10-30-2020

What Trying to Forget Tells Us About Trying to Remember: A Link Between Associative Memory and Directed Forgetting

Brette E. Lansue Burns
University of Windsor

What Trying to Forget Tells Us About Trying to Remember: A Link Between Associative Memory and Directed Forgetting

By

Brette E. Burns (Lansue)

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts
at the University of Windsor

Windsor, Ontario, Canada

2020

What Trying to Forget Tells Us About Trying to Remember: A Link Between Associative Memory and Directed Forgetting

by

Brette E. Burns (Lansue)

APPROVED BY:

P. Weir
Department of Kinesiology

J. Casey
Department of Psychology

L. Buchanan, Advisor
Department of Psychology

August 12, 2020

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Much of directed forgetting research has pitted two competing hypotheses against each other: selective rehearsal and retrieval inhibition. Primarily, the current work explores a novel link between directed forgetting and associative memory, such that the item and cue are unitized and encoded together at study and at test, participants make their old/new judgement based on this association. Specifically, a more typical directed forgetting procedure where participants were told to remember specific items (R-cue) and forget others (F-cue) was compared with a procedure where participants were told to remember which cue (Blue-cue vs. Yellow-cue) was associated with each item. At test, participants were required to indicate which cue items were presented with at study (i.e. R, F, or a new item or Blue, Yellow, or new). As expected, a directed forgetting effect was only found using the typical procedure, with better discriminability for R-cued items than F-cued items. Although the overall level of accuracy in tagging was relatively low for all R-cued, F-cued, Yellow-cued, and Blue-cued items, there were no differences found between correctly tagged R- and F-cued items and correctly tagged Yellow- and Blue-cued items. Despite a directed forgetting effect, participants were equally able to tag R-cued and F-cued items (or Yellow- and Blue-cued items). Our results suggest that conceptualizing directed forgetting as a type of tagging through associations is worthy of follow-up research. The present work is also the first to directly compare the effect of cue timing on directed forgetting. Results demonstrate that directed forgetting is most efficient when cues are placed before the study items, as participants tend to remember R-cued information and forget F-cued information most with this cue timing thus implicating a role for rehearsal as opposed to retrieval processes in the directed forgetting effect.

ACKNOWLEDGEMENTS

To my mentor, Dr. Lori Buchanan: Thank you for your unwavering support throughout my journey in academia thus far. You have an incredible ability to empower and inspire young women to achieve their goals; I feel lucky to have the opportunity to work with you.

To my committee members, Dr. Patricia Weir and Dr. Joseph Casey: Thank you for fostering my ideas and giving constructive and insightful feedback along the way.

To my research assistants, Christopher McCoy, Katarina Kolobaric, and Cassidy Fahringer: Thank you for the many hours you devoted to this project, I could not have done it without you!

To my labmates: Thank you for keeping me sane with endless amounts of laughs and delicious food; I am so grateful to be part of such a loving and supportive group. A special thanks to Julia Borsatto, Tara McAuley, and Jessica Hurtubise – you have all provided me with endless support and happiness, and have helped to keep me calm during difficult times throughout the past two years. I am unbelievably grateful to have you in my life! Another special thank you to Daniela Wong Gonzalez – I cannot express how much your guidance and mentorship has impacted me. Thank you for allowing me to work with you, showing me the ropes, and promoting my skills as a researcher. I am so appreciative of the time you have invested in me!

To Grant: Thank you for your unconditional love and encouragement. Thank you for being my voice of reason and never letting me give up. You motivate me to keep pushing for success. I truly could not have made it here without you.

TABLE OF CONTENTS

CHAPTER 1

REVIEW OF LITERATURE

Introduction to Directed Forgetting

The act of forgetting information is often perceived in a negative light. For instance, we use unfavourable terms such as *failure*, *loss*, *impairment,* or *deficit* to describe the act of forgetting (Sheard & MacLeod, 2005). In contrast, individuals who can remember large amounts of information are often highly regarded in society as extreme intellectuals, prodigies, or even geniuses. There is no doubt that the benefits of remembering are more salient than those of forgetting.

It is easy to think that all of our lapses in memory occur unintentionally, as we often do not "try" to forget, especially the important things like our mother's birthday or where we parked our car. In light of this, it is common to overlook the importance of the ability to forget. Our environment is extremely intricate, forcing us to process a colossal amount of information simultaneously and often unconsciously. Imagine our memory as a hard-drive and how much space it would require to remember every single detail in our surroundings attained in just in a five-minute span of walking down the street. Therefore, our ability to forget irrelevant, or outdated, information is crucial for the successful memory of material that we deem to be important (Sheard & MacLeod, 2005). There are also some instances when we need to purposefully push memories aside, either because the memory is embarrassing or traumatic, or because it is causing interference with other important cognitive processes (Geiselman et al., 1983). As illustrated, the ability to intentionally forget information can be adaptive, and as a result, cognitive psychologists have shown interest in understanding the underlying mechanisms.

It is important to distinguish between intentional forgetting, that which will be discussed in this thesis, and the spontaneous forgetting that often happens by chance in everyday life. Intentional forgetting is a motivated effort to limit the future expression of some specific stimuli or memory, whereas spontaneous forgetting is unmotivated and occurs despite the relevance of that stimuli (Johnson, 1994). In order to tap into this type of intentional forgetting in laboratory studies, researchers instruct participants to remember or forget specific information using a technique called *directed forgetting* (Basden & Basden, 1998).

In the classic directed forgetting paradigm, participants are presented with stimuli (typically words) and are instructed to either remember or forget the specific stimulus (Bjork, 1970). There are two popular variations of this paradigm: *item-method* and *list-method* directed forgetting. In the item-method procedure, a cue to either remember (R-cue) or forget (F-cue) is presented after each item, and the participant's memory is later tested, whereas in the list-method procedure, a single cue is given following a list of items (typically 10-20 words) and the memory task is generally completed after the presentation of multiple lists (Anderson, 2005; MacLeod, 1998). A *directed forgetting effect* is found when memory is worse for items presented with the to-be-forgotten cue (F-cued) compared to items that are presented with the to-be-remembered cue (R-cued; Basden & Basden, 1998).

Bjork and colleagues (1968) have been credited among the first researchers to use a variant of the directed forgetting method. Participants were presented with 48 lists consisting of a series of digits and one or two consonant quadrigrams (e.g., CKRT). Participants read each digit and quadrigram aloud as they were presented and were asked to recall the consonant items after each list. Some lists required participants to recall the one or two quadrigrams presented, while another list provided a signal to forget the first quadrigram and recall only the second. Bjork et al. (1968) found that recall was worst when participants had to remember both quadrigrams presented, whereas accuracy was similar for when they had to recall just one quadrigram presented and when they could forget the first of the two. Bjork and colleagues (1968) attributed these results to the proactive interference of the first item to the second, and in the condition where participants were signalled to forget the first item, they were able to limit the influence of interference. Follow-up studies found complementary results demonstrating that an instruction to forget reduces proactive interference on R-cued items (Bjork, 1970; Epstein, 1969; Muther, 1965; Turvey & Wittlinger, 1969). Bjork and his colleagues used the directed forgetting method mainly in the study of "memory updating," or how we work with incoming information that makes previous knowledge irrelevant or unnecessary (MacLeod, 1998).

The item-method directed forgetting manipulation is also an effective method to study intention to learn, distinguishing between intentional and incidental encoding. It is posited that the instruction to remember elicits intentional or purposeful encoding of information, whereas the instruction to forget evokes more incidental or spontaneous encoding (Johnson, 1994).

Theoretical Perspectives of Directed Forgetting

As briefly described above, directed forgetting research has a rich history and is a well-establish technique. There have been several proposed explanations to describe the underlying cognitive mechanisms of the successful instruction to forget, although they are still under debate. Bjork (1970) was one of the first to propose a theory of intentional forgetting. Based on a series of three experiments, he suggested that participants use the forget instruction in two ways: (1) they differentially group the R- and F-cued items (differential grouping); and (2) as soon as an F-cue is presented, the participants dedicate all resources to rehearsing the previously presented R-cued items (selective rehearsal). Bjork (1970) further suggested that neither of these two mechanisms alone could fully explain intentional forgetting. Specifically, in Experiment 3, the location of a list of F-cued items was manipulated listwise, such that they appeared either before or after the list of R-cued items, and was compared to a condition in which participants were responsible for remembering all items. At test, performance was much better when the F-cued stimuli preceded the R-cued stimuli. Bjork (1970) posited that if differential grouping was solely responsible, it would not matter where the F-cued items were located because participants would only need to search in the set of R-cued items, but this was not the case. Selective rehearsal alone was also ruled out through results from a study conducted by Reitman et al. (1973), who demonstrated that the "forgotten" information is not completely erased. In fact, this information was still remembered at better than chance levels, even though participants should not have been rehearsing it. Consistent with this, Bjork (1970) suggested that both differential grouping and selective rehearsal must work together to explain intentional forgetting, where the items are first differentially grouped, and those R-cued items are allocated all rehearsal resources making them more readily available, and the F-cued items are left unattended and thus, less available.

Although the selective rehearsal hypothesis seems to be a parsimonious account of directed forgetting, the results of early studies also suggest that there may be another mechanism involved at retrieval. For instance, using pupillary dilation as a measure of cognitive load, Johnson (1971) found a distinct pattern consisting of a fast increase followed by a sizeable decrease in participants' pupil dilation following a listwise cue to forget. Further, Jongeward and colleagues (1975) modified the procedure to limit the opportunity for rehearsal by presenting sets of words, one at a time, followed by a 3 s rehearsal period, which was then followed by a 1 s cue indicating which of the words, if any, were to be remembered. Thus, there was no opportunity to selectively rehearse items, as they were not made aware which of these words would need to be remembered until the very quick presentation of the memory cue. Regardless, performance was much better for R-cued items (35%) than F-cued items (5%) on a free recall test (Jongeward et al., 1975). Taken together, these results led Geiselman et al. (1983) to suspect that a retrieval-based inhibition process, whether conscious or unconscious, was operating in response to a forget instruction.

Basden et al. (1993) argued that different processes were likely responsible for the item- and list-method paradigms. As such, they conducted a series of four experiments directly comparing the two methods and examining differences in recognition and recall on implicit and explicit memory tests. Using the list-method, the directed forgetting effect was only present in the recall test; with the item-method, however, the effect was found in both recognition and recall tests (Basden et al., 1993). In a similar comparison, Woodward et al. (1974) found better recall and recognition of F-cued items in the list-method than in the item-method, suggesting the to be forgotten items in the latter method are less elaborately rehearsed. Taken together, these results support a dual-process claim and provide evidence that list-method directed forgetting was likely due to retrieval inhibition, whereas selective rehearsal was likely responsible for the effect found with the item-method (Basden et al., 1993; Woodward et al., 1974).

To further differentiate the two directed forgetting methods, performance in explicit and implicit memory tasks were compared. MacLeod (1989) posited that if selective rehearsal were responsible and R-cued items were encoded more strongly than F-cued items, the directed forgetting effect should only be found in explicit memory tests; if retrieval inhibition is responsible, then the effect should occur in both implicit and explicit tests, as retrieval-based processes are assumed to operate similarly on both types of tests. Using the item-method, Paller (1990) found a directed forgetting effect in an explicit stem-cued recall test but not in an implicit stem completion task. This was the case even when the cue and item were presented simultaneously in attempt to limit the opportunity for rehearsal. Further, Basden et al. (1993) found a directed forgetting effect in an implicit word association task with the list-method, but not with the item-method. However, in the explicit memory task, the magnitude of the directed forgetting effect was larger in the item-method than in the list-method (Basden et al., 1993). These distinct differences in performance on implicit and explicit tests between methods provides further evidence that these methods recruit different underlying cognitive mechanisms.

Other early theories that have largely and quickly been discounted include deletion or erasure (Muther, 1965), repression (Weiner, 1968), and selective search (Epstein, 1969); however, much of the debate has pitted selective rehearsal and retrieval inhibition hypotheses against each other, despite these early studies suggesting of multiple underlying mechanisms.

The Selective Rehearsal Hypothesis. As discussed above, the generally accepted explanation for the directed forgetting effect found using the item-method procedure has been selective rehearsal, such that the items are initially held in working memory until the cue is given, at which time the items followed by the R-cue are given more elaborate rehearsal and those followed by the F-cue are dropped (Johnson, 1994). According to this hypothesis, F-cued items will have only been encoded with maintenance rehearsal, whereas the R-cued items will have been more extensively rehearsed. There seems to be an obvious link between better memory for R-cued items and selective rehearsal. Electrophysiological evidence supports this association, with greater event-related potentials (ERP) positivity during the 200-800 ms time window over parietal scalps following R-cues compared with F-cues (Paller, 1990; Paz-Caballero et al., 2004). Paz-Caballero and colleagues (2004) suggest that this activity in the parietal area implies that prior to any instruction, the information is kept on "stand-by," at which point an R-cue would activate the continued processing of that information. However, Paz-Caballero (2004) et al. also found that F-cues activate mechanisms in frontal and prefrontal areas – areas that have been associated more with inhibition. Therefore, although mechanisms surrounding F-cues might be better explained by inhibition, selective rehearsal appears to be contributing to R-cued information. Research and results using ERP data is more thoroughly explained in the following section.

To confirm the selective rehearsal hypothesis, both Wetzel and Hunt (1977) and Lee et al. (2007) increased the opportunity for rehearsal of R-cued items by varying the delay of the cue, the duration of the cue presentation (Wetzel & Hunt, 1977), and the post-cue interval (Lee et al., 2007), hypothesizing that only R-cued items should benefit from the increase in rehearsal time. Although the increase in duration provided a larger advantage to R-cued items, memory for both R- and F-cued items improved with the increase (Lee et al., 2007; Wetzel & Hunt, 1977). Interestingly, Bancroft et al. (2013, Experiment 1) obtained a directed forgetting effect with cue durations as brief as 300, 600, and 900 ms (much faster than the typical cue presentation speed, which ranges from 1s to 3s), but both R- and F-cued items benefitted equally from an increase in cue duration (i.e., 300 to 900ms). In all studies described above, the magnitude of the directed forgetting effect was independent of the increased rehearsal time, meaning that better recognition with increase in cue delay or duration was similar for R- and F-cued items. directly contrasting the selective rehearsal hypothesis. Based on the effect being present even when very brief cue durations were used, Bancroft et al. (2013) suggest that there may be an initial burst of rehearsal (as brief as a single rehearsal) following an R-cue that is sufficient enough to give an advantage over F-cued items.

It has commonly been suggested that participants cannot immediately stop processing the F-cued items, even though they may not be consciously rehearsing them (Bancroft et al., 2013; Lee et al., 2007; Lee, 2012). Interestingly, both Gao and colleagues (2016) and Zwissler et al. (2015) found that when they added a condition in which no cue was given (i.e., R-cue vs. F-cue vs. no cue), memory for F-cued items was better than no-cued items, suggesting that directed or intentional forgetting is actually not as efficient as non-directed forgetting (Gao et al., 2016; Zwissler et al., 2015). Gao and colleagues (2016) propose a similar explanation to that described above, whereby the F-cued items are processed for a second time (which could be considered a single rehearsal, as suggested by Bancroft et al. (2013)) when the cue is presented, which ultimately strengthens the memory trace over those items with no cue that are only processed once.

Recent work conducted by Tan et al. (2020) used a unique variant of the directed forgetting procedure to support the selective rehearsal hypothesis. Rather than presenting a single item, two items were presented and participants were either told to remember both or forget both (pure condition) or to remember one but forget the other (mixed condition). An underlying assumption of the selective rehearsal hypothesis is that remembering an item is more effortful, or requires more cognitive resources, than forgetting an item. Thus, the pure remember condition would require more resources (and therefore would lead to worse performance) than the mixed condition, which would require more resources than the pure forget condition. In contrast, the retrieval inhibition hypothesis posits that forgetting is an active process, and therefore requires more cognitive resources than remembering. Therefore, according to this hypothesis, the pure forgetting condition would require the most cognitive resources (Tan et al., 2020). The authors found that recognition performance was best in the mixed conditions for both remember and forget items. Further, recognition for forget items was similar to neutral items with no instruction attached (Tan et al., 2020).

The Retrieval Inhibition Hypothesis. As discussed previously, the retrieval inhibition hypothesis rejects the notion that forgetting is a passive process and instead suggests that an instruction to forget triggers an active, inhibitory mechanism that disposes of F-cued information from memory (Anderson, 2003). In an attempt to test their retrieval inhibition hypothesis, Geiselman et al. (1983) conducted a series of four experiments using a modified cued-list directed forgetting procedure that eliminated participants' ability to consciously suppress or withhold the recall of F-cued items. Specifically, participants were told that they were involved in two studies simultaneously: they were asked to either learn certain items or to judge items based on pleasantness. At the halfway point, half of the participants were told to forget all prior items as they were just for practice (F-cue condition), and the others were told that they were half-way done and to continue on (R-cue condition). In all four experiments, participants recalled more learned words than judged words. Geiselman and colleagues (1983) obtained a directed forgetting effect for both types of items, even though participants had no incentive to remember the judged words. Additionally, at test, participants were able to accurately discriminate which type of item (learn vs. judge) they were recalling, with less than 5% of total words misclassified (Geiselman et al., 1983, Experiment 1). Due to the F-cue operating similarly on the judge items that should not have been rehearsed at all, the authors suggest some mechanism whereby the F-cue inhibited access to those items with a forget instruction. Specifically, this inhibition seems to occur as retrieval blocking, as there were no differences between conditions or item types in a recognition memory test, which provided retrieval-based cues. Further corroborating this retrieval inhibition hypothesis, Geiselman et al. (1983) made a comparison of their results to those of studies examining posthypnotic amnesia, in which retrieval-based processes have been implicated. Specifically, participants in the F-cue condition demonstrated list-half source amnesia, in that they had difficulty categorizing the words presented prior to the F-cue into either the first or second half of the list. Also similar to posthypnotic amnesia, there was a significantly reduced correlation between the input-output order for F-cued items compared to the R-cued items (Geiselman et al., 1983).

Based on this hypothesis, Bjork (1989) along with Geiselman and Bagheri (1985) were able to demonstrate that participants could undergo a release from the initial retrieval inhibition. Bjork (1989), using the list-method, had participants complete an initial recognition memory test followed by a free-recall test. The re-exposure of the F-cued items through the prior recognition test eliminated the directed forgetting effect in the free recall test, a finding which Bjork (1989) reported as a release from inhibition. Further, Geiselman and Bagheri (1985) used the item-method and after an initial recall test, the unrecalled R- and F-cued items were presented again with R-cues. The repeated F-cued items benefitted much more than the repeated R-cued items. The authors attributed this advantage to a release from the inhibition caused by the initial F-cue (Geiselman & Bagheri, 1985).

Based on a previously established observation that older adults have reduced inhibition abilities, (Hasher et al., 1991), Zacks and colleagues (1996) compared the directed forgetting performance of younger and older adults. In further support of the retrieval inhibition hypothesis, Zacks and colleagues (1996) discovered that older adults have more difficulty inhibiting the F-cued information. Older adults produced more F-cued items on an immediate recall memory test, had longer reaction times (RTs) when rejecting F-cued items on a recognition test, and had relatively better memory (both recall and recognition) of F-cued information on a delayed retention task compared to younger adults (Zacks et al., 1996).

An Interplay of Theories. Memory research using ERPs has consistently found an old/new effect, such that correctly recognized old items evoke more positive ERPs than correctly recognized new items (Johnson, 1995). The location and timing of these ERP signals suggests distinct underlying processes. For instance, an early frontal old/new effect, occurring in the 300-500 ms time window is associated with familiarity in recognition tests; a parietal old/new effect, beginning 400 ms after the stimuli presentation, suggests conscious recollection and the strength of a memory trace; and right-frontal old/new effect, with similar onset to the parietal effect, is thought to be related to retrieval-based processes (Ullsperger et al., 2000). In directed forgetting research, the old/new effect is appreciably different for successfully recognized R- vs. F-cued items. Specifically, the F-cued items have a completely absent parietal old/new effect and much less early frontal activity suggesting inhibition of these items (Ullsperger et al., 2000). Additionally, a *reversed* old/new effect has been found, such that more negative ERPs are elicited by successfully forgotten F-cued items than correctly recognized new items (Nowicka et al., 2009). The reversed old/new effect has been considered evidence for intentional and successful inhibition of the forgotten F-cued items (Nowicka et al., 2009; Van Hooff et al., 2009).

As discussed in the previous section, there has been evidence using ERP data to support a selective rehearsal-type mechanism for R-cued information (Paller, 1990; Paz-Caballero et al., 2004). However, as Bjork (1970) discussed early on, and Paz-Caballero et al. (2004) more recently stated, it does not account for the whole story and there is likely an interplay of both theories. Many researchers have agreed with Paz-Caballero and colleagues (2004) in suggesting that this distinct ERP activity between R- and F-cues is actually due to an active inhibitory mechanism operating on F-cued items (e.g., Patrick et al., 2015; Van Hooff et al., 2009; Yang et al., 2012), similar to those used to stop a motor response (Fawcett & Taylor, 2010). More specifically, F-cues produce a larger N2 amplitude than R-cues in healthy participants, which is a component thought to be an electrophysiological correlate of cue-induced memory inhibition or goal-directed inhibition (Patrick et al., 2015; Yang et al., 2012). Although this inhibitory response cannot immediately halt processing, as previously mentioned, it seems to draw attentional resources away from the F-cued items in working memory (Fawcett & Taylor, 2010; Thompson et al., 2014; Thompson & Taylor, 2015).

More recent work by Rummel and colleagues (2016) expanded on the speculation of an interplay of theories through the use of a multinomial modelling approach. They found that the better memory for R-cued items was, in fact, due to better storage (i.e., rehearsal) of this information in both the item- and list-methods. The authors also linked the poorer memory for F-cued items in the list-method solely to retrieval-based processes. However, they were also able to demonstrate that this poorer memory for F-cued items in the item-method was due both to poorer storage and to inhibited retrieval (Rummel et al., 2016).

Similarly, Gao and colleagues (2016) examined the underlying neural mechanisms and time course of directed forgetting, where participants engaged in a modified item-method directed forgetting task that required either high or low cognitive loads, while ERP and behavioural data were measured. Similar to results reported by Rummel and colleagues (2016), Gao et al. (2016) reported findings consistent with two-stages being involved in directed forgetting: (1) task-relevance identification, whereby more attentional resources are allocated to

R-cued information than F-cued information; and (2) information discarding, whereby cognitive control resources are recruited to actively inhibit F-cued information (Gao et al., 2016).

Marevic and Rummel (2018) found that providing item-based semantic cues facilitated retrieval for both F- and R-cued information, and through the use of multinomial modelling, a storage-retrieval model confirmed that this effect was retrieval-mediated. In contrast, Taylor and colleagues (2018) report that providing recognition cues as to which type of item was being presented (R- or F-cued) at test did not aid in the recognition of either remember or forget items. When explicit feedback was given on the accuracy of participants' responses, a conservative bias for F-cued words was eliminated. The discriminability advantage for R-cued words over F-cued words, however, remained, providing contrasting evidence that retrieval-based strategies, such as the semantic cues presented by Marevic and Rummel (2018), did not aid the recognition of the F-cued words.

Jing and colleagues (2019) examined the effect of maintenance rehearsal on directed forgetting in two experiments. First, participants were presented with items and then shown a maintenance cue prior to seeing the memory cue (i.e., R or F). As a result, there were four conditions: maintenance followed by remember (M-R), maintenance followed by forget (M-F), only maintenance (M), or only forgetting (F). Jing et al. (2019) found a typical directed forgetting effect, in that the M-R items yielded better recognition performance than the M-F items. However, items in both M-R and M-F conditions had better recognition accuracy compared to items in the F condition. Moreover, there was no difference in recognition performance between M items and M-F items; the authors suggest that this finding provides evidence that any inhibition or processes that are triggered by the F-cues during the prolonged interval in the M-F condition did not necessarily promote forgetting (Jing et al., 2019).

Interestingly, recognition was better for M items than F items, demonstrating that an F-cue promotes reduced or halted maintenance rehearsal. Jing et el. (2019) found supporting evidence for the retrieval inhibition hypothesis in their second experiment using ERPs time-locked to cue (M-R, M-F, M, F) presentation. Specifically, there was an enhanced fronto-central P3 component demonstrated for F cues compared to M cues – which was taken to suggest attentional inhibition was evoked by an F cue. Together, Jing et al. (2019) posit that when triggered by an F cue, attentional inhibition works to terminate maintenance rehearsal.

Another important aspect of memory is contextual information. Associations between situational or environmental contexts are often incidentally encoded with purposefully studied information (Godden & Baddeley, 1975; Hockley, 2008). Burgess et al. (2017) examined whether study context had an effect on the magnitude of directed forgetting when context was encoded incidentally (Experiment 1) or intentionally (Experiment 2). Study items were presented against different natural landscape backgrounds, which provided "context" to the item, and were followed by either R- or F-cues. At test, items were either presented with their original context, a rearranged context, or a new context. Burgess et al. (2017) found an effect of context-dependent recognition, such that memory was better when items were presented in their original context. A directed forgetting effect was demonstrated in both experiments; however, in both incidental and intentional encoding, context did not interact with directed forgetting, as it affected both R-cued and F-cued information similarly. As such, the authors consider a "one-shot" hypothesis, whereby context is encoded within the first few seconds following the presentation of the stimulus. After this initial time period, further processing of contextual information is not enhanced with an R-cue, nor interrupted with an F-cue (Burgess et al., 2017). A null interaction between context and memory instruction was also found when sentences (Taylor & Hamm,

2018) and irrelevant aspects of schematic faces (Orghian et al., 2018) were used as contextual information. These results do not disentangle selective rehearsal or retrieval inhibition, but they suggest that regardless of the processes underlying directed forgetting, they are target-focused and do not depend on context.

Early on, MacLeod (1975) suggested that participants may actually store the instructional cue with the stimulus representation during encoding. Interestingly and of relevance to the current study, participants are generally quite accurate at identifying the cue associated with each item at test (Bancroft et al., 2013; Davis & Okada, 1971; Horton & Petruk, 1980; MacLeod, 1975; Thompson et al., 2011). It has been argued that the storage of the relationship between instructional cue and item is not direct, but rather that this decision is made based on the amount of rehearsal or encoding processes used during encoding that lead to either a stronger or weaker memory trace of the item (Bancroft et al., 2013; Johnson, 1988; MacLeod, 1975).

A study by Thompson and colleagues (2011) assessed the utility of a “tagging task” to examine source attribution errors in directed forgetting. At test, participants tagged the presented items as from either the R, F, or N (new) categories. Thompson et al. (2011) compared this task to the typical yes/no recognition task and demonstrated no differences in the magnitude of the directed forgetting effect, demonstrating the value of using the method. Importantly, participants were reasonably accurate in their source attributions, correctly identifying the source or tag (i.e., R/F) 75% of the time. An analysis of the errors indicated that F and N words were more likely to be confused compared to R words, leading the authors to suggest the strength of memory traces for these two sources are similar (Thompson et al., 2011).

Similarly, Bancroft and colleagues (2013) found that participants ability to accurately identify the associated cue increased with longer cue presentation durations. Further, their findings support the idea that this decision is based on the strength of the memory for the items, as participants were also more likely to confuse F and N items, similar to Thompson et al., (2011). As such, they claim that memory traces are stronger for R-cued items compared to F-cued items, which are in turn stronger than for new items (Bancroft et al., 2013).

Associative Memory

Associative memory is an aspect of episodic memory that involves representations of relationships between items (Murdock, 1974), such as the association between words and their meaning. Traditional associative recognition tasks are similar to the single-item tasks presented previously, but they require participants to form associations between random pairs of items and later discriminate between intact (i.e., old pairs) or rearranged pairs (i.e., new pairs consisting of old items from different study pairs; Hockley, 1992). For example, both DEATH PICK and PEPPER POLE may be presented at study, but DEATH POLE may be presented at test, which would require participants to indicate that this pair is new. As such, the memory is not based on single items, but rather the associations formed between the items as both items need to be present for the response to be considered correct (Hockley & Consoli, 1999).

Several differences between item and associative recognition have been identified. Associative recognition requires a greater amount of conscious recollection (remembering contextual information about the episode), whereas item recognition is based more on familiarity or the feeling of knowing without additional information (Clark, 1992; Hockley & Consoli, 1999; Yonelinas, 2002). If the associative information is unitized during encoding, then familiarity can play a larger role (Ahmad & Hockley, 2014; Bader, Mecklinger, & Hoppstädter, 2010). Unitization occurs when two items are encoded as a coherent whole, rather than as two individual constituents, so that this unitized whole is more familiar than its two constituent items or a new recombined association (Bastin et al., 2010; Graf & Schacter, 1989). Using an example provided previously, DEATH PICK could be unitized making the combination of those two items more familiar than both DEATH and PICK individually and the new combination of DEATH POLE. Further, item information is forgotten at a faster rate than associative information (Hockley, 1992); however, associative memory is more likely to decline in older age than memory for single items (Old & Naveh-Benjamin, 2008).

Studies using the directed forgetting paradigm with associative information have produced similar effects to those mentioned previously, with better memory for R-cued word pairs than F-cued word pairs (Bancroft et al., 2013; Hockley et al., 2016; Lansue, 2017). This effect is similarly robust to variations in presentation intervals as it has been found with cue presentations as brief as 300 ms and as long 3 s (Bancroft et al., 2013). Using associative information in the directed forgetting procedure provides a relatively pure measure of intentional and incidental encoding, because as mentioned previously, memory for the individual item information is not beneficial because both intact and rearranged pairs consist of old items. Thus, memory for the R-cued associations in this task is strategic and intentional (Ahmad & Hockley, 2014).

In an examination of intentional and incidental encoding, Hockley et al. (2016) presented word pairs that had a pre-experimental association (compound word (CW) pairs; e.g., HOME SICK) and unrelated (noncompound word (NCW) pairs; e.g., PLAY MAN) word pairs. For the R-cued pairs, participants were instructed to create an association between the words in the pair in order to better remember them. A directed forgetting effect was observed for both the CW and NCW pairs, such that hit rates were greater for R-cued CW and NCW pairs than for the F-cued pairs. However, this effect was reduced for the CW pairs. In both Hockley et al. (2016) and

Brancroft et al. (2013), accuracy rates for F-cued pairs were still above chance levels. Taken together, these results suggest that both intentional and incidental encoding occurs for associative information. More specifically, participants demonstrate more incidental encoding for word pairs that had pre-experimental associations (Hockley et al., 2016).

The effects found by Hockley and colleagues (2016) were replicated and extended a study by Lansue (2017), where the role of depth of processing was also examined. Research has consistently found that deeper processing of information during encoding creates stronger memory representations leading to easier retrieval (Craik & Lockhart, 1972; Craik & Tulving, 1975; Prior & Bentin, 2003). Although overall memory was better with deeper processing, intentional and incidental encoding of associative information occur both when participants are asked to use a surface level encoding strategy (i.e., letter counting) and a deeper level of processing (i.e., semantic elaboration). Additionally, a directed forgetting effect was found under both levels of processing, suggesting that this effect for associative information is robust to the quality of encoding of the information (Lansue, 2017).

The Present Study

The primary purpose of the current study was to examine a novel aspect of directed forgetting and associative memory, such that the instructional cue and the item are unitized and encoded as an association and at test, participants make decisions based on this association. As demonstrated, participants are generally able to remember whether the item was presented with an R-cue or an F-cue (Bancroft et al., 2013; Davis & Okada, 1971; Horton & Petruk, 1980; MacLeod, 1975; Thompson et al., 2011). Rather than MacLeod (1975) and Johnson's (1988) explanation that participants make this judgement on the basis of memory strength, the current study is examining the possibility that participants "tag" the items during encoding as either R-or F-cued and are aware of this association between the item and cue; thus, at test, they make their decisions based on the instruction associated with the item.

Additionally, a second purpose of the current study was to examine the timing of the instructional cue. To ensure that items are encoded, researchers typically present the memory instruction after the item. Although some early studies have presented a memory cue during the presentation of the items (Paller, 1990; Roediger & Crowder, 1972; Weiner & Reed, 1969), in real life learning settings we often know in advance whether we care to remember the information we encounter. There have been no studies of directed forgetting with the cue presented prior to the presentation of the items. Therefore, the current study examined how being informed of this cue before, during, or after encoding affects the directed forgetting effect commonly found in the typical procedure. Additionally, introducing a temporal manipulation (cues before, during, or after the items) will further disentangle initial encoding from rehearsal effects in the to-be-forgotten items.

Research Objectives and Hypotheses

The current experiment was designed foremost to assess the idea that the directed forgetting paradigm is essentially the same as asking participants to form associations between two study items, with the only difference being that one "item" is an instruction to follow an action (remember or forget an item) rather than another word. As such, two procedures were used in which participants are asked to remember an association between a colour and an item, with the only difference in procedures being that in one, the colours have no meaning and in the other, the colours signal an R- or an F-cue. The experiment was also designed to directly assess our "tagging" theory using the method used by Thompson et al. (2011), whereby participants are asked to indicate whether an item was R-cued, F-cued, or new (or whether an item was Blue-cued, Yellow-cued, or new). We hypothesized that participants would be able to accurately identify the source of the item at test at better than chance levels in both instructional conditions, and, more specifically, participants in the remember/forget condition would be more likely to confuse F-cued and new items, as in Thompson et al. (2011).

As previously mentioned, a secondary purpose of the current experiment was to examine the location of the instructional cue. As such, we used a temporal manipulation to assess how cue timing interacts with the above manipulations. We hypothesized that in the remember/forget association condition, an interaction would be present between the cue and the cue timing such that it would not matter when the R-cue is presented, but memory for F-cued items would be worst when the forget instruction was given before the study items and best when the cue was given after the items, with accuracy falling somewhere in between when the F-cue is given at the same time as the presentation of the study items. Moreover, we hypothesized that there would not be differences between cues or cue timing in the item-colour association condition.

Signal Detection Theory

Signal detection theory (SDT) provides a theoretical understanding of accuracy as behavioural data, in such a way that takes the individual's decision-making process into account (Macmillan & Creelman, 2004; Swets, 1964). In a basic Yes-No Design, as employed here in the context of recognition memory, participants are deciding between "signals" and "noise," or in this case, a word they previously studied and a new word. Thus, participants should try to endorse the signals and deny the noise, and this leads to four possible outcomes: a hit (i.e., they correctly said "yes" when a signal was present); a false alarm (i.e., they incorrectly said "yes" when noise was present); a correct rejection (i.e., they correctly said "no" when noise was present); and a miss (i.e., they incorrectly said "no" when a signal was present) (Macmillan, 2002). SDT posits that individuals will set a criterion value to make this distinction, and they will respond “yes” to events that exceed that criterion, and “no” to events that do not. However, a critical part of SDT is the stipulation that individuals have unique criteria, and therefore some participants will be more conservative in their decision making and some will be more liberal (Macmillan & Creelman, 2004; Swets, 1964). For example, if an individual is extremely liberal with their responses, always responding with “yes” regardless of the stimuli presented, they are going to achieve both a high hit rate and a high false alarm rate. Therefore, in a recognition memory task where accuracy is important, only analyzing the hits would provide misleading results. This is a clear example of extreme response bias and in practice it can be subtler but bias can still influence results in large ways (Macmillan & Creelman, 1990).

In recognition memory tasks, participants are asked to *discriminate* between old and new items and *d'* (index of discriminability) is a popular statistic used in this research because it provides an unbiased measure of discriminability, taking both hits and false alarms into account. It is calculated per participant using the following formula: $d' = Z_{hit} - Z_{FA}$ (where Z represents the Z transformation of the probability value based on a normal distribution; Macmillan & Creelman, 2004). Thus, higher *d'* values indicate a better ability to discriminate between old and new items (Macmillan, 2002).

CHAPTER 2

DESIGN AND METHODOLOGY

Stimulus Development

The stimulus set comprises 192 nouns carefully matched on important variables as described below. Concreteness values were obtained from Brysbaert and colleagues (2014), where concrete and abstract words are operationalized as ratings above and below 3, respectively; only concrete words (i.e., words that refer to physically tangible entities) were used in the present study. Word length was kept between 5 to 7 letters ($M = 5.87$). Orthographic frequency ($M = 48.19$), orthographic neighbourhood size ($M = 1.18$), and semantic neighbourhood density (SND; $M = 0.31$) were also controlled, as they have been recognized as being influential in both memory and language processing (Coltheart et al., 1979; Glanzer & Adams, 1990; Wong Gonzalez, 2018). Orthographic frequency refers to how often a word is used in language and orthographic neighbourhood size refers to the number of words that are exactly one letter different than the target word (Coltheart et al., 1979); these values were obtained from Wordmine2 (Durda & Buchanan, 2006). Semantic neighbourhood density is defined as the variability in the distribution of semantically related words surrounding a target word's semantic neighbourhood (Durda & Buchanan, 2008), and these values were obtained from WINDSORS (Durda & Buchanan, 2008). See appendix A for the full stimulus set.

The 192 nouns were randomly divided into 96 study items and 96 foils (i.e., distractor) items. Several t-tests were conducted to ensure the lexical variables stated above did not differ between the subsets of items. As such, the subsets did not differ in concreteness, $t(95) = -0.41$, $p > .05$, word length, $t(95) = 1.63$, $p > .05$, orthographic frequency, $t(95) = -0.35$, $p > .05$, orthographic neighbourhood size, $t(95) = -0.93$, $p > .05$, or SND, $t(95) = -0.32$, $p > .05$.

Participant Recruitment and Inclusion Criteria

Participants were undergraduate students at the University of Windsor who volunteered to participate in exchange for bonus credits toward their eligible psychology courses. All participants were required to meet the following inclusion criteria: reported English as their first language and normal or corrected-to-normal vision, and no diagnosed deficits in colour vision. Ninety-Nine participants were recruited (86% female, M_{age} = 20.86).

Task Software and Display Details

The experimental tasks were administered on a Dell PC with Windows XP operating system using Direct RT (Pearson v2012; Empirisoft Corporation). All words were written in capital letters, in size 32, Times New Roman font and presented in the center of the computer screen against a solid light gray background. All instructions were presented in black text. The font colour of the experimental items was dependent on the experimental condition.

Method

Stimuli. The stimulus set described above was used to create the study list of 96 items and the test list of all 192 items (i.e., the 96 study items and the 96 foil items). Of the 96 items in the study list, half were cued with an R or the colour blue and the other half were cued with F or the colour yellow (these item-cue pairs were counterbalanced across participants). The colour associated with each instruction (i.e., R-blue and F-yellow) were also counterbalanced across participants. The mazes used in the distraction phase were obtained from an online source called Maze Generator.

Procedure. The experiment was approved by the University of Windsor Research Ethics Board. Participants volunteered to participate through the University of Windsor Psychology Department Participant Pool and were randomly assigned to one of six conditions (two

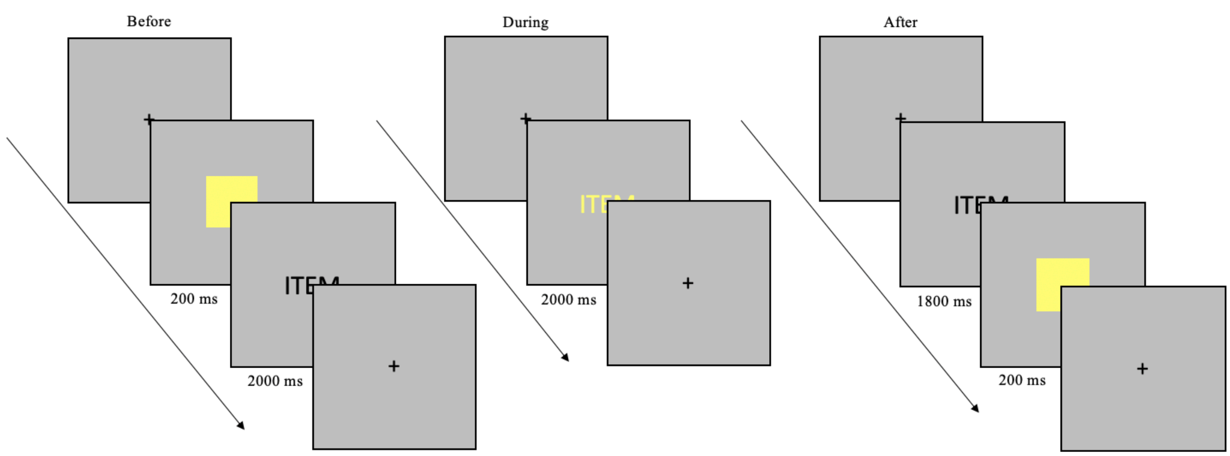
Instruction X three Timing). The entire task was completed on a computer screen in an individual testing room and lasted 30 minutes. Regardless of condition, there were three phases to the experiment: a study phase, a distraction phase, and a test phase. All participants were explicitly made aware that their memory would be tested. The varying presentations of the study phase are outlined below.

Instruction Conditions. In both instruction conditions, participants were presented with 96 items individually, 48 were presented with the colour blue and 48 were presented with the colour yellow. If assigned to the typical directed forgetting condition, participants were instructed to remember items presented with a specific colour and forget those presented with the other colour. If assigned to the colour-item association condition, participants were presented with the same 96 items, but they were instructed to remember which item was presented with which colour. The timing of the cue was dependent upon which cue timing condition the participant is assigned. The order of the presentation of items was random for each participant.

Cue Timing Conditions. Participants that were assigned to the “before” cue condition were presented with the cue prior to seeing the item. Participants that were assigned to the “after” cue condition were presented with the cue following the item presentation. Specifically, participants saw a square in the appropriate colour positioned in the center of the computer screen. Participants that were assigned to the “during” cue condition were presented with the items in the corresponding cue colour. To keep constant the amount of rehearsal time available across conditions (at 2000 ms), the items and cues were presented for differing lengths of time depending on the cue timing condition: in the *after* cue condition, the items were presented for 1800 ms and the cue was presented for 200 ms; in the *before* cue condition, the cue was presented for 200 ms and the items were presented for 2000 ms; and in the *during* cue placement condition, the item and the cue were presented together for 2000 ms. Refer to Figure 1 for a visual representation of these conditions. To ensure participants were encoding the items, they were asked to read the items aloud into a recorder.

Figure 1.

Visual Representation of Cue Timing Conditions



Note. The black fixation crosses demonstrate where a single trial begins and ends.

Following the study phase, all participants completed mazes for three minutes as the distraction phase. Next, in the test phase, the participants were presented with all 192 items and were asked to indicate whether the item was presented with the colour blue, the colour yellow, or if it was a new item. Participants used the keys Z (covered with a sticker indicating BLUE or YELLOW), / (covered with a sticker indicating BLUE or YELLOW), and B (covered with a NEW sticker) – the Z and / keys were counterbalanced across participants to control for effects of handedness. Each test trial began with an item appearing in the center of the screen in black text and the item remained on the screen until the participant responded. All 48 R/blue-cued items and 48 F/yellow-cued items were presented, along with the 96 new items. The presentation of the test items was random for each participant.

Multiple analysis of variances (ANOVA) were conducted to examine participants discriminability (*d'*), tagging accuracy, and RT. Data was cleaned based on participants' *d'* scores and RT. *d'* scores too close to 0 or 1 were removed, as these values render the index invalid. For the RT analysis, only correct responses were statistically analyzed and responses that were faster or slower than 300ms or 4000ms were removed to ensure participant effort and accuracy.

CHAPTER 3

DATA ANALYSIS AND RESULTS

Prior to any analyses being conducted, six participants' data were removed due to: (1) answering only R-, F-, or Blue-, Yellow-cued with zero Foil responses (four participants); (2) not completing the full task (one participant); (3) a coding error when the participant was run (one participant); and (4) a *d'* hit rate too close to 1.00, which renders the score invalid (one participant). The final sample included in the analysis was $N = 92$ participants (86% female, M_{age} = 20.76) with 49 participants in the colour-item association condition ($n = 13$ in after cue placement condition, $n = 14$ in before condition, and $n = 22$ in during condition) and 43 participants in the remember-forget association condition ($n = 15$ in after cue placement condition, $n = 13$ in before condition, and $n = 16$ in during condition).

Part 1: Remember/Forget Association (RFA) Analyses

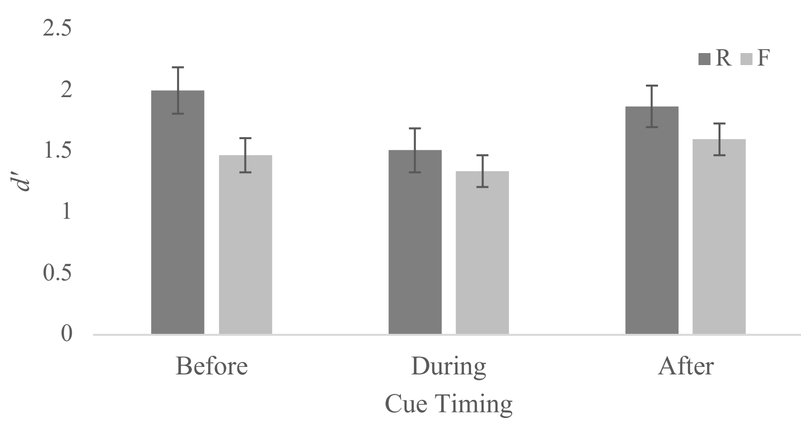
See appendix B for a discussion of outlier removal and testing of assumptions for all analyses contained in Part 1.

Old/New Analysis. First, a 2 x 3 ANOVA was conducted to examine the effect of cue, as well as cue timing on discrimination of old vs. new items. Thus, the cue (remember vs. forget) was manipulated as a within-subjects variable and the cue timing (before vs. during vs. after) was manipulated as a between-subjects variable. The dependent variable is the *d'* statistic. To calculate *d'*, all R and F responses, whether correct or incorrect, were initially coded as old and the NEW responses remained coded as new. For each participant per condition, a hit rate was calculated with the mean proportion of accurate responses to old items, and a false alarm rate was calculated with the mean proportion of inaccurate responses to new items. *d'* was then calculated per participant for each condition using the formula: ($d' = Z_{hit} – Z_{FA}$; Leeuw, 2015;

Macmillan & Creelman, 2004). Results reveal an effect of cue type on discrimination, $F(1, 39) = 30.06$, $p < .001$, partial $\eta^2 = .44$, such that participants discriminated R-cued items ($M = 1.79$, $SE = 0.10$) better than F-cued items ($M = 1.47$, $SE = 0.08$). There was also an interaction of cue and cue timing, $F(2, 39) = 3.13$, $p = 0.055$, $\eta^2 = .14$, whereby remember cues placed before **(**$M = 2.00$, $SE = 0.19$) yielded better discrimination accuracy than remember cues placed after ($M = 1.87$, $SE = 0.17$). However, forget cues placed before ($M = 1.47$, $SE = 0.14$) yielded lower discrimination accuracy than those placed after ($M = 1.60$, $SE = 0.13$). See Figure 2.

Figure 2.

Mean d' scores for R- and F-cues per cue timing condition.



Tagging Accuracy Analysis. A second set of analyses were conducted to examine participants' accuracy in their ability to tag items appropriately with their cues (i.e., R/F/Foil). This analysis examines what tag participants actually responded with and also demonstrates the accuracy for each condition beyond the old/new endorsement. These proportions examine R/R+F+Foil, F/R+F+Foil and Foil/R+F+Foil for R-cued items, F-cued items, and Foil items; thus, there are nine different proportions for each participant, three of which are used in each analysis. The first analysis is proportion of R responses calculated by R/R+F+Foil for R-cued items vs. F-cued items vs. Foil words, the second is proportion of F responses calculated by F/R+F+Foil for R-cued items vs. F-cued items vs. Foil words, and the third is proportion of Foil responses calculated by Foil/R+F+Foil for R-cued items vs. F-cued items vs. Foil words. Three one-way repeated measures ANOVAs were run per item type to examine differences in the proportions of items endorsed as being R-cued, F-cued or as New. The first analysis was conducted to determine if there were significant differences in the proportions of tags (whether the participant responded to the item as being R-cued, F-cued, or Foil) across R-cued items. The proportion of responses to R-cued items varied by tag, $F(1.64, 69.03) = 22.97$, $p < .001$, partial $\eta^2 = .35$. Post hoc analysis with a Bonferroni adjustment for multiple comparisons revealed that R-cued items were correctly tagged as being R-cued ($M = .48$, $SE = .03$) more than F-cued ($M = .28$, $SE = .02$, $p < .001$) or Foil items ($M = .23$, $SE = .02$, $p < .001$). A difference in response proportions was not found between F-cue endorsement and Foil endorsement, $p > .05$ (See Table 1).

The second ANOVA was run to determine if there were differences in the proportions of tags across F-cued items. The proportion of responses to F-cued items varied by tag, $F(2, 84) = 22.76$, $p < .001$, partial $\eta^2 = .35$. Post hoc analysis with a Bonferroni adjustment revealed that F-cued items were correctly tagged as being F-cued ($M = .47$, $SE = .02$) more than R-cued ($M = .20$, $SE = .02$, $p < .001$) or as Foil items ($M = .34$, $SE = .02$), $p < .05$. Further, F-cued items were tagged as being Foil items ($M = .34$, $SE = .02$) more than as being R-cued items ($M = .20$, $SE = .02$), $p < .05$.

A third one-way repeated measures ANOVA was conducted to examine if there were significant differences in the proportions of tags across Foil (i.e., new) items. The proportion of responses to Foils varied by tag, $F(1.38, 55.35) = 430.24$, $p < .001$, partial $\eta^2 = .91$. Post hoc analysis with a Bonferroni adjustment revealed that Foil items were correctly tagged as being Foils ($M = .81$, $SE = .02$) more as R-cued ($M = .05$, $SE = .01$, $p < .001$) or F-cued items ($M = .14$, $SE = .02$, $p < .001$). Further, Foils items were incorrectly tagged as F-cued items more than as R-cued items, $p < .001$.

Table 1.

Mean accuracy rates of tags per correct item type.

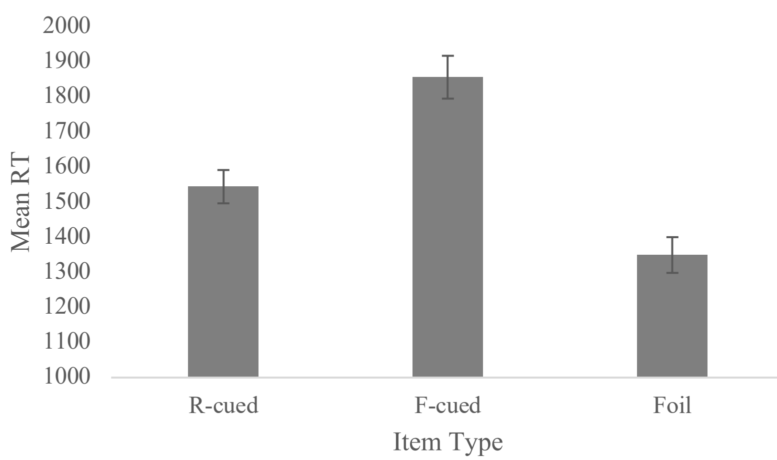
	Tag		
Item Type	R	F	Foil
R-cued	48%	28%	23%
F-cued	20%	47%	34%
Foil	5%	14%	81%

To examine if there were differences in the proportions of correctly tagged items across each cue type (i.e. R-cue, F-cue or Foil), an additional one-way repeated measures ANOVA was conducted. After correcting for multiple comparisons using a manual Bonferroni correction, the proportion of correctly tagged items varied significantly by cue type, $F(2, 84) = 69.33$, $p < .001$, partial $\eta^2 = .62$. Post hoc analysis with a Bonferroni adjustment and an additional manual Bonferroni correction for multiple analyses revealed that Foils ($M = .81$, $SE = .02$) were correctly tagged more than R-cued ($M = .48$, $SE = .03$, $p < .001$) or F-cued items ($M = .47$, $SE = .02$, $p < .001$). A difference was not found between proportions of correctly tagged R-cued and F-cued items, $p > .999$.

Reaction Time Analysis. A 3 x 3 mixed-factorial ANOVA was conducted to examine differences in mean RT as a function of item and cue timing. Similar to previous analyses, the item presented (R-cued vs. F-cued vs. Foil) were manipulated as a within-subjects variable and the cue timing (before, during, after) was manipulated as a between-subjects variable. The dependent variable was RT. First, all incorrect responses were removed. Next, responses that were faster or slower than a pre-selected specified cut-off of 300ms or 4000ms, respectively, were removed. Mean RTs were calculated across each participant for each condition. Results reveal a main effect of item type, $F(2, 80) = 60.83$, $p < .001$, partial $\eta^2 = .60$. Post hoc analysis with a Bonferroni adjustment demonstrated that Foil items ($M = 1349.30$ms, $SE = 50.90$) were correctly responded to faster than R-cued items ($M = 1544.10$ms, $SE = 47.49$, $p = .001$) or F-cued items ($M = 1856.12$ms, $SE = 61.09$, $p < .001$). Further, R-cued items were correctly responded to faster than F-cued items, $p < .001$. There was no effect of cue timing, $F(2, 40) = .88$, $p = .42$, partial $\eta^2 = .04$, nor an interaction between item type and cue timing, $F(4, 80) = .63$, $p = .64$, partial $\eta^2 = .03$. See Figure 3.

Figure 3.

Mean RTs for R-cued, F-cued, and Foil items in the RFA condition.



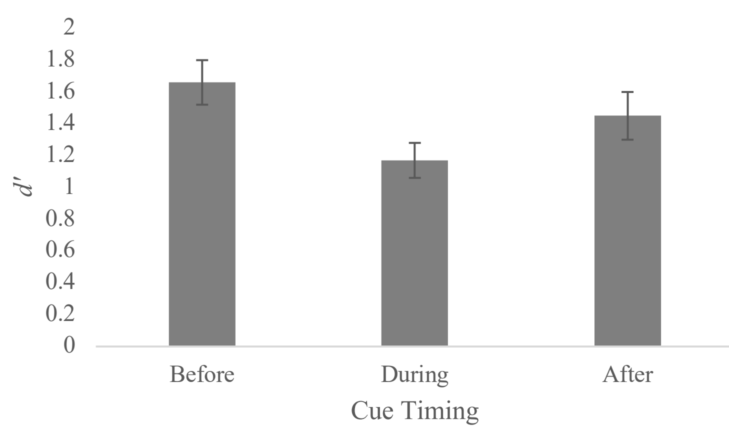
Part 2: Colour-Item Association (CIA) Analysis

See appendix C for a discussion of outlier removal and testing of assumptions for all analyses contained in Part 2.

Old/New Analysis. First, a 2 x 3 mixed-factorial ANOVA was conducted to examine the effect of cue colour and cue timing on discrimination of old vs. new items. Thus, the cue presented (blue vs. yellow) was manipulated as a within-subjects variable and the cue timing (before vs. during vs. after) was manipulated as a between-subjects variable. The dependent variable is the *d'* statistic, calculated the same way as described above. Results demonstrated an effect of cue timing on *d'*, $F(2, 46) = 3.85$, $p < .05$, partial $\eta^2 = .14$. Post hoc analysis with a Bonferroni correction revealed that cues placed before ($M = 1.66$, $SE = 0.14$) yielded better discrimination than cues placed during ($M = 1.17$, $SE = 0.11$, $p > .05$). No differences were found between cues placed after ($M = 1.45$, $SE = 0.15$) and other cue timings. See Figure 4. There was no difference in discrimination of cue colour (blue vs. yellow), $F(1, 46) = 2.89$, $p > .05$, partial $\eta^2 = .06$. There was not an interaction found between cue colour and item type, $F(2, 46) = 0.33$, $p > .05$, partial $\eta^2 = .01$.

Figure 4.

Mean d' scores per cue timing condition.



Tagging Accuracy Analysis. Similar to the RFA analysis, another set of analyses were conducted to examine participants' accuracy in their ability to tag items appropriately with their cues (i.e., Blue/Yellow/Foil). This analysis was conducted in the same way as described in the previous section.

A one-way repeated measures ANOVA was conducted to determine if there were differences in the proportions of items tagged as being Blue-cued across item types.
The proportion of responses to Blue-cued items varied significantly by tag, $F(1.74, 78.07) = 22.66$, $p < .001$, partial $\eta^2 = .34$. Post hoc analysis with a Bonferroni adjustment for multiple comparisons revealed that Blue-cued items were correctly tagged as being Blue-cued ($M = .44$, $SE = .02$) more than as Yellow-cued ($M = .31$, $SE = .01$, $p < .001$) or Foil items ($M = .25$, $SE = .02$, $p < .001$). There was no difference in response proportions was between incorrect Yellow-cue tags and incorrect Foil tags, $p > .05$ (See Table 2).

Another one-way repeated measures ANOVA was run to determine if there were differences in the proportions of tags across Yellow-cued items. Post hoc analysis with a Bonferroni adjustment revealed that Yellow-cued items were correctly tagged as being Yellow-cued ($M = .39$, $SE = .02$) more than as being Foils ($M = .28$, $SE = .02$), $p < .05$. There was not a difference between Yellow-cued items tagged correctly and those tagged incorrectly as Blue-cued ($M = .33$, $SE = .02$, $p > .05$), nor were there differences in response proportions found between items incorrectly tagged as Blue and items incorrectly tagged as Foils ($M = .28$, $SE = .02$, $p > .05$).

A third one-way repeated measures ANOVA was conducted to examine if there were differences in the proportions of tags across Foil items. The proportion of responses to Foils varied by tag, $F(1.36, 66.85) = 183.66$, $p < .001$, partial $\eta^2 = .79$. Post hoc analysis with a Bonferroni adjustment revealed that Foil items were correctly tagged as being Foils ($M = .73$, $SE = .03$) more than as being Blue-cued ($M = .14$, $SE = .02$, $p < .001$) or Yellow-cued items ($M =$

.14, *SE* = .01, $p < .001$). There was no difference in response proportions for Foils items that were incorrectly tagged as Blue-cued items or that were incorrectly tagged as Yellow-cued items, $p > .999$.

Table 2.

Mean accuracy rates of tags per correct item type.

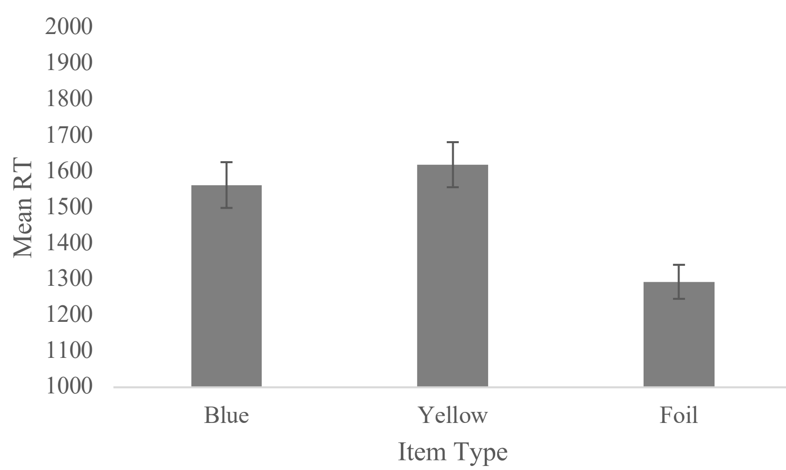
	Tag		
Item Type	Blue	Yellow	Foil
Blue-cued	44%	31%	25%
Yellow-cued	33%	39%	28%
Foil	14%	14%	73%

A final one-way repeated measures ANOVA was conducted to examine if there were differences in the proportions of correctly tagged items across each cue type (i.e. R-cue, F-cue or Foil). Correcting for multiple comparisons using a manual Bonferroni correction, the proportion of correctly tagged items varied significantly by cue type, $F(1.53, 71.97) = 82.84$, $p < .001$, partial $\eta^2 = .64$. Post hoc analysis with a Bonferroni adjustment and an additional manual Bonferroni correction for multiple analyses revealed that Foils (*M* = .73, *SE* = .02) were correctly tagged more than Blue-cued (*M* = .43, *SE* = .02, $p < .001$) or Yellow-cued items (*M* = .39, *SE* = .02, $p < .001$).

Reaction Time Analysis. A 3 x 3 mixed-factorial ANOVA was conducted to examine any differences in RT as a function of item or cue timing. Similar to previous analyses, the item presented (blue-cued vs. yellow-cued vs. foil) was manipulated as a within-subjects variable and the cue timing (before, during, after) was manipulated as a between-subjects variable. The dependent variable was RT. First, all incorrect responses were removed. Next, responses that were faster or slower than a pre-selected specified cut-off of 300ms or 4000ms, respectively, were removed. Mean RTs were calculated across each participant for each condition. Results demonstrated an effect of item type, $F(1.68, 77.22) = 34.68$, $p < .001$, partial $\eta^2 = .43$. Analysis with a Bonferroni adjustment revealed that Foil items ($M = 1293.21$, $SE = 47.36$) were correctly responded to faster than Blue-cued items ($M = 1563.05$, $SE = 63.70$, $p < .001$) and Yellow-cued items ($M = 1619.75$, $SE = 62.75$, $p < .001$). No difference in mean RT was found between Blue-cued items and Yellow-cued items, $p > .05$. There was not an effect of cue timing on RT, $F(2, 46) = 0.36$, $p > .05$, partial $\eta^2 = .02$. There was not an interaction between item type and cue timing, $F(3.36, 77.22) = 1.24$, $p > .05$, partial $\eta^2 = .05$.

Figure 5.

Mean RTs for Blue-cued, Yellow-cued, and Foil items in the CIA condition.



CHAPTER 4

DISCUSSION

General Discussion

The primary purpose of the current study was to use unique variants of the item-method paradigm to examine a novel explanation for the underlying processes of directed forgetting, such that the item and cue are unitized and encoded together at study. Specifically, two procedures were used: (1) a more typical directed forgetting procedure where participants were told to remember specific items and forget others; and (2) a procedure where participants were told to remember which cue was associated with each item. Further, Thompson and colleagues' (2011) tagging method was used at test which required participants to indicate which cue items were presented with at study (i.e. R, F, or a new item). We hypothesized that a directed forgetting effect would be found in the remember-forget association condition and that there would be no difference in cues in the colour-item association condition. Further, we expected that participants would be able to accurately tag items with their appropriate cue at better than chance levels regardless of instructional condition, as it is theorized that participants are encoding the item-cue pair as one. A secondary purpose of this study was to assess the effect of cue timing, in relation to study items, on directed forgetting. To accomplish this objective, a temporal manipulation was conducted whereby participants were either presented with the item first, cue first, or both together. We hypothesized that cue timing would have the largest effect on F-cued items – specifically, recognition performance would be worst for F-cued items that were in the before condition, followed by during condition, with the best recognition in the after condition.

Foremost, a directed forgetting effect was observed for participants in the RFA condition, such that participants were better able to discriminate items they were told to remember compared to those they were told to forget. Although this finding was expected, it should be noted that our presentation times (both cue and item) were much quicker than those typically used. The briefest cue presentation used had been 300ms and the more typical cue presentation length ranges from 1s to 3s (Bancroft et al., 2013). Thus, although not a primary purpose of this experiment, finding a directed forgetting effect with a very brief cue presentation of 200ms (in the before and after cue timing conditions) is particularly salient and should be noted. Further, it was found that when the R-cue was placed before the study item, participants' discrimination ability was greater than when the R-cue was placed after the item. Importantly, and as hypothesized, F-cues had an opposite effect – participants were able to discriminate better when F-cues were placed after the study items compared to before. That is, directed forgetting is most successful when the participants know in advance that they can forget the following item. Interestingly, in the RFA condition, there were no differences found between the during cue timing condition and before/after conditions.

As hypothesized, there were no differences found between cues in the CIA condition, demonstrating that it is specifically the instruction to remember or forget that creates the difference in discrimination between cues. Notably, in the CIA condition, all cues placed before the study items yielded better discrimination than those placed during the item, but there were no differences between cues placed before and after. There appears to be a burden of processing in the during condition, such that participants must encode both the item and the cue simultaneously and this might lead to a lack of quality encoding. Importantly, this seems to be more cumbersome in the CIA condition than in the RFA condition, as participants in the RFA condition are able to forget half of the items presented, whereas participants in the CIA condition must attempt to remember all items and the cues associated with each item. Therefore, the CIA condition required more cognitive resources overall. This seems to be in line with work by Tan et al. (2019) described previously, that demonstrated that when participants needed to remember two items presented, it required more resources than when only needed to remember one and could forget the other. Recent research conducted by Popov et al. (2019) further supports this idea, as they found that participants had worse memory for items, regardless of their cue, when the previous item was R-cued. Specifically, the more R-cued items in a row, the worse the participants memory was, suggesting a cumulative effect of cognitive load.

Using Thompson and colleagues' (2011) tagging method allowed us to examine participants' accuracy in their ability to tag items as either R-cued, F-cued, or New (or Blue-cued, Yellow-cued, or New). This analysis allowed an examination of finer details regarding the types of errors participants make in tagging items. In the RFA condition, participants correctly tagged R-cued items with 48% accuracy, F-cued items with 47% accuracy, and Foil items with 81% accuracy. In the CIA condition, participants correctly tagged Blue-cued items with 44% accuracy, Yellow-cued items with 39% accuracy, and Foil items with 73% accuracy. These results contrast those found by Thompson et al. (2011) where participants were able to correctly tag items, both R-cued and F-cued, with 75% accuracy. Although our hypothesis that participants would be able to tag items correctly at better than chance levels was not supported, it is important to note that there was no difference found between the proportion of correctly tagged R-cued and F-cued items and between the proportion of correctly tagged Blue-cued and Yellow-cued items. Differences between tagging performance in this and Thompson et al (2011) may reflect the brevity of the cue presentation in this study. Although this finding does not fully support the proposed tagging theory of directed forgetting, it is not inconsistent with it. Participants were equally able to correctly tag R-cued items and F-cued items, but as mentioned previously, a directed forgetting effect was found when all “old” data was combined as explained in the results section. Thus, although participants are better able to discriminate R-cued items, they are not better able to tag them. In both the CIA and RFA conditions, participants performed quite well at tagging Foil items correctly. This finding supports the idea that participants might make the judgement based on the strength of the memory trace (Bancroft et al., 2013; Johnson, 1988; MacLeod, 1975; Thompson et al., 2011); in other words, it appears to be easier to identify an item that has a weak memory trace, or that the participant has never seen.

In terms of errors, in the RFA condition, participants correctly tagged each item more than they tagged those items as coming from the other two sources (e.g., participants correctly tagged R-cued items as R-cued more than they incorrectly tagged them as F-cued or Foils). However, consistent with results found by Thompson et al. (2011) and Bancroft et al. (2013), participants were more likely to confuse F-cued items and Foil items. This further supports the theory that the judgement is being made based on the strength of the memory trace. In theory, R-cued items should have the strongest memory trace, followed by F-cued items, and Foil items having weakest trace. The RT analysis also provides corroborating evidence for this idea, as participants made their judgements fastest for Foil items and slowest for F-cued items. Thus, it appears that participants could be reasonably confident in their ability to say a Foil item was New and an R-cued item was R-cued, but F-cued items required more deliberation. In the CIA condition, participants were also able to accurately tag Foil items quite well, but were more likely to confuse Yellow-cued items with Blue-cued items. Participants in the CIA condition also made their decision fastest for Foil items, but there was no difference in RT for Blue- and

Yellow-cued items. These CIA results support the strength of the memory trace argument: Blue- and Yellow-cued items should have similarly strong memory traces which could contribute to confusion in identifying their source. However, it should be noted that the confusion could be also due to the overall higher level of resources required for the task, making it more difficult to accurately identify the source of the old items.

These results do not necessarily support or disprove either the selective rehearsal or retrieval inhibition theories. The findings related to cue timing seem to lend some evidence toward selective rehearsal, rather than retrieval inhibition, or at the least an interplay of both – as participants had better discrimination when R-cues were placed before and when F-cues were placed after. It seems plausible that this better performance could be due to participants preparing to rehearse (R-cue) or preparing to divert attention elsewhere (F-cue), or in the after condition, engaging in maintenance rehearsal until an F-cue was presented and then diverting attention (or as some may argue, actively engaging in inhibition); however, more research is needed to confirm this, especially that using ERP data. As previously mentioned, these results are also not inconsistent with a tagging theory – particularly when taking results from previous research into account where participants were extremely accurate in being able to correctly tag items (Bancroft et al., 2013; Thompson et al., 2011). The fact that participants were not able to correctly tag items at better than chance levels in this study could be due to how quick the cues and items were presented.

Limitations and Future Directions

The current study has provided several novel findings with regard to directed forgetting; although it is not without limitations to be noted. Most notable, the cue timing conditions involved a mix of visual stimuli: the before and after conditions used a coloured square to identify which type of cue was presented, whereas the during condition used the item IN the specific colour – therefore in two conditions a shape was used as cues and in one condition the letters of the item were used as cues. Although the directed forgetting effect has been found using a variety of stimuli (Ahmad et al., 2019; Taylor et al., 2018), the lack of any salient findings surrounding the during condition could be due to the mixing of stimuli. It could be possible to use a coloured square indicating the cue located slightly above the study item. Thus, a follow-up study addressing this issue should be conducted. Relatedly, future research should use line drawings or shapes that are less able to be rehearsed, rather than words, as stimuli with the coloured squares as cues. Ahmad et al. (2019) demonstrated that the directed forgetting effect can be applied to pictures of complex scenes, and specifically affected the fine details in those scenes more than the overall gist. Importantly, it is still possible to rehearse both the gist and details of pictures of scenery. For example, a participant might recite to themselves what they see in the picture (e.g., a woman riding a bike on a path through the trees). Using meaningless symbols or line drawings could limit the possibility to rehearsal to further disentangle selective rehearsal and other theories.

As mentioned, although we cannot completely confirm a tagging theory of directed forgetting, our results do lend some evidence toward its validity. As such, future research should focus on substantiating or disproving this as an explanation for directed forgetting findings. One such study could provide participants with recognition cues at test (e.g., ITEM-R, ITEM-F, ITEM-N) and require participants to decide whether the pair is correct or incorrect based on what cue was presented with each item during study. Further, it is also important to conduct a follow-up study that continues to use Thompson and colleagues' (2011) tagging method, but with a more typical cue presentation speed (i.e., 1 to 3s) to determine if performance more closely aligns with previous findings (Bancroft et al., 2013; Thompson et al., 2011). Removing the cue timing conditions would allow for more flexibility in cue presentation speed, as it would not be as necessary to ensure static rehearsal time. Similarly, due to the design used, we could not directly compare the two procedures used (i.e., CIA and RFA conditions) in a way that would produce meaningful and cogent results, as participants in the colour-item association condition were required to remember all 96 items, whereas participants in the remember-forget association condition only needed to remember 48 (R-cued) items. Therefore, it may be worthy of conducting a follow-up study that creates a more similar comparison, for example using a colour-item association condition that only requires participants to remember 48 associations, instead of all 96.

As described in Chapter 1, both associative memory and directed forgetting are susceptible to aging and cognitive decline (Old & Naveh-Benjamin, 2008; Zacks et al., 1996). These two age related deficits have yet to be linked in such a way that suggests they are essentially coming from the same mechanism. Older adults perform similar to patients with Alzheimer's disease on a directed forgetting task that requires source memory, or memory for the source from which a particular item was encoded, further linking the two processes in this manner (Haj & Allain, 2015). If the tagging theory considered in the current study is deemed appropriate, the fact that older adults are less able to demonstrate directed forgetting and source monitoring may stem from the decline in associative memory. Future research should begin to examine this possibility.

Importantly, the finding that participants are better able to remember stimuli when they are told in advance that they need to remember it can be important in education or occupational settings. For instance, if there is a particularly important concept or skill that students or employees need to remember, telling that person ahead of time might directly promote stronger encoding and therefore better retrieval. Future research could incorporate more ecologically relevant stimuli into directed forgetting research, such as instructions or directions. Moreover, using more implicit methods of telling participants to remember or forget (e.g., accidentally providing incorrect information) information may shed light onto to how to incorporate these findings into real-world settings.

It is also important to highlight the limitations of the statistical analyses conducted. The numerous violations of the assumptions of the repeated measures ANOVA, and the associated increase in type I error, is important to consider when interpreting these results. However, ANOVA is quite robust to violations of normality, especially with roughly equal between-group sample sizes. Further, it is important to note that we interpreted the interaction between cue (R- and F-cues) and cue timing as significant despite the *p* value being .055, as we found it worthy to report due to the meaningful effect size. Although these are limitations of the analyses, after applying several corrections for violations of sphericity, as well as for multiple comparisons, the effect sizes remain quite large for the significant effects and therefore, these results do not appear to be spurious and should be interpreted as valid.

Summary and Conclusions

In sum, the present study was the first to explore a link between directed forgetting and associative memory such that the item and cue are unitized and encoded as a pair. Although the overall level of accuracy in tagging was quite low for all R-cued, F-cued, Yellow-cued, and Blue-cued items, there were no differences between correctly tagged R- and F-cued items and correctly tagged Yellow- and Blue-cued items. Thus, despite a directed forgetting effect, it appears that participants are equally able to tag R-cued and F-cued items (or Yellow- and Blue-cued items). The data trends suggest that this way of conceptualizing directed forgetting as a type of tagging is worthy of follow-up.

The current study is also the first to directly compare the effect of cue timing on directed forgetting. Our results demonstrate that directed forgetting is most efficient when cues are placed before the study items, as participants tend to remember R-cued information and forget F-cued information most with this cue timing thus implicating a role for rehearsal as opposed to retrieval processes in the directed forgetting effect.

REFERENCES

Ahmad, F. N., & Hockley, W. E. (2014). The role of familiarity in associative recognition of unitized compound word pairs. *The Quarterly Journal of Experimental Psychology, 67*(12), 2301-2324. doi: 10.1080/17470218.2014.923007

Ahmad, F. N., Tan, P., & Hockley, W. E. (2019). Directed forgetting for categorised pictures: recognition memory for perceptual details versus gist. *Memory*, *27*(7), 894–903. doi: 10.1080/09658211.2019.1591456

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49,* 415–445. https://doi-org.ledproxy2.uwindsor.ca/10.1016/j.jml.2003.08.006

Anderson, M. C. (2005). The role of inhibitory control in forgetting unwanted memories: A consideration of three methods. In N. Ohta, C. M. MacLeod & B. Uttl (Eds.), *Dynamic cognitive processes* (3rd ed., pp. 130-189). New York, NY: John Wiley & Sons, Inc.

Bader, R., Mecklinger, A., Hoppstädter, M., & Meyer, P. (2010). Recognition memory for one-trial unitized word pairs: Evidence from event-related potentials. *NeuroImage, 50*(2), 772-781. doi: 10.1016/j.neuroimage.2009.12.100

Bancroft, T. D., Hockley, W. E., & Farquhar, R. (2013). The longer we have to forget the more we remember: The ironic effect of post-cue duration in item-based directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 691-699. doi: 10.1037/a0029523

Basden, B. H., & Basden, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 139-172). Mahwah, NJ: Lawrence Erlbaum Associates.

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 603-616. doi: 10.1037/0278-7393.19.3.603

Bastin, C., van der Linden, M., Schnakers, C., Montaldi, D. & Mayes, A. R. (2010). The contribution of familiarity to within- and between-domain associative recognition memory: Use of a modified remember/know procedure. *European Journal of Cognitive Psychology, 22,* 922-943. doi: 10.1080/09541440903007834

Burgess, N., Hockley, W. E., & Hourihan, K. L. (2017). The effects of context in item-based directed forgetting: Evidence for "one-shot" context storage. *Memory and Cognition*, *45*(5), 745–754. doi: 10.1016/j.actpsy.2018.05.009

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior, 9*, 255-268. doi: 10.1016/S0022-5371(70)80059-7

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork, R. A., LaBerge, D., & Legrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science, 10*(2), 55-56. doi: 10.3758/BF03331404

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911. doi:10.3758/s13428-013-0403-5

Burgess, N., Hockley, W. E., & Hourihan, K. L. (2017). The effects of context in item-based directed forgetting: Evidence for "one-shot" context storage. *Memory and Cognition, 45*(5), 745-754. doi: 10.3758/s13421-017-0692-5

Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition, 20,* 231-243. doi: 10.3758/BF03199660

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology*, *31*(3), 489–507. doi: 10.1080/14640747908400741

Craik, F. I. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour, 11*(6), 671-684. doi: 10.1016/S0022-5371(72)80001-X

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294. doi: 10.1037/0096- 3445.104.3.268

Davis, J. C., & Okada, R. (1971). Recognition and recall of positively forgotten items. *Journal of Experimental Psychology, 89*(1), 181-186. doi: 10.1037/h0031183

Durda, K., & Buchanan, L. (2006). WordMine2. Available from: http://web2.uwindsor.ca/wordmine

Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods, 40*(3), 705-712. doi: 10.3758/BRM.40.3.705

Ensor, T. M., Bancroft, T. D., & Hockley, W. E. (2019). Listening to the picture-superiority effect: Evidence for the conceptual-distinctiveness account of picture superiority in recognition. *Experimental Psychology*, *66*(2), 134–153. doi: 10.1027/1618-3169/a000437

Epstein, W. (1969). Poststimulus output specification and differential retrieval from short-term memory. *Journal of Experimental Psychology, 82*(1), 168-174. doi: 10.1037/h0028045

Fawcett, J., & Taylor, T. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not with stop-signal inhibition. *Memory and Cognition, 38*(6), 797-808. doi: 10.3758/MC.38.6.797

Gao, H., Cao, B., Qi, M., Wang, J., Zhang, Q., & Li, F. (2016). Two stages of directed forgetting: Electrophysiological evidence from a short-term memory task. *Psychophysiology, 53*(6), 806-813. doi: 10.1111/psyp.12628

Gao, H., Cao, B., Zhang, Q., Qi, M., Li, F., & Li, H. (2016). Intending to forget is not easy: Behavioral and electrophysiological evidence. *International Journal of Psychophysiology*, *104*, 1–9. doi: 10.1016/j.ijpsycho.2016.03.007

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: evidence for retrieval inhibition. *Memory & Cognition, 13*(1), 57-62. doi: 10.3758/BF03198444

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112*(1), 58-72. doi: 10.1037/0096-3445.112.1.58

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 5–16. doi: 10.1037/0278-7393.16.1.5

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*, 325–331. doi: 10.1111/j.2044-8295.1975.tb01468.x

Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *Journal of Experimental Psychology: Learning Memory, and Cognition, 15,* 930-940. doi: 10.1037/0278-7393.15.5.930

Green, R. L. (1992). *Human memory: Paradigms and paradoxes.* Hillsdale, NJ: Erlbaum.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(1), 163-169. doi: 10.1037/0278-7393.17.1.163

El Haj, M., Fasotti, L., & Allain, P. (2015). Directed forgetting of source memory in normal aging and Alzheimer's disease. *Aging Clinical and Experimental Research*, *27*(3), 329–336. doi: 10.1007/s40520-014-0276-1

Hockley, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1321-1330. doi: 10.1037/0278-7393.18.6.1321

Hockley, W. E. (1994). Reflections of the mirror effect for item and associative recognition. *Memory & Cognition, 22*(6), 713-722. doi: 10.3758/BF03209256

Hockley, W. E. (2008). The effects of environmental context on recognition memory and claims of remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1412–1429. doi: 10.1037/a0013016

Hockley, W. E., Ahmad, F. N., & Nicholson, R. (2016). Intentional and incidental encoding of item and associative information in the directed forgetting procedure. *Memory & Cognition, 44*(2), 220-228. doi: 10.3758/s13421-015-0557-8

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*(4), 657-664. doi: 10.3758/BF03211559

Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24*(2), 202-216. doi: 10.3758/BF03200881

Horton, K. D., & Petruk, R. (1980). Set differentiation and depth of processing in the directed forgetting paradigm. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 599-610. doi: 10.1037/ 0278-7393.6.5.599

Jing, J., Qi, M., Gao, H., & Zhang, Q. (2019). The role of forgetting cues in directed forgetting: Ceasing maintenance rehearsal. *Acta Psychologica*, *199*(Complete). https://doi.org/10.1016/j.actpsy.2019.102922

Johnson, D. A. (1971). Pupillary responses during a short-term memory task: Cognitive processing, arousal, or both? *Journal of Experimental Psychology, 90*(2), 311-318. doi: 10.1037/h0031562

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 27-36. doi: 10.1037/0096-3445.117.4.390

Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin, 116*, 274-292. doi: 10.1037/0033-2909.116.2.274

Johnson, R. (1995). Event-related potential insights into the neurobiology of memory systems. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (pp. 135-163). Amsterdam, The Netherlands: Elsevier.

Jongeward, R. H., Woodward, A. E., & Bjork, R. A. (1975). The relative roles of input and output mechanisms in directed forgetting. *Memory & Cognition, 3*(1), 51-57. doi: 10.3758/BF03198210

Lansue, B. (2017). *Forget about it: Intentional and incidental encoding of associative information*. (Unpublished undergraduate thesis). University of Windsor, Windsor, ON.

Lee, Y. S. (2012). The cognitive load hypothesis of item-method directed forgetting. *The Quaterly Journal of Experimental Psychology, 65*(6), 1110-1122. doi: 10.1080/17470218.2011.644303

Lee, Y. S., Lee, H. M., & Tsai, S. H. (2007). Effects of post-cue interval on intentional forgetting. *British Journal of Psychology, 98*(2), 257-272. doi: 10.1348/000712606X120410

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory, 1*(3), 271-279. doi: 10.1037/0278-7393.1.3.271

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(1), 13-21. doi: 10.1037/0278-7393.15.1.13

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Lawrence Erlbaum Associates.

Macmillan, N. A. (2002). Signal detection theory. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology* (3rd ed., pp. 43-90). New York, NY: John Wiley & Sons, Inc.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107*(3), 401-413. doi: 10.1037/0033-2909.107.3.401

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Psychology Press.

Marevic, I., & Rummel, J. (2018). Retrieval-mediated directed forgetting in the item-method paradigm: the effect of semantic cues. *Psychological Research*, *0*(0), 0. doi: 10.1007/s00426-018-1085-5

Murdock, B. B., JR. (1974). Human memory: Theory and data. Hillsdale, NJ: Erlbaum.

Muther, W. S. (1965). Erasure or partitioning in short-term memory. *Psychonomic Science, 3*, 429-430. doi: 10.3758/BF03343215

Nowicka, A., Jednoróg, K., Wypych, M., & Marchewka, A. (2009). Reversed old/new effect for intentionally forgotten words: An ERP study of directed forgetting. *International Journal of Psychophysiology, 71*(2), 97-102. doi: 10.1016/j.ijpsycho.2008.06.009

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104-118. doi: 10.1037/0882-7974.23.1.104

Orghian, D., Garcia-Marques, L., Marques, P., & Braga, J. (2018). Memory and conceptual learning of relevant and non-relevant items in item-method directed forgetting. *Memory*, *26*(9), 1233–1243. doi: 10.1080/09658211.2018.1441424

Paller, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology, 16*(6), 1021-1032. doi: 10.1037/0278-7393.16.6.1021

Patrick, R., Kiang, M., & Christensen, B. (2015). Neurophysiological correlates of emotional directed-forgetting in persons with Schizophrenia: An event-related brain potential study. *International Journal of Psychophysiology, 98*(3), 612-623. doi: 10.1016/j.ijpsycho.2015.01.006

Paz-Caballero, M., Menor, J., & Jiménez, J. (2004). Predictive validity of event-related potentials (ERPs) inrelation to the directed forgetting effects. *Clinical Neurophysiology, 115*(2), 369-377. doi: 10.1016/j.clinph.2003.09.011

Popov, V., Marevic, I., Rummel, J., & Reder, L. M. (2019). Forgetting Is a Feature, Not a Bug: Intentionally Forgetting Some Things Helps Us Remember Others by Freeing Up Working Memory Resources. *Psychological Science*, *30*(9), 1303–1317. doi: 10.1177/0956797619859531

Prior, A., & Bentin, S. (2003). Incidental formation of episodic associations. The importance of sentential context. *Memory & Cognition, 31*(2), 306-316. doi: 10.3758/BF03194389

Reitman, W., Malin, J. T., Bjork, R., Higman, B. (1973). Strategy control and directed forgetting. *Journal of Verbal Learning and Verbal Behavior, 12*(2)*,* 140-149. doi: 10.1016/S0022-5371(73)80003-9

Rummel, J., Marevic, I., & Kuhlmann, B. G. (2016). Investigating storage and retrieval processes of directed forgetting: A model-based approach. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 42(10), 1526-1543. doi: 10.1037/xlm0000266

Sheard, E. D., & MacLeod, C. M. (2005). List method directed forgetting: Return of the selective rehearsal account. In N. Ohta, C. M. MacLeod & B. Uttl (Eds.), *Dynamic cognitive processes* (3rd ed., pp. 219-248). New York, NY: John Wiley & Sons, Inc.

Swets, J. A. (1964). *Signal detection and recognition in human observers: Contemporary readings.* New York, NY: John Wiley & Sons, Inc.

Tan, P., Ensor, T. M., Hockley, W. E., Harrison, G. W., & Wilson, D. E. (2020). In support of selective rehearsal: Double-item presentation in item-method directed forgetting. *Psychonomic Bulletin & Review, 27,* 529-535. doi: 10.3758/s13423-020-01723-w

Taylor, T. L., Cutmore, L., & Pries, L. (2018). Item-method directed forgetting: Effects at retrieval? *Acta Psychologica, 183,* 116-123. doi: 10.1016/j.actpsy.2017.12.004

Taylor, T. L., & Hamm, J. P. (2018). A grand memory for forgetting: Directed forgetting across contextual changes. *Acta Psychologica*, *188*(March), 39–54. doi: 10.1016/j.actpsy.2018.05.009

Thompson, K. M., Fawcett, J. M., & Taylor, T. L. (2011). Tag, you're it: Tagging as an alternative to yes/no recognition in item method directed forgetting. *Acta Psychologica, 138*(1), 171-175. doi: 10.1016/j.actpsy.2011.06.001

Thompson, K. M., Hamm, J. P., & Taylor, T. L. (2014). Effects of memory instruction on attention and information processing: Further investigation of inhibition of return in item-method directed forgetting. *Attention, Perception, and Psychophysics*, *76*(2), 322–334. doi: 10.3758/s13414-013-0584-0

Thompson, K. M., & Taylor, T. L. (2015). Memory instruction interacts with both visual and motoric inhibition of return. *Attention, Perception, and Psychophysics*, *77*(3), 804–818. doi: 10.3758/s13414-014-0820-2

Turvey, M. T., & Wittlinger, R. P. (1969). Attenuation of proactive interference in short-term memory as a function of cue to forget. *Journal of Experimental Psychology, 80*, 295-298. doi: 10.1080/00335557043000078

Ullsperger, M., Mecklinger, A., & Müller, U. (2000). An electrophysiological test of directed forgetting: The role of retrieval inhibition. *Journal of Cognitive Neuroscience, 12*(6), 924-940.

Van Hooff, J., Whitaker, T., & Ford, R. (2009). Directed forgetting in direct and indirect tests of memory: Seeking evidence of retrieval inhibition using electrophysiological measures. *Brain and Cognition, 71*(2), 153-164). doi: 10.1016/j.bandc.2009.05.001

Weiner, B., (1968). Motivated forgetting and the study of repression. *Journal of Personality, 36*(2)*,* 213-234. doi: 10.1111/j.1467-6494.1968.tb01470.x

Wetzel, C. D., & Hunt, R. E. (1977). Cue delay and the role of rehearsal in directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory, 3*(2), 233-245. doi: 10.1037/0278-7393.3.2.233

Wong Gonzalez, Daniela, "The relationship between semantic and episodic memory: Exploring the effect of semantic neighbourhood density on episodic memory" (2018). Electronic Theses and Dissertations. 7585. https://scholar.uwindsor.ca/etd/7585

Woodward, A. E. & Bjork, R. A. (1971). Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology, 89*(1), 109-116. doi: 10.1037/h0031188

Woodward, A. E., Park, D. C., & Seebohm, K. (1974). Directed forgetting as a function of explicit within-list cueing and implicit postlist cuing. *Journal of Experimental Psychology, 102*(6), 1001-1006. doi: 10.1037/h0036387

Yang, W., Liu, P., Xiao, X., Li, X., Zeng, C., Qiu, J., & Zhang, Q. (2012). Different neural substrates underlying directed forgetting for negative and neutral images: An event-related potential study. *Brain Research*, *1441*, 53–63. doi: 10.1016/j.brainres.2011.10.042

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441-517. doi: 10.1006/jmla.2002.2864

Zacks, R. T., Hasher, L., Radvansky, G. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 143-156. doi: 10.1037/0278-7393.22.1.143

Zwissler, B., Schindler, S., Fischer, H., Plewnia, C., & Kissler, J. M. (2015). “Forget me (not)?” - Remembering forget-items versus un-cued items in directed forgetting. *Frontiers in Psychology*, *6*, 1–14. doi: 10.3389/fpsyg.2015.01741

APPENDICES

Appendix A. Study and foil items with all controlled variables and values.

Item	Word Length	Frequency	Orthographic Neighbourhood Size	Concreteness	SND
TEACHER	7	55.31	1	4.52	0.28
BRANCH	6	44.08	2	4.90	0.22
HARNESS	7	8.96	0	4.14	0.34
DOCTOR	6	142.25	0	4.69	0.27
NURSERY	7	14.08	0	4.48	0.27
CHANNEL	7	29.65	2	4.18	0.35
ATTIC	5	7.29	0	4.53	0.31
STORM	5	52.01	3	4.70	0.40
MARKET	6	154.35	2	4.70	0.26
VESSEL	6	29.66	0	4.66	0.36
ALTAR	5	19.63	2	4.85	0.35
GUARD	5	54.20	0	4.04	0.31
CAMERA	6	14.95	0	5.00	0.37
GLASS	5	111.04	3	4.82	0.27
CASTLE	6	56.43	1	4.96	0.28
LAWYER	6	29.26	1	4.70	0.28
THIEF	5	16.42	1	4.37	0.31
APRON	5	9.11	0	4.87	0.34
BOTTLE	6	43.36	3	4.91	0.36
MIRROR	6	38.42	0	4.97	0.31
WAGON	5	21.29	0	4.89	0.35
MAYOR	5	18.50	2	4.37	0.26
VIOLIN	6	8.17	0	4.96	0.40
ARROW	5	17.21	0	4.97	0.27
BATTERY	7	14.44	2	4.67	0.29
BALLOT	6	6.19	1	4.73	0.37
WITNESS	7	37.83	3	4.07	0.30
MUSCLE	6	14.13	0	4.50	0.31
RAINBOW	7	8.88	0	4.57	0.26
EARTH	5	200.12	0	4.80	0.31
SPIDER	6	7.06	0	4.97	0.32
OFFICE	6	198.87	0	4.93	0.24
STREET	6	204.85	0	4.75	0.26
TWINS	5	13.97	4	4.57	0.35

INSECT	6	8.65	3	4.89	0.38
ACADEMY	7	12.52	0	4.29	0.25
FRAME	5	40.92	1	4.30	0.25
VETERAN	7	6.77	0	4.58	0.25
BLANKET	7	14.28	1	5.00	0.32
TEMPLE	6	42.76	0	4.53	0.32
STROKE	6	24.57	3	4.10	0.30
BISCUIT	7	5.39	0	4.90	0.45
ARTICLE	7	73.46	0	4.33	0.28
CORAL	5	9.40	1	4.40	0.34
BAMBOO	6	5.28	0	4.86	0.29
CARPET	6	22.58	1	4.96	0.31
AIRPORT	7	11.69	0	4.87	0.34
SURGEON	7	11.35	1	4.54	0.33
ARTIST	6	38.90	0	4.24	0.28
TURKEY	6	14.72	0	4.89	0.37
PARADE	6	9.94	0	4.54	0.32
PICTURE	7	114.95	0	4.52	0.30
CABIN	5	39.30	0	4.92	0.30
HUSBAND	7	144.85	0	4.11	0.25
WIZARD	6	6.85	1	4.43	0.40
FEATHER	7	11.75	3	4.90	0.30
NEEDLE	6	13.84	0	4.93	0.29
CHEEK	5	34.97	4	4.83	0.32
PRISON	6	58.49	2	4.68	0.35
ANCHOR	6	12.05	0	4.77	0.36
TRAFFIC	7	35.28	0	4.67	0.30
BRIDGE	6	63.13	2	4.97	0.27
MANOR	5	11.38	3	4.70	0.37
PEPPER	6	12.74	2	4.59	0.36
WIDOW	5	25.94	0	4.33	0.32
BALLOON	7	13.65	0	4.92	0.27
SERVER	6	14.23	2	4.55	0.30
BEDROOM	7	37.84	0	4.90	0.32
VEHICLE	7	24.89	1	4.64	0.31
TUTOR	5	8.31	1	4.28	0.27
RAILWAY	7	50.04	0	4.63	0.31
BLOOD	5	163.23	4	4.86	0.25
BEETLE	6	5.26	0	4.83	0.33
VILLAGE	7	114.66	2	4.89	0.26

RANCH	5	13.22	1	4.73	0.39
PENCIL	6	14.83	0	4.88	0.32
MAPLE	5	5.02	0	4.46	0.36
WOMAN	5	423.59	2	4.46	0.24
FENCE	5	26.10	2	4.82	0.27
INFANT	6	16.70	0	4.93	0.33
PRINCE	6	108.33	1	4.44	0.27
PARENT	6	28.12	1	4.56	0.24
WRIST	5	13.65	4	4.93	0.37
POLISH	6	15.23	2	4.23	0.36
COACH	5	27.77	4	4.12	0.26
RECORD	6	105.77	1	4.15	0.24
MEDAL	5	7.03	3	4.89	0.30
MANAGER	7	100.37	2	4.14	0.26
SUNSET	6	19.80	1	4.54	0.30
MOUTH	5	145.64	4	4.74	0.28
WINDOW	6	168.64	1	4.86	0.27
UNIFORM	7	26.08	0	4.67	0.35
ROBBER	6	5.34	4	4.31	0.34
DRAGON	6	14.48	0	4.39	0.32
VIDEO	5	46.56	2	4.67	0.27
CHAPEL	6	24.15	0	4.60	0.33
GLOVE	5	6.49	3	4.97	0.37
RADIO	5	37.67	2	4.74	0.33
MOTOR	5	36.74	1	4.84	0.30
BAGGAGE	7	9.31	0	4.43	0.32
OCEAN	5	36.47	0	4.86	0.32
STOMACH	7	27.91	0	4.89	0.32
ALBUM	5	9.38	0	4.69	0.22
MANSION	7	11.12	0	4.89	0.28
RIBBON	6	10.09	1	4.89	0.30
BARREL	6	12.43	3	4.86	0.37
BREAST	6	56.19	0	4.89	0.31
MARBLE	6	23.49	2	4.85	0.36
ALLEY	5	8.18	2	4.82	0.27
WRITER	6	39.10	4	4.32	0.28
SUGAR	5	33.31	0	4.87	0.31
VAULT	5	7.07	2	4.62	0.29
QUEEN	5	80.29	3	4.45	0.27
DOLLAR	6	33.00	1	4.93	0.36

WORKER	6	19.85	4	4.59	0.35
THUMB	5	14.34	1	4.96	0.31
WORLD	5	638.10	1	4.36	0.20
GLOBE	5	15.96	2	4.59	0.30
TUNNEL	6	16.76	2	4.82	0.31
SCHOOL	6	250.56	0	4.79	0.26
ANIMAL	6	79.07	0	4.61	0.29
HELMET	6	7.53	1	4.92	0.29
ORPHAN	6	5.05	0	4.04	0.29
CANDLE	6	22.16	2	4.86	0.31
VELVET	6	15.88	0	4.44	0.34
STRING	6	34.25	4	4.76	0.33
VISITOR	7	27.30	0	4.25	0.29
UNCLE	5	86.25	0	4.24	0.25
PICNIC	6	7.29	0	4.83	0.36
SILVER	6	78.81	2	4.52	0.31
SHIELD	6	16.16	0	4.66	0.30
TARGET	6	36.60	0	4.11	0.25
ELBOW	5	17.34	0	5.00	0.39
CHAIR	5	129.61	2	4.58	0.27
WAIST	5	23.62	2	4.72	0.41
PILLOW	6	14.31	2	5.00	0.29
STRAW	5	22.56	4	4.77	0.28
ALARM	5	29.49	0	4.47	0.32
ADULT	5	26.06	0	4.40	0.25
PIANO	5	23.90	0	4.90	0.38
WARRIOR	7	11.39	0	4.17	0.27
VOICE	5	382.45	1	4.13	0.29
WARDEN	6	6.30	4	4.27	0.29
NOVEL	5	36.08	3	4.21	0.32
WRECK	5	11.88	2	4.07	0.33
ACTOR	5	14.77	0	4.57	0.28
PACKAGE	7	29.25	0	4.72	0.25
VANILLA	7	9.60	0	4.68	0.43
PERFUME	7	9.93	0	4.66	0.33
SHOWER	6	17.82	2	4.89	0.33
BLADDER	7	5.56	2	4.48	0.38
STOOL	5	10.52	3	4.90	0.31
PUPPY	5	5.01	4	4.78	0.48
BACON	5	14.75	2	4.90	0.36

ENGINE	6	37.96	0	4.86	0.32
SATIN	5	8.82	2	4.57	0.39
FURNACE	7	7.67	0	4.69	0.34
SADDLE	6	26.07	2	4.85	0.33
HAMMER	6	14.08	3	4.77	0.30
THUNDER	7	24.23	0	4.34	0.26
TRUNK	5	20.10	2	4.71	0.26
FABRIC	6	14.35	0	4.63	0.38
JEWEL	5	9.84	1	4.96	0.27
STATION	7	88.19	0	4.32	0.28
GARMENT	7	10.64	0	4.78	0.43
PLAYER	6	29.43	3	4.15	0.24
APPLE	5	34.45	2	5.00	0.31
FIELD	5	144.95	3	4.26	0.23
ANKLE	5	8.05	1	4.81	0.38
GARLIC	6	5.33	2	4.89	0.42
SCRIPT	6	26.60	1	4.72	0.29
MONEY	5	402.71	2	4.54	0.30
THEATER	7	5.38	2	4.92	0.31
JOURNAL	7	26.05	0	4.63	0.32
TABLE	5	263.29	4	4.90	0.25
CHAIN	5	37.91	1	4.55	0.33
SHERIFF	7	13.54	0	4.50	0.29
PALACE	6	59.41	1	4.57	0.26
WILLOW	6	7.08	3	4.35	0.32
BEARD	5	26.34	3	4.96	0.36
STREAM	6	61.38	2	4.50	0.31
BOARD	5	118.86	3	4.57	0.23
CORPSE	6	13.25	0	4.89	0.32
ORGAN	5	18.92	0	4.77	0.31
SHADOW	6	65.77	0	4.54	0.25
RABBIT	6	18.82	1	4.93	0.38
MUSIC	5	136.11	0	4.31	0.22
ALCOHOL	7	16.48	0	4.76	0.34
COMPANY	7	245.58	0	4.11	0.23
STUDENT	7	48.32	0	4.92	0.28
WHISKY	6	13.23	1	5.00	0.41
MEADOW	6	12.69	0	4.86	0.32

Note. Foil items are highlighted in grey.

Appendix B. Review of all outlier removal and testing of assumptions for analyses contained in Part 1: RFA Analysis.

Old/New Analysis

An outlier analysis revealed one participant's responses to be an outlier, based on their *d'* statistic. According to a Shapiro-Wilk test, the assumption of normality was met after the removal of this outlier, with $p > .05$. Levene's test was also nonsignificant for all groups, $p > .05$, indicating that the assumption of homogeneity of variance was also met.

Tagging Accuracy Analysis

R-cued ANOVA. According to a Shapiro-Wilk test, the response proportions for Foils were not normally distributed, $p < .05$. However, due to the sample size and robustness of ANOVA to violations of normality, the analysis was conducted. One case was identified as an outlier in the proportions of responses endorsing the item as R-cued; therefore, the analysis was run with and without it included. Removal of this outlier produced similar results, $F(1.77, 72.68) = 21.57$, $p < .001$, partial $\eta^2 = .35$, with a greater violation of normality for Foil responses, $p = .011$; therefore, the case was left in for completeness and to maximize use of participant data. Mauchly's test of sphericity indicated that the assumption of sphericity had also been violated, $\chi^2(2) = 10.02$, $p < .05$. However, the results remained significant when using each of the three estimates of epsilon (i.e., Greenhouse-Geisser, Huynh-Feldt, Lower Bound) after the corrections of degrees of freedom for the test of within-subjects effects. Therefore, Greenhouse-Geisser, a more conservative epsilon, correction was reported ($\varepsilon = .82$).

F-cued ANOVA. The proportions for R-cue endorsements are not normally distributed, according to a Shapiro-Wilk test, $p < .05$. Two cases were identified as marginal outliers for F-cue endorsements. Again, analyses were run with and without them included and removal of these outliers created more outliers and produced similar results, $F(2, 80) = 20.98$, $p < .001$, partial $\eta^2 = .34$, with a greater violation of normality, $p < .05$, and were thus included in the analysis for reasons stated above. Mauchly's test of sphericity indicated that the assumption of sphericity had been met, $\chi^2(2) = 0.86$, $p > .05$.

Foils ANOVA. According to a Shapiro-Wilk test, normality was violated across all item types, $p < .05$. Two extreme outliers and four marginal outliers were identified for the R endorsement proportions and one nonextreme outlier for the Foil endorsement proportions. Removal of the extreme outliers does not allow for the assumption of normality to be met. Therefore, all outliers were included in the analysis for reasons stated above (i.e., robustness of ANOVA to normality violations and completeness). Mauchly’s test indicated that sphericity was violated, $\chi^2(2) = 29.89$, $p < .001$. Again, the Greenhouse-Geisser correction was used ($\varepsilon = .659$).

Correctly Tagged ANOVA. According to a Shapiro-Wilk test, normality was violated in the Foil group, $p < .05$. One slight outlier existed in the Remember group and two marginal outliers exist in the Forget group. These outliers were not removed for reasons previously stated. Mauchly’s test indicated that the assumption of sphericity was met, $\chi^2(2) = 2.27$, $p = .32$.

Reaction Time Analysis

According to a Shapiro-Wilk test, normality was met for all conditions, $p > .05$. One nonextreme outlier was identified for R-cued items and one marginal outlier for Foil items; however, removal of outliers created more outliers, thus these cases were included in the analysis. Both Levene’s and Box’s M tests were nonsignificant, $p > .05$, indicating that the assumptions of homogeneity of variance and homogeneity of covariance were met. Mauchly's test of sphericity indicated that the assumption of sphericity was also met, $p > .05$.

Appendix C. Review of all outlier removal and testing of assumptions for analyses contained in Part 2: CIA Analysis.

Old/New Analysis

According to a Shapiro-Wilk test, the assumption of normality was met, $p > 0.5$. Levene's test was also nonsignificant, $p > 0.5$, indicating the assumption of homogeneity of variance was also met. The assumption of homogeneity of covariances was met, as assessed by Box's M test, $p > .05$.

Tagging Accuracy Analysis

Blue-cued ANOVA. An outlier analysis revealed two slight outliers for Blue response proportions, one slight outlier for yellow response proportions, and two slight outliers for Foil response proportions. While removal of these outliers creates more outliers, it fixed a normality violation (Shapiro-Wilk test, $p = .31$, prior to removal) and created a less severe violation of sphericity Mauchley's test of sphericity, $\chi^2(2) = 9.95$, $p < .05$, prior to removal). Thus, outliers were removed from the analysis. After removal, the assumption of normality was met across all conditions according to a Shapiro-Wilk test, $p >.05$. Mauchly's test of sphericity indicated that the assumption of sphericity was still violated, $\chi^2(2) = 7.30$, $p < .05$. Greenhouse-Geisser correction was reported ($\varepsilon = .87$).

Yellow-cued ANOVA. The proportions for yellow-cue endorsements are all normally distributed, according to a Shapiro-Wilk test, $p > .05$. Two cases were identified as slight outliers for the Blue-cue endorsement group. There were no assumptions violated, so analysis proceeded with the outliers included. Mauchly's test of sphericity indicated that the assumption of sphericity had been met, $\chi^2(2) = 3.25$, $p > .05$. The proportion of responses to F-cued items varied significantly by endorsement type, $F(2, 98) = 6.09$, $p < .05$, partial $\eta^2 = .11$.

Foil ANOVA. According to a Shapiro-Wilk test, normality was violated across all item types, $p < .05$. Four slight outliers and one extreme outlier were identified for Blue responses and one slight outlier for Foil responses. Removal of these cases creates two more outliers. The analysis was run with and without the extreme outlier included. Removal of this outlier produced similar results, $F(1.24, 59.42) = 249.45$, $p < .001$, partial $\eta^2 = .84$ and did not fix assumption violations; therefore, the case was left in for completeness and to maximize use of participant data. Mauchly's test indicated that sphericity was violated, $\chi^2(2) = 30.12$, $p < .001$. Again, the Greenhouse-Geisser correction was used ($\varepsilon = .68$).

Correctly Tagged ANOVA. According to a Shapiro-Wilk test, normality was violated in the Foil group, $p < .05$. Two slight outliers were identified for Blue-cued endorsements and one slight outlier for Foil endorsement. Removal of both outliers restores normality across all groups according to a Shapiro-Wilk test, $p > 0.5$, and leaves no outliers, thus both outliers were excluded from analysis. Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(2) = 19.78$, $p < .001$. The Greenhouse-Geisser correction was used ($\varepsilon = .75$).

Reaction Time Analysis

According to a Shapiro-Wilk test, normality was violated for blue-cued items with cues presented in the before condition, $p < .05$. The assumption of normality was met for all other conditions, $p > .05$. For blue-cued items, one nonextreme outlier was identified in the before condition; for yellow-cued items, one nonextreme outlier was identified in the after condition; and for foil items, two nonextreme outliers were identified in the before and during conditions. Due to one case being an outlier in multiple groups, this case was removed from the analysis, correcting the violation of normality according to a Shapiro-Wilk test, $p > .05$. Both Levene's and Box's M tests were nonsignificant, $p > 0.5$, indicating that the assumptions of homogeneity of variance and homogeneity of covariance were met. Mauchly's test of sphericity indicated that the assumption of sphericity was violated, $\chi^2(2) = 9.56$, $p < .001$. The Greenhouse-Geisser correction was used ($\varepsilon = .84$).

VITA AUCTORIS

NAME:	Brette (Lansue) Burns
PLACE OF BIRTH:	Windsor, ON
YEAR OF BIRTH:	1994
EDUCATION:	General Amherst High School, Amherstburg, ON, 2012
	University of Windsor, B.A. Hons., 2017
	University of Windsor, M.A., 2020