

**Retrieval-Induced Forgetting:
Testing the Competition Assumption of Inhibition Theory**

by

Tanya R. Jonker

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Arts
in
Psychology

Waterloo, Ontario, Canada, 2011

© Tanya R. Jonker 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Practicing the retrieval of some information can lead to poorer retrieval of other related information; this phenomenon is called retrieval-induced forgetting. This pattern has been explained as the result of inhibition of the related information during retrieval practice (Anderson, 2003). A core assumption of this inhibition account is that, to be suppressed, the related information must compete with the target information at the time of practice. Four experiments are reported that test this competition assumption. Two experiments showed that retrieval-induced forgetting did not occur without specific retrieval practice of the target items, replicating and extending prior findings. Two further experiments then showed that retrieval-induced forgetting did occur, however, when competition between target information and related information during retrieval practice was eliminated, undermining the competition assumption and hence the inhibition account. A new explanation of retrieval-induced forgetting is introduced that emphasizes context change between study, retrieval practice, and test.

Keywords: retrieval-induced forgetting; memory; inhibition theory; competition assumption

Acknowledgements

I am grateful to my supervisor, Dr. Colin M. MacLeod, for his insightful comments into methodology development, for providing me with opportunity to learn from his exceptionally clear writing style, and for his Socratic guidance through the interpretation of results.

I extend further thanks to my thesis readers, Drs. Jonathan Fugelsang and Myra Fernandes, for their thoughtful comments.

Finally, I would also like to acknowledge the financial support of my Masters work by the Natural Sciences and Engineering Research Council, in the form of a postgraduate scholarship.

Table of Contents

I.	Introduction	p. 1
II.	Pilot Experiment	p. 9
III.	Experiment 1	p. 13
IV.	Experiments 2 and 3	p. 17
V.	General Discussion	p. 20
	References	p. 28

Forgetting can be a distressing event. Much time is wasted trying (often unsuccessfully) to recall the name of an acquaintance or the items written on the grocery list that was left at home. However frustrating these lapses of memory may be, our processing system functions to reduce the influence of interfering information so as to facilitate the processing and recall of relevant information. Indeed, constantly remembering previous experiences that are not relevant to our present situation would be quite distracting and even disturbing. Our ability to forget is an adaptive function (see Bjork, 1989), one that we may not become aware of until it fails us.

So, if forgetting is adaptive, what mechanisms aid in the forgetting of irrelevant (and remembering of relevant) material? Perhaps forgetting occurs because of the activation and strengthening of other material. For example, when you want to remember where you parked your car this morning, you rehearse it or tie the location to some contextual cues to help you remember. Thus, today's parking spot may be more memorable (due to recency, contextual cues, or other mnemonic cues), and therefore it may be more difficult to remember where you parked your car yesterday. In this case, forgetting of yesterday's parking spot is a by-product of the strong memory for today's spot. This is passive forgetting. Alternatively, perhaps when you attempt to remember today's parking spot, the location of yesterday's spot is suppressed to reduce its interference. In this case, forgetting of yesterday's parking spot is part of the process of remembering. This is a more active form of forgetting.

The retrieval-induced forgetting (RIF) paradigm has become popular among memory researchers because it provides an arena for testing these two possible forms of forgetting. This type of forgetting—RIF—seems rather paradoxical. Not surprisingly, the act of retrieval improves later memory for the retrieved items (e.g., Darley & Murdock, 1971; McDaniel & Masson, 1985; McDaniel, Roediger, & McDermott, 2007; Roediger & Karpicke, 2006). But it

turns out that this act of retrieval also influences memory for related but non-retrieved items. Such retrieval actually can result in a cost for related items (Anderson, Bjork, & Bjork, 1994).

Investigating Retrieval-Induced Forgetting

The paradigm typically used to investigate RIF was developed by Anderson et al. (1994). It involves studying several sets of category-exemplar pairs taken from different categories—for example, six exemplars from each of six categories. Thus, participants might have studied *FRUIT – apple*, *FRUIT – orange*, *DRINK – wine*, and *DRINK – rum* at random points throughout a 36-item list. This study phase is followed by retrieval practice of half of the items from half of the categories. So, participants might then practice retrieving *apple* from the *FRUIT* category, but not *orange*; in this case, they would not practice any of the items from the *DRINK* category. To practice retrieving *apple*, participants would see the category label along with a two-letter word stem and would be asked to complete the word (*FRUIT – ap__*).

This retrieval practice phase therefore creates three distinct groups of items. By convention, the items that are practiced during the retrieval practice phase are given the shorthand Rp+ (e.g., in the preceding example, *apple* would be an Rp+ item). Unpracticed items from practiced categories are given the shorthand Rp- (e.g., *orange* would be an Rp- item). Items from the categories in which no items were practiced are given the shorthand Nrp and provide a baseline because no retrieval occurs in those categories (e.g., *rum* and *wine* both would be Nrp items).

The experiment by Anderson et al. (1994), as well as other variations on the paradigm (see Anderson, 2003, for a review), have produced two basic findings. The first finding is entirely intuitive—enhanced recall for Rp+ items relative to Nrp item. Practice helps the practiced items. The second finding is the one of interest—impaired recall for Rp- items relative

to Nrp items. Practice hurts the unpracticed items. For example, practicing the item *FRUIT – apple* during a retrieval practice phase will impair the later recall of the unpracticed item *FRUIT – orange* relative to the recall of *DRINK – rum* (i.e., $Rp+ > Nrp > Rp-$).

Since the original RIF demonstration by Anderson et al. (1994), researchers have investigated many variations of the paradigm, contributing to a by now quite extensive body of literature. RIF has been demonstrated using many different types of stimulus sets, such as coloured shapes (Ciranni & Shimamura, 1999), personality traits (Macrae & M. MacLeod, 1999), stolen items from a crime scene (M. MacLeod & Saunders, 2008; Shaw, Bjork, & Handal, 1995), and second-language learning (Levy, McVeigh, Marful, & Anderson, 2007). Regardless of the stimulus type, these studies have all demonstrated that RIF occurs among related exemplars. Furthermore, Rp- items have been shown to be especially vulnerable to misinformation effects (M. MacLeod & Saunders, 2008), and tend to be impaired relative to Nrp items even on recognition tests (Hicks & Starns, 2004; M. MacLeod & Saunders, 2008; Starns & Hicks, 2004).

The first and still dominant explanatory framework for RIF is the theory of memory inhibition (for reviews, see Anderson, 2003; Anderson et al, 1994; Anderson, Bjork, & Bjork, 2000; Anderson & Spellman, 1995; Bäuml, Pastötter, & Hanslmayr, 2010). According to inhibition theory, the cue given during the retrieval practice phase activates not only the target exemplar but also the other studied exemplars from the same category. These related exemplars then compete for retrieval (e.g., upon seeing *FRUIT*, the exemplars *apple*, *orange*, *plum*, etc. all become activated and compete for retrieval). This competition causes significant interference; to restrict retrieval to the target exemplar, the mental representations of competing exemplars are suppressed, or inhibited. This reduces the interference that they cause and allows for successful

retrieval of the target. Critically, however, this inhibition is enduring such that later recall of Rp-items is impaired; the inhibited competitors have become less accessible in memory.

The strongest empirical support for inhibition theory comes from demonstrations involving exemplar strength, similarity, and especially independent cuing. Anderson et al. (1994) found RIF among Rp- exemplars with high taxonomic frequency (e.g., *orange* from the category *FRUIT*), but not among Rp- exemplars with lower taxonomic frequency (e.g., *guava*); and this was the case even when output interference was controlled. They explained their differential strength findings by suggesting that competition mediates the impairment of Rp-items. That is, if Rp- items are strongly associated exemplars from the category, they will compete intensely for retrieval and will therefore be inhibited. On the other hand, weakly associated exemplars do not compete as strongly, if at all, so they do not need to be inhibited. This notion of “reactive inhibition” goes back to Wundt (1902).

Manipulations of item similarity, whether semantic or episodic, also support the competition component of inhibition theory. On the semantic side, Bäuml and Hartinger (2002) demonstrated that RIF did not occur when Rp+ and Rp- exemplars were highly similar, and Goodmon and Anderson (2011) showed that RIF occurred when there were few associations between the Rp- and Rp+ items but not when there were many associations between them. Likewise, on the episodic side, Anderson and McCulloch (1999) and Anderson, Green, and McCulloch (2000) demonstrated that RIF does not occur when exemplars are strongly interrelated. According to inhibition theory, in these situations of item inter-relatedness, the mental representations of the competitors and the target items overlap and the unpracticed competitors therefore benefit from the strengthened features of the practiced targets (see Anderson & Spellman, 1995, for a theoretical explanation of feature overlap).

The third—and most compelling—vein of support for inhibition theory is the demonstration of independent cuing. Independent cues are test cues given at final test that were not presented during the study phase (for the origin of this manipulation, see Anderson & Spellman, 1995). Johnson and Anderson (2004) executed this method by presenting implicit, unstudied category cues in conjunction with one-letter word-stems during the final test phase. For example, the target word *salt* was studied with the category *SEASONING*, but during the final test phase, the recall cue *POPCORN – s_____* was given. Even though the original study cue was not used at test, RIF occurred (see also Saunders & M. MacLeod, 2006). This cost for an independent cue has thus far been the strongest evidence that inhibition—not just interference on test—plays a central role in RIF (see Anderson 2003, Anderson & Levy, 2007, for more on this argument), although it should be noted that independent cue data are in fact only rarely reported in the RIF literature.

It is difficult for other classic explanations of forgetting, like blocking theory, to account for these findings (see Anderson, 2003, for a review). Furthermore, inhibition theory is supported by a number of other empirical findings. For example, RIF has been demonstrated using recognition tests (e.g., Hicks & Starns, 2004; Spitzer & Bäuml, 2007; Starns & Hicks, 2004), suggesting that mental representation itself is suppressed as opposed to the association between the category and exemplar. Furthermore, a positive correlation between working memory capacity and susceptibility to RIF supports the assertion that executive control suppresses competitors during practice (Aslan & Bäuml, 2010; see also Roman, Soriano, Gómez-Ariza, & Bajo, 2009). The role of executive control is further corroborated by neuropsychological evidence (e.g., Kuhl, Dudukovic, Kahn, & Wagner, 2007)

Although inhibition theory is prominent in RIF literature, some researchers have challenged the theory (e.g., C. MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003). For example, the findings regarding exemplar strength and independent cuing have not always replicated (see Williams & Zacks, 2001). Furthermore, some researchers have challenged the idea of independent cuing, suggesting that the cues used are not in fact “independent” but instead can be related through covert cuing (Camp, Pecher, & Schmidt, 2007; Camp, Pecher, Schmidt, & Zeelenberg, 2009). Others have found that RIF is not always present and seems to be cue-dependent based on the retrieval practice and test cues presented (Jonker, Seli, & C. MacLeod, under review; Perfect, Stark, Tree, Moulin, Ahmed, & Hutter, 2004). Moreover, others have provided manipulations of strength by varying the number of presentations during study and examining primacy and recency strengthening, and have concluded that an experimentally controlled manipulation of strength as opposed to taxonomic strength does not support the inhibition account (Jakab & Raaijmakers, 2009).

These findings challenge inhibition theory, but it is of course difficult to evaluate a theory when the conflicting evidence is based primarily on null findings. This thesis takes the complementary approach by seeking to demonstrate RIF under conditions where inhibition would predict none.

The series of experiments in this thesis tests the competition assumption of inhibition theory, central to the theory since its beginning (Anderson et al., 1994). In his theoretical article, Anderson (2003, p. 420) describes competition as a necessary condition for inhibition: “Inhibition is driven by the need to override interference from competing memories during the selective retrieval of target items.” To explain this process during retrieval practice, Anderson et al. (p. 1079) postulate that “presenting a cue should activate all associated responses in parallel;

this initial spread of activation may then need to be focused to isolate the target response from interfering competitors,” the mechanism of focus being inhibition. Thus, when a category cue is presented, strongly-associated exemplars will become activated and compete for retrieval, causing interference. To facilitate retrieval, the competitors must be suppressed. This inhibition is, however, reactive, so a mental representation that does not cause competition will not require suppression and therefore should not show RIF.

A test of the competition assumption required a manipulation that would not cause competition between the exemplars. The manipulation developed here involves generation of a subordinate item from a studied exemplar. Critically, this manipulation is performed during the retrieval practice phase at the level of the exemplar without displaying the category cue. Bäuml (2002) used a generation task in the place of retrieval practice, but his task involved generating additional exemplars when given the category label. Thus, following the study of *PET – dog*, participants in Bäuml’s study might be given the category word *PET* and asked to come up with additional—yet unstudied—exemplars, like “fish” or “snake.” During Bäuml’s task, this type of generation should, ostensibly, activate studied exemplars in response to the category cue; these studied exemplars should then need to be suppressed.

In contrast, during the generation task used here, participants might see the exemplar *dog* (instead of the category word *PET*) and be asked to generate a type of dog, like “beagle.” During this task, there is no need for other studied exemplars of the category *PET* to become activated and compete for retrieval. Yet retrieval is still involved, in that an instance of that exemplar—a subordinate such as beagle—must be retrieved from memory. This manipulation is used to test the competition assumption. Experiment 1 employs just the generation task, while Experiments 2 and 3 employ the generation task in conjunction with a category retrieval task. In the absence

of competition among the studied exemplars, inhibition theory predicts no RIF. Indeed, Anderson (2003, p. 428) states that “any type of retrieval practice that minimizes the need to resolve interference between competing items is unlikely to produce inhibition.” The experiments in this thesis provide a direct test of this assertion.

Pilot Experiment

This pilot experiment was a close approximation of the non-competition condition of Anderson et al. (2000). In their study, they showed that when participants were prompted to retrieve the category (e.g., *Fr___ – apple*) instead of the exemplar (e.g., *Fruit – ap___*), there was no cost of Rp- items, though there was a benefit for the items serving as retrieval cues (*apple*). The goal was to provide converging evidence for the finding that category retrieval alone does not result in RIF, while ensuring generalizability to the (only slightly modified) version of their task used in this thesis.

Method

Participants. Participants were 5 males and 22 females, with ages ranging from 18 to 23 ($M = 19.9$). They were recruited from the University of Waterloo's Research Experience Group and offered bonus course credit for their participation. All had normal or corrected-to-normal vision and English was their first and most fluent language.

Materials. The same stimulus set was used for this and all subsequent experiments. Stimuli consisted of six categories of words (e.g., *FOOD* and *PET*), each with six exemplars (e.g., *dog*, *horse*), quite analogous to previous RIF studies. Two filler categories were also selected, each with three exemplars. Thus, there was a total of 42 category-exemplar word pairs (36 experimental, 6 filler). No exemplars within any category began with the same letter. Stimuli for all tasks in all experiments were presented on a 17-in CRT monitor using E-Prime programming software. They were presented against a black background in white 24-point Times New Roman font at the center of the screen.

Procedure. During the initial study phase, participants saw category-exemplar word pairs individually on the computer screen and were asked to study the pairs for a later memory

test. Each pair was presented for 5 s with an inter-stimulus interval of 250 ms between successive word pairs. Stimuli were presented in a constrained random order such that category-exemplar pairs from the same category were never presented in succession. Three of the filler word pairs were presented before the experimental stimuli, and three after, to limit primacy and recency effects (Murdock, 1962).

Following study, participants performed a category retrieval task (i.e., the retrieval practice phase). This task was similar to the non-competitive condition in Anderson et al. (2000); participants retrieved the category word instead of an exemplar. On each trial, participants were shown a studied exemplar and were asked to type in the studied category to which that exemplar belonged. This task should not cause competition between related exemplars because the retrieval cue is an exemplar for which there is only one category name (Anderson, 2003; Anderson et al., 2000). This task was performed for half of the exemplars from half of the categories. The selection of categories and exemplars for category retrieval was random for each participant, and the order of presentation was constrained such that exemplars from the same category were never presented in succession. Thus, three of the categories became retrieval categories, and three of the exemplars in each of these categories were presented, totaling nine exemplars. The category-retrieval task was completed on these nine exemplars three times, totaling 27 retrieval practice trials. The presented exemplars were given the typical notation Rp^+ , the unpracticed exemplars from the same category as the presented exemplars were labeled Rp^- , and the exemplars from categories with no category retrieval were labeled Nrp .

A distractor task followed the category-retrieval phase. All experiments reported in this article used a distractor task borrowed from Macrae and Roeseveare (2002), in which participants were given 5 min to make a list of as many countries as possible.

The final cued recall test involved a control for output interference (see Roediger, 1974). Output interference controls have been used in RIF studies to rule out the possibility that the effect occurs because of the output of Rp+ items before Rp- items on a category-cued free recall test (M. Anderson et al., 1994). To control for output interference, all Rp- items within a category were tested before Rp+ items. The final recall test employed category cues along with one-letter word stems (*PET - d ___*) to control the order of recall. Cues were presented individually and remained in view until the participant responded or until 10 s elapsed; presentation was blocked by category.

Results and Discussion

The alpha level for statistical significance for all experiments was set at .05. All recall data were analyzed using a one-way repeated-measures analysis of variance (ANOVA) along with two planned comparisons using repeated-measures *t*-tests to examine the potential benefit (Rp+ > Nrp) and cost (Rp- < Nrp) of retrieval practice.

Participants retrieved the studied category on 93.0% of the retrieval practice trials.

Overall, recall for Rp+, Rp-, and Nrp items differed, $F(2,52) = 10.61$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .29$. The first planned comparison revealed a significant retrieval practice benefit, $t(26) = 4.06$, $SE = .03$, $p < .001$, $\eta_p^2 = .39$; that is, participants were able to recall a greater proportion of Rp+ items ($M = .74$) than of Nrp items ($M = .60$). However, the second planned comparison showed that there was no RIF, $t(26) = 0.50$, $SE = .04$, $p = .62$. The proportion of Rp-

items recalled ($M = .58$) did not differ from the proportion of Nrp items recalled. These results are displayed in the upper left panel of Figure 1.

The category retrieval task proved successful in benefiting the Rp+ items relative to the Nrp items, but there was no cost for the Rp- items (i.e., no RIF). These findings replicate those of Anderson et al. (2000) and are consistent with inhibition theory. There should be no RIF because the category-retrieval task should not cause any competition among exemplars when the exemplar itself is given as the retrieval cue for the category.

Experiment 1

With a demonstration of equivalent recall of Rp- and Nrp items when the category-retrieval task is used, rather than the standard exemplar retrieval task, the next step is to introduce the generation task to the RIF paradigm. Here, participants are shown half of the studied exemplars from half of the studied categories and have to generate a subordinate for each exemplar on each practice trial. The rationale for these experiments was that generation of subordinates involves retrieval, but that it does not involve retrieval of other studied exemplars. Consequently, other exemplars should not be competitors in the generation task. According to inhibition theory, there should be no need to inhibit the other exemplars in a category when subordinate generation occurs, and therefore there should be no RIF.

Method

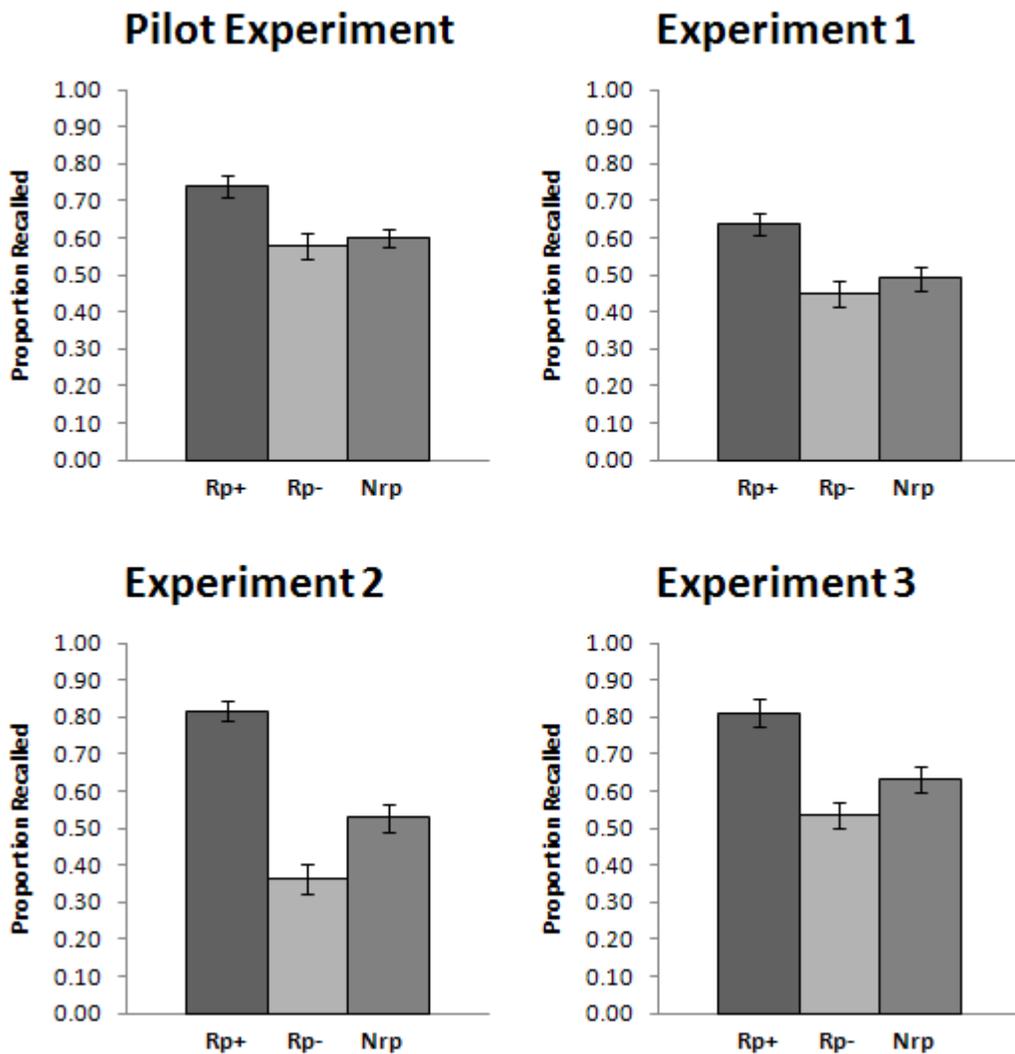
Participants. There were 5 males and 25 females, with ages ranging from 18 to 39 ($M = 19.7$). All participants were recruited from the same pool and with the same constraints as in the pilot experiment.

Materials. Stimuli and testing equipment were identical to those used in the pilot experiment.

Procedure. Participants first studied items in a manner identical to that of the pilot experiment. Following the study phase, participants performed a subordinate generation task (in place of the standard retrieval practice task). They were shown a studied exemplar on the screen and asked to produce a subordinate item. So, if they had studied the pair *PET* – *dog*, during the generation task they might see the exemplar *dog* and then would generate a type of dog, like “beagle.” Participants were given 10 s for each generation and were asked to produce a unique generation each time that they saw the same exemplar (e.g., on the second presentation of *dog*,

Figure 1

Figure 1. The mean proportions of exemplars correctly recalled on the final cued recall test in the pilot experiment and Experiments 1, 2, and 3. The error bars represent one standard error of their respective means. Nrp exemplars are those from categories where no semantic generation (or category retrieval in the pilot study) occurred; these represent the baseline condition. For categories where semantic generation did occur, Rp+ refers to those exemplars involved in subordinate generation and Rp- refers to those exemplars not involved in subordinate generation.



they should generate “poodle” and not repeat “beagle”). They did this generation for half of the exemplars from half of the categories, following the standard procedure for the retrieval practice phase. So, participants were doing retrieval, but it was retrieval of subordinates of exemplars, and not of the exemplars themselves.

The distractor phase followed the generation task; the procedure was identical to that described in the pilot experiment.

The final memory test was a category-cued recall test—one commonly used in RIF studies (Anderson et al., 1994; Anderson & McCulloch, 1999; Anderson & Spellman, 1995; M. MacLeod & Macrae, 2001). Each category label (e.g., *PET*) from the studied pairs was presented on the computer screen one at a time and participants were to write down as many of the studied exemplars (e.g., *dog*, *horse*) from that category as they could. Each category cue was presented for 30 s; a tone alerted participants when the time was up and then the next category appeared on the screen.

Results and Discussion

Participants generated a unique subordinate of a studied exemplar on 92.1% of the trials.

Overall, recall for Rp+, Rp-, and Nrp items differed, $F(2,58) = 15.58$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .35$. The first planned comparison revealed a significant benefit of practice for the practiced items, $t(29) = 4.49$, $SE = .03$, $p < .001$, $\eta_p^2 = .41$; that is, participants recalled more Rp+ items ($M = .64$) than Nrp items ($M = .49$). However, the second planned comparison showed that there was no RIF, $t(29) = 1.12$, $SE = .03$, $p = .27$. The proportion of Rp- items recalled ($M = .45$) did not differ from the proportion of Nrp items recalled. The results are displayed in the upper right panel of Figure 2.

As predicted by inhibition theory, no RIF was observed in this experiment. When the standard retrieval practice task was replaced with a semantic generation task that should not cause competition between related exemplars, no RIF was observed. Perhaps, though, the absence of category cues in the generation task reduced or prevented the strengthening of the association between the category cue and the exemplar. Thus, the memory representation of the exemplar was strengthened because of the generation task, but the exemplar did not cause interference with Rp- items on the final test because the association between the category cue and the Rp+ items was not strengthened. Experiments 2 and 3 test this possibility.

Experiments 2 and 3

To test the hypothesis that the category must appear during generation for RIF to occur, the semantic generation task from Experiment 1 was paired with the category-retrieval task from the pilot experiment. Following generation, participants also had to retrieve the studied category for that exemplar. Neither of these manipulations should result in competition at the level of the exemplars, so inhibition theory would still predict no RIF. So far, the results of these two manipulations support inhibition theory, but combining these manipulations provides a stronger test of the competition assumption and can test the associative explanation.

Method

Participants. In Experiment 2, there were 9 males and 17 females, with ages ranging from 18 to 22 ($M = 18.8$). In Experiment 3, there were 4 males and 22 females, with ages ranging from 18 to 24 ($M = 19.9$; noting that 6 participants were not included in Experiment 2 due to misunderstanding instructions). All participants were recruited from the same pool and with the same constraints as in the pilot experiment.

Materials. Stimuli and testing equipment were identical to those in the pilot experiment.

Procedure. Experiments 2 and 3 were identical with respect to the study phase, the generation plus category-retrieval phase, and the distractor task. They differed only in the testing procedure.

Following the standard study phase, the semantic generation task used in Experiment 1 was performed together with the category-retrieval task used in the pilot experiment (both in place of the standard retrieval practice task). Participants first generated a subordinate item for the exemplar shown to them, and then immediately retrieved the original category name for the exemplar. Thus, having studied *PET – dog* earlier, when shown the exemplar *dog*, participants

first produced a subordinate, “beagle,” and then reported the studied category, “pet.” They did this generation and category retrieval for half of the exemplars from half of the categories, following the standard procedure for the retrieval practice phase. The distractor phase followed immediately.

The memory test for Experiment 2 was a category-cued recall test, identical to that of Experiment 1. That is, each category label (e.g., *PET*) from the studied pairs was presented on the computer screen one at a time and participants were to write down as many of the studied exemplars (e.g., *dog, horse*) from each category as they could. The testing procedure for Experiment 3 was identical to that of the pilot experiment, effectively manipulating testing order to control for output interference. That is, participants were prompted to output Rp- items before Rp+ items, with blocked presentation of items from the same category.

Results and Discussion

Experiment 2. Participants generated a unique subordinate of a studied exemplar on 94.7% of the trials and successfully retrieved the studied category on 92.7% of the trials.

The data are presented in the lower left panel of Figure 1. Overall, recall for Rp+, Rp-, and Nrp items differed, $F(2,50) = 109.96$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .82$. The first planned comparison revealed a significant benefit of practice for the practiced items, $t(25) = 9.71$, $SE = .03$, $p < .001$, $\eta_p^2 = .79$; that is, participants recalled more Rp+ items ($M = .82$) than Nrp items ($M = .53$). More importantly, the second planned comparison showed that RIF was also significant, $t(25) = 5.38$, $SE = .03$, $p < .001$, $\eta_p^2 = .54$; participants recalled fewer Rp- items ($M = .36$) than Nrp items.

Experiment 3. Participants generated a subordinate of a studied exemplar on 90.2% of the trials and successfully retrieved the studied category on 89.7% of the trials.

The data are presented in the lower right panel of Figure 1. Overall, recall for Rp+, Rp-, and Nrp items differed, $F(2,50) = 21.12$, $MSE = 0.52$, $p < .001$, $\eta_p^2 = .46$. The first planned comparison revealed a significant benefit of practice for the practiced items, $t(25) = 4.66$, $SE = .04$, $p < .001$, $\eta_p^2 = .47$; that is, participants recalled more Rp+ items ($M = .81$) than Nrp items ($M = .63$). More importantly, the second planned comparison showed that RIF was also significant, $t(25) = 2.25$, $SE = .04$, $p = .03$, $\eta_p^2 = .17$; participants recalled fewer Rp- items ($M = .53$) than Nrp items.

Experiments 2 and 3 both showed reliable RIF despite there being no need to suppress other exemplars during the modified retrieval practice phase. These two experiments provide replication of each other and also rule out any role for output interference in the observed pattern of results. Because there was no competition among exemplars, there should not have been any cost for Rp- items (i.e., there should not have been RIF). The cost to Rp- items in these experiments calls into question the fundamental competition assumption of the inhibition explanation of RIF.

General Discussion

This series of experiments set out to test a fundamental assumption of the inhibition theory of RIF—that inhibition operates to suppress related competitors that would otherwise undermine retrieval. A pilot experiment replicated the findings of Anderson et al. (2000), where category retrieval failed to produce RIF. According to inhibition theory, retrieving the category name when given the exemplar does not require suppression of other potentially competing exemplars, so impairment does not occur. Then Experiment 1 used a semantic generation task, which involved generating an item subordinate to a studied exemplar; this also failed to produce RIF again. By the inhibition theory, no RIF occurred because other potentially competing exemplars did not need to be suppressed when the exemplar was actually provided. Yet, when the semantic generation task and the category retrieval task were carried out together in Experiments 2 and 3, RIF was observed, even when controlling for output interference (Experiment 3). The presence of RIF without competition at the same level conflicts with the critical competition assumption of inhibition theory.

Inhibition theory has been the dominant explanatory framework for RIF since its initial demonstration by Anderson et al. (1994). Inhibition theory has also been a leading theory for other cognitive effects, like negative priming (e.g., Tipper, 1985, 2001), inhibition of return (e.g., Posner & Cohen, 1984), task switching (e.g., Mayr & Keele, 2000), and directed forgetting (e.g., Basden, Basden, & Gargano, 1993; Weiner & Reed, 1969). However, in spite of its wide-spread popularity, mounting evidence against inhibition theory as an explanation for various cognitive phenomena challenges the validity and reliability of inhibition theory for many of these effects (for a review, see C. MacLeod et al., 2003), at the very least providing plausible non-inhibitory accounts.

The body of literature relating to RIF is no exception. According to Anderson (2003), there are four properties of RIF that provide unique support for an inhibitory mechanism over other strength-based or interference accounts. These properties are cue independence, strength independence, retrieval specificity, and interference dependence. However, as discussed earlier, many researchers have failed to replicate these “pillars” of evidence for inhibition theory.

This line of research targets interference dependence, here called the competition assumption. This assumption holds that competition between related exemplars during the retrieval practice task is necessary for producing inhibition; that is, if an exemplar does not compete—and cause interference—with the target exemplar during retrieval practice, it will not be inhibited (Anderson, 2003; Anderson et al., 1994). To test this competition assumption directly, the standard retrieval practice task was replaced with a semantic generation task that, while requiring semantic retrieval, causes no competition between related exemplars. Indeed, while performing the generation task in Experiment 1, 2, and 3, participants rarely produced one of the other studied exemplars or the studied category (< 1%)¹ instead of a subordinate item.

When this generation task was paired with a category retrieval task (neither of which produced RIF on its own) RIF occurred. This finding also replicated with a control for output interference. In this situation, inhibition theory predicts no RIF because of the absence of

¹ In Experiment 1, 30 participants produced a total of 746 semantic generations (averaging 24.9 generations each); five of these were a production of a studied exemplar in place of a subordinate item (0.7%), and two were a production of the studied category in place of a subordinate item (0.3%). In Experiment 2, 26 participants produced a total of 665 generations (averaging 25.6 generations each); five of these were a production of a studied exemplar in place of a subordinate item (0.8%); participants never produced the studied category in place of a subordinate item. In Experiment 3, 26 participants produced a total of 633 generations (averaging 24.4 generations each); three of these were a production of a studied exemplar in place of a subordinate item (0.5%), and four were a production of the studied category in place of a subordinate item (0.6%).

retrieval interference between exemplars. The conclusion is that competition between exemplars during retrieval practice is not a necessary condition for producing RIF, and therefore that inhibition theory, as it currently stands, fails to account for the present findings. What could instead be responsible for RIF? The earliest articles on RIF quickly ruled out any interference or strength-based accounts (e.g., Anderson et al., 1994; Anderson & Spellman, 1995), however, based on the findings in this thesis, combined with the mounting evidence that challenges inhibition theory, it is time to take a fresh look at this approach, proposing a variation.

It is critical to acknowledge the difference between the role of interference in inhibition theory and the role of interference in an interference-based account of RIF. The difference can be seen most clearly in examining the origins of the impairment. Under inhibition theory, the impairment occurs at the retrieval practice phase and is simply measured during the final test phase. In contrast, under an interference or strength-based account, the impairment occurs at test because the test cues favor the practiced items, causing them to obstruct the recall of the unpracticed items (cf. response competition theory; McGeoch, 1942). An interference or strength-based account takes a very simple approach to RIF. According to this approach, the act of retrieval during the retrieval practice phase strengthens the R_{p+} items. During the final recall test, the test cue (e.g., the category cue) easily triggers the recall of the practiced items. These items dominate recall and are difficult to overcome when trying to access R_{p-} items (see Anderson, 2003, for a review), hence interfering with their retrieval during test.

It is easy to see how an interference theory can explain the findings of Experiment 2. The generation task paired with category retrieval strengthens both the R_{p+} items themselves and the associations between the R_{p+} items and the category cue. Thus, when provided with a category cue on the final cued recall test, R_{p+} items, which are tied strongly to the cue, are easily recalled

and interfere with recall of Rp- items. However, it is less clear how an interference theory would account for the findings of the pilot experiment and Experiment 1, because in both of these studies the Rp+ items were strengthened, as is revealed by the significant benefit for Rp+ items relative to Nrp items, yet the Rp+ items do not seem to be interfering with Rp- items at final test.

Context-Change Hypothesis

To account for this entire set of experiments, a modified interference account is suggested, derived from the list-method directed forgetting literature. During a list-method directed forgetting experiment, participants study one list of items, and are then told either to remember the list for a later test (Remember condition) or to forget the list because it was a practice trial (Forget condition). Participants then study a second list of items which is always to be remembered. Similar to RIF, this manipulation results in both a cost and a benefit: There is a cost for List 1 in the Forget condition relative to List 1 in the Remember condition, and there is a benefit for List 2 in the Forget condition relative to List 2 in the Remember condition (for reviews, see Bjork, Bjork, & Anderson, 1998; C. MacLeod, 1998).

The cost to List 1 in the Forget condition has been explained as the product of an inhibitory mechanism (e.g., Bjork & Bjork, 1996), and for some time this was the dominant explanation. A more recent explanation put forward by Sahakyan and Kelley (2002) looks to contextual factors (see also Bäuml & Samenieh, 2010; Mulij & Bodner, 2010; Pastötter & Bäuml, 2007; Sahakyan, 2004; Sahakyan & Goodmon, 2010). The contextual-change account of list-method directed forgetting suggests that the forget instructions that separate List 1 from List 2 encourage participants to shift their mental context for the presentation of the second list. The impairment for List 1 in the Forget condition results because the test context more closely matches the context of List 2 than that of List 1. Thus, the distinct instructions between List 1

and List 2 make the List 1 items less accessible because these items are associated with a context that is no longer in place. Indeed, when Sahakyan and Kelley induced a context reinstatement manipulation for List 1, they showed a significant reduction in the cost and benefit of directed forgetting, providing support for a contextual-change account.

Based on this account, the postulation is that interference in RIF might come not from the individual items themselves, but from the contextual shift that occurs between the study phase and the subsequent phase (retrieval practice or, in the present experiments, semantic generation). This shift creates two separate learning contexts. The retrieval practice phase—the second context—will have a benefit over the study context when each item practiced during that phase is stronger, and each item is more strongly associated to its category cue. As a result, this second context causes interference, making it difficult to access the study phase.

To make the second context dominant—and hence interfering—the present argument is that the task in the second context must both strengthen the association between category and exemplar and be sufficiently retrieval-demanding (as in the act of retrieval practice or semantic generation, but not study; see Anderson & Bell, 2001; Bäuml, 2002; Ciranni & Shimamura, 1999). When the retrieval practice or semantic generation context is dominant over the study context, this second context causes interference, making it difficult to access the study context. Thus, when presented with a cue for a practiced category, the second context will be cued and the participant will have easy access to the Rp+ items. Because the Rp- items only exist in the study context, however, they are then difficult to access for recall. On the other hand, category cues for Nrp items will only cue the study phase context because no Nrp items exist in the retrieval/generation phase. This account nicely explains the standard RIF findings.

Critically, this account predicts that if the test cues favor the study phase, no RIF will occur. Perfect et al. (2004) used two cues during study, both a category name and a picture of a face. When both cues were used for both retrieval practice and test, RIF did occur. However, when both cues were presented at test, but only one cue (i.e., the category word) had been used for retrieval practice, RIF did not occur. According to the contextual cuing account, in the case of a mismatch between retrieval practice and test cues, test cues will preferentially access the study phase, and the retrieval phase will not be a dominant, interfering context.

This explanation also nicely fits the data reported in this thesis. In Experiment 1, there are no category cues in the semantic generation context, so, although retrieval-strengthening occurs, the category cues presented at test do not preferentially access the semantic generation context (see Figure 2B). Consequently, the context of the semantic generation phase should not interfere with that of the study phase. Indeed, the initial study context may be more similar to the final test context than is the intervening semantic generation context. It is important to note, however, that the semantic generation context is still accessible (as evidenced by the benefit to Rp+ items).

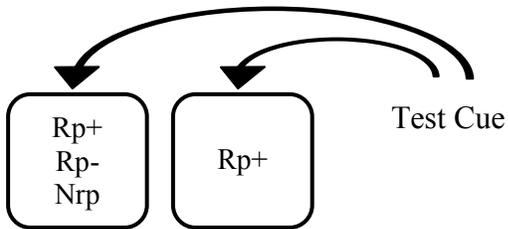
In Experiments 2 and 3, the combination of the semantic generation and category-retrieval tasks during the second phase makes this phase the dominant context due to sufficient strengthening both of the exemplar and of its association to the category cue (see Figure 2C). The test cues for practiced categories then cue this second context, causing interference with the initial study context. As a result, participants have difficulty accessing the context of the study phase containing the Rp- items.

It is less clear what might be happening in the pilot experiment, but it is likely that the study context and the category retrieval context are both quite accessible such that neither causes

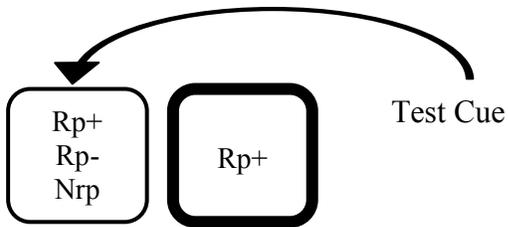
Figure 2

Figure 2. The contextual account of RIF, applied to the experiments reported in this article. Arrows represent cuing, as determined by the use of category cues in the second phase, and bolded boxes represent strengthening during the second phase.

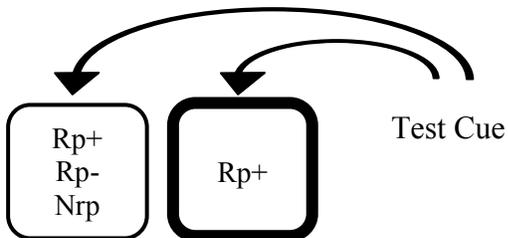
A – Pilot Experiment



B – Experiment 1



C – Experiments 2 and 3



interference with the other because the task of category retrieval is not challenging enough to make the retrieval phase a dominant, interfering context (see Figure 2A). Indeed, Anderson et al. (2000) found that participants were much more successful at category retrieval (99.3%) than at exemplar retrieval (82.9%), suggesting that the category retrieval task is very simple and low in retrieval demands. Category retrieval may promote enough exemplar strengthening to result in a benefit for Rp+ items relative to Nrp items, but not enough category-exemplar strengthening to make the category retrieval phase a dominant—and interfering—context.

In conclusion, the set of experiments reported in this thesis provides a test of a key component of inhibition theory. This thesis has taken the direct approach of testing the competition assumption—a necessary condition—of inhibition theory. In so doing, RIF has been observed even under conditions of no inter-exemplar interference, where inhibition would predict no RIF. These findings add to a body of evidence that challenges the inhibition explanation of RIF. I instead suggest that RIF may be better explained by a contextual interference account, an account that has already been successful in explaining the related phenomenon of directed forgetting.

References

- Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanism of forgetting. *Journal of Memory and Language, 49*, 415-445.
- Anderson, M. C. & Bell, T. (2001). Forgetting our facts: The role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General, 130*(3), 544-570.
- Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7*(3), 522-530.
- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1063-1087.
- Anderson, M. C., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: Evidence for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1141-1159.
- Anderson, M. C. & Levy, B. J. (2007). Theoretical issues in inhibition: Insights from research on human memory. In D. S. Gorfein and C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 81-102). Washington DC: American Psychological Association.
- Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(3), 608-629.
- Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102*(1), 68-100.

- Aslan, A. & Bäuml, K. (2010). Individual difference in Working Memory capacity predict retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Online first publication.
- Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 603-616.
- Bäuml, K. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*, *13*(4), 356-360.
- Bäuml, K. & Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory*, *10*(3), 215-224.
- Bäuml, K., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neuroscience & Biobehavioural Reviews*, *34*(7), 1047-1054.
- Bäuml, K. & Samenieh, A. (2010). The two faces of memory retrieval. *Psychological Science*, *21*(6), 793-795.
- Bjork, E. L., Bjork, R. A., & Anderson, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding and C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103-137). Mahwek, NJ: Lawrence Erlbaum Associates.
- Bjork, E. L. & Bjork, R. A. (1996). Continuing influence of to-be-forgotten information. *Consciousness and Cognition*, *5*(1-2), 176-196.
- Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 950-958.
- Camp, G., Pecher, D., Schmidt, H. G., & Zeelenberg, R. (2009). Are independent probes truly independent? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 934-942.
- Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1403-1414.
- Darley, C. F. & Murdock, B. B. (1971). Effects of prior free recall testing on final recall and recognition. *Journal of Experimental Psychology*, 91(1), 66-73.
- Goodmon, L. B & Anderson, M. C. (2011). Semantic integration as a boundary condition on inhibitory processes in episodic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 416-436.
- Hicks, J. L., & Starns, J. J. (2004). Retrieval-induced forgetting occurs in tests of item recognition. *Psychonomic Bulletin & Review*, 11(1), 125-130.
- Jakab, E. & Raaijmakers, J. G. W. (2009). The role of item strength in retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 607-617.
- Johnson, S. K. & Anderson, M. C. (2004). The role of inhibitory control in forgetting semantic knowledge. *Psychological Science*, 15(7), 448-453.
- Jonker, T. R., Seli, P., & MacLeod, C. M. (under review). Less we forget: Retrieval cues and release from retrieval-induced forgetting. Submitted to *Psychological Science*.

- Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience, 10*, 607-617.
- Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18*(1), 29-34.
- MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding and C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. *The Psychology of Learning and Motivation, 43*, 163-212.
- MacLeod, M. D. & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12*(2), 148-152.
- MacLeod, M. D. & Saunders, J. (2008). Retrieval inhibition and memory distortion: Negative consequences of an adaptive process. *Current Directions in Psychological Science, 17*(1), 26-30.
- Macrae, C. N., & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77*(3), 463-473.
- Macrae, C. N., & Roseveare, T. A. (2002). I was always on my mind: The self and temporary forgetting. *Psychonomics Bulletin & Review, 9*(3), 611-614.
- Mayr, U. & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129*, 4-26.

- McDaniel, M. A. & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(2), 371-385.
- McDaniel, M.A., Roediger, H.L., & McDermott, K.B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14, 200-206.
- McGeoch, J. A. (1942). *The psychology of learning*. New York: Longman, Green.
- Mulij, R. & Bodner, G. E. (2010). Wiping out memories: New support for a mental context change account of directed forgetting. *Memory*, 18(7), 763-773.
- Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482-488.
- Pastötter, D. & Bäuml, K. H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 977-982.
- Perfect, T. J., Stark, L., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language*, 51, 399-417.
- Posner, M. I. & Cohen, Y. (1984). Components of visual orienting. In D. Bouma and D. Bonwhuis (Eds.), *Attention and performance* (Vol. X, pp. 531-566). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition*, 2(2), 261-269.
- Roediger, H. L., & Karpicke, J.D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249-255.

- Román, P., Soriano, M. F., Gómez-Ariza, C. J., & Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychological Science, 20*(9), 1053-1058.
- Sahakyan, L. (2004). Destructive effects of “forget” instructions. *Psychonomic Bulletin & Review, 11*(3), 555-559.
- Sahakyan, L. & Goodmon, L. B. (2010). Theoretical implications of extralist probes for directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 920-937.
- Sahakyan, L. & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1064-1072.
- Saunders, J. & MacLeod, M. D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition, 34*(2), 307-322.
- Shaw, J. S., Bjork, R. A., & Handal A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin & Review, 2*(2), 249-253.
- Spitzer, B. & Bäuml, K.-H. (2007). Retrieval-induced forgetting in item recognition: Evidence for a reduction in general memory strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(5), 863-875.
- Starns, J. J. & Hicks, J. L. (2004). Episodic generation can cause semantic forgetting: Retrieval-induced forgetting of false memories. *Memory & Cognition, 32*(4), 602-609.
- Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37A*, 571-590.

- Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review of integration of conflicting views. *Quarterly Journal of Experimental Psychology*, *54A*, 321-343.
- Weiner, B. & Reed, H. (1969). Effects of the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology*, *79*, 226-232.
- Williams, C. C. & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process? *American Journal of Psychology*, *114*(3), 329-354.
- Wundt, W. (1902). *Grundzüge de physiologischen Psychologie* [Principles of physiological psychology] (5th ed.). Leipzig, Germany: Engelmann.