shortcoming regards the testing procedure we used to assess the context specificity of the putative inhibitory mechanism underlying RIF. At test, subjects were allowed to recall the exemplars (Rp+, Rp-, and NRP) in any order. Because they likely first activated Rp+ items prior to Rp- items, the activation of the Rp+ item first could have disrupted recall of Rp- items as a result of output interference, a mechanism that would not be expected to be affected by the similarity of the test context to any context used during training. Nevertheless, several reports that controlled for output interference have found RIF effects equivalent to those obtained in the studies from which we borrowed our procedures and parameters (for a review, see Storm and Levy, 2012).

In summary, there are both theoretical and empirical reasons to expect renewal following RIF. Empirical evidence includes prior reports of spontaneous recovery from RIF (over 24 h intervals) and the observation of renewal in numerous retroactive outcome interference preparations. Nevertheless, we failed to observe any renewal following RIF. Moreover, our Bayesian statistics favored there being no renewal following RIF. Our data speak to the necessity of further research in order to better understand the nature of RIF mechanisms and the appropriateness of an analogy between RIF and conventional retroactive outcome interference. A better understanding of the processes underlying RIF and retroactive outcome interference will help researchers condense numerous analogous paradigms into more parsimonious theories with greater explanatory power across various preparations.

Acknowledgments

Gonzalo Miguez was supported by the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT-Chile) and the Fulbright Program. The authors would like to thank Julia Soares for her comments on an earlier version of this manuscript, and Henry Cham for his help in collecting the data.

References

- Anderson MC. Rethinking interference theory: Executive control and the mechanisms of forgetting. *J Mem Lang*. 2003; 49:415–445.
- Anderson MC, Bjork RA, Bjork EL. Remembering can cause forgetting: retrieval dynamics in long-term memory. *J Exp Psychol Learn Mem Cogn*. 1994; 20:1063–1087. [PubMed: 7931095]
- Anderson MC, Bjork EL, Bjork RA. Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychol Bull & Rev*. 2000; 7:522–530.
- Anderson, MC.; Levy, BJ. On the relation between inhibition and interference in cognition. In: Benjamin, A., editor. *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork*. London, England: Psychology Press; 2010. p. 107-132.
- Anderson MC, Spellman BA. On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychol Rev*. 1995; 102:68–100. [PubMed: 7878163]
- Bjork, RA.; Bjork, EL. A new theory of disuse and an old theory of stimulus fluctuation. In: Healy, AF.; Kosslyn, SM.; Shiffrin, RM., editors. *Essays in honor of William K. Estes, Vol. 2: From learning processes to cognitive processes*. Erlbaum; Hillsdale, NJ: 1992. p. 35-67.
- Bouton ME. Context, time, and memory retrieval in the interference paradigms of Pavlovian conditioning. *Psychol Bull*. 1993; 114:80–99. [PubMed: 8346330]
- Bouton, ME. The multiple forms of “context” in associative learning theory. In: Mesquita, B.; Barrett, LF.; Smith, ER., editors. *The Mind in Context*. The Guilford Press; New York: 2010. p. 233-258.
- Bouton ME, Bolles RC. Contextual control of the extinction of conditioned fear. *Learn Motiv*. 1979; 10:445–466.
- Bouton ME, King DA. Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *J Exp Psychol: Anim Behav Process*. 1983; 9:248–265. [PubMed: 6886630]
- Bouton ME, Peck CA. Spontaneous recovery in crossmotivational transfer (counterconditioning). *Anim Learn Behav*. 1992; 20:313–321.

- Carroll M, Campbell-Ratcliffe J, Murnane H, Perfect T. Retrieval-induced forgetting in educational contexts: Monitoring, expertise, text integration, and test format. *Eur J Cog Psychol.* 2007; 19:580–606.
- Chan JCK. When does retrieval induce forgetting and when does it induce facilitation? implications for retrieval inhibition, testing effect, and text processing. *J Mem Lang.* 2009; 61:153–170.
- Ebbinghaus, H. *Memory: A contribution to experimental psychology.* Ruger, HA.; Bussenius, CE., translators. New York: Dover; 1885/1964. (Original work published 1885)
- García-Bajos E, Migueles M, Anderson MC. Script knowledge modulates retrieval-induced forgetting for eyewitness events. *Memory.* 2009; 17:92–103. [PubMed: 19058092]
- Jenkins JG, Dallenbach KM. Obliviscence during sleep and waking. *A J Psychol.* 1924; 35:605–612.
- Jonker TR, Seli P, MacLeod CM. Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review.* (in press).
- González-Martín E, Cobos PL, Morís J, López FJ. Interference between outcomes, spontaneous recovery, and context effects as measured by a cued response reaction time task: Evidence for associative retrieval models. *J Exp Psychol: Anim Behav Process.* 2012; 38:419–432. [PubMed: 22905829]
- Laborda MA, Miller RR. Preventing return of fear in an animal model of anxiety: additive effects of massive extinction and extinction in multiple contexts. *Behav Ther.* 2013; 44:249–261. [PubMed: 23611075]
- León SP, Abad MJF, Rosas JM. Giving contexts informative value makes information context-specific. *Exp Psychol.* 2010; 57:46–53. [PubMed: 20176550]
- McGeoch JA. Forgetting and the law of disuse. *Psychol Rev.* 1932; 39:352–370.
- MacLeod MD, Macrae CN. Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychol Sci.* 2001; 12:148–152. [PubMed: 11340924]
- Migueles M, García-Bajos E. Selective retrieval and induced forgetting in eyewitness memory. *App Cog Psychol.* 2007; 21:1157–1172.
- Miller RR, Escobar M. Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behav Process.* 2002; 57:163–185.
- Miller RR, Laborda MA. Preventing recovery from extinction and relapse: A product of current retrieval cues and memory strengths. *Curr Dir Psychol Sci.* 2011; 20:325–329.
- Morris CD, Bransford JD, Franks JJ. Levels of processing versus transfer appropriate processing. *J Verb Learn Verb Behav.* 1977; 16:519–533.
- Nairne JS. The myth of the encoding-retrieval match. *Memory.* 2002; 10:389–395. [PubMed: 12396651]
- Ortega-Castro N, Vadillo MA. Retrieval-induced forgetting and interference between cues: Training a cue-outcome association attenuates retrieval by alternative cues. *Behav Process.* 2013; 94:19–25.
- Pavlov, IP. *Conditioned Reflexes.* Clarendon Press; London: 1927.
- Peck CA, Bouton ME. Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learn Motiv.* 1990; 21:1–31.
- Pineño O, Matute H. Interference in human predictive learning when associations share a common element. *Int J Comp Psychol.* 2000; 13:16–33.
- Polack CW, Laborda MA, Miller RR. On the differences in degree of renewal produced by the different renewal designs. *Behav Process.* in press.
- Raaijmakers JGW, Jakab E. Retrieval-induced forgetting without competition: Testing the retrieval specificity assumption of the inhibition theory. *Mem Cog.* 2012; 40:19–27.
- Raaijmakers JGW, Jakab E. Is forgetting caused by inhibition? *Curr Dir Psychol Sci.* 2013; 22:205–209.
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical conditioning: II. Current theory and research.* Appleton-Century Crofts; New York, United States: 1972. p. 64-99.
- Rosas JM, Bouton ME. Context change and retention interval can have additive, rather than interactive, effects after taste aversion extinction. *Psychon Bull Rev.* 1998; 5:79–83.

- Rosas JM, Callejas-Aguilera J. Context switch effects on acquisition and extinction in human predictive learning. *J Exp Psychol: Learn Mem Cog*. 2006; 32:461–474.
- Rosas JM, Vila MJ, Lugo M, Lopez L. Combined effect of context change and retention interval on interference in causality judgement. *J Exp Psychol: Anim Behav Process*. 2001; 27:153–164. [PubMed: 11296490]
- Rouder JN, Speckman PL, Sun D, Morley RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev*. 2009; 16:225–237. [PubMed: 19293088]
- Saunders J, Fernandes M, Kosnes L. Retrieval-induced forgetting and mental imagery. *Mem Cog*. 2009; 37:819–828.
- Saunders J, MacLeod MD. New evidence on the suggestibility of memory: The role of retrieval-induced forgetting in misinformation effects. *J Exp Psychol: App*. 2002; 8:127–142.
- Storm, BC. Retrieval-induced forgetting and the resolution of competition. In: Benjamin, A., editor. *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork*. London, England: Psychology Press; 2010. p. 89-105.
- Storm BC, Bjork EL, Bjork RR. On the durability of retrieval-induced forgetting. *J Cog Psychol*. 2012; 24:617–629.
- Storm BC, Bjork EL, Bjork RA, Nestojko JF. Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychol Bull Rev*. 2006; 13:1023–1027.
- Storm BC, Levy BJ. A progress report on the inhibitory account of retrieval-induced forgetting. *Mem Cog*. 2012; 40:827–843.
- Tulving, E. Relation between encoding specificity and levels of processing. In: Cermak, LS.; Craik, FIM., editors. *Levels of processing in human memory*. Hillsdale, NJ: Erlbaum; 1979. p. 405-428.
- Tulving E, Psotka J. Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *J Exp Psych*. 1971; 87:1–8.
- Vadillo MA, Orgaz C, Luque D, Cobos PL, López FJ, Matute H. The role of outcome inhibition in interference between outcomes: a contingency-learning analogue of retrieval-induced forgetting. *Br J Psychol*. 2013; 104:167–180. [PubMed: 23560664]
- Verde, MF. Retrieval-induced forgetting and inhibition: A critical review. In: Ross, BH., editor. *Psychology of learning and motivation*. Vol. 56. Waltham, MA: Academic Press; 2012. p. 47-80.
- Verde MF. Retrieval-induced forgetting in recall: Competitor interference revisited. *J Exp Psychol: Learn Mem Cog*. 2013; 1037/a0032975
- Wixted JT. A theory about why we forget what we once knew. *Curr Dir Psychol Sci*. 2005; 14:6–9.

Highlights

- Retrieval-induced forgetting is similar to associative retroactive outcome interference.
- Recovery from retroactive outcome interference is seen when testing occurs outside the interference context (i.e., renewal).
- We failed to see renewal following RIF despite Bayesian analysis indicating that we were sensitive.

Figure 1. Proportion of Rp+, Rp-, and Nrp words recalled in Experiment 1. Vertical bars denote standard error of the means. See text for details.

Figure 2. Proportion of Rp+, Rp-, and Nrp words recalled by Groups Immediate (Imm) and Delay (Del) in Experiment 2. Vertical bars denote standard error of the means. See text for details.

Table 1

Groups	---Phase 1 (1 block)---	---Phase 2 (3 blocks)---	-----Test-----
ABA	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp+] pairs (Context B)	Exemplar cued recall by category (Context A)
ABB	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp+] pairs (Context B)	Exemplar cued recall by category (Context B)

Note: Design summary of Experiment 1

Table 2

Groups	---Phase 1 (1 block)---	---Phase 2 (3 blocks)---	Delay	-----Test-----
ABA-Imm	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp +] pairs (Context B)	No delay	Exemplar cued recall by category (Context A)
ABB-Imm	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp +] pairs (Context B)	No delay	Exemplar cued recall by category (Context B)
ABA-Delay	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp +] pairs (Context B)	20 min	Exemplar cued recall by category (Context A)
ABB- Delay	Category - Exemplar [Rp+, Rp-, Nrp] pairs (Context A)	Category - Stem cued exemplar [Rp +] pairs (Context B)	20 min	Exemplar cued recall by category (Context B)

Note: Design summary of Experiment 2