RETRIEVAL INDUCED FORGETTING IN MENTAL ARITHMETIC

A Thesis

by

CHELSEA MARGARET BRADLEY

Submitted to the College of Graduate Studies of
Tarleton State University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Thomas Faulkenberry, PhD
Committee Members,	Kimberly Rynearson, PhD
	Robert Newby, PhD
Head of Department,	Kimberly Rynearson, PhD
Dean, College of Graduate Studies,	Barry D. Lambert, PhD

August 2018

Major Subject: Applied Psychology

ACKNOWLEDGEMENTS

I would like to first thank Dr. Thomas Faulkenberry, who has spent the past five years serving as my instructor, advisor, sounding board, fountain of encouragement, and consistent reminder that I am not an utter failure (even at times when I believed I might be). Thank you for countless hours of lab work, editing, and grading. Most importantly, thank you for never giving up on me. Your future students will be lucky to learn under you as I did.

Thank you to Dr. Kimberly Rynearson, my department head and committee member. You are a bold and talented scientist with leadership skills to boot. Women like you are so very important to our field. Thank you for your encouragement and for being the kind of person young female researchers can look to as an example of just how far a woman can go when she puts her brilliance to the task.

Thank you to Dr. Robert Newby, whose humor and wit keep our department alive. Thank you for the wisdom you lend to your students, and for being transparent and honest at all times. Your decades of service to this institution form a legacy that will last much longer than awards and grants. Most of all, thank you for serving on my committee and providing much-needed perspective as I moved through this process.

To the rest of the Psychological Sciences faculty and staff – thank you for being the friendly faces, supportive instructors, and incredible coworkers that have shaped me into the student and instructor I am today. Each of you have influenced me in more ways than I can list (because they only allow me two pages for acknowledgements). A special thank you to Carrie Baughn and Katie Holden, my favorite officemates, for the hours of life advice and gentle reassurance that life will be okay in the end.

I owe a massive debt of gratitude to my mother, who paved the way for me to become an educated, hardworking woman from the very beginning. Thank you, mom, for the immeasurable support you provided. You are the reason I am who I am today. All that I have accomplished, all that I have earned, and all that I have yet to experience is because of you.

To the rest of my family and closest friends – thank you for being my retreat from the madness of life. You all keep me sane. Thank you for listening to my fretting over experiments and assignments even when you had no idea what I was talking about. I love all of you, and praise God for each of you.

Finally, thank you to Crowley, my 75 pound mutt whose soft fur and puppy eyes kept me company on my longest and hardest nights. I would not have completed this research without my oversized lap warmer and many, many mugs of hot coffee.

ABSTRACT

Bradley, Chelsea M. Retrieval Induced Forgetting in Mental Arithmetic, MASTER OF SCIENCE (Applied Psychology), August, 2018, 52 pp., 8 tables, 13 figures, bibliography, 24 titles.

Retrieval induced forgetting (RIF) is a memory phenomenon in which the retrieval of one piece of information leads to the forgetting of similar information. RIF has been studied frequently in word recall and, more recently, mental arithmetic. Many experiments in this realm have found evidence for an inhibition mechanism's role in RIF – competition between responses requires the incorrect response to be repressed in favor of the correct response, leading to forgetting of the repressed response. In the present study, we investigated whether the retrieval practice of multiplication problems leads to forgetting of their addition counterparts. Response times were recorded as participants practiced simple addition problems (i.e. 2 + 3 = ?) and then either studied (2 x 3= 6) or retrieved (2 x 3 = ?) their multiplication counterparts. The participants then practiced the addition problems again, and response times for this posttest were compared to those of the pretest to identify whether RIF occurred. While we found substantial problem size effects, we found no evidence of RIF for simple arithmetic problems for either mean RT or median RT. Further, we used ex-Gaussian modeling of the response time distributions to infer potential shifts in strategy (i.e., from direct memory retrieval to a counting procedure), but this analysis also revealed no evidence for RIF. Analysis methods and possible explanations for the lack of effect are discussed.

TABLE OF CONTENTS

LIST OF GRAPHICS

TABLE **Page**

FIGURE **Page**

CHAPTER I

INTRODUCTION

Retrieval Induced Forgetting

There are instances where recall of some items can cause forgetting of similar items, an event called retrieval-induced forgetting (Anderson et al., 1994). The purpose of the present study is to investigate this phenomenon in the context of mental arithmetic. Retrieval-induced forgetting (RIF) in word recall has been observed in multiple studies, perhaps most notably in the research of Anderson, Bjork, and Bjork (1994). In their study, Anderson et al. developed what is now considered a standard practice/retrieval paradigm for identifying RIF in word recall.

When RIF occurs, the act of retrieving stored information inhibits the participant's ability to recall similar information afterward. Anderson et al. (1994) postulated that this forgetting was a result of weakening of potential retrieval cues. When one item is strengthened through recall, the semantically related words are weakened by that strengthening. The researchers referred to this as a strength-dependent competition model of interference, with three assumptions: (1) items that share the same cue will compete for recollection when the cue is presented (the competition assumption); (2) an increase in the strength of its competitor will jeopardize an item's recall (the strength-dependence assumption); and (3) retrieval of an item is a form of learning, because it improves later recall of an item – the retrieval-based learning assumption (Anderson et al., 1994).

To induce RIF in the laboratory, Anderson et al. (1994) constructed a paradigm that would foster competition between items by associating them with a specific cue. In the first phase of the experiment (the practice phase), participants studied word pairs that included a category and exemplar (e.g., *fruit-orange*, *fruit-banana, insects-roach, insects-hornet, metal-*

copper, metal-silver). In the second phase (retrieval practice), participants were presented with a cue (described below) to recall a specific item. In this phase, pairs fell into one of three categories:

- •RP+ pairs: pairs whose category were studied and practiced. These items were cued with the presentation of the category and the first two letters of the intended exemplar (*fruit-or*_____).
- •NRP pairs: pairs whose category and exemplar were studied, but not practiced at all during the retrieval practice phase.
- •RP- pairs: pairs whose category and exemplar were studied, but whose *exemplars* were not practiced in the second phase. For example, participants were cued to recall *orange* but not *banana*, although both shared the *fruit* category association.

Following retrieval practice, participants were given a 20 minute distractor task before being surprised with a final test phase in which they were asked to recall all of the pairs. As expected, RP+ pairs were recalled best, as they had been both studied and strengthened through practice. The NRP pairs were recalled less accurately. These pairs had been studied, but their exemplars were not strengthened through retrieval practice of any kind. Lowest recall, however, came in the RP- category. These pairs had been practiced in category, but not in exemplar. Participants could, therefore, remember *fruit-orange* but forgot *fruit-banana*. This lower recall of RP- compared to NRP (which were, again, items that had not been practiced at all) was evidence of the retrieval-induced forgetting effect (Anderson et al., 1994).

The implications of RIF carry over into multiple areas, including recall of personality traits (Macrae & MacLeod, 1999), inhibiting reliance on stereotypes (Dunn & Spellman, 2003), facial recognition (Schooler & Engstler-Schooler, 1990), and episodic memory for objects and experiences (Ciranni & Shimamura, 1999; Anderson, 2003). First-language attrition, or forgetting of native-language words following extended immersion in a second, foreign language, has also been linked to the inhibition mechanism thought to be associated with retrieval-induced forgetting (Levy, McVeigh, Marful, & Anderson, 2007). The generalizability of the retrieval induced forgetting effect, therefore, makes clear that further research into the mechanisms that underlie RIF is pertinent.

An Inhibition Mechanism

A number of studies have attempted to find a reliable predictor for when RIF will occur, and to identify a cognitive mechanism responsible for the effect. It is reasonable to assume, based on the current body of research, that practice strength is an insufficient explanation – if RP+ items were remembered more accurately simply because they had been practiced more, then there should be no difference in recall between NRP and RP- pairs. Anderson, Bjork, and Bjork (2000) proposed a recall-specific inhibition mechanism in which the competition between items is necessary to induce RIF.

In a study following the same paradigm as their 1994 study, Anderson et al. (2000) again had participants study, practice, and then take a final recall test over category-exemplar pairs. This time, however, the practice phase was manipulated so that one group experienced competition and one did not. In the competitive group, participants received category-partial exemplar pairings like normal. In contrast, the noncompetitive group received exemplars and were asked to generate the correct category. Because all exemplars practiced were associated with the same specific categorical cue, there was no competition between exemplars. Anderson et al. (2000) found RIF in the competitive group, but not in the noncompetitive group. The results indicate that during recall, when multiple items competed for association with a specific category cue, an inhibitory mechanism weakens the incorrect items in favor of recalling the correct item, resulting in forgetting of the inhibited item (Anderson et al., 2000).

Anderson et al. (1994) also identified an effect of exemplar strength on RIF. Taxonomic frequency was higher for some exemplars than others – for example, *banana* is more commonly associated with the category *fruit* than, say, *guava*. RP- exemplars with high taxonomic frequency were more susceptible to RIF than exemplars with low taxonomic frequency. Anderson et al. (1994) suggested that high-frequency exemplars present greater competition during retrieval, and so require inhibition; in comparison, low-frequency exemplars do not compete as hard for retrieval and therefore do not require inhibition, in line with the strength-dependence assumption of RIF.

Johnson and Anderson (2004) demonstrated further evidence for an inhibition mechanism by inducing RIF using new, independent cues (for example, *seasonings-salt* would be studied during the study phase, but the retrieval practice phase would cue the exemplar using *popcorn-s* instead). Additionally, participants generated the exemplars for the initial categories themselves so that the task would involve semantic retrieval rather than episodic retrieval. The independent cues were still able to produce RIF, implying that the generation of semantic knowledge impairs retrieval of similar concepts even without the initial cue for the target word, and that semantic forgetting can be at least partially credited to an inhibition mechanism that suppresses related information in favor of the target memory (Johnson & Anderson, 2004).

Alternatives to the Competition Assumption

The inhibition account of RIF carries with it the broader implication that forgetting itself may be closely linked to the act of recall (Story & Levy, 2012). However, further research suggests that this may not be the single most reliable predictor of RIF, and that competition may not be an explicit requirement for the effect to occur. In fact, a growing number of studies have demonstrated just the opposite – many researchers have manipulated the standard RIF paradigm to successfully exhibit the phenomenon in the absence of competition between responses.

Jonker and MacLeod (2012) were able to demonstrate RIF without competition using a generation task in which participants first studied category-exemplar pairs, then produced subordinates based off presented exemplars in lieu of the standard retrieval practice phase. For example, if the participant studied *pet-dog* during phase one, they might see *dog* during phase two and be asked to generate a subordinate in the form of a specific breed of dog, such as *husky*; following generation, they were instructed to report the original category associated with *dog*, thereby strengthening the association between exemplar and category without inducing competition. In this manipulation, participants were retrieving information related to only half the studied exemplars, but were not directly retrieving a particular exemplar. According to inhibition theory, this should not have produced RIF because there was no competition between exemplars of the same category. However, the experiment was successful in demonstrating RIF consistently despite the lack of competition (Jonker & MacLeod, 2012).

One of multiple theories opposing the competition assumption is the strength-based interference account (Raaijmakers & Jakab, 2012), which hypothesizes that retrieval practice strengthens the target exemplar's association with the given category, so much so that the final test results in interference from the RP+ (strengthened) exemplar and the inability to recall the RP- (non-strengthened) exemplar. Raaijmakers and Jakab (2012) argued that the results of experiments supporting the inhibition account are also compatible with the strength-based interference account, and demonstrated RIF without competition by instructing participants to recall the associated category rather than a targeted exemplar. Subjects could more easily recall the category of strengthened RP+ exemplars than non-strengthened RP- exemplars, despite the task not inducing competition between them.

Another alternative is the context shift account (Jonker, Seli, & MacLeod, 2013), which also proposes that competition is not a requirement for RIF to occur. Jonker, Seli, and MacLeod (2013) hypothesized that the change in task between study and retrieval practice creates two internal contexts (study and practice), either of which may be triggered by the final test phase, and that RIF will occur even in the absence of competition assuming two conditions are met: (1) a context shift must occur between the initial study phase and the subsequent retrieval practice phase, and (2) the target exemplars must benefit from sharing a context with the retrieval practice phase in a way that RP- items do not. According to this account, RIF occurs because of two factors: NRP items reinstate the study context and experience a memorial benefit during testing, and RP+ items reinstate the practice context and experience a similar benefit. The only items that do not experience this benefit are RP-, causing lower recall. Similar to the strength-based interference account, this context-shift account also does not require competition to produce RIF (Jonker et al., 2013). Additionally, in a later study when Jonker et al. (2015) guided participants into reinstating the original study context during the final test phase, the RIF effect disappeared despite the target exemplars still having benefited from additional practice during the retrieval practice phase.

Camp and Sander (2016) demonstrated RIF by designing a generation task that did not require competition. In lieu of the standard retrieval practice phase, participants were given the category and full target exemplar with the first two letters of the exemplar transposed (*fruit-roange*, for example). In this manipulation, the presentation of the full, transposed target exemplar still required practice and, thus, activation of the exemplar in working memory, but did not induce competition from related responses because the given exemplar was mostly intact. Despite this lack of competition, the results of the study still produced RIF (Camp & Sander, 2016).

Clearly, the inhibition explanation of RIF cannot account for every instance in which the effect occurs, as multiple studies have successfully replicated the RIF effect without inducing competition between exemplars. However, studies in favor of the inhibition account have demonstrated that it is a consistently reliable predictor of RIF (Anderson et al., 2000; Johnson & Anderson, 2004; Storm & Levy, 2012). Overall, it seems reasonable to assume RIF may not be limited to a specific mechanism, but rather a range of potential mechanisms that are capable of producing the effect.

RIF in Mental Arithmetic

Nearly all of the experiments mentioned above were performed using word recall as the primary indicator of RIF. More recently, research has shown that this effect generalizes to the realm of numerical cognition, specifically in mental arithmetic. Difficulty in retrieving multiplication facts among children had previously been theorized to be related to an overload of associations for each number, with Norem and Knight (1930) proposing the issue initially. Further investigation of this hypothesis and its relationship with retrieval-induced forgetting has taken place more recently in an attempt to identify the underlying cognitive processes associated with these numerical associations.

Several studies have been conducted in recent years to further explore whether RIF of mental arithmetic follows the same principles as word RIF. In 2009, Campbell and Phenix found support for an inhibition mechanism and an effect of target strength in this type of RIF using multiplication retrieval practice. Participants studied simple addition problems (e.g. 2 + 3 = ?)

and then either studied or retrieved the answers to their multiplication counterparts (2 x 3 = 6 or 2 x 3 = ?, for example). Participants were then tested on the addition counterparts again. Similar to the analysis of word RIF, arithmetic RIF is indicated by increased errors and slower response times on addition problems following the retrieval practice of their multiplication counterparts (Campbell & Phenix, 2009).

Phenix and Campbell (2004) utilized a true or false verification task with multiplication facts to demonstrate whether the presence of RIF would provide support in opposition to a contrasting theory, the integrated structures model of simple multiplication (Manly & Spoehr, 1999). The integrated structures model argues that retrieving a multiplication fact should strengthen memory for the operands of that fact (for example, 3 x 2 = ? should strengthen memory for facts using 6, 9, 12, etc. and 4, 6, 8, etc.) whereas RIF's inhibition account would assume that related facts should be suppressed. Participants practiced multiplication facts before being tested on related facts using operands of the original set, and as predicted by RIF, the related facts were resolved less accurately because the operands had been suppressed during retrieval practice (Phenix & Campbell, 2004).

Critical to the present study are the results of Campbell and Thompson (2012), who demonstrated RIF of simple addition facts using their multiplication counterparts. Participants either studied or practiced multiplication facts that corresponded to practiced addition problems. The study supported the inhibitory account of RIF, with participants only experiencing RIF in the practice group, not the study group. Similar to the results from studies of word pair RIF, these results indicated that increased practice strength was not a reliable explanation; participants in both groups performed well on the multiplication facts they practiced, and neither group presented more improvement on MU (unpracticed) addition problems in comparison to MP (practiced) problems.

The Campbell and Thompson study (2012) also demonstrated an important concept for arithmetic RIF – interference dependence, which was first proposed by Anderson (2003). This principle predicts that a) RIF occurs only when there is retrieval competition (so studying a multiplication fact without retrieval does not trigger inhibition of competitors), and b) the strength of the target addition problem and competing multiplication problem will influence the strength of the RIF effect; that is, a stronger competitor will attract more direct inhibition and produce a stronger RIF effect. In line with this prediction, Campbell and Thompson's results indicated that small addition problems (whose sums were less than 10) produced stronger RIF than large addition problems, possibly because their higher memory strength required greater action from the inhibition mechanism (Phenix & Campbell, 2004). In contrast, large addition problems did not produce RIF because their weaker memory associations did not require inhibition during retrieval of their multiplication counterparts (Campbell & Thompson, 2012).

The Present Study

The present study had two aims. First, the study sought to replicate the results of Campbell and Thompson (2012), who demonstrated that practice of multiplication facts (2 x 5 =) affects subsequent performance of addition counterparts (2 + 5 =). Second, response time modeling (Luce, 1986; Faulkenberry, 2017) was used to analyze characteristics of the RT distributions that are potentially indicative of strategy shifts (e.g. Campbell & Penner-Wilger, 2006), which was not part of the original 2012 study. This study is the first known attempt to use response time modeling on an RIF experiment in mental arithmetic.

CHAPTER II.

METHOD

Participants

Fifty undergraduate psychology students participated for extra credit in their respective psychology courses at a mid-sized public university in Texas. The recruitment description invited students to participate in a study of simple mental arithmetic. The sample group included 13 men and 37 women between the ages of 18 and 26. Participants were assigned to a condition by groups of 10, and participation took approximately 20 minutes.

Materials

Experimental stimuli were presented on a 20” iMac computer running Superlab 5.0. A wearable microphone headset worn by the participant was connected to a Cedrus SV-1 voice key device, which recorded vocal onset latency in milliseconds. The investigator input participant’s responses after each arithmetic problem using a standard computer keypad and a small key-response box on screen. Prior to participation, individuals were given the option of completing a short demographic survey, administered on paper and completed using a pen.

Procedure

Participants were given a consent form and demographic survey to complete upon entering the lab. They were seated in front of the computer screen and fitted with the headset so that the microphone was positioned to capture their responses. They were then informed that they would be completing a study of simple mental arithmetic, and given instructions on how to give their responses.

Phase 1: Addition pretest. In Phase 1, participants were presented with one block of 36 addition problems, ranging from 2 + 2 to 9 + 9, which consisted of two types of problems: MP (problems whose multiplication counterparts would be practiced during Phase 2) and MU (problems whose multiplication counterparts would not be practiced during Phase 2). On each trial, participants were presented with the full equation (2 + 2 = ?) and asked to answer aloud quickly, without sacrificing accuracy. Response time was recorded in milliseconds using the microphone headset and voice key. The addition pretest results provided a baseline analysis to ensure there were no significant differences between initial performance on MP and MU problems.

Each trial began with the presentation of a central fixation dot for 1000 milliseconds. A randomly presented equation then appeared so that the addition sign was in the spot of fixation. Timing began with the onset of the equation and continued until the participant's verbal response triggered the voice key timer. When the response was detected, the equation disappeared from the screen and the investigator input the participant's response via keyboard, after which the fixation dot appeared to signal the next trial. After completion of the addition block, the participant moved on to Phase 2.

The set of 36 addition problems was divided into two sets, designed by Campbell and Thompson (2012) to include non-overlapping operand pairs (2, 5, 7, 8 and 3, 4, 6, 9) and was balanced so that each digit occurred in four problems per set. Both sets included an equal number of tie problems (e.g. 2+2) and non-tie problems (e.g. 2+3), as well as equal number of odd-plus-odd, even-plus-even, and odd-plus-even operand pairs (e.g. 3+5, 2+6, and 2+5). Set 1 included the problems 2+2, 2+5, 3+4, 2+6, 3+5, 4+4, 3+6, 2+9, 4+7, 3+9, 4+8, 5+7, 5+8, 6+7, 6+8, 7+7, 8+9, and 9+9. Set 2 included the problems 2+3, 2+4, 3+3, 2+7, 4+5, 2+8, 3+7, 4+6, 5+5, 3+8, 5+6, 6+6, 4+9, 5+9, 6+9, 7+8, 7+9, and 8+8.

Phase 2: Multiplication practice. In this phase, participants were presented with six blocks of multiplication problems, each of which contained the same operands as either Set 1 or Set 2 of the addition problems presented in Phase 1. Half the participants received Set 1 and half received Set 2. These multiplication counterparts were designed by replacing the addition (+) sign with the multiplication (x) sign. The set practiced by the participant (Set 1 or 2) was designated as MP (multiplication practiced) and the set not practiced was designated as MU (multiplication unpracticed).

Participants were again presented with a fixation dot for 1000 milliseconds to signal the start of a new trial. Timing began with the onset of the equation and continued until the participant's verbal response triggered the voice key timer. When the response was detected, the equation was removed from the screen and the investigator input the participant's response, after which the fixation dot appeared to signal the next trial.

This phase contained a critical between-subjects manipulation. Half of the participants were assigned to a *retrieval-study group*, in which they were presented with multiplication equations, including the correct answer (2 x 5 = 10), and instructed to silently read the equation before stating the answer out loud. The remaining participants were assigned to a *retrieval-practice group*, in which they were presented with multiplication problems that did not include the correct answer (2 x 5 = ?) and instructed to state the correct answer aloud. The differentiation between retrieval-study and retrieval-practice was designed to help confirm whether the RIF, if found, supported the inhibition account, as it did for Campbell and Thompson (2012).

Phase 3: Addition and multiplication posttest. In the final phase, participants were presented with a block of all 36 addition problems (including both MP and MU problems) and asked to answer aloud quickly, without sacrificing accuracy. As before, response time was recorded using the microphone headset and voice key. Following completion of the addition block, participants were presented with a block of all 36 multiplication problems and given the same instructions. Finally, participants were presented with a final block of the 36 addition problems, and response times were again recorded using the microphone. The purpose of the middle multiplication block was to eliminate RIF effects in the final addition block by moving all problems to the “MP” category.

CHAPTER III

RESULTS

Participants completed 12,600 experimental trials. A total of 1,004 pretest, practice phase, or posttest RTs (7.97%) were excluded from analysis due to experimenter, software, or participant error and marked as mistrials. Of the remaining 11,596 trials, we filtered 5,400 multiplication practice or study trials from Phase 2 as our analysis was over the blocks of problems before and after practice. Finally, we discarded a total of 205 outlier trials using the following procedure: for each block's set of response times (e.g. "Phase 1 addition pretest block") we computed the mean and standard deviation of response times for correct responses and discarded trials that were a) less than 200 ms or b) greater than the mean plus 3 standard deviations. This left 6,565 trials for analysis between pretest/posttest addition, posttest multiplication, and the final addition block. For the purposes of this study, problems with a sum $\leq$10 were categorized as small and problems with a sum >10 were categorized as large.

To facilitate the comparison of the present study's results to those of Campbell and Thompson (2012), we analyzed both mean and median RTs using ANOVA. However, we noted that density plots of the response times for each block revealed a positively skewed distribution (see Figure 1). To better account for this skew, we fit each data set to an ex-Gaussian distribution (Campbell and Penner-Wilger, 2006; Balota & Yap, 2011; Luce, 1986). This approach allowed us to test how the *distributions* of RTs behaved under our experimental manipulations, rather than collapsing each participant's distribution to a single mean or median RT. Further, the use of the ex-Gaussian model allowed us to observe indirectly whether participants relied most often on memory retrieval or some mental procedure to answer problems. This procedure was in lieu of relying on self-report data of procedure following each trial as the original study (Campbell & Thompson, 2012) did.

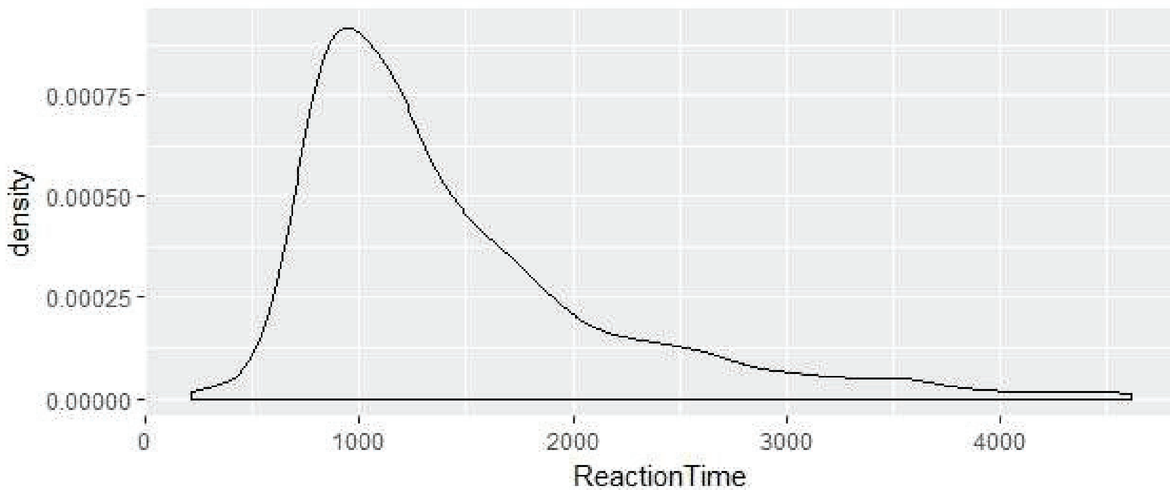

Figure 1. Response time density for the Phase 1 addition pretest. Density plots for all analyzed blocks had a similar positive skew.

In an ex-Gaussian model, response time distributions are decomposed into three components – μ, which corresponds to the mean of the assumed "normal" component of the distribution of response times, σ, the standard deviation of the normal component, and τ, which corresponds to the exponential "tail" of the distribution made up of slower response times (Campbell and Penner-Wilger, 2006). μ, the mean of the normal component, increases or decreases based on uniform shifts in RTs whereas τ, the mean of the exponential component, changes when the skew changes (i.e., problem size effects and shifts in strategy). Therefore, by comparing μ and τ values for each block, we can capture shifts in strategy between blocks. If participants predominantly use memory retrieval strategies (reflected in a higher μ value) for MP problems during the addition pretest block but shift to predominantly using procedural strategies (reflected in an increased τ value) during the addition posttest block, for example, this could be considered evidence of RIF that may not be reflected just by comparing the overall mean RTs of the two blocks. This type of analysis for response time modeling has been used in other studies (Luce, 1986; Faulkenberry, 2017) to better analyze response time distributions, which tend to be positively skewed (Balota & Yap, 2011).

Mean Response Times

For each block, mean RTs were analyzed using a Group (retrieval vs study) X Set (MP addition vs MU addition) X Size (small vs large) mixed factorial ANOVA. The Phase 1 addition pretest block was analyzed to ensure there were no differences in performance based on group or problem sets. Mean and median RTs for this phase can be found in Table 1. There was no main effect of group or set on mean or median reaction time, all $p > 0.2$, so study and retrieval participants responded with similar speed on all pretest addition problems. There was a standard problem size effect, $F(1, 48) = 128.34$, $p < 0.001$. Large problems took longer to solve than small across all participants, which was expected. Otherwise, there was no statistical difference between the groups (see Figure 2).

Table 1
Mean and Median Correct RT for Addition Pretest

Problem Size	Mean RT (ms)		Median RT (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	1182	1214	1080	1050
Large	1721	1704	1512	1511
Study Group				
Small	1178	1120	1001	1015
Large	1588	1579	1312	1324

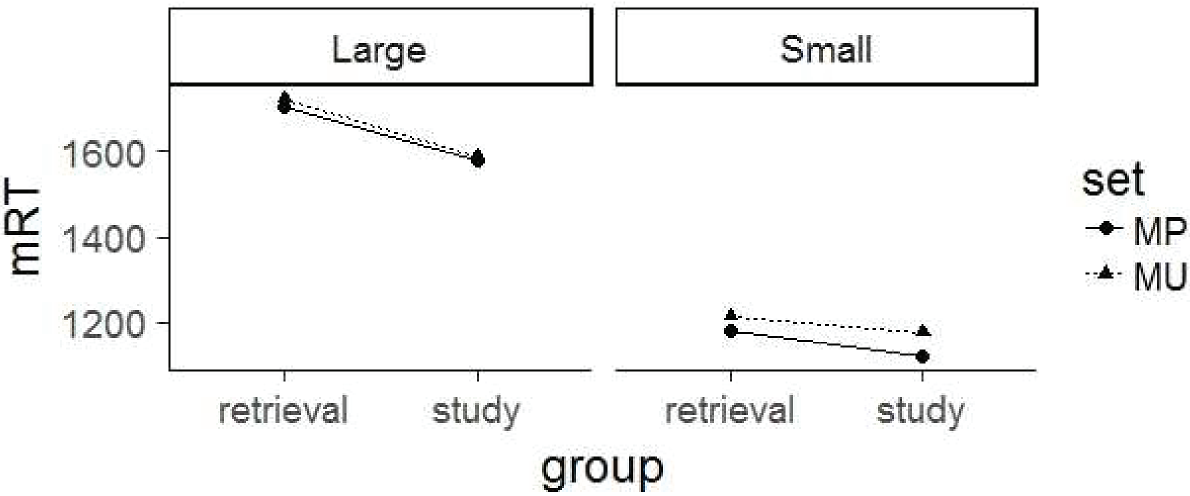

Figure 2. Mean RTs as a factor of group, set, and size for Phase 1 addition pretest.

For the Phase 3 addition posttest there was a problem size effect $F(1, 47) = 152.05$, $p < 0.001$. There was also a significant group x size interaction, $F(1, 47) = 4.77$, $p = 0.034$ (see Figure 3). The problem size effect was larger for the retrieval group (551ms on MP and 501ms on MU) than for the study group (308ms on MP and 411ms on MU). Means and medians for this block can be found in Table 2.

Table 2

Mean and Median Correct RT for Addition Posttest

Problem Size	Mean RT (ms)		Median RT (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	1237	1231	1113	1081
Large	1738	1782	1590	1628
Study Group				
Small	1119	1194	1018	1048
Large	1530	1502	1249	1256

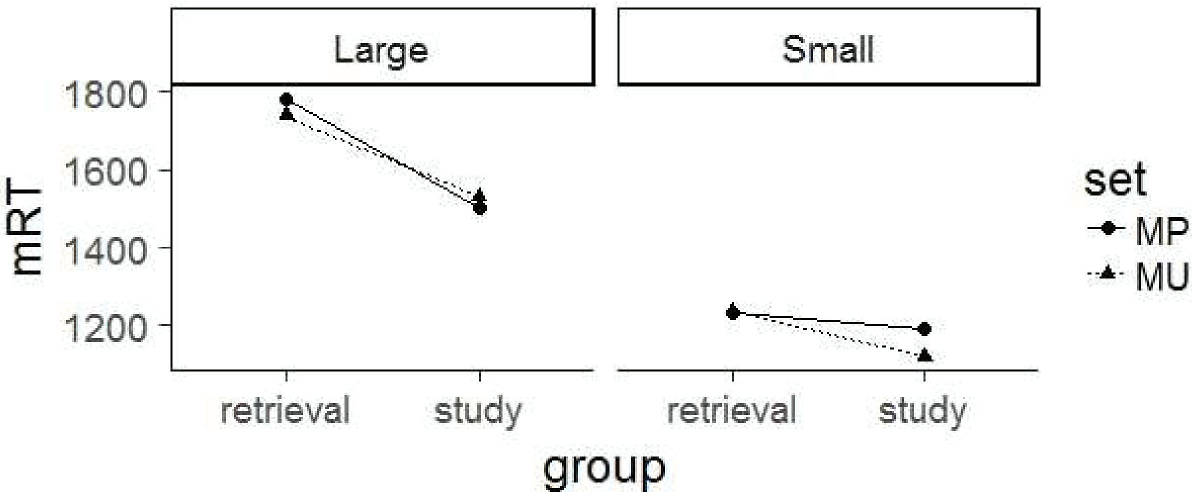

Figure 3. Mean RTs as a factor of group, set, and size for Phase 3 addition posttest. As a reminder, MP addition were problems whose multiplication counterparts were practiced or studied during Phase 2.

For the final addition block of Phase 3, there was a problem size effect, $F(1, 47) = 107.38$, $p < 0.001$, but no other significant interactions, all $p > 0.2$ (see Figure 4). Means and medians for the final addition block can be found in Table 3.

Table 3

Mean and Median Correct RT for Final Addition Block

Problem Size	Mean RT (ms)		Median RT (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	1238	1304	1066	1143
Large	1728	1720	1520	1545
Study Group				
Small	1168	1169	1080	1064
Large	1533	1583	1315	1354

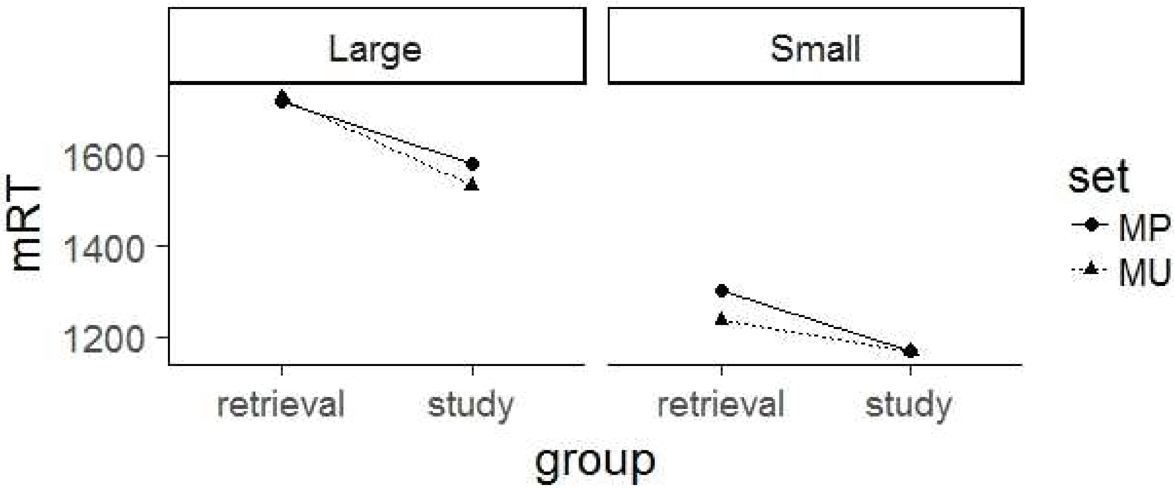

Figure 4. Mean RTs as a factor of group (retrieval vs study), set (MP vs MU), and size, for the final addition block.

For the Phase 3 multiplication block, there was a problem size effect, $F(1, 47) = 79.94$, $p < 0.001$. Additionally, there was a main effect of set, $F(1, 47) = 21.07$, $p < 0.001$, a Set x Size interaction, $F(1, 47) = 8.72$, $p = 0.004$, and a Group x Set x Size interaction, $F(1, 47) = 5.39$, $p = 0.024$ (see Figure 5). These interactions, while not specifically meaningful in terms of our overall hypothesis, do reveal a greater problem size effect on MU problems, particularly for the retrieval group, that can most likely be attributed to practice effects. Mean and median RTs of this block can be found in Table 4.

Table 4

Mean and Median Correct RT for Final Multiplication Block

Problem Size	Mean RT (ms)		Median RT (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	1587	1410	1353	1192
Large	2315	1801	1832	1432
Study Group				
Small	1448	1344	1186	1175
Large	1988	1789	1545	1273

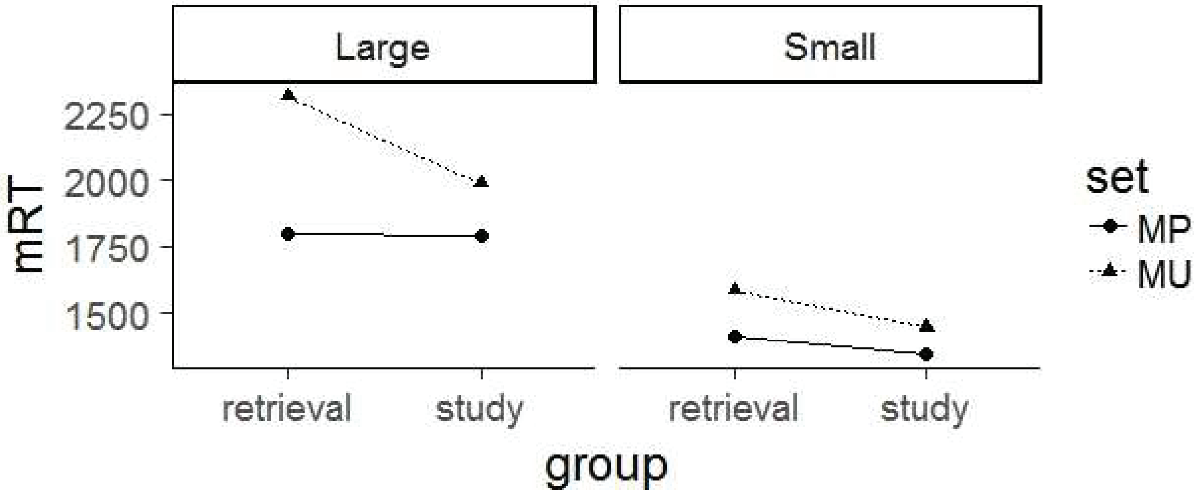

Figure 5. Mean RTs as a factor of group, set, and size for Phase 3 multiplication block.

Median Response Times

For each block, median RTs were analyzed using a Group (retrieval vs study) X Set (MP addition vs MU addition) X Size (small vs large) mixed factorial ANOVA. In the Phase 1 addition pretest, there was a problem size effect, $F(1, 48) = 89.83$, $p < 0.001$, but no other significant effects, all $p > 0.2$ (see Figure 6).

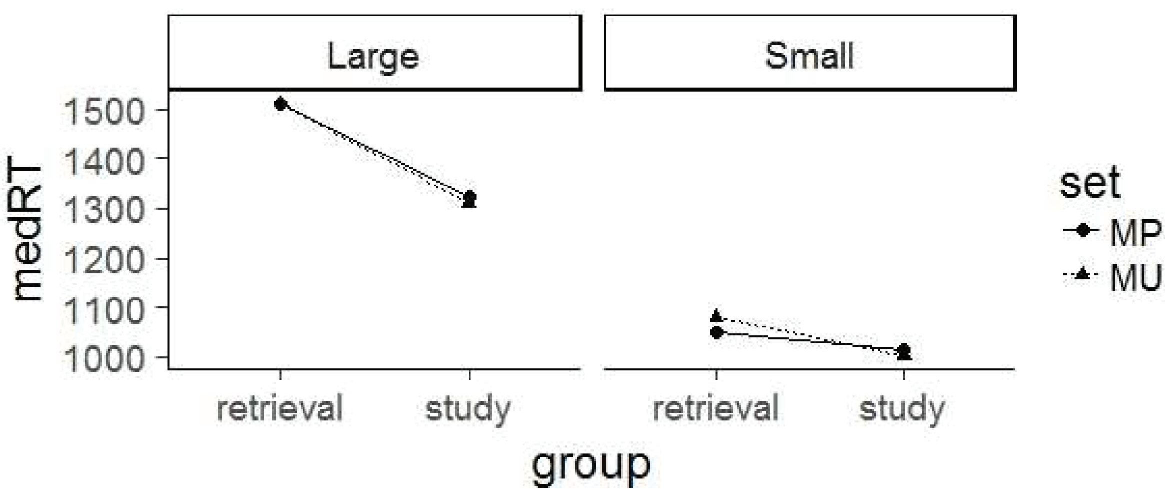

Figure 6. Median RTs as a factor of group, set, and size for Phase 1 addition pretest.

For the Phase 3 addition posttest, there was a main effect of group, $F(1, 47) = 4.78$, $p = 0.033$. Across all conditions, median RTs for the retrieval group were slower than those of the study group (see Figure 7). There was also a problem size effect, $F(1, 47) = 96.93$, $p < 0.001$, and a Group x size interaction, $F(1, 47) = 7.73$, $p = 0.008$. The problem size effect was larger for the retrieval group (547ms on MP and 477ms on MU) than for the study group (208ms on MP and 228ms on MU).

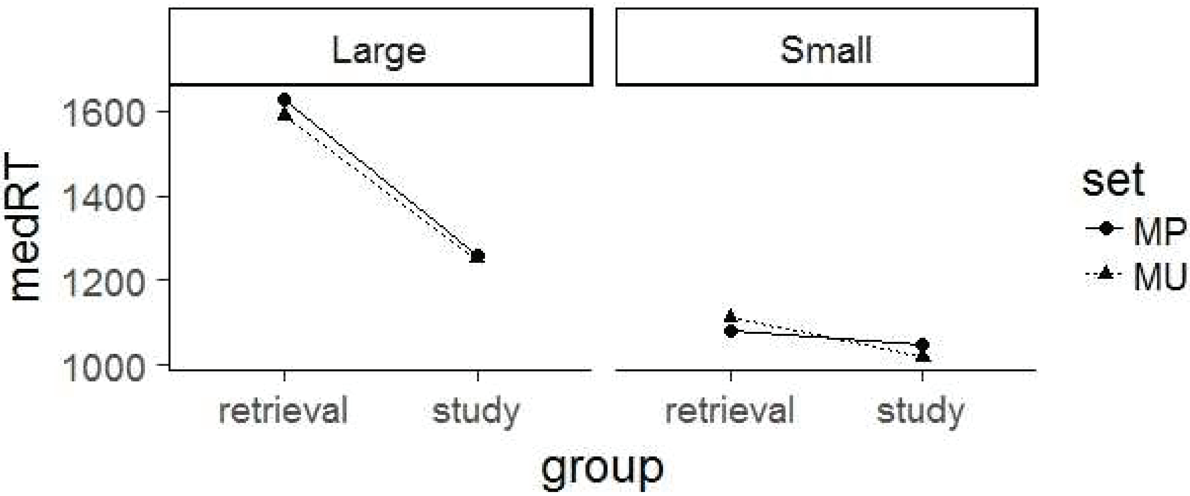

Figure 7. Median RTs as a factor of group, set, and size for Phase 3 addition posttest. As a reminder, MP addition were problems whose multiplication counterparts were practiced or studied during Phase 2.

In the Phase 3 final addition block, there was a problem size effect, $F(1, 47) = 78.34$, $p < 0.001$. There were no other significant effects, all $p > 0.3$ (see Figure 8).

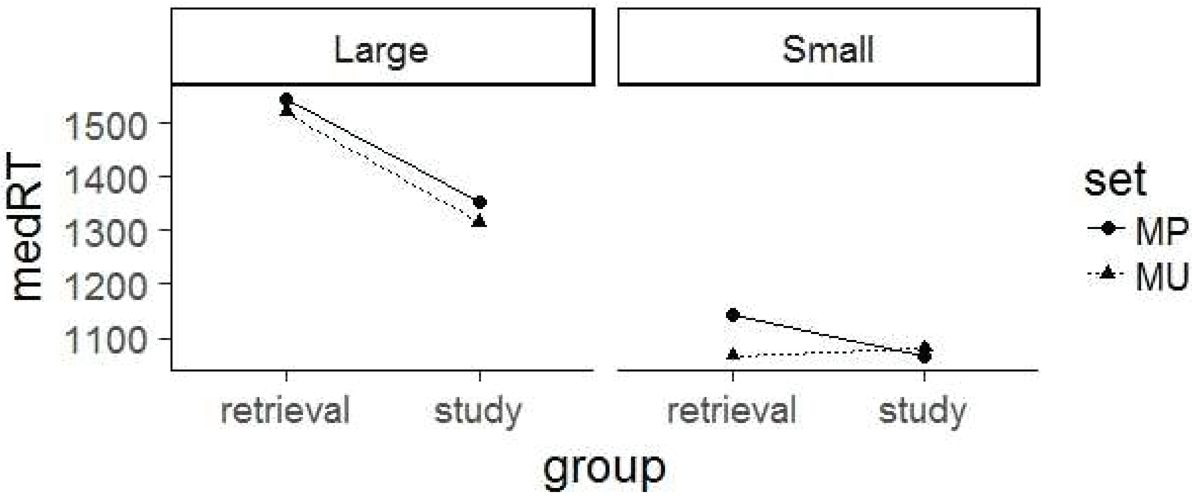

Figure 8. Median RTs as a factor of group, set, and size for Phase 3 final addition block.

For the Phase 3 multiplication block, there was a main effect of set, $F(1, 47) = 31.50$, $p < 0.001$, and of problem size, $F(1, 47) = 57.55$, $p < 0.001$. There was also a Set x Size interaction, $F(1, 47) = 16.63$, $p < 0.001$ (see Figure 9).

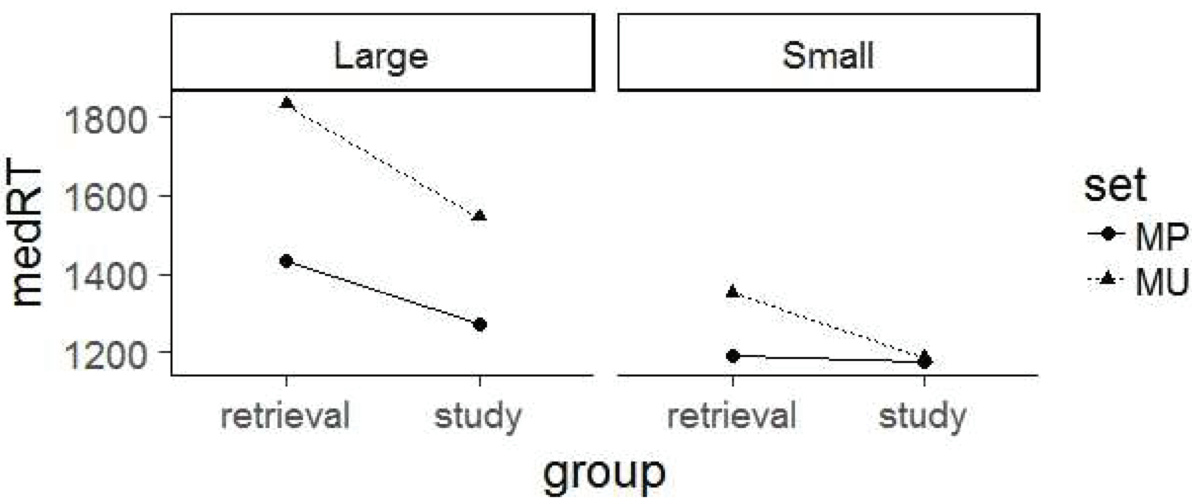

Figure 9. Median RTs as a factor of group, set, and size for Phase 3 multiplication block.

Ex-Gaussian Modeling

Response times for each block were fit to an ex-Gaussian model. Specifically, for each participant's distribution of response times in each experimental condition, we used maximum likelihood estimation (Myung, 2011) to compute an estimate for μ and τ. These values of μ and τ were then submitted to the same mixed-factorial ANOVAs as were the mean and median RTs. The advantage of this approach is that it allows us to potentially infer strategy shifts between phases. These strategy shifts, if present, would indicate RIF effects where participants were able to retrieve answers for MP addition from memory during the pretest but following multiplication practice in Phase 2, moved to a procedural strategy such as counting to solve MP addition problems in the posttest. The ex-Gaussian model compares values for μ (where RTs fall on a normal curve) in which participants were retrieving from memory, and τ (where RTs fall on an exponential curve) in which participants were using procedural strategy (Campbell & Penner-Wilger, 2006).

For the Phase 1 addition pretest, the μ ANOVA revealed a problem size effect, $F(1, 48) = 24.19$, $p < 0.001$, and a Group x Set x Size interaction, $F(1, 48) = 4.67$, $p = 0.04$; the problem size effect differed somewhat between study and retrieval and between MP and MU addition problems (see Figure 10). The τ analysis indicated a problem size effect, $F(1, 48) = 14.61$, $p < 0.001$, but no other significant values. Values of μ and τ for the addition pretest can be found in Table 5.

Table 5

μ and τ of Correct RT for Addition Pretest

Problem Size	μ (ms)		τ (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	749	739	465	442
Large	848	840	873	864
Study Group				
Small	690	711	487	409
Large	789	799	799	780

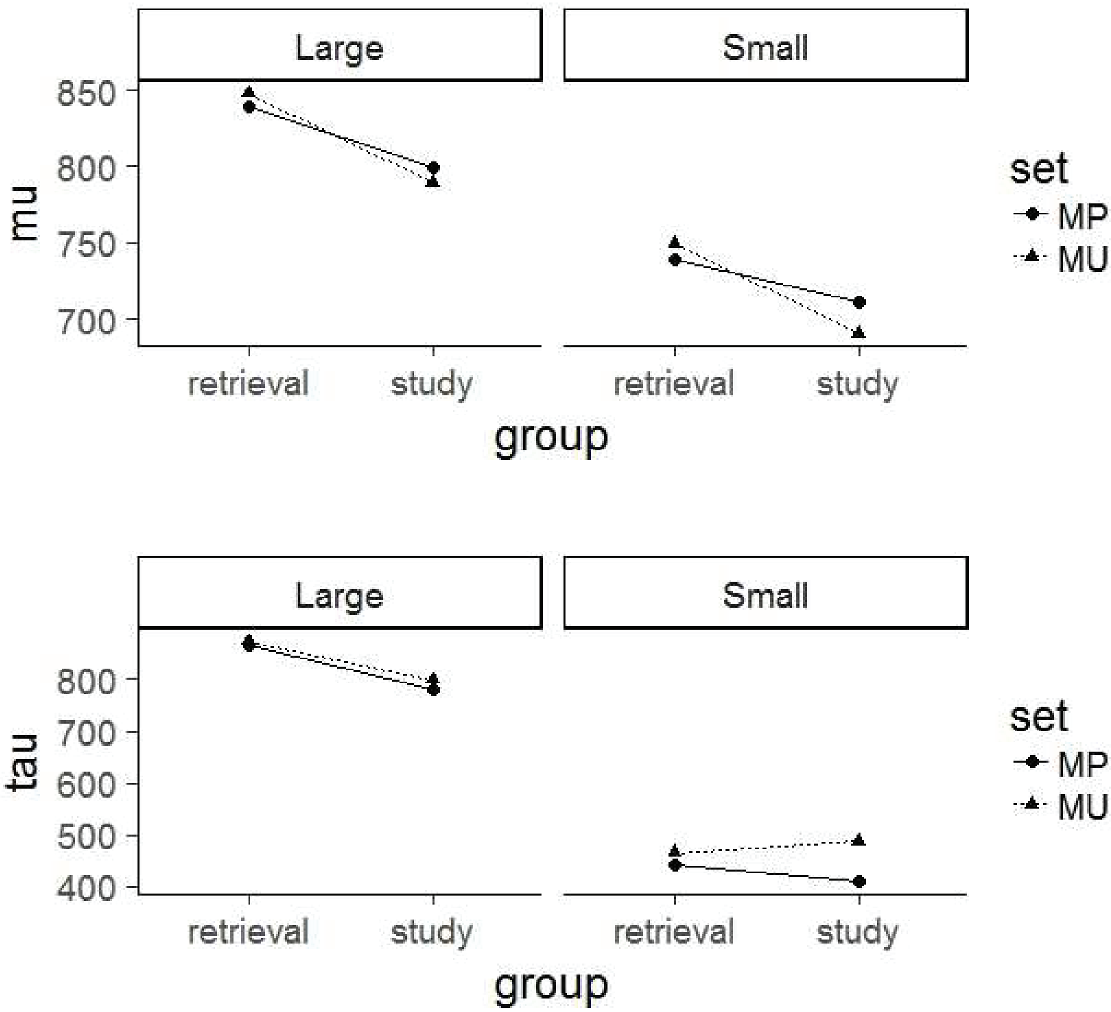

Figure 10. μ and τ of correct RTs as a factor of group, set, and size for addition pretest.

For the Phase 3 addition posttest, the μ ANOVA showed a main effect of group, $F(1, 47) = 5.85$, $p = 0.02$, and a problem size effect, $F(1, 47) = 54.51$, $p < 0.001$.

There was also a Group x Size interaction, $F(1, 47) = 6.50$, $p = 0.014$; in general, response times were slower and the problem size effect was larger for the retrieval group. The τ analysis showed only a problem size effect, $F(1, 47) = 6.70$, $p = 0.013$ (see Figure 11). Values of μ and τ for this block can be found in Table 6.

Table 6

μ and τ of Correct RT for Addition Posttest

Problem Size	μ (ms)		τ (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	710	726	527	505
Large	947	945	790	837
Study Group				
Small	687	688	433	506
Large	712	688	819	814

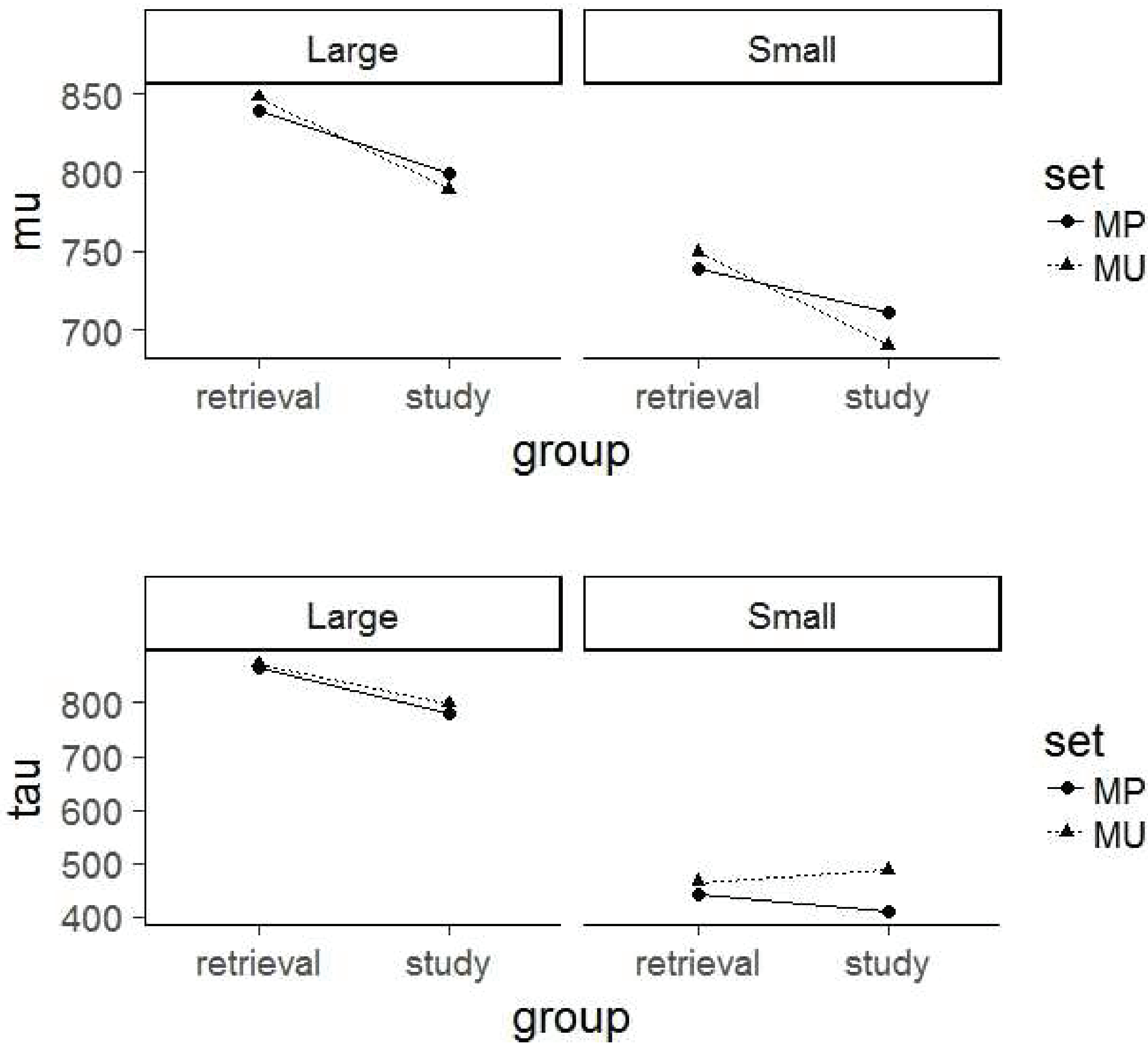

Figure 11. μ and τ of correct RTs as a factor of group, set, and size for addition posttest.

The μ and τ ANOVAs for the Phase 3 final addition block both showed a problem size effect – $F(1, 47) = 28.51$, $p < 0.001$ for μ and $F(1, 47) = 8.30$, $p = 0.006$ for τ (see Figure 12). There were no other significant results for this block. Values of μ and τ for this block can be found in Table 7.

Table 7

μ and τ of Correct RT for Final Addition Block

Problem Size	μ (ms)		τ (ms)	
	MU Addition	MP Addition	MU Addition	MP Addition
Retrieval Group				
Small	689	734	549	570
Large	765	867	963	853
Study Group				
Small	724	719	443	450
Large	764	725	769	859

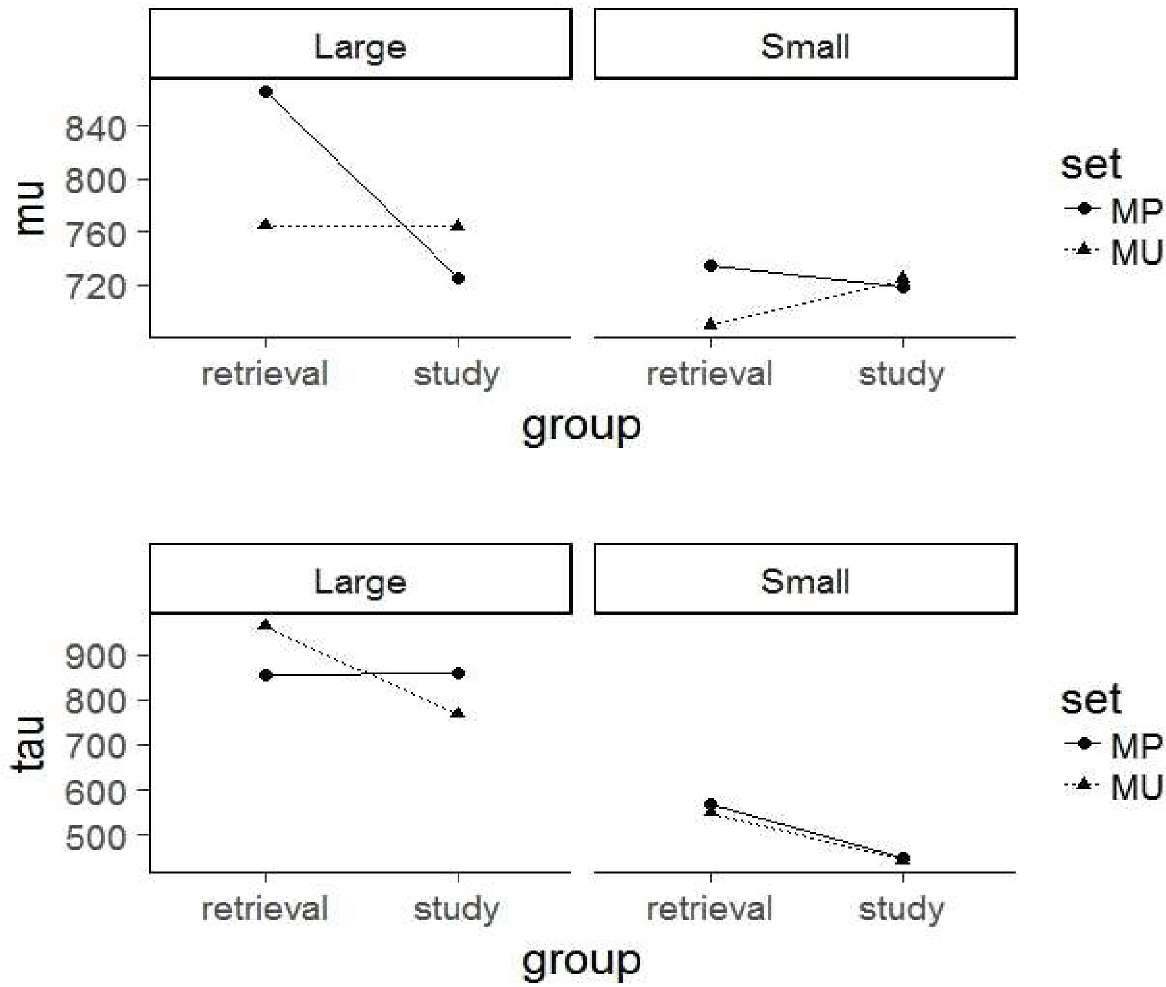

Figure 12. μ and τ values as a function of group, set, and size for final addition block.

Finally, the μ ANOVA for the Phase 3 multiplication block showed a main effect of set, $F(1, 47) = 20.98$, $p < 0.001$, and of size, $F(1, 47) = 24.16$, $p < 0.001$. There was also a Set x Size interaction, $F(1, 47) = 12.71$, $p < 0.001$; the problem size effect was larger for MP problems than MU problems. The τ ANOVA showed a problem size effect, $F(1, 47) = 14.39$, $p < 0.001$ (see Figure 13). Values of μ and τ for this block can be found in Table 8.

Table 8

μ and τ of Correct RT for Final Multiplication Block

Problem Size	μ (ms)		τ (ms)	
Retrieval Group	MU Addition	MP Addition	MU Addition	MP Addition
Small	692	724	895	687
Large	948	767	1366	1034
Study Group				
Small	762	756	685	589
Large	830	732	1157	1058

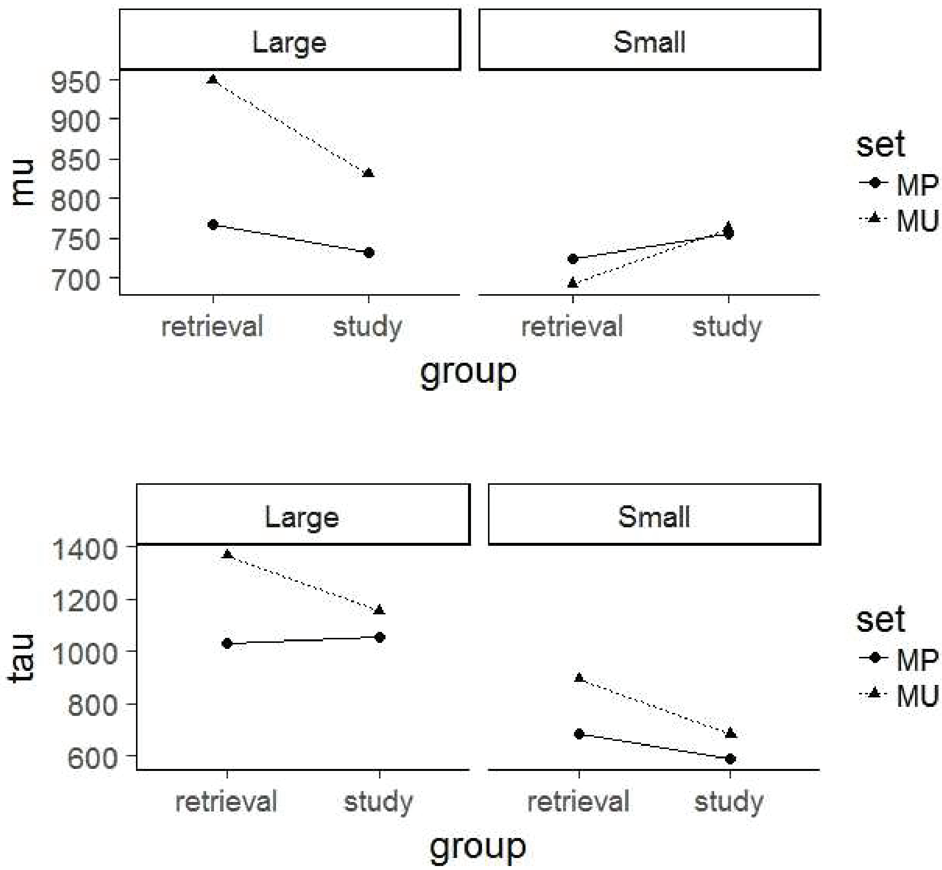

Figure 13. μ and τ values for final multiplication block.

CHAPTER IV

DISCUSSION

There were two goals of this study. The first was to replicate the results of Campbell and Thompson (2012), whose participants experienced RIF of simple addition problems following retrieval practice of their multiplication counterparts. The second was to implement response time modeling to potentially identify evidence of RIF in strategy shifts. Our initial hypothesis predicted that, upon analysis of the addition posttest block of Phase 3, we would find evidence of retrieval-induced forgetting in small problems for the retrieval group only. This was based on Campbell and Thompson's (2012) results, as well as the assumption that RIF in mental arithmetic should only occur when retrieval competition is present (i.e. when the problem must be answered, not simply studied, Anderson, 2003) and the notion that small problems have higher memory strength than large problems, which are more likely to be solved using a procedural method like mental counting, and should therefore require greater inhibition (Phenix & Campbell, 2004). Using multiple methods of analysis designed to find evidence of RIF, we assessed whether our results supported this hypothesis.

RIF Evidence in Means and Medians

The crucial problems for evidence of retrieval-induced forgetting were the small MP addition problems in the Phase 3 addition posttest (block 1). Theoretically, we should have found stronger evidence of RIF in this block and weaker evidence of RIF in the final addition block of Phase 3, with RIF effects having been "cleared" by the practicing of all 36 multiplication problems between blocks. Campbell and Thompson (2012) found a Group X Set effect on small problem median RTs and a -94ms RIF effect in the first block; in the final block, they found no Group X Set interaction and only a -25ms RIF effect. In contrast, our study showed no RIF effects based on mean or median RTs for small MP problems in the retrieval group, and there were no significant interactions or main effects suggesting the presence of RIF in this block. In the final addition block, there was a mean RT difference of -66ms and a median RT difference of -77ms on small MP problems for the retrieval group, suggesting greater evidence of RIF in the final addition block than in the addition posttest, though the results were still not significant. Overall, we found no evidence of RIF in our analysis of mean and median response times.

Ex-Gaussian Analysis

We wondered whether RIF could be present but obscured by our collapsing of the data to standard measures of central tendency (e.g., mean and median RT). The original study utilized self-report following each question to determine what arithmetic strategy participants utilized to solve each problem. When participants switched from a memory retrieval strategy to an alternate strategy (for example, counting), this was considered evidence of the presence of RIF (Campbell & Thompson, 2012). In our study, we used a different approach, instead applying an ex-Gaussian model to detect strategy shifts (Campbell & Penner-Wilger, 2006). The purpose of the change in analysis technique from the self-report approach used in Campbell & Thompson (2012) was to identify strategy shifts of which the participants may not be aware. Asking individuals to report their strategy after every single problem could become tedious or result in attrition of a participants' self-report accuracy. Using this method, the amount of time required for each trial was reduced and we were still able to collect data on strategy shifts that could indicate RIF.

As mentioned earlier, the ex-Gaussian model decomposes the distribution of response times into both a "normal" distribution (the μ component of the curve) that indicates memory retrieval, and an exponential curve (the τ component) where increased RT indicates use of a procedure (Campbell & Penner-Wilger, 2006); when participants were using memory more than procedure to answer problems, μ would be greater than τ, and vice versa if procedure was used more than memory. Therefore, by comparing μ and τ of the addition pretest to the addition posttest, we would be able to identify whether or not practicing multiplication counterparts caused participants to shift strategies in the addition posttest, using a procedural method like counting to solve a problem they had just been able to answer from memory during the addition pretest.

The presence of a problem size effect in the τ ANOVA for all four blocks told us there were high levels of procedure use for large problems regardless of set (MP vs MU) or group, in line with the theory that memory strength for large arithmetic is weaker than that of small arithmetic (Phenix & Campbell, 2004). Therefore, it was unsurprising that we did not find RIF on large problems. However, we were interested in whether we would see a strategy shift on small MP problems for the retrieval group because, again, the study group did not have to retrieve answers during multiplication practice in Phase 2 and the retrieval group only retrieved answers for the MP set. However, as the results above indicate, we did not see a strategy shift from memory to procedure on either block of addition from Phase 3, and in fact, both retrieval and study participants relied on memory retrieval for small MP problems more often than procedural strategies in all of the blocks analyzed. This suggests that practicing multiplication counterparts for problems they had answered from memory during the pretest did not affect participants' memory retrieval in the posttest, and significant evidence of RIF was not present.

Evaluation of the Hypothesis

We expected to find support for our hypothesis that participants would experience RIF of simple, small addition problems after practicing their multiplication counterparts. Much of the research on RIF in mental arithmetic has found evidence of the phenomenon in this area, the majority of available literature points to the reliability of the effect, and the study we replicated found robust evidence (Campbell & Thompson, 2012). However, our study provided no evidence for RIF. There are a few of possible explanations for the lack of significant results, the simplest of which is that our participants were not retrieving from memory to begin with. The ex-Gaussian analysis allows us to observe shifts in procedure, but does not guarantee that participants were ever primarily using memory to begin with. If indeed they were not, there would be no retrieval to interfere with and therefore no retrieval-induced forgetting.

Another possibility is that the six blocks of MP problems did not strengthen them enough to induce significant inhibition during the posttest addition block. One of the assumptions of RIF is that the effect is dependent on the MP (or RP+ for word recall) items being significantly strengthened through retrieval practice. Campbell and Phenix (2009) were able to produce RIF after presenting participants with 40 blocks of multiplication; our replication of Campbell and Thompson (2012), however, only called for six blocks. It is possible the retrieval practice was not sufficient to induce competition between responses.

Campbell and Thompson (2012) found that the problem size effect was eliminated in the addition posttest, and suggested two explanations for this: that RIF slowed small MP problem RTs so greatly that they became even with large problem RTs, and that multiplication practice of large MP problems primed the answer to their addition counterparts and actually helped participants answer them faster. In contrast, we found significant problem size effects for every block. This suggests that perhaps, rather than slowing small MP and “speeding up” large MP, the multiplication practice strengthened the addition answers for both sets of problems at a similar rate, leading to a lack of RIF but still producing a problem size effect. Again, this explanation relates to the strength dependence assumption of RIF.

CHAPTER V

CONCLUSION

Despite the lack of significant results, we believe retrieval-induced forgetting warrants further study in the realm of mental arithmetic. The reliability of the RIF effect in dozens of other studies suggests that our results do not contradict its validity but, rather, suggest that greater strengthening of target items is required to produce the effect. The present experiment should encourage future study of the varying situations in which RIF could be found.

As addressed above, we believe our results may have been significant if we had strengthened the multiplication items sufficiently enough to trigger stronger inhibition of the MP problems. Future studies could increase the number of blocks until a sufficient amount of practice is achieved. Further, the self-report of strategy (Campbell & Thompson, 2012) that we chose not to use in favor of the ex-Gaussian response time modeling (Campbell & Penner-Wilger, 2006) may have given us a more concrete picture of when participants used memory retrieval and when they relied on procedures. If a replication of this study were to combine the two methods, it would potentially capture both conscious and unconscious strategy use in participant responses and help determine whether shifts in strategy occurred.

Study of this area is still young, but the study of memory and procedure in mental math is not. The field of RIF is rich with possibilities to explore whether multiple associations of numbers interfere with each other to the point that a person's ability to retrieve facts from memory is inhibited, a fundamental theoretical question in the field of mental arithmetic.

REFERENCES

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49*(4), 415-445.

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., Bjork, R. A., & Bjork, E. L. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7*(3), 522-530.

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

Camp, G., & Sander, D. (2016). Competitive retrieval is not a prerequisite for forgetting in the retrieval practice paradigm. *Canadian Journal of Experimental Psychology, 70*(3), 248-252.

Campbell, J. I., & Penner-Wilger, M. (2006). Calculation latency: The μ of memory and the τ of transformation. *Memory and Cognition, 34*(1), 217-226.

Campbell, J. I., & Phenix, T. (2009). Target strength and retrieval-induced forgetting in semantic recall. *Memory and Cognition, 37*(1), 65-72.

Campbell, J. I., & Thompson, V. A. (2012). Retrieval-induced forgetting of arithmetic facts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 118.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1403-1414.

Dunn, E. W., & Spellman, B. A. (2003). Forgetting by remembering: Stereotype inhibition through rehearsal of alternative aspects of identity. *Journal of Experimental Social Psychology, 39*, 420-433.

Faulkenberry, T. J. (2017). A single-boundary accumulator model of response times in an addition verification task. *Frontiers in Psychology, 8*, 1-12.

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., & MacLeod, C. M. (2012). Retrieval-induced forgetting: Testing the competition assumption of inhibition theory. *Canadian Journal of Experimental Psychology, 66*(3), 204-211.

Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review, 120*(4), 852-872.

Jonker, T. R., Seli, P., & MacLeod, C. M. (2015). Retrieval-induced forgetting and context. *Current Directions in Psychological Science, 24*(4), 273-278.

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.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Macrae, C. N., & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77*, 463-473.

Manly, C. F., & Spoehr, K. T. (1999). Mental multiplication: Nothing but the facts? *Memory & Cognition, 27*, 1087-1096.

Myung, I. J. (2011). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology, 47*(1), 90-100.

Phenix, T. L., & Campbell, J. I. D. (2004). Effects of multiplication practice on product verification: Integrated structures model or retrieval-induced forgetting? *Memory and Cognition, 32*(2), 324-335.

Raaijmakers, J. G. W., & Jakab, E. (2012). Retrieval-induced forgetting without competition: Testing the retrieval specificity assumption of the inhibition theory. *Memory & Cognition, 40*(1), 19-27.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22*, 36-71.

Storm, B. C., & Levy, B. J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition, 40*, 827-843.

APPENDIX

Provided below is a portion of the R script developed for data analysis.

```
library(tidyverse)
library(stringr)
library(moments)

# code for ex-Gaussian modeling
# first, define fit functions
# fit and extract mu

muEG = function(dat){
  # step 1 -- define ex-Gaussian density function
  dexg <- function(x, mu, sigma, tau){
    return((1/tau)*exp((sigma^2/(2*tau^2))-
                         (x-mu)/tau)*pnorm((x-mu)/sigma-
     (sigma/tau)))
  }

  # step 2 -- define negative log-likelihood function
  nll.exg <- function(data,par){
    return(-sum(log(dexg(data,
                         mu=par[1],
                         sigma=par[2],
                         tau=par[3]))))
  }

  # step 3 -- define initial EG parameters (from Lacouture &
     Cousineau, 2008, TQMP)
  EGinit <- function(x) {
    tau = 0.8*sd(x)
    mu = mean(x) - skewness(x)
    sig = sqrt(var(x)-tau^2)
    return(c(mu, sig, tau))
  }

  # Step 4: find parameters (mu, sigma, tau) that minimize NLL
  fit <- optim(EGinit(dat), fn=nll.exg, data = dat)

  # step 5: extract mu
  return(fit$par[1])
}
```

```
######################
# fit and extract tau

tauEG = function(dat){
  # step 1 -- define ex-Gaussian density function
  dexg <- function(x, mu, sigma, tau){
    return((1/tau)*exp((sigma^2/(2*tau^2))-
                     (x-mu)/tau)*pnorm((x-mu)/sigma-
     (sigma/tau)))
  }

  # step 2 -- define negative log-likelihood function
  nll.exg <- function(data,par){
    return(-sum(log(dexg(data,
                     mu=par[1],
                     sigma=par[2],
                     tau=par[3]))))
  }

  # step 3 -- define initial EG parameters (from Lacouture &
     Cousineau, 2008, TQMP)
  EGinit <- function(x) {
    tau = 0.8*sd(x)
    mu = mean(x) - skewness(x)
    sig = sqrt(var(x)-tau^2)
    return(c(mu, sig, tau))
  }

  # Step 4: find parameters (mu, sigma, tau) that minimize NLL
  fit <- optim(EGinit(dat), fn=nll.exg, data = dat)

  # step 5: extract tau
  return(fit$par[3])
}


#################################
# data preparation
# step 1 -- load and reshape data from silly SuperLab format
     into a useable form
setwd("C:/Users/Batgirl/Desktop/chelseaData/chelseaData/")
fileList = dir("data", full.name=TRUE)
all = data.frame()
for (file in fileList){
  d_head = read_tsv(file, skip=3, n_max=1)
  d = read_tsv(file, skip=5)
```

```
  # a bit of wizardry to collapse two-row variable names into
     single name
  colnames(d) = str_c(colnames(d_head) %>%
     str_replace('_[:digit:]*',''),
                       colnames(d) %>%
     str_replace('_[:digit:]*','')
  )
  colnames(d)[12] = "pressedOrReleased" # this fixes duplicate
     variable names

  all = rbind(all,d)
}

# step 2 - cleaning and further reshaping
rawdata = all %>%
  filter(ErrorCode != "E") %>% # get rid of SuperLab's coded
     mistrials
  filter(!str_detect(EventName,'Instruction')) %>% # throw out
     "Instruction" trials
  mutate(answer=lead(InputString)) %>% # map experimenter-
     entered answer onto same line as voice-key event
  filter(str_detect(ParticipantResponse,'SV-1')) %>% # isolate
     only voice-key events
  filter(!str_detect(TrialName,'Study')) %>% # throw out "Study"
     trials
  mutate(operand1 = as.numeric(str_sub(EventName, start=-6,
     end=-6)),
         operand2 = as.numeric(str_sub(EventName, start=-2,
     end=-2))) %>%  # extract operands
  mutate(correctAnswer = ifelse(SessionOperation=="Addition",
     operand1+operand2, operand1*operand2)) %>% # calculate
     correct answer
  mutate(error = ifelse(as.numeric(answer)==correctAnswer,0,1))


# look at errors by operation (note: NA's represent mistrials
     and experimenter keyboard errors)
table(rawdata$error, rawdata$SessionOperation, useNA="ifany")
```

```
################################
# model 1 - addition pretest
# restrict to BlockName = Phase 1
# for other models, restrict different BlockName

data = rawdata %>%
  filter(str_detect(BlockName,'Phase 1')) %>%
  mutate(group =
     ifelse(str_detect(ParticipantGroup,'Retrieval'),
     "retrieval", "study")) %>%
  mutate(size = `ParticipantProblem Size`) %>%
  filter(ReactionTime > 200 & ReactionTime < mean(ReactionTime)
     + 3*sd(ReactionTime)) # removes 3SD outliers

# density plot for RTs

data %>%
  ggplot(aes(x=ReactionTime)) +
  geom_density()

set = numeric(dim(data)[1])
for (i in 1:length(set)){
  if (str_detect(data$ParticipantGroup[i], 'Set 1') &
     data$BlockSet[i]=='Set 1'){
    set[i] <- "MP"
  }
  else if (str_detect(data$ParticipantGroup[i], 'Set 2') &
     data$BlockSet[i]=='Set 2'){
    set[i] <- "MP"
  }
  else{
    set[i] <- "MU"
  }
}

data$set=set

# now do the ANOVA
# model 1 - addition pretest

# Mean RT ANOVA
anova1 = data %>%
  filter(error==0) %>%
  group_by(ParticipantName, group, set, size) %>%
  summarize(mRT = mean(ReactionTime))
model1.aov=aov(mRT ~ group*set*size +
     Error(ParticipantName/(set*size)), data=anova1)
```

```
summary(model1.aov)

# construct table and plot of means

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(mRT = mean(ReactionTime), sd = sd(ReactionTime))

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(mRT = mean(ReactionTime)) %>%
  ggplot(aes(x=group, y=mRT, group=set)) +
  geom_line(aes(linetype=set)) +
  geom_point(aes(shape=set), size=2) +
  facet_grid(~size) +
  theme_classic(18)

# Median RT ANOVA

anova1 = data %>%
  filter(error==0) %>%
  group_by(ParticipantName, group, set, size) %>%
  summarize(medRT = median(ReactionTime))

model1.aov=aov(medRT ~ group*set*size +
     Error(ParticipantName/(set*size)), data=anova1)
summary(model1.aov)

# median table/plot

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(medRT = median(ReactionTime), sd = sd(ReactionTime))

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(medRT = median(ReactionTime)) %>%
  ggplot(aes(x=group, y=medRT, group=set)) +
  geom_line(aes(linetype=set)) +
  geom_point(aes(shape=set), size=2) +
  facet_grid(~size) +
  theme_classic(18)
```

```
# fit mu for each experimental condition and participant

anova1mu = data %>%
  filter(error==0) %>%
  group_by(ParticipantName, group, set, size) %>%
  summarize(mu = muEG(ReactionTime))

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(mu = muEG(ReactionTime), sd = sd(ReactionTime))

# analyze values of mu in ANOVA model
model1mu.aov=aov(mu ~ group*set*size +
     Error(ParticipantName/(set*size)), data=anova1mu)
summary(model1mu.aov)

# plot values of mu
data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(mu = muEG(ReactionTime)) %>%
  ggplot(aes(x=group, y=mu, group=set)) +
  geom_line(aes(linetype=set)) +
  geom_point(aes(shape=set), size=2) +
  facet_grid(~size) +
  theme_classic(18)


# fit tau for each experimental condition and participant

anova1tau = data %>%
  filter(error==0) %>%
  group_by(ParticipantName, group, set, size) %>%
  summarize(tau = tauEG(ReactionTime))

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(tau = tauEG(ReactionTime), sd = sd(ReactionTime))


# analyze values of mu in ANOVA model
model1tau.aov=aov(tau ~ group*set*size +
     Error(ParticipantName/(set*size)), data=anova1tau)
summary(model1tau.aov)
```

```
# plot values of tau

data %>%
  filter(error==0) %>%
  group_by(group, set, size) %>%
  summarize(tau = tauEG(ReactionTime)) %>%
  ggplot(aes(x=group, y=tau, group=set)) +
  geom_line(aes(linetype=set)) +
  geom_point(aes(shape=set), size=2) +
  facet_grid(~size) +
  theme_classic(18)
```