Does Self-Efficacy Contribute to the Retrieval Practice Effect?

BY

ANDREA N. FRANKENSTEIN
B.S., University of North Florida, 2014
M.S., University of North Florida, 2016

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Arts in Psychology (Cognitive Psychology)
in the Graduate College of the
University of Illinois at Chicago, 2019

Chicago, Illinois

Defense Committee:

Eric D. Leshikar, Advisor
Karl K. Szpunar, Chair
Daniel P. Cervone

ACKNOWLEDGMENTS

I would like to thank my thesis committee – Eric Leshikar, Karl Szpunar, and Dan Cervone – for their support in this project. I would also like to thank my lab mates Allison Sklenar and Matt McCurdy for their input on this project.

ANF

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The retrieval practice effect is a memory benefit for information that is tested or retrieved one or more times before a final test. This memory improvement is seen when compared to instances when information is simply re-read or re-studied prior to final test. The benefits of retrieval practice are robust, extending to different types of learning materials (e.g., word pairs, text passages, lecture videos) and different settings (e.g., lab, classroom, online). There are several potential explanations for why retrieval practice benefits memory over re-studying, most of which hinge on cognitive mechanisms. However, it is possible that other, social-cognitive factors could also underlie this memory effect. One such factor could be self-efficacy, which has been shown to be positively related to memory performance. This study sought to replicate previous work showing a retrieval practice effect and examined whether a social-cognitive factor, self-efficacy, could partially explain improved performance due to retrieval practice. Final test performance for the retrieval practice and re-study conditions were compared, and self-efficacy scores were analyzed to determine if they mediated the relationship between condition (re-study, retrieval practice) and final test performance. Overall, there was a significant retrieval practice effect, with the retrieval practice group scoring significantly higher on the final test than the re-study group. This finding replicates prior work showing the memory benefits of retrieval practice. Self-efficacy was not a significant mediator of the retrieval practice effect. However, change in self-efficacy was positively related to final test performance: the larger the increase in self-efficacy, the better memory performance on the final test. These findings are in line with previous work showing that higher self-efficacy is associated with better memory. While self-efficacy was not found to be a significant mediator in this study, further research can focus on the effects of direct manipulations of self-efficacy on memory performance.

I. INTRODUCTION

A. **Background**

Research on the topic of memory has focused on discovering methods to improve memory. One such procedure that has gained traction in recent years is the retrieval practice effect. In a typical retrieval practice paradigm, the comparison of interest is between material studied under a retrieval condition relative to a re-study condition. In the retrieval practice condition, individuals answer questions about to-be-learned material (e.g., multiple choice, short answer, free recall), which requires them to retrieve the information from memory. In contrast, the re-study condition involves participants re-reading or re-viewing the material. After these intervening conditions, participants are then assessed on a final test. Retrieval practice leads to improved memory on a final test relative to re-study (see Delaney, Verkoeijen, & Spirgel, 2010; Roediger & Karpicke, 2006; Rohrer & Pashler, 2010; Rowland, 2014 for reviews), for many different types of stimuli (e.g., word lists, paired associates, prose passages, online statistics lectures; Szpunar, Khan, & Schacter, 2013; for a review see Adesope, Trevisan, & Sundararajan, 2017). These findings have immense practical implications for educational settings where students are predominately assessed via exams over course materials.

Researchers have proposed different mechanisms to explain why retrieval practice improves memory compared to re-study. Most of these accounts focus on the cognitive processes underlying the retrieval practice effect, such as strengthening semantically related information (elaborative retrieval hypothesis; Carpenter, 2009) or promoting similar processing at both encoding and retrieval (transfer-appropriate processing; Morris, Bransford, & Franks, 1977). Although these cognitive mechanisms have received a large amount of attention, less work has investigated non-cognitive mechanisms underlying the benefits of retrieval practice. One such mechanism may be self-efficacy. Self-efficacy is an individual's personal belief in their ability to succeed in specific situations or to accomplish a task in a novel, unexpected, or potentially stressful context (Bandura, 1977; Schunk, 1985). I propose that retrieval practice will improve individuals' self-efficacy, which may in turn improve their memory performance. A person's self-efficacy can influence how they approach goals and challenges, the degree of effort they put into a task, and their persistence when they encounter obstacles. Higher self-efficacy is related to increased effort, persistence, and better performance in a particular domain (Bandura 1995, 1997). Two lines of evidence point to a connection between self-efficacy and memory. First, work in academic settings shows an indirect connection between self-efficacy and memory. This line of work has shown that academic self-efficacy influences student motivation, learning, and overall achievement (Betz & Hackett, 1983; Pajares, 1996; Pajares & Schunk, 2001). In a meta-analysis, Multon, Brown, and Lent (1991) found that academic self-efficacy was positively related to both academic persistence and academic performance as measured by standardized achievement tests, classroom-related measures (e.g., course grades, GPA), and basic skills tasks, where higher levels of academic self-efficacy were associated with higher persistence and better academic performance. Robbins and colleagues (2004) also found that higher academic self-efficacy was associated with higher student GPA across numerous studies. Because academic success is tied to memory performance on educational assessments (such as tests), this points to a relationship between self-efficacy and memory. Second, self-efficacy and memory have been directly linked in studies with older adults. Older adults that had higher memory self-efficacy (higher belief in their ability to remember information) remembered more words in a free recall task (Beaudoin & Desrichard, 2017; Wells & Esopenko, 2008). The authors reported that task persistence was one of the ways in which memory self-efficacy contributed to higher memory scores. The goal of the present work is to examine the extent retrieval practice effect might increase self-efficacy, which in turn may improve memory.

Self-efficacy is a domain-specific rather than global construct in the sense that a person can have high self-efficacy in some domains (e.g., statistics, chemistry, art) but low self-efficacy in others (e.g., physics, biology, history; Bandura, 2006). Any study on the relationship between retrieval practice effect and self-efficacy should be targeted to a specific knowledge domain. In the case of the current study, participants' research methods self-efficacy will be measured, which is the knowledge domain I will be using in this study. Because this study is aimed at determining whether increased self-efficacy leads to better memory, it is important to understand how self-efficacy can be improved. Work on self-efficacy states that individuals appraise their sense of self-efficacy using information gained from different sources, some of which are mastery experience, vicarious experience, social persuasion, and physiological or emotional arousal (Bandura, 1977). Most relevant to the current study is mastery experience. A mastery experience involves the individual completing a task on their own, showing direct evidence that the goal can be reached. As such, it is the most effective at increasing a person's self-efficacy as it shows that the individual can achieve their goal or complete their task (Bandura, 1977). In an academic setting, retrieval practice could serve as an opportunity for mastery experience as it simulates the final test. Therefore, retrieval practice should increase students' sense of self-efficacy. In this work, we will investigate the extent retrieval practice influences self-efficacy, and how this in turn might lead to improved memory on a final test.

Studies have shown that self-efficacy can be improved over time and that some self-efficacy manipulations can improve learning; however, this work has not focused on retrieval practice. For example, Huang and Meyer (2018) experimentally introduced self-efficacy manipulations into an online statistics lesson to investigate whether self-efficacy would increase and whether changes in self-efficacy would improve learning outcomes. These self-efficacy manipulations included a variation of mastery experience (e.g., mentally rehearsing the solving process), vicarious experience (e.g., watching an example problem being solved), social persuasion (e.g., receiving effort-based feedback), and affective state (e.g., math anxiety coping messages). In this experiment, students rated their baseline self-efficacy, received instructions, solved practice statistics problems, and were either exposed to all four embedded self-efficacy features or no self-efficacy features (e.g., mastery, vicarious experience, persuasion, etc.). Results showed that post-lesson self-efficacy was higher for participants in the group exposed to the self-efficacy manipulations. Additionally, those in the self-efficacy experimental group performed better on practice problems, final retention performance, and learning-transfer test as compared to the control group. This evidence points to the idea that self-efficacy can be improved and leads to enhanced learning in an educational setting.

B. **Specific Aims and Hypotheses**

To date, there has been no published research exploring the relationship between retrieval practice and self-efficacy. The current study used a typical retrieval practice paradigm to investigate the effect of retrieval practice on self-efficacy and final test performance, as well as the influence of self-efficacy change on final test performance. Participants watched online lecture videos on research methods, engaged in either retrieval practice or re-study, and completed a cumulative test over all lecture content 48 hours later. Self-efficacy was measured three times during Session 1: at baseline, halfway through the learning procedure, and at the end of the learning procedure. Participants returned 48 hours later for the final test (Session 2). There were three specific aims for this project:

Aim 1: Determine whether self-efficacy would increase over the course of the experiment, and whether the degree of change would be different based on condition (re-study versus retrieval practice).

Prior work has shown that self-efficacy can be changed by engaging in a mastery experience (Bandura, 1977) and that self-efficacy manipulations can increase self-efficacy over the course of an educational lesson (e.g., Huang & Mayer, 2018). The current study extends this work by focusing specifically on retrieval practice's influence on self-efficacy. Retrieval practice should serve as a mastery experience and thus be more effective at increasing self-efficacy than re-studying (Bandura, 1977). Therefore, self-efficacy should increase more for participants who engage in retrieval practice compared to those who re-study the material.

Aim 2: Determine whether memory performance on the final test is influenced by condition (re-study versus retrieval practice).

Previous work shows that retrieval practice is an effective way to improve memory on a final test (see Adesope et al., 2017 and Rowland, 2014 for reviews). Therefore, I expect to find a typical retrieval practice effect in this study, with participants in the retrieval practice group scoring significantly better on the final test than the re-study group.

Aim 3: Determine whether changes in self-efficacy account for the retrieval practice effect.

Researchers have developed possible explanations for why retrieval practice improves memory (e.g., elaborative retrieval, transfer-appropriate processing), but these theories focus on the cognitive mechanisms underlying the retrieval practice effect. However, there may be another social-cognitive factor (self-efficacy) that may account for this memory benefit. Three lines of research hint at a relationship between self-efficacy and memory. First, meta-analyses have shown that higher academic self-efficacy is associated with better academic performance that rely at least in part on memory (e.g., test performance, course grades, GPA; Multon et al., 1991; Robbins et al., 2004). Second, adding self-efficacy manipulations to educational lessons has been shown to improve learning as measured by performance on practice problems, final retention test, and learning-transfer test (Huang & Mayer, 2017). Third, the relationship between self-efficacy and memory has been explored in older adults, showing that higher memory self-efficacy is related to better memory performance (Beaudoin & Desrichard, 2017; Wells & Esopenko, 2008). Based on this work, I predict that self-efficacy will account for at least part of the retrieval practice effect. By conducting a mediation analysis, I can determine whether the relationship between retrieval practice and final test performance decreases after accounting for changes in self-efficacy. If results show that self-efficacy is a significant mediator, this would suggest a novel mechanism underlying the retrieval practice effect.

II. METHOD

A. Participants

One hundred thirty-four students were recruited from University of Illinois at Chicago's introductory psychology Subject Pool (63.43% female, Age: M = 19.17, SD = 1.38). Participants gave their informed consent in accordance with the University of Illinois at Chicago Institutional Review Board and were awarded course credit for their participation in the study. Ten participants were removed from analysis because they failed to complete the study, leaving a total of 124 participants with complete data.

B. Stimuli

1. Lecture videos

Lecture videos from an Introduction to Research Methods in Psychology course served as the learning materials for this experiment. The four videos ranged from approximately 8:50 – 9:50 each. These videos cover topics such as operationalization of variables, reliability and validity, main types of research designs, and sources of bias (see **Table 1** for full list of topics and video lengths).

2. Retrieval practice and final test items

Thirty-two multiple-choice items were developed based on the lecture video content, with half being definitional questions and the other half being application questions. Definitional questions required participants to simple recognize the correct term, definition, or fact, whereas application questions required participants to apply previously learned knowledge to a novel example or scenario. Twenty-four items (6 per video) were used during the retrieval practice and restudy opportunities. At final test, participants answered both previously encountered retrieval items and new items (75% old, 25% new). Sample items can be found in **Appendix A**.

3. <u>Self-efficacy questionnaire</u>

Since self-efficacy is a domain-specific construct, self-efficacy items for the current study were tailored to the lecture content (e.g., psychological research methods) in accordance with Bandura's (2006) guidelines. I did this because Bandura (2006) argues against a "one measure fits all" approach when assessing self-efficacy. For the self-efficacy questionnaire, participants were asked to rate their current degree of confidence in their ability to complete tasks related to research methods (16 items; 0 "cannot do at all" to 100 "highly certain can do"). For example, "How confident are you that you can explain at least one limitation of correlational research?" (see **Appendix B** for full list of items.) The same questionnaire was used at baseline, Time 1 (T1), and Time 2 (T2) measurements.

C. <u>Procedure</u>

This experiment is a mixed design with retrieval practice as the between-subjects factor and self-efficacy as the within-subjects factor. There was a control group (re-study) and an experimental group (retrieval practice). The study was composed of two main phases: encoding phase (Session 1) and a final test phase 48 hours later (Session 2). **Figure 1** depicts the experimental procedure. Upon arrival for Session 1, participants reviewed and signed an informed consent document and filled out a demographics questionnaire. Next, they were given information about the different phases of the experiment and completed a baseline measurement of self-efficacy. Following the initial measurement of self-efficacy, participants completed the encoding phase. They returned 48 hours later for the Session 2 final test.

1. **Encoding phase**

Participants were randomly assigned to one of two conditions: re-study (control) or retrieval practice (experimental). Both groups of participants watched a total of four research methods videos and completed four cycles of re-study or retrieval (see **Figure 1**). After each video, all participants engaged in a distractor task for approximately 2-3 minutes (i.e., forward digit span, backward digit span, digit-symbol substitution, verbal fluency). After completing the distractor task, participants either engaged in retrieval practice or re-study. In both conditions, participants were exposed to 6 multiple-choice items per video (24 total). For the retrieval practice group, participants selected their answer for each item and were immediately shown the correct answer. Those in the re-study condition were shown the same multiple-choice questions but with the answers already indicated in bold-face type. After the second and fourth lecture videos, both groups rated their self-efficacy using the same self-efficacy questionnaire. All parts of the encoding phase were experimenter-paced to maintain uniformity across participants. The distractor tasks were administered on paper while the videos and retrieval/re-study materials were presented via Qualtrics on a PC computer. Upon completion of the encoding phase, participants were dismissed.

a) **Distractor tasks**

Participants completed four distractor tasks, one after each lecture video: forward digit span, backward digit span, digit-symbol substitution (WAIS-R, 1981), and phonemic F-A-S fluency task (Spreen & Benton, 1977). The forward digit span task involved participants listening to the experimenter read out series of numbers that increased in length. After each string of numbers, participants wrote down the numbers in the same order they heard them. Backward digit span involved a similar procedure: participants heard strings of numbers of increasing length, but their task was to write down each series of numbers in reverse order that they heard them. The forward and backward digit spans were scored together as the total number of digit strings that were written down correctly. In the digit-symbol substitution task (Wechsler Adult Intelligence Scale; see **Appendix C**), participants saw digits from 1 to 9 that were paired with different symbols. Below were a series of boxes with digits above and blank boxes below. Participants had 90 seconds to fill in as many blank boxes as they could with the appropriate symbols. This task was scored as the total number of boxes that were filled in correctly. For the F-A-S fluency task, participants heard a letter of the alphabet and had 1 minute to write down as many words as they could that started with that letter. They completed this task for three letters: F, A, S. The number of valid words listed for each letter were totaled, excluding proper nouns or proper adjectives. See **Table 2** for descriptive statistics for these tasks. Independent t-tests were conducted to ensure that scores on these distractor tasks did not vary by condition. Results showed that scores on the tasks were not significantly different for the re-study and retrieval practice groups, ps > .05. This indicates the two groups did not differ in fluid intelligence, and thus should be comparable.

2. **<u>Final test phase</u>**

After a 48-hour delay, participants returned to the lab to complete a self-paced, pencil-and-paper final cumulative test. Participants answered 32 multiple-choice questions covering all four lecture videos. As described above, 75% of the items were previously seen by participants during encoding while 25% were new items.

III. RESULTS

In this section, the results from several analyses will be presented. To analyze whether self-efficacy changed over time and whether the degree of change varied by condition (re-study versus retrieval practice), I conducted a mixed-effects model. Next, multiple regression analyses were used to investigate whether there was a standard retrieval practice effect and whether change in self-efficacy predicted memory as measured by final test performance. Next, final test performance was analyzed separately by item type (previously seen items, new items) to determine if the retrieval practice effect extended to new items. Finally, I conducted a mediation analysis to determine whether changes in self-efficacy accounted for the retrieval practice effect.

Participants' self-efficacy was scored by totaling their responses (0-100) for all items on the self-efficacy questionnaire. The internal consistency of these items was .94, which is above the standard cut-off of .70 (DeVellis, 2012). Self-efficacy change scores were obtained by subtracting participants' baseline self-efficacy scores from their final self-efficacy scores. Participants' final test performance was scored as the percent of items they answered correctly (out of 32), while their score for previously seen items was calculated as the percent of previously seen items they answered correctly (out of 24) and their score for new items was calculated as the percent of new items they answered correctly (out of 8). All analyses were conducted using the statistical computing software R, more specifically RStudio (RStudio Team, 2016).

A. <u>Mixed-Effects Models: Self-Efficacy</u>

To assess the effect of retrieval practice on self-efficacy, data were analyzed using a mixed-effects model. Participants' final measure of self-efficacy was the dependent variable, with retrieval condition (between-subjects factor: re-study, retrieval practice) and time (repeatedmeasure: baseline, T1, and T2) as fixed predictors. I also included each subject's slope of change over time as a random factor to account for individual differences. At baseline, individuals in the re-study group had an average self-efficacy rating of 852.87 out of 1600, and this did not differ significantly from the baseline ratings of the retrieval practice group (M = 885.90, t = 0.58; see **Figure 2**). Self-efficacy ratings significantly increased over time, t = 6.09, but the rate of change did not differ significantly between re-study (M = 183.05) and retrieval practice groups (M = 154.95), as indicated by the non-significant interaction term, t = 0.07 (see **Table 3**).

B. **Regression Models: Retrieval Practice Effect**

I next conducted a series of regression analyses to determine whether the change in self-efficacy was a significant predictor of final test performance and whether there was a typical retrieval practice effect. See **Table 4** for a summary of the results. Final test score was the outcome variable, with condition and change in self-efficacy as predictors. Overall, the regression analysis was significant, $F(3,120) = 5.28$, $p = 0.002$, $R^2 = .095$. Results showed that the average final test score for the re-study condition was 75.01%, with participants in the retrieval practice condition scoring significantly higher (6.77%) than those in the re-study condition, t = 2.70, p = .008. This is evidence of the typical retrieval practice effect. Interestingly, change in self-efficacy was a significant predictor of final test score, with larger increases associated with higher final test scores. With every 1 SD unit increase in self-efficacy, participants scored 4.42% higher on the final test, t = 2.45, p = .02. The interaction between condition and change in self-efficacy was not significant, t = -0.476, p = .63, meaning that the degree of change in self-efficacy did not differ between re-study and retrieval practice conditions.

Next, I examined the retrieval practice effect as a function of previously seen items and new items using change in self-efficacy as a co-variate (see **Table 5** for a summary of results). Past work on retrieval practice has shown mixed effects on whether performance on new items improves under retrieval practice conditions (see Adesope et al., 2017 and Rowland, 2014 for reviews). The results of this analysis will provide further insight into the effect of retrieval practice on un-tested (or new) items. The overall model was significant, $F(2,121) = 6.45$, $p = .002$, $R^2 = .08$. For previously seen items, the average score for the re-study group was 76.01% (see **Figure 3**). There was a significant retrieval practice effect, with those in the retrieval practice group scoring 6.96% higher on previously seen items than the re-study group, $t = 2.63$, $p = .01$. Change in self-efficacy was a significant predictor of participants' scores on previously seen items, with an increase of 1 SD unit in self-efficacy associated with a 3.43% increase in memory performance for previously seen items, $t = 2.58$, $p = .01$. Next, I analyzed performance on new items with retrieval condition and change in self-efficacy as predictors, and the overall model was significant, $F(2,121) = 6.23$, $p = .003$, $R^2 = .08$. Participants in the re-study condition scored an average of 71.91% on the new items (see **Figure 3**). The retrieval practice effect for new items was marginally significant, $t = 1.93$, $p = .056$, with those in the retrieval practice group scoring 6.18% higher than those in the re-study group. Interestingly, the relationship between change in self-efficacy and test performance was larger in magnitude for new items (4.93% increase per 1 SD unit increase in self-efficacy, $t = 3.06$, $p = .003$) than for previously seen items (3.43% increase per 1 SD unit increase in self-efficacy).

C. <u>**Mediation Model: Change in Self-Efficacy**</u>

To investigate whether change in self-efficacy partially mediated the relationship between condition (re-study, retrieval practice) and final test score, I conducted a causal mediation analysis using the BCa method. This method involves bootstrapping confidence intervals for the effects in the mediation analysis (direct, indirect, total effect) to determine whether mediation has occurred. The bootstrapped confidence intervals (simulations: 2,000) indicated that there was no significant mediating effect of change in self-efficacy, $p = 0.471$, CI[-0.02,0.01]. This result shows that self-efficacy did not significantly contribute to the memory benefit from retrieval practice.

IV. DISCUSSION

I was interested in whether a social-cognitive factor, self-efficacy, contributes to the memory benefit from retrieval practice (in addition to the cognitive factors that have been identified). I also sought to replicate previous work showing the effectiveness of retrieval practice in improving memory over re-study (e.g., Adesope et al., 2017; Rowland, 2014). Overall, there were three main findings in this study: a) self-efficacy increased over time; b) there was a significant retrieval practice effect consistent with past work; and c) change in self-efficacy did not mediate the relationship between retrieval condition and final test performance, but it did significantly predict final test performance. This provides evidence that there is a relationship between self-efficacy and memory.

First, I was interested in testing whether self-efficacy would improve over time and if it would increase more for those in the retrieval practice group than re-study. Although self-efficacy did increase over time, it increased by the same magnitude for participants in both the re-study and retrieval practice groups, which does not support my hypothesis. There are two potential explanations for why self-efficacy levels were similar for both groups. First, it is possible that the act of re-studying over-inflated individuals' self-efficacy so that their scores mirrored that of the retrieval practice group. Participants who engaged in retrieval practice may have had more accurate self-efficacy ratings, but with the potentially distorted ratings of the re-study group, the scores of the two groups were not reliably different. If it is the case that participants in the re-study condition had over-inflated self-efficacy ratings, this might have obscured our ability to detect a relationship between self-efficacy and the retrieval practice effect. Research has shown that re-studying material may lead to overconfidence, and that retrieval practice is one tool that can help reduce over-inflated confidence (Miller & Geraci, 2014). Further research is necessary to determine how self-efficacy is differentially affected (inflated or accurate) by re-study and retrieval practice. Second, retrieval practice may not have acted as a sufficiently strong mastery experience to increase self-efficacy more than re-studying. Although the retrieval practice items mirrored the items on the final test, participants may not have received enough feedback to make retrieval practice a strong enough mastery experience. It may take more elaborated feedback such as knowing the overall percentage of items answered correctly, or how they perform on the task relative to other people, before retrieval practice might reflect useful mastery experience information to affect self-efficacy. Although self-efficacy was not improved more by retrieval practice compared to re-study, it is still possible that such a relationship exists. It may take further experimental work to tease apart this relationship. To provide a stronger test between the relationship between retrieval practice and self-efficacy, future work might manipulate the feedback that participants get to either increase or decrease self-efficacy regardless of actual performance on retrieval practice questions.

I was interested in examining how a social-cognitive factor (self-efficacy) contributes to the retrieval practice effect, but it is also worth mentioning that the self-efficacy measure could be related to another well-studied cognitive factor: judgments of learning (JOLs). In JOL experiments, participants might study lists of words or other materials and are subsequently asked (either immediately or after a delay) to rate how likely they would be to remember *that* stimuli (e.g., Arbuckle & Cuddy, 1964; Rhodes & Castel, 2008). Accuracy of JOLs (how well participants can predict whether they will later remember an item) is then measured by a final memory test. Although it is possible that there is overlap between JOLs and self-efficacy, we argue that self-efficacy substantively differs from JOLs because self-efficacy is not focused on predicting accuracy on future memory tasks, per se. Rather, self-efficacy is an individual's *confidence* in their ability to do well on specific tasks. Another reason that self-efficacy and JOLs can be considered different constructs is that work in these respective domains tend to differ in terms of their relationship with task performance. For example, the correlation between JOLs and memory performance tends to relatively moderate, with individual study estimations anywhere between .09 and .48 (Vesonder & Voss, 1985), .21 and .30 (Leonesio & Nelson, 1990), and .24 and .39. On the other hand, a meta-analysis investigating academic self-efficacy's relationship with GPA found that the true population correlation between these variables (after correcting for measurement error) was close to .496 (Robbins et al., 2004). Although the upper-bound of the correlations between JOLs and memory performance is close to the correlation between academic self-efficacy and GPA, on average self-efficacy is a better predictor of its outcomes than JOLs. Based on these lines of evidence, we argue JOLs and self-efficacy are separate constructs. Future work could directly test the relationship of these constructs (self-efficacy and JOLs) to memory performance to provide stronger evidence that these are distinct psychological phenomena. Next steps in my research can also include a more direct investigation of JOLs, self-efficacy, and memory performance in the context of a retrieval practice paradigm, to conclude that JOLs are different constructs that differentially relate to memory performance.

Second, we found a significant retrieval practice effect, which replicates previous research on retrieval practice (e.g., Adesope et al., 2017; Rowland, 2014). Consistent with prior retrieval practice studies, I found that the memory benefit of retrieval practice extended to both previously tested items and new items, though the effect for new items was less robust. This is in line with work showing that the retrieval practice effect is consistent for previously seen items, while the results are more mixed for new items (e.g., Adesope et al., 2017). The fact that we saw a marginally significant effect for new items suggests that there may have been transfer of learning to previously untested material due to retrieval practice. This might lend additional support to the elaborative retrieval hypothesis (Carpenter, 2009), which states that retrieving information from memory strengthens semantically related information, making it easier to retrieve at final test.

Third, the primary interest in this experiment was to examine whether self-efficacy might account for a portion of the memory improvement attributable to retrieval practice. To evaluate this hypothesis, I performed a mediation analysis to determine if changes in self-efficacy accounted for the difference in memory performance between re-study and retrieval practice. The results showed that participants' change in self-efficacy did not significantly mediate the relationship between retrieval condition and final test performance; however, it was clear that change in self-efficacy was independently predictive of final test performance. Larger increases in self-efficacy were associated with better memory performance on the final test in both the retrieval practice and re-study conditions. Interestingly, the relationship between change in self-efficacy and memory performance was slightly higher for new items than for previously seen items (4.93% versus 3.43% increase in performance per 1 SD unit increase in self-efficacy, respectively). As there was no interaction between self-efficacy change and retrieval condition, this stronger relationship between change in self-efficacy and memory performance for new items was similar for all participants. Although not causal, this relationship may provide insight into how self-efficacy is related to memory performance. Prior research has found that higher levels of self-efficacy are associated with increased persistence and academic achievement (Multon et al., 1991; Robbins et al., 2004), as well as with better memory performance in older adults (Beaudoin & Desrichard, 2017; Wells & Esopenko, 2008).

The current study provides further evidence that retrieval practice is an effective technique for improving memory for educational materials. As research has shown, retrieval practice can be easily incorporated in applied settings (Bangert-Drowns, Kulik, & Kulik, 1991; Leeming, 2002; Screbo, Warm, Dember, & Grasha, 1992; Szpunar et al., 2013), and the evidence from this study indicates that it is worthwhile to implement this technique in learning environments. Additionally, this study shows that self-efficacy is a significant predictor of memory performance for both previously seen test items and new test items. The next step in this line of research is to directly manipulate self-efficacy in a retrieval practice paradigm. For example, different types of feedback may differentially affect self-efficacy in the context of retrieval practice: false feedback either positive or negative in nature, feedback on performance or effort, etc. are potential avenues to explore. If directly increasing or decreasing self-efficacy leads to differences in the retrieval practice effect, then there would be more direct and causal evidence that self-efficacy plays a role in the retrieval practice effect.

V. CONCLUSION

This study replicated previous research that showed retrieval practice as an effective technique for improving memory performance compared to re-study. Self-efficacy did not mediate the relationship between condition (re-study, retrieval practice) and memory performance. However, change in self-efficacy was positively related to memory performance: the more individuals' self-efficacy improved, the better their memory performance. This suggests that self-efficacy is an important factor when considering memory, but more research is needed to establish a causal link.

CITED LITERATURE

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701.

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology, 81*(1), 126-131.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies*. New York: Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*, *5*(307-337).

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research, 85,* 89–99.

Beaudoin, M., & Desrichard, O. (2017). Memory self-efficacy and memory performance in older adults: The mediating role of task persistence. *Swiss Journal of Psychology, 76*(1), 23-33.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23,* 329-345.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1563-1569. doi:http://dx.doi.org/10.1037/a0017021

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53*, 64-111.

DeVellis, R.F. (2012). *Scale development: Theory and applications*. Los Angeles: Sage.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Huang, X., & Mayer, R. E. (2018). Adding self-efficacy features to an online statistics lesson. *Journal of Educational Computing Research*, 1-35.

Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology, 29,* 210–212.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 464-470.

McDaniel, M. A., & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 371–385.

Miller, T. M., & Geraci, L. (2014). Improving metacognitive accuracy: How failing to retrieve practice items reduces overconfidence. *Consciousness and Cognition: An International Journal, 29*, 131-140.

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

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology*, *38*(1), 30-38.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66,* 543-578.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426-443.

Pajares, F., & Schunk, D. (2001). The development of academic self-efficacy. *Development of achievement motivation.* United States.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615-625.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis.*Psychological Bulletin, 130*(2), 261-288.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181-210.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39,* 406-412.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432-1463.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc. Boston, MA.

Screbo, M. W., Warm, J. S., Dember, W.N., Grasha, A. F. (1992). The role of time and cuing in a college lecture. *Contemporary Educational Psychology, 17,* 312–328

Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools,* 208-223.

Spaniol, S. R. (2018). Students' mathematics self-efficacy, anxiety, and course level at a community college*, Information & Learning* (in press).

Spreen, O., & Benton, A. L. (1977). *Neurosensory Center Comprehensive Examination for Aphasia*. Victoria BC: University of Victoria, Neuropsychology Laboratory.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 110*(16), 6313-6317.

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

Vesonder, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language, 24*(3), 363-376.

Wells, G. D., & Esopenko, C. (2008) Memory self-efficacy, aging, and memory performance: The roles of effort and persistence. *Educational Gerontology, 34*(6), 520-530.

Table I. Research methods lecture video length and video topics.

	Length	Topics
Video 1	8:59	Scientific Method Operationalizing Variables Reliability and Validity
Video 2	9:50	Research Designs Overview Observational Research Survey Research
Video 3	9:08	Correlational Research Experimental Research
Video 4	8:53	Random Assignment Controlling Bias

Table II. Descriptive statistics for distractor tasks by condition.

	Digit Span Total	**Digit-Symbol Substitution Total**	**F-A-S Fluency Total**
Condition	M (SD)	M (SD)	M (SD)
Re-Study Condition	18.45 (4.07)	61.44 (11.70)	37.98 (9.63)
Retrieval Practice Condition	19.02 (3.07)	63.73 (9.69)	38.64 (9.46)

Table III. Summary of mixed-effects model assessing change in self-efficacy by condition, with co-efficient estimates and standard error (in parentheses).

	Interaction Model
Fixed Effects	
(Intercept: Baseline Self-Efficacy for Re-Study)	852.87* (40.47)
Time	183.05* (30.07)
Condition: Retrieval Practice	33.03 (57.24)
Time x Retrieval Practice Condition	-28.10 (42.53)
Random Effects	
Random Slope of Time per Subject	15444
Goodness of Fit	
AIC	3370.3
BIC	3391.3
Deviance	3358.3

Table IV. Summary of regression model assessing overall final test performance by condition, with co-efficient estimates and standard error (in parentheses).

	Interaction Model
(Intercept: Final Test Performance for Re-Study)	75.010***
	(1.770)
Condition: Retrieval Practice	6.769**
	(2.504)
Change in Self-Efficacy	4.422*
	(1.805)
Retrieval Condition x Change in Self-Efficacy	-1.197
	(2.515)
R^2	0.117
Adj. R^2	0.095
Num. obs.	124
RMSE	13.918

Table V. Summary of regression models assessing final test performance by item type, with co-efficient estimates and standard error (in parentheses).

	Previously Seen Items	**New Items**
(Intercept: Test Performance for Re-Study)	76.089***	71.908***
	(1.868)	(2.268)
Condition: Retrieval Practice	6.961**	6.184†
	(2.644)	(3.210)
Change in Self-Efficacy	3.431*	4.929**
	(1.328)	(1.611)
R^2	0.096	0.093
Adj. R^2	0.081	0.078
Num. obs.	124	124
RMSE	14.701	17.845

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, †$p < 0.1$

Figure 1. Experimental design including encoding and final test phases.

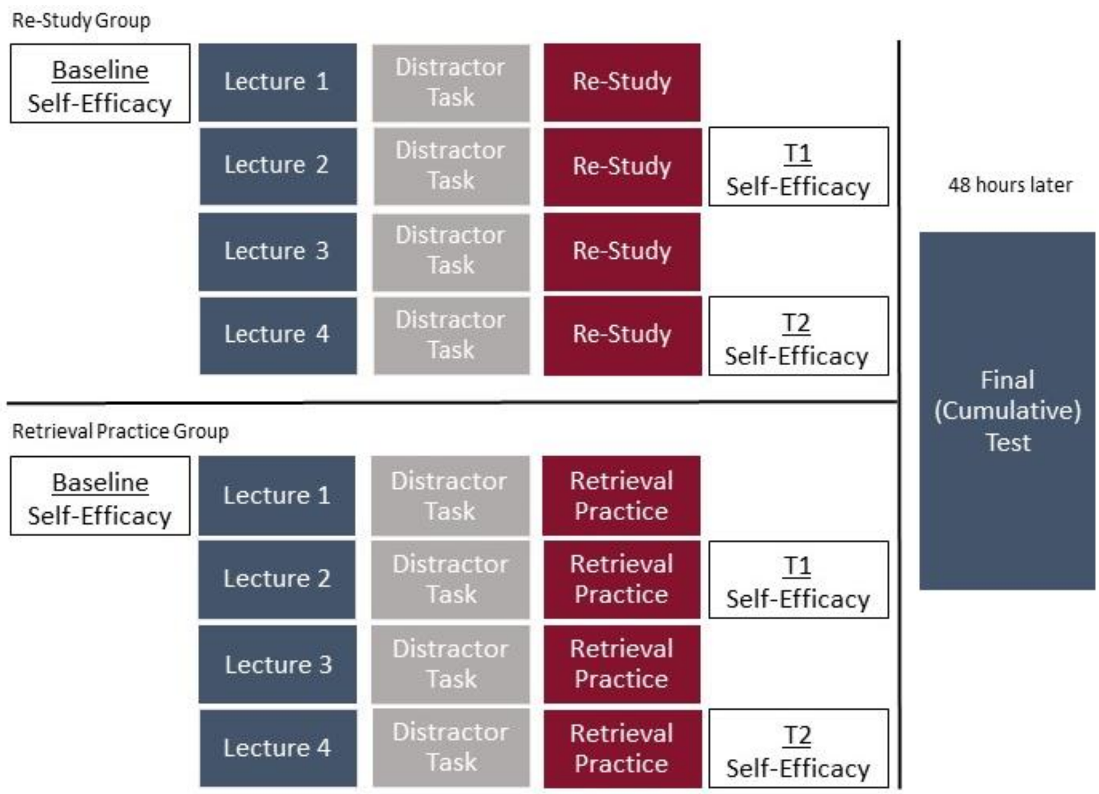


Figure 2. Total self-efficacy scores as a function of condition (re-study, retrieval practice) and time (Baseline, T1, T2). Self-efficacy increased significantly over time, but these increases did not differ as a function of condition. Error bars represent standard error of the mean.

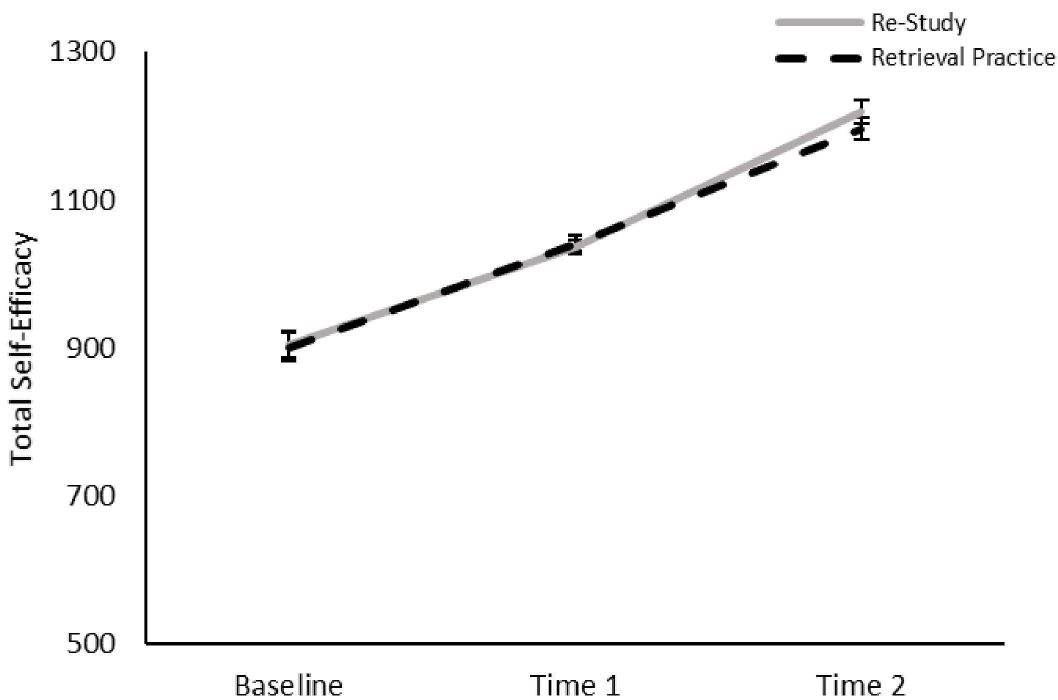
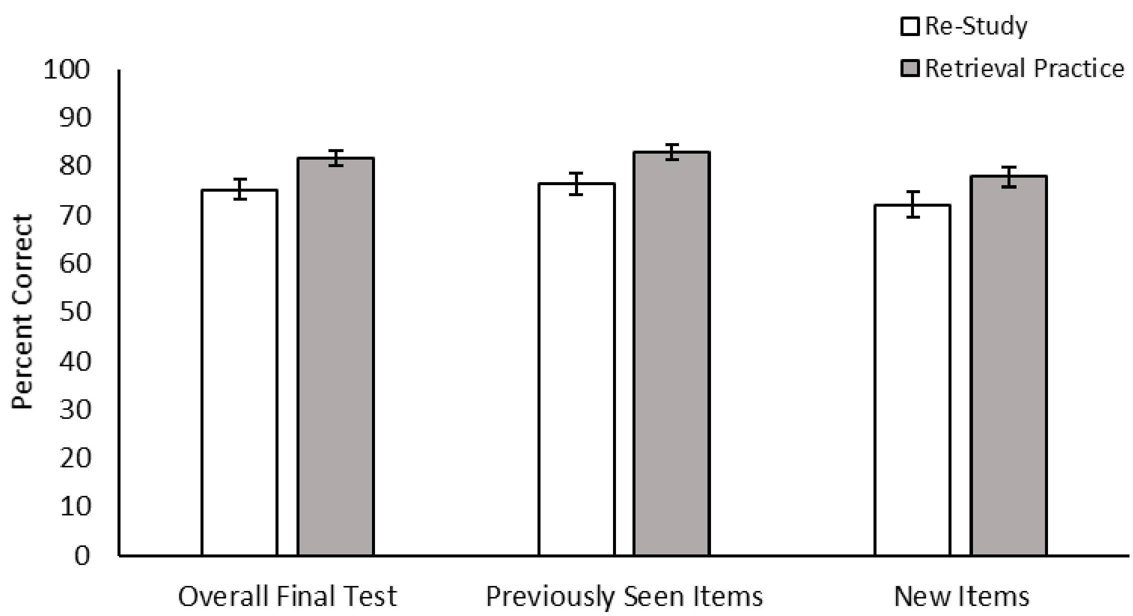


Figure 3. Memory performance as a function of condition (re-study, retrieval practice) and item type (overall final test, previously seen items, and new items). The retrieval practice group performed significantly better than the re-study group on the overall final test and on previously seen items. The difference in memory performance between the retrieval practice group and re-study group for new items was marginally significant. Error bars represent standard error of the mean.



APPENDIX A

(Definitional)

What term refers to making an abstract concept more concrete so that it can be measured?

- Independent variable
- Dependent variable
- **Operationalizing**
- Manipulating

(Application)

Which of the following is a way to manipulate fatigue (tiredness)?

- Keep half the participants up all night and let the other half sleep normally
- Ask participants how many hours they slept the night before
- Ask half the participants to exercise for 30 minutes and ask the other half sit quietly for 30 minutes
- All of the above
- **A & C**

(Definitional)

Reliability is...

- Whether a measure is measuring what it's supposed to
- **Whether a measure is consistent**
- Whether a measure is abstract enough
- Whether a measure is concrete enough

(Application)

A researcher created a new measure and wants to make sure it's reliable. To do this, she asks a group of people to complete her measure one day, then come back a week later to complete the same measure again. What type of reliability is this?

- Split-half reliability
- **Test-retest reliability**
- Reversed forms reliability
- Alternate forms reliability

(Application)

A researcher is interested in learning about how university students in the library will react to disturbing noise. The researcher sits in a crowded area of the library and starts playing music from her laptop speakers. She then records students' reactions to the music. What type of research design is this?

- Experiment
- Observation
- **Participant observation**
- Correlation

(Application)

A researcher collected data on students' sleep schedules and their exam grades. Data showed that the more students slept, the better their exam scores. Which type of correlation does this illustrate?

- **Positive**
- Negative
- Zero correlation

(Definitional)

What is the main advantage of experimental research?

- Experiments in the lab are very similar to the real world
- **Experiments can show cause and effect relationships**
- Experiments can be used to study all research questions
- Experiments do not require a lot of resources

(Definitional)

Why is random assignment important for determining cause and effect?

- It makes sure people do not have an equal chance of being chosen
- It adds confounding variables to the experiment
- **It eliminates alternative reasons for changes in the dependent variable**
- It influences the participants' responses

APPENDIX B

The following questions assess your confidence in your ability to complete certain tasks related to research methods. Please answer these questions based on your **current** levels of confidence, rather than how confident you think you might be later.

Rate your degree of confidence by recording a number from 0 to 100 using the scale given below:

0	10	20	30	40	50	60	70	80	90	100
Cannot do at all					Moderately can do					Highly certain can do

	Confidence (0-100)
If you were given an example research study, how confident are you that you would <u>know what research design</u> was being used?	
If you were given an example research study, how confident are you in your ability to <u>identify at least one strength and one weakness</u> of the study?	
If you were asked to <u>point out the main characteristics of a research study</u>, how confident are you that you would be able to do so?	
If you were told about a <u>potential bias</u> in a research study, how confident are you that you could <u>come up with a solution</u> to fix the bias?	
How confident are you in your ability to <u>explain when researchers can make cause-and-effect claims</u>?	
How confident are you in your ability to <u>explain the importance of random assignment</u> in a research study?	
If you were given an example research study, how confident are you that you could <u>identify at least one potential confound</u> in that study?	
If you were given an example research study, how confident are you that you could <u>identify the independent and dependent variables</u> in that study?	
If you were given a psychological construct (ex: happiness, anxiety), how confident are you in your ability to <u>come up with at least one way to measure</u> that construct?	
If you were given an example study, how confident are you that you could <u>determine whether a variable is being measured or manipulated</u>?	

If you were given a psychological measure, how confident are you in your ability to come up with at least one way to test the reliability of the measure?	
If you were given a psychological measure, how confident are you in your ability to tell whether the measure is valid?	
Given an example correlation coefficient (ex: r = 0.5, r = -0.3), how confident are you in your ability to explain what the correlation coefficient means?	
How confident are you that you can explain at least one limitation of correlational research?	
Given an example of an observational study, how confident are you that you could determine what type of observation was being used?	
How confident are you in your ability come up with at least one strength and one weakness of survey research?	

APPENDIX C

DIGIT	1	2	3	4	5	6	7	8	9
SYMBOL	—	⊥	⊐	L	U	O	Λ	X	=

SCORE

SAMPLES

2	1	3	7	2	4	8	1	5	4	2	1	3	2	1	4	2	3	5	2	3	1	4	6	3

1	5	4	2	7	6	3	5	7	2	8	5	4	6	3	7	2	8	1	9	5	8	4	7	3

6	2	5	1	9	2	8	3	7	4	6	5	9	4	8	3	7	2	6	1	5	4	6	3	7

9	2	8	1	7	9	4	6	8	5	9	7	1	8	5	2	9	4	8	6	3	7	9	8	6

VITA

NAME: Andrea Nicole Frankenstein

EDUCATION: B.S., Psychology, University of North Florida, Jacksonville, Florida, 2014

M.S., General Psychology, University of North Florida, Jacksonville, Florida, 2016

M.A., Cognitive Psychology, University of Illinois at Chicago, Chicago, Illinois, 2019 (in progress)

PUBLICATIONS: Leshikar, E. D., Leach, R. C., McCurdy, M. P, Trumbo, M. C., Sklenar, A. M., Frankenstein, A. N., & Matzen, L. E. (2017). Transcranial direct current stimulation during encoding of dorsolateral prefrontal cortex improves recall but not recognition memory. *Neuropsychologia*, 106, 390-397.

Alloway, T. P., Robinson, T., & Frankenstein, A. N. (2016). Educational application of working memory training. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications*. New York: Springer.

PRESENTATIONS: Frankenstein, A.N., McCurdy, M.P., Sklenar, A.M., & Leshikar, E.D. (February 2019). *That Was Unexpected: How Expectancy Confirmation and Violation Affect Memory for Social Behaviors*. Poster presented at the Annual Meeting of the Society for Personality and Social Psychology, Portland, OR, USA.

Frankenstein, A. N., McCurdy, M.P., Sklenar, A.M., & Leshikar, E.D. (November 2018). *Predicting Social Behaviors: Do We Remember Better When We're Wrong?* Poster presented at the Annual Meeting of the Psychonomic Society, New Orleans, LA, USA.

Frankenstein, A. N., McCurdy, M. P., Sklenar, A. M., Leshikar, E. D. (November 2017). *Future Thinking: Influence of Prediction Outcome on Memory.* Poster presented at Psychonomic Society Annual Meeting, Vancouver, British Columbia.

SERVICE: Graduate Assistant. *Cognitive Psychology Program,* University of Illinois at Chicago. (2018 – 2019)

TEACHING: Teaching Assistant, *Department of Psychology*, University of Illinois at Chicago. (2016 – present)

Instructor of Record, *Department of Psychology*, University of North Florida. (2015 – 2016)

EXPERIENCE: Graduate Student, *Functional Aging Brain Laboratory*. University of Illinois at Chicago. (2016 – present)

Graduate Student, *Working Memory Lab*, University of North Florida. (2014 – 2016)

Undergraduate Research Assistant, *Working Memory Lab*, University of North Florida. (2013 – 2014).