GARLITCH, SYDNEY M., M.A. The Role of Attentional Fluctuation in Recollecting Episodic Changes. (2019)
Directed by Dr. Christopher N. Wahlheim. 64 pp.

Detecting and recollecting change can counteract proactive interference, and, in some cases, lead to proactive facilitation. The Memory-for-Change account assumes that attention is involved in detecting change, but that has not been directly tested. The current study is the first to investigate the role of attention in change detection and its consequences for change recollection. Participants studied a list of word pairs comprised of four seamless blocks. In each block there were three sets of word pairs: one set repeated across all four blocks (A-B, A-B), one set repeated in the first three blocks and then had the same cue with a changed response in the fourth block (A-B, A-D), and one set was unique to each block (C-D). Attention during encoding was measured using a probe-caught procedure. Thought probes asking participants to indicate whether they were “on-task” or “off-task” appeared throughout the study phase. Participants then completed a cued recall test for responses from the fourth block. Participants were also asked to indicate if each pair changed during the study phase, and to report the earlier response if there was a change. Results showed that recollecting change was associated with higher memory accuracy at test compared to when change was not recollected. In both between and within subject analyses, “on-task” reports were associated with higher memory accuracy and change recollection compared to “off-task” reports. These findings implicate a critical role for attention in change detection and recollection, and recall performance under conditions that could lead to proactive interference.

THE ROLE OF ATTENTIONAL FLUCTUATION IN RECOLLECTING EPISODIC CHANGES

by

Sydney M. Garlitch

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
2019

Approved by

____
Committee Chair

APPROVAL PAGE

This thesis written by SYDNEY M. GARLITCH has been approved by the following committee of the Faculty of The Graduate School at the University of North Carolina at Greensboro.

Committee Chair ______________________________

Committee Members ______________________________

____

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In our daily lives, we sometimes experience moments of inattention, where our focus drifts from the current task we are completing to our own internal thoughts. In many routine activities, there are typically minor, if any, consequences associated with this lapse in attention because you can complete the activity automatically. However, the consequences to these attentional lapses are greater for novel activities that require new learning. This may be especially important under conditions that can lead to interference in memory. Imagine that you have been told about the positive side effects of a drug on several occasions. Later on, information is presented to you about how the drug poses significant negative side effects, but you were thinking about something else when you were given this information. Since you were not fully attending to the negative side effects of the drug, you may fail to accurately represent its side effects. Then, when someone asks your opinion on the drug, you experience interference for the most recent information on the side effects of the drug, but decide to suggest the drug to them anyway. This inability to process the change and update your memory representation for the side effects of the drug could have negative consequences for the person to whom you recommended it. As outlined below, one way to avoid these negative consequences is to detect and recollect the change that occurred, but this might require that you paid attention to that changed information.

The goal of the current experiment was to investigate the consequences of natural fluctuations in attention on the ability to recall recent information when that information changed from an earlier episode. More specifically, the current study examined how attention to changes influenced the likelihood that people could later recollect such changes, along with how that would have an associated effect on memory for recent information. To accomplish this, I used a probe-caught mind wandering technique to capture fluctuations in attention during study in a variant of an A-B, A-D paradigm. This allowed me to observe the downstream consequences associated with fluctuations in attention on the ability to later recollect changes and recall recent information. By examining the role of attention in change processing during study, the current experiment will further our understanding of the factors that influence memory performance under conditions that can lead to proactive interference.

Memory-for-Change Account

As in the drug example above, interference is likely to occur when two stimuli have shared overlapping features. Traditionally, interference has been induced using the A-B, A-D paired associate learning paradigm, where participants are asked to study two lists of word pairs and are later tested on only one of the lists (for a review, see Anderson & Neely, 1996). The lists typically contain a mixture of word pairs that either repeat across lists (A-B, A-B), are only on one study list (C-D) or have the same cue with two different responses on each list (e.g., pearl-*harbor*; pearl-*jewelry;* A-B, A-D). Competition arises for A-B, A-D items because the two responses share the same retrieval cue (e.g., Postman & Underwood, 1973). Proactive interference is shown by lower correct recall of A-B, A-D items relative to C-D items, and a tendency to produce A-B intrusions at a rate above baseline when asked to recall A-D.

Although it is typical to observe proactive interference in such paradigms, there are several ways in which one can avoid such interference. Most often, interference can be counteracted by using a method to differentiate the lists from each other in order to reduce the competition between responses (for a discussion on list differentiation, see Postman & Underwood, 1973). Factors such as other effective cues about list membership or shorter retention intervals can allow one to resolve the interference that would occur from the overlap in retrieval cues by knowing the list membership of the responses. More recently, it has been suggested that the integration of responses can also serve to reduce interference that can occur in A-B, A-D paired associate learning paradigms. The Memory-for-Change account (e.g., Jacoby, Wahlheim, & Kelley, 2015; Wahlheim & Jacoby, 2013; MFC) posits that whether interference or facilitation is observed for A-B, A-D items depends on how often participants engage in change detection and recollection. Building on Hintzman's (2004, 2010, 2011) recursive reminding hypothesis, the MFC account proposes that a reminding can be triggered when two stimuli have an overlap in features, such as the same retrieval cue. This reminding can lead to change detection and allow for the representation of the previous stimulus to be embedded with a representation of the current stimulus. This configural representation is assumed to preserve the temporal order of the stimuli, since it can be inferred at retrieval that the stimulus that triggered the reminding had to occur more recently than the stimulus that was the object of the reminding. Furthermore, this configural representation provides a route to establish the list membership of responses (Jacoby, Wahlheim, & Yonelinas, 2013). Bringing the configural representation to mind later on allows one access to both stimuli and the temporal order in which they occurred, thereby providing a memorial benefit. This recollective process, coined change recollection, can be measured by asking participants at test if any other words came to mind when recalling the recent response (Wahlheim & Jacoby, 2013; Wahlheim, 2014) or by having participants explicitly indicate if they remembered a change (Jacoby et al., 2013; Negley, Kelley, & Jacoby, 2018; Wahlheim & Zacks, 2019; Wahlheim, 2015).

Wahlheim and Jacoby (2013) provided support for the MFC account in a series of studies using a variant of the A-B, A-D paired associate learning paradigms as described above. Participants completed a cued recall test for List 2 that included a change recollection measure. The results across the three experiments converged, showing that correct recall of List 2 responses was higher when change was recollected than when it was not. Critically, Experiment 1 showed that counteracting proactive interference was dependent on the ability to recollect change at test after it had been detected in List 2. The MFC account proposes that change detection involves the retrieval of the List 1 response during encoding of List 2, and without remembering that this change occurred, one may incorrectly report the List 1 response as the most recent at test. This intrusion occurs because retrieval practice of the List 1 response leads to it being highly accessible and therefore, mistaken as more recent. However, the high accessibility of List 1 can be offset with change recollection at test because this retrieval process provides access to the configural representation that includes both responses and their temporal relationship.

These studies provided evidence for the prediction from the MFC account that performance on A-B, A-D items is a combination of both proactive interference and proactive facilitation. Proactive facilitation emerged for items when change was recollected, and proactive interference occurred for items when change was not recollected.

Further work in this line has identified several factors that influence how often participants engage in change detection and recollection. These include: semantic associations (Wahlheim, 2014), age (Wahlheim & Zacks, 2019; Wahlheim, 2014), task instructions (Jacoby et al., 2015), interpolated testing (Wahlheim, 2015), List 1 repetitions (Wahlheim & Jacoby, 2013, Experiment 1; Wahlheim 2014, Experiment 2) and study time (Negley et al., 2018). Most relevant to the current study, Jacoby et al. (2015) showed that using instructions to manipulate attention towards changes during List 2 had a causal influence on the rate of change detection. Their experiments employed a variant of the A-B, A-D paradigm that included pairs that changed between List 1 and List 2, and also within List 2. Participants in the N-back group were told to indicate changes that occurred anywhere in the experiment (between List 1 and List 2 and within List 2) while participants in the Within-list back group were told to indicate only changes that occurred within List 2. This looking-back manipulation of controlled attention was successful because those in the N-back group detected more changes between List 1 and List 2 than those in the Within-list back group. Furthermore, those in the N-back group experienced proactive facilitation on the final test for between-list changed pairs, as recall performance for List 2 responses was higher on those items than on control items, while those in the Within-back group did not show this effect (Experiments 2 & 3).

By instructing participants on where to look for changes, the experimenters influenced how often people attended to changes that occurred between lists, and in turn, influenced later recall performance. This suggests that people can be guided in how often they look back to prior episodes. However, there are likely to be both inter- and intra-individual differences in the extent to which people direct their attention to previous episodes without instruction to do so. Rather than focusing on exogenous factors that can be manipulated or controlled, the current experiment will be the first to directly examine how natural fluctuations in attention during encoding influence the change detection and recollection processes proposed by the MFC account. Based on the results of Jacoby et al. (2015), I expected that when participants in the current experiment are attending to changed pairs during study, they will be more likely to detect change and recollect them. However, when their attention is focused on something other than the changed pairs during study, they will be less likely to detect change and recollect the changed pairs.

Measuring Attention Fluctuation through Mind Wandering Episodes

The general tenets of the MFC account have been supported by the work reviewed in the previous section. However, the assumption that natural variations in attention to changes can influence change processing has not been directly examined yet. To test this assumption, I applied a technique used to study more specific lapses in attention: mind wandering. Mind wandering, also known as task unrelated thoughts (Giambra, 1989) or stimulus independent thoughts (Antrobus, Singer & Greenberg, 1966), occurs when a person's thoughts move off from the current task to their own internal state (for a review, see Smallwood & Schooler, 2015). Based on studies using experience sampling, mind wandering has been estimated to occur between 30% (Kane et al., 2007) and 50% (Killingsworth & Gilbert, 2010) of the time during an average day. To capture these mind wandering episodes as they occur, researchers must rely on the ability of participants to report their own thought content or level of attention.

One laboratory-based method for measuring these episodes is the self-caught method, which requires participants to make a response if, or when, they become aware that they are mind wandering (Smallwood & Schooler, 2006). Due to the somewhat spontaneous nature of mind wandering (e.g., Smallwood, 2013), and the fact that mind wandering episodes may not always reach conscious awareness (e.g., Smallwood & Schooler, 2015), researchers have also employed the probe-caught method (Smallwood & Schooler, 2006). This involves spontaneously inserting probes during a task to ask people to categorize their thoughts just before that probe appeared. These probes can vary in form, such as giving a simple dichotomous on-task or off-task choice, as is used in the current study, or providing categorical options about what participants were thinking about (e.g., Kane et al., 2007; Kane et al., 2017). There is some inconsistency on the specifics of the probe design (for a recent discussion, see Weinstein, 2018), but unlike the self-caught method, the probe caught method does not require that participants make a response when they become aware of their own mind wandering.

Using these methods, researchers have gained valuable insight into the correlates of mind wandering. In the laboratory and in everyday life, individual differences in executive control abilities and working memory capacity have been shown to significantly predict the frequency of mind wandering (e.g., Kane et al., 2007; Kane & McVay, 2012; Kane et al., 2016; McVay & Kane, 2009). In addition, some external variables, such as the type of task being performed, can also influence how likely one is to mind wander (e.g., Smallwood & Schooler, 2006). Mind wandering typically occurs more often in less demanding tasks (e.g., Smallwood, Nind, & O'Connor, 2009) and increases with the time on the task (e.g., Metcalfe & Xu, 2016; McVay & Kane, 2012; Teasdale, 1995, Experiment 3; Thomson, Seli, Besner & Smilek, 2014). It may be best to characterize the frequency of mind wandering as dependent on the interaction of internal and external factors. Mind wandering episodes occur when executive control fails to maintain focus in the face of internal thoughts that are often spurred by the environment (e.g., Kane & McVay, 2012; McVay & Kane, 2009, 2010).

Mind Wandering and Memory

It has been argued that the route through which mind wandering has its negative effects on memory is primarily through encoding (e.g., Thomson, Smilek, & Besner, 2014). As posited by the decoupling hypothesis of mind wandering, when one has a mind wandering episode, the focus of attention goes from taking in sensory input from the external environment to input from one's internal thoughts and feelings. This is then associated with a decrease in sensitivity to further sensory input coming from the external environment, and therefore, poorer encoding of task demands (Smallwood & Schooler, 2006; Smallwood, 2011; Smallwood & Schooler, 2015). Researchers have found suggestive evidence that mind wandering may result in disruptions in encoding that can lead to more familiarity-based retrieval than a full recollective process (e.g., Smallwood, Riby, Heim & Davies, 2006; Riby, Smallwood & Gunn, 2008; Smallwood, Baracaia, Lowe & Obonsawin, 2003). When the mind wanders, it becomes more difficult to pay attention to the task at hand, and the consequences to that lapse in attention may be more magnified in situations where elaborative encoding and retrieval is involved.

Evidence for the detrimental effects of mind wandering when elaborative encoding is necessary has been shown by researchers studying the consequences of mind wandering on different levels of processing (e.g., Maillet & Rajah, 2013). In this study, the level of processing was manipulated within subjects, where participants were asked to make shallow encoding judgements for half of the experiment, and deep encoding judgements for the other half. The results showed that mind wandering predicted recognition memory performance, such that higher rates of mind wandering led to lower performance, but only for the deep encoding condition. One limitation of the study was that mind wandering was only assessed at the end of each encoding block. This could lead to imprecision in measuring the frequency of mind wandering since was not an “online” measure and instead required participants to rely on their memory of their internal thoughts from the prior encoding phase.

Improving on this limitation, Thomson et al. (2014) conducted a similar study that implemented thought probes throughout the encoding phase while participants made deep or shallow encoding judgements. In their study, when participants were in the deep encoding condition, they had to indicate if the size of the object the word represented was bigger than a computer monitor or not. When in the shallow encoding condition, they had to indicate if the word was presented in uppercase or not. The results converged with earlier work, showing there was a significant negative correlation between mind wandering and recognition memory performance in the deep encoding task only. Interestingly, there was no relationship between mind wandering and task performance after controlling for how accurately participants made the size judgement. This led the authors to argue that mind wandering disrupted the subjects' ability to make a correct encoding judgment, and therefore, reduced the effectiveness of the deep encoding manipulation, and subsequently, memory performance.

In addition to studies on the level of processing that participants engage in during study, mind wandering has been found to disrupt the ability of participants to infer or induce information that is necessary for completion of a task. In one such study, participants were asked to study paintings made by famous artists, and then were tested on the ability to induce which previously studied artist had completed new paintings. The results showed that mind wandering negatively predicted inductive abilities (Metcalfe & Xu, 2016). Furthermore, mind wandering has also been shown to reduce the ability for one to update a situation model, which had later consequences for comprehension of the associated narrative (Smallwood, McSpadden, & Schooler, 2008). The authors of this study reasoned that mind wandering during critical points in the narrative prevented participants from retrieving vital information that would have helped them integrate information in order to make correct inferences. Finally, mind wandering has also been shown to reduce student learning outcomes in both classroom (Risko, Anderson, Sarwal, Engelhardt & Kingstone, 2012; Wammes, Seli, Cheyne, Boucher & Smilek, 2016) and laboratory settings (Farley, Risko, & Kingstone, 2013; Kane et al., 2017; Loh, Tan, & Lim, 2016; Risko, Buchanan, Medimorec & Kingstone, 2013). Greater mind wandering during the video lecture was associated with poorer learning of the lecture material, presumably because the ability to integrate and retrieve knowledge was reduced when attention was not focused on the lecture.

Across these different contexts in which mind wandering impaired memory, there is evidence to suggest that mind wandering impacts the ability of one to encode the information necessary to perform well on the memory test. In particular, mind wandering was detrimental when the encoding was more elaborative in nature, such as during deep encoding or when integrating information to update a situation model. As described earlier, the process of change detection and recollection also requires elaborate integration in order to form a configural representation that can then be recollected at test. In accordance with the work by Smallwood, McSpadden, and Schooler (2008), it is expected that when participants in the current study are mind wandering when studying a changed pair, they will be less likely to be reminded of the earlier response, which will have downstream consequences for later memory for changes.

The Present Study

Based on previous studies, lapses in attention that are associated with mind wandering appear to play a critical role in memory, particularly through the encoding of information. Given that the MFC account outlines one way in which interference can be avoided, it is necessary to directly examine how attention influences the processes outlined by the MFC account. Attention is posited to play a role in the various processes outlined by the MFC account. First, the subjective experience of studying a changed word pair is akin to a recognition memory test or a categorization judgement where participants must decide if the word pair they are studying is the same as it was before, or if it has changed. If they have not given enough attention to effectively encode the original pair, this may lead them to experience the changed word pair as brand new. This could lead to memory performance on that pair to be similar to that of a control item, eliminating the potential for facilitation that occurs from detecting and recollecting change. Second, even if one has encoded the original presentation of the word pair, attention must also be given to the changed word pair when it is presented. Without attending to the changed word pair, the reminding that enables the formation of a configural memory representation containing both pairs will not occur. Bringing both stimuli into one configural representation creates the potential to access it later and to improve memory for the changed pair. Without this, one could experience a memory error whereby the original word pair is thought to be more recent because the changed word pair was encoded poorly. Change detection is a necessary first step in the process, as recall performance does not differ based on whether change recollected or not if change is not originally detected (Wahlheim & Zacks, 2019). Although the current study does not employ a measure of attention at retrieval, it is assumed that attention is also necessary for change recollection to occur at test.

Goals and Hypotheses

The main goal of the current experiment was to investigate these hypothesized downstream effects of attentional fluctuation on change detection, change recollection, and associated recall performance by examining both inter- and intra- individual differences in self-reported attention during encoding. To accomplish this goal, I designed a variant of an A-B, A-D paradigm that included thought probes during the study phase to examine how attention during encoding influences later memory. To my knowledge, this is the first study to use the thought probe method in a paradigm designed to induce proactive interference. Consequently, the present experiment is the first to characterize the consequences of variations in attention for memory under conditions that could lead to proactive interference. In the present experiment, the study phase contained seamless blocks in which items either repeated across blocks (A-B, A-B), were unique to each block (C-D), or had the same cue from earlier blocks paired with a different response in the last block of the study phase (A-B, A-D). During study, participants were periodically asked to report if they were "on-task" or "off-task". After completing the study phase, participants were tested on all responses that appeared in the last block, and were asked to indicate change and recall the earlier response if there was a change (i.e., the change recollection measure). I did not include an explicit change detection measure during study to minimize the exogenous influences on attention. However, I am still able to infer differences in change detection rates during study from the change recollection measure at test. This assumption is based on findings from prior research showing that manipulations that influence change detection also have similar influences on change recollection at test. For example, in a retroactive effects of memory study, Negley et al., (2018) found that longer study time for competing information increased the rate of change detection, and there was also an associated increase in change recollection for those items.

During the study phase, I expected that the proportion of "on-task" reports would decrease across blocks of the study phase because mind wandering typically increases as the duration of the task increases (e.g., Metcalfe & Xu, 2016; Teasdale, 1995, Experiment 3; Thomson, Seli, Besner & Smilek, 2014). During the test phase, I expected that overall recall performance would be highest for A-B, A-B items, and, consistent with the MFC account, performance for recall of the most recent response for A-B, A-D items would reflect both proactive facilitation and proactive interference. The balance of these effects was expected to depend on how often participants recollected change. I expected proactive interference, where recall for the recent response of A-B, A-D items is lower than recall for C-D items, when participants did not recollect change and proactive facilitation, where recall for the recent response of A-B, A-D items is higher than recall for C-D items, when participants did recollect change (e.g., Wahlheim, 2015). I also expected that participants would report the highest rate of intrusions for the A-B, A-D items as compared to the other two item types (for which intrusions represent a baseline of how likely participants are able to guess what the response would have changed from if it had been an A-B, A-D item).

Importantly, I expected positive association between attention to the task and later recall, such that participants with a greater proportion of "on-task" reports would have higher recall than participants with a greater proportion of "off-task" reports. Furthermore, I expected that within participants, when an "on-task" report was given, recall performance would be higher than when an “off-task” report was given. This difference in performance should be larger for recall of A-B, A-D items than A-B, A-B items, since the former requires encoding of both the original and the changed response while the latter provides more chances for effective encoding due to repeated presentations. I also expected higher rates of intrusions for A-B, A-D items when an “off-task” report was given compared to an “on-task” report because participants had more opportunities to encode the A-B, A-B item throughout the study blocks and only one opportunity to encode the A-B, A-D item. Therefore, the A-B, A-B item should have a higher probability of being reported if participants were “off-task” during the last block of the study phase. Regarding change classifications made at test, I expected that when participants gave an “on-task” report, they would have higher change classifications than when they report “off-task”. Being off-task means that attention was shifted away from encoding that item, making it less likely that participants would detect the change and report it as changed later at test.

In addition to the hypotheses regarding the relationship between attention and memory, I also explored whether the rate at which participants were on-task in the current experiment would relate to how often they make inattention errors in everyday life. To do so, I made a computerized version of the Revised Attention Related Cognitive Errors Scale (ARCES; Carriere, Cheyne & Smilek, 2008). The measure contains 12 statements about errors in everyday life that stem from a lapse in attention. As an example, the second question on the measure states “I go into a room to do one thing (e.g., brush my teeth) and end up doing something else (e.g., brush my hair), and participants are asked to indicate how often this happens for them on a scale from 1 (*never*) to 5 (*very often*). This scale has been shown to significantly correlate with other measures of a general propensity have attention lapses and has been validated in a large college sample (Carriere et al., 2008; Cheyne, Carriere, & Smilek, 2006). I expected that there would be a negative relationship between attention to the task and the ARCES measure, indicating that more "on-task" reports were associated with a lower frequency of attention related errors in everyday life.

CHAPTER II

METHOD

In what follows, I report how I determined sample size, all data exclusions, all manipulations, and all measures in this study (Simmons, Nelson & Simonsohn, 2012). Data files and analysis scripts can be found here: https://osf.io/56t9k/.

Participants

My final sample consisted of 132 younger adults (95 female), ages 18-29 (*M* = 19.02, *SD* = 1.70), who were recruited through an online participant pool for undergraduates at the University of North Carolina at Greensboro (UNCG). The sample had an average of 12.77 years of education (*range* = 12-16, *SD* = 1.19). The sample composition was as follows: 44% identified as African American (59), 40% Caucasian (53), 5% Hispanic or Latino (7), 5% as more than one race (7), and 3% Asian or Pacific Islander (4). I determined the sample size for the experiment based on the number of participants needed to examine the interaction within-subjects between task reports and item type for recall performance, which was the main interest for the study. Prior work manipulating external variables to influence change recollection and recall performance have found small to medium sized effects ranging from $\eta_p^2 = .06 - .09$ (Negley et al., 2018; Wahlheim, 2015). According to G*Power Version 3.1.9.2 (Faul, Erdfelder, Buchner, & Lang, 2009), a total sample size of 128 is sufficient to detect a small to medium effect ($\eta_p^2 = .06$, which translates to Cohen's $f = .25$), at power = .80 and $\alpha = .05$

(two-tailed). With this sample size, the study is also powered at 80% to detect a between-subjects correlation of $r = .25$. Given that the current study requires 12 formats to complete the counterbalance (described in the next section), I collected data from 132 participants to allow each format of the experiment to appear equally often across participants. Participants received partial course credit as compensation. Two participants were replaced, one due to an interruption from a fire drill, and one for falling asleep during the session (total of 134 participants tested).

Design and Materials

This experiment employed a within-subjects design, with the primary independent variable being Item Type (A-B, A-B vs. C-D vs. A-B, A-D). The materials consisted of 156 word sets (144 critical and 12 buffers) taken mostly from Jacoby (1996). The design of the study required more word sets than were available from the original study, so sets from the Nelson, McEvoy, and Schreiber (1998) free association norms were also included. Each set contained a cue (e.g., knee) and two responses (e.g., bone, bend). The two responses could complete the same word fragment because they have an orthographic relationship (e.g., knee-b_n_), but the fragments were not used here. For counterbalancing, the critical word sets were divided into 6 groups of 24. Each group appeared as each item type equally often across participants. For the first 6 formats, the response arbitrarily labeled as Response 1 (e.g., bone) was the target word while the response labeled as Response 2 (e.g., bend) was the target word for the other 6 formats. The average lengths of cues ($M = 5.26$, $SD = 1.60$, range 2-9 characters) and responses ($M = 4.76$, $SD = 1.08$, range 3-8 characters) were matched across groups. The frequency of responses was assessed according to hyperspace analog to language (HAL), which is used as a representational model of semantic memory (Lund & Burgess, 1996) and was catalogued by the English Lexicon Project (Balota et al., 2007). The average frequency was matched across groups for the cues (*M* = 9.44, *SD* = 1.45, range = 6-14) and the responses (*M* = 9.34, *SD* = 1.60, range = 5-14). Associations between cues and targets, indexed by Nelson, McEvoy and Schreiber (1998), were low on average for both forward associative strength (*M* = .06, *SD* = .08, range = .03-.10) and backward associative strength (*M* = .08, *SD* = .14, range = .03-.15). In addition, targets that were paired with the same response had weak backward and forward associations to each other (*M* = .02, *SD* = .06, range = .001-.07).

A schematic for the study phase is shown in Figure 1. The study list was divided into 4 seamless blocks with 72 word pairs in each block. One set of word pairs (24 in each block) repeated in all four blocks (A-B, A-B). Another set of word pairs (24 in each block; 96 total) were new in each block and had no relationship to any pairs from the previous blocks (C-D). The last set of word pairs (24 in each block) repeated in the first three blocks and then had the same left-hand word paired with a different right-hand word in the fourth block (A-B, A-D). For example, the participant would see throat-tonsil in the first, second and third block and then see throat-tongue in the fourth block. Buffer items appeared at the beginning and end of the study phase, with 4 buffer items that represented each of the three item types (12 buffer items total). Word pairs appeared in a fixed random order in each block of the study phase, with the stipulation that no item type appeared more than three times consecutively. The average serial position for each item type was also equated within each block to control for serial position effects on later recall performance.

Nine thought probes were inserted into each block of the study phase (36 total). The thought probes were pseudo-randomly placed to occur an equal number of times after each item type (three probes after each item type in each block). The probe screen appeared directly after the word pair disappeared. Probes were assigned to the same item type as the word pair that appeared directly before the probe. To reduce the systematicity of the probes, the interval at which the probes appeared was either 46, 54, 62, 70, or 78 seconds (corresponding to 6, 7, 8, 9, or 10 word pairs, respectively). The average duration between thought probes was 62 seconds (*SD* = 12.09). Each thought probe consisted of a discrete on-task or off-task judgment.

The test phase, which came immediately after the study phase, was self-paced and included cues from all 72 pairs that were presented in the fourth block. The cues appeared in a fixed random order for each format during the test phase, with the stipulation that cues representing the same item type did not appear more than three times consecutively and that serial position was relatively equal for all item types.

Participants completed the ARCES measure (Carriere, Cheyne, & Smilek, 2008) on the computer. Each question from the paper and pencil version was converted to appear individually on the computer screen. All 12 statements appeared in the exact order as the paper and pencil version, and were given the same scaling options that ranged from 1 (*never*) to 5 (*very often*).

Procedure

All participants were tested individually in rooms with a white noise machine running to reduce external distractions. Once consent was obtained, participants completed a computerized demographic questionnaire. All experimental stimuli were administered through E-prime software (Version 3, Psychology Software Tools, Inc) and appeared in white Arial size 24 font on a black background. Participants were asked to study a list of word pairs for an upcoming memory test. Word pairs appeared for 6 s each with a 2 s interstimulus interval (ISI) between each presentation. Participants were told that they would periodically be asked about their attention to the task, and were given an explanation about what was meant by “on-task” and “off-task” reports (see Appendix B for study instructions). On these thought probe trials, the probe screen appeared after the 6 s study duration for the word pair (before the ISI). Participants were told to indicate if they were “On-task” or “Off-task” by using the mouse to click on the corresponding button. These responses were self-paced. After answering any questions that participants had, and monitoring the first few trials (the buffer items), the experimenter left the room to allow participants to experience natural variation in attention without the influence of the experimenter.

After completing the study phase, participants notified the experimenter, who returned and remained in the room for the test phase. To begin, the same six of the twelve buffer items were chosen for all participants to familiarize them with the test procedure. A cue appeared with a question mark (e.g., throat-?), and participants were asked to type in the response that the cue was paired with most recently (e.g., tongue). After entering their response, a question appeared asking if the right word paired with the cue changed during the study phase. Participants were given the option to either press the “1” key to indicate that the word pair changed during the study phase, or to press the “0” key if the right word did not change during the study phase. If participants indicated change by pressing “1”, they were then asked to type in what the cue was paired with earlier in the study phase (e.g., tonsil). If participants indicated that the word pair did not change, they moved on to the next trial (see Appendix C for test instructions). Participants could guess or pass on trials when they could not remember a response. After the test, participants were administered the Revised Attention Related Cognitive Errors scale (Carriere et al., 2008) computerized version that I programed. Finally, participants were debriefed, escorted out of the laboratory, and given partial course credit. The session lasted approximately 1.5 hours.

CHAPTER III

RESULTS

All analyses were conducted using R software (R Core Team, 2019). Logistic mixed effects models were fit to the data using the glmer function from the lme4 package to account for random effects of subjects and items (Bates, Maechler & Dai, 2011). All models in the analyses below include subjects and items as random effects and experimental manipulations as fixed effects unless otherwise noted. Hypotheses were tested using the Anova function from the car package (Fox & Weisburg, 2011). Post hoc tests were done on the estimated marginal means produced by the emmeans package, which uses the Tukey method to control for the family wise error rate (Lenth, 2018). In addition, output from standard ANOVAs and t-tests along with partial eta squared and Cohen's d values as estimates of effect size are included in Appendix D. All significant findings were based on $\alpha = .05$.

Study

Attention to Task. To quantify the extent of attentional lapses during study, I calculated the proportion of "on-task" reports made by participants as a function of Block (1-4) and Item Type (Figure 2). I fit a model to the on-task proportion that included Block and Item Type as fixed effects. The model indicated a significant effect of Block, χ^2 (3) = 40.94, $p < .001$, no significant effect of Item Type, χ^2 (2) = .83, $p = .66$, and a significant Item Type × Block interaction, χ^2 (6) = 26.00, *p* < .001. Pairwise comparisons showed that for A-B, A-B items, the on-task proportion did not differ between Block 1 and the other 3 blocks, *largest z* ratio = 2.50, *p* = .06. The on-task proportion was higher in Block 2 than Blocks 3 and 4, *smallest z* ratio = 3.48, *p* = .003, and did not differ between Block 3 and 4, *z* ratio = 1.17, *p* =.64. For C-D items, the on-task proportion in Block 1 did not differ from Block 2, *z* ratio = .62, *p* = .93, but was significantly higher than in Blocks 3 and 4, *smallest z* ratio = 3.42, *p* = .004. The on-task proportion in Block 2 was significantly higher than in Blocks 3 and 4, *smallest z* ratio = 2.80, *p* = .03. The on-task proportion did not differ in Block 3 and Block 4, *z* ratio = .18, *p* = 1.00. For A-B, A-D items, the on-task proportion was significantly higher in Block 1 than in Blocks 2 and 3, *smallest z* ratio = 2.79, *p* = .03, but did not differ from the on-task proportion in Block 4, *z* ratio = .53, *p* = .95. The on-task proportion in Block 2 did not differ from Blocks 3 and 4, *largest z* ratio = 2.27, *p* = .11. Notably, the on-task proportion was significantly higher in Block 4 than in Block 3, *z* ratio = 3.53, *p* = .002. These results suggest that attention increased from the third to the fourth block for items that changed, but this effect was not shown for items that repeated or had never been studied before. I address the relationship between the on-task proportion and recall performance, which is the main interest of this study, in a later section.

Test

Recall Performance. Correct recall as a function of Item Type is displayed in Figure 3 (black points). To examine how recall performance varied as a function of Item Type, I fit a model to the correct recall data with a fixed effect of Item Type. As expected, the model indicated a significant effect, χ^2 (2) = 1167.60, $p < .001$. Post hoc analyses revealed that recall for A-B, A-B items was higher than for the other two item types, *smallest z* ratio = 28.81, $p < .001$. Recall for A-B, A-D items was also significantly higher than for C-D items, z ratio = 2.67, $p = .02$. Together these results show that memory was highest when there were spaced repetitions of study items and memory was poorest when an item was only seen once. In addition, proactive facilitation was observed in overall recall because recall for A-B, A-D items was higher than C-D items.

Intrusions rates are displayed in Figure 4 (black points). Note that intrusions for A-B, A-B and C-D items represent baseline estimates of how often participants produced what the response would have been, presumably by guessing, had the item been in the A-B, A-D condition. To examine intrusion rates as a function of Item Type, I fit a model with a fixed effect of Item Type. The model indicated a significant effect, χ^2 (2) = 982.26, $p < .001$, showing that intrusions were the highest for A-B, A-D items compared to the other two item types, *smallest z* ratio = 23.18, but did not differ between A-B, A-B and C-D items, z ratio = .11, $p = .99$. These results show that participants experienced proactive interference on A-B, A-D items, in the form of intrusions.

Change Classification. The change classification rates at test as a function of Item Type are shown in Table 1. I fit a model to the change classification data that included a fixed effect of Item Type. The model indicated a significant effect, χ^2 (2) = 1379.40, $p < .001$, with the highest change classifications rates for A-B, A-D items compared to the other two item types, *smallest z* ratio = 29.06, $p < .001$. Change classifications were also higher for A-B, A-B items than for C-D items, z ratio = 2.73, p =

.02. These findings suggest that participants could identify a reasonable proportion of changed items as such. Since I was most interested in change classifications made for A-B, A-D items, I also calculated the joint probabilities of the three possible outcomes for change classifications at test for those items: classifying an item as changed and correctly reporting the earlier response (change recollection), classifying an item as changed without correct recall of the earlier response, and not identifying an item as changed. As shown in Table 2, participants were able to report the earlier pairing some, not but all of the time after they indicated change.

Recall Performance Conditionalized on Change Classification. To examine the association between change recollection and later recall performance, I used the change classifications outlined above to conditionalize correct recall for A-B, A-D items (Figure 3, colored points). The model fit to the conditionalized recall data included a fixed effect of Change Classification (including the three levels listed above) along with recall performance on C-D items so that proactive effects of memory could be examined. The model indicated a significant effect of Change Classification, χ^2 (3) = 669.37, $p <$ .001. Recall performance was significantly higher when participants recollected change than when they did not, *smallest z* ratio = 15.62, $p < .001$, but did not differ between change classifications made without recall of the earlier response and no change classification, z ratio = 2.01, $p = .18$. Regarding proactive effects of memory, participants experienced proactive facilitation when change was recollected, as recall for A-B, A-D items was higher than recall for C-D items, z ratio = 20.42, $p < .001$. When participants did not recollect change, proactive interference occurred, where recall for A-B, A-D items was lower than recall for C-D items, *smallest z* ratio = 4.65, $p < .001$.

I also fit a model to intrusion data that were conditionalized based on whether change was indicated or not (Figure 4, colored points). I did not conditionalize intrusions based on change classification along with correct recall of the earlier response because making that response would negate producing it as an intrusion. The model indicated a significant effect of Change Classification, χ^2 (1) = 126.15, $p < .001$, with higher intrusion rates when change was not indicated compared to when it was, *z* ratio = 11.32, $p < .001$.

Relationship between Study and Test

Between-Subjects Correlations between Recall Performance and Proportion On-Task.

Given that participants were tested on responses that were presented in Block 4, I limited the correlations I performed to the proportion of "on-task" reports given in Block 4. This allowed for the most direct link between attention to the task and memory performance on the final test. I conducted Pearson product-moment correlation coefficients for the on-task proportion in Block 4 and the associated correct recall on the test separately for each item type (Figure 5). For all three item types, there were significant positive correlations, such that recall performance increased as the on-task proportion increased (A-B, A-B: $r(130) = .34$, $p < .001$; C-D: $r(130) = .41$, $p < .001$; A-B, A-D: $r(130) = .45$, $p < .001$). As shown in Figure 6, the correlation between the on-task proportion in Block 4 and intrusion rates for A-B, A-D items was not significant, $r(130) = -.12$, $p = .16$. Finally, as shown in Figure 7, there was a significant positive correlation between the on-task proportion in Block 4 and change recollection for the A-B, A-D items, $r(130) = .38$, $p < .001$. Together these results show that participants who were on-task more often had higher correct recall and higher rates of change recollection compared to those who were off-task more often.

As an exploratory measure, I also conducted a correlation between the on-task proportion across the experiment (collapsing across blocks) and recall performance on the C-D items (Figure 8). Although the on-task proportion for Blocks 1-3 does not align with items that participants were tested on, this could reveal how attention to the task in general was related to recall performance. There was a significant positive correlation, $r(130) = .50$, $p < .001$, showing that as the on-task proportion increased, recall performance for C-D items increased.

Recall Performance Conditionalized on Task Reports. In addition to examining the relationship between attention during study and later memory performance between subjects, I also examined this relationship within subjects. If attention is necessary for recall performance, then participants should perform worse when they make an “off-task” report than when they make an “on-task” report. This may be especially true for A-B, A-D items because attention would be required to encode changes. To test this, I conditionalized correct recall for each Item Type on whether participants gave an “on-task” or “off-task” reports during Block 4 (Figure 9). Consistent with the previous analyses, I limited this comparison to task reports made in Block 4 only since participants were tested on Block 4 items (and this is where the change for A-B, A-D items occurred). To be included in this analysis, participants must have had at least one “on-task” report and one “off-task” report in Block 4. Restricting the analysis in this way resulted in an unequal number of participants included in the comparisons across each of the Item Types, but the same participants being compared across on and off-task responses within each Item Type.

I fit a model to the recall data that included fixed effects of Item Type and Task Report. I do not report results that are redundant with those reported before. The model indicated a significant effect of Task Report, χ^2 (1) = 10.26, *p* = .001, showing that participants had higher recall when giving “on-task” compared to “off-task” reports. Although the interaction between Item Type and Task Report was not significant, χ^2 (2) = 2.77, *p* = .25, I examined the post hoc analyses for this interaction because it was the main interest for the current study. Post hoc analyses showed that the difference between recall performance when participants were on or off-task for A-B, A-B items was not significant, *z* ratio = .69, *p* = .49, but that this difference was significant for both C-D items, *z* ratio = 2.58, *p* = .001, and A-B, A-D items, *z* ratio = 2.44, *p* = .01. Although inconclusive, the pattern of results does suggest that attention during encoding in Block 4 influenced recall performance for changed and new items, but not for repeated items.

For the A-B, A-D items, I also conditionalized intrusions (Figure 10) and change recollection rates (Figure 11) based on whether participants made “on-task” or “off-task” reports for probes following A-B, A-D items in Block 4. I fit separate but comparable models for conditionalized intrusions and change recollection rates that included a fixed effect of Task Report. The model fit to intrusions indicated no significant effect of Task Report, χ^2(1) = .20, *p* = .65, showing that intrusion rates did not differ between “on-task” and “off-task” reports. The model for change recollection rates indicated a significant effect of Task Report, $\chi^2(1) = 5.98$, $p = .01$, showing that participants were more likely to recollect change when they reported being “on-task” compared to “off-task”.

ARCES

Between-Subjects Correlations between ARCES and Proportion On-Task.

To examine the relationship between attention to the task and how often participants may experience everyday attention errors, I examined the correlation between the score on the ARCES measure and the proportion of on-task reports. Note that this measure is potentially contaminated given that it was administered once participants had already made on and off-task judgements during the study phase. I calculated an item mean for each participant (Cheyne, Carriere, & Smilek, 2006). As discussed in the Introduction, possible scores ranged from 1-5 on the measure, with higher scores indicating a greater likelihood of experiencing everyday cognitive failures related to attention. The average score on the measure across all participants was $M = 3.33$, $CI = [3.28, 3.38]$. The correlation between the ARCES score and the proportion on-task was not significant, $r(130) = -.02$, $p = .79$ (Figure 12). There was no evidence for a relationship between the propensity to experience errors in everyday life due to lack of attention and the attention given during encoding in the current task.

CHAPTER IV

DISCUSSION

The present experiment examined the consequences of natural fluctuations in attention on memory under conditions that could lead to proactive interference. Results showed that, in general, attention decreased as the study phase continued. However, this pattern was different for A-B, A-D items: there was an increase in the self-reported attentional engagement from the third to the fourth (and last) block of the study phase, which is when the changes appeared. This increase did not occur for A-B, A-B or C-D items. The results from the cued recall test showed that proactive facilitation was observed in overall recall, suggesting that participants were able to detect and recollect changes often enough to drive up performance. Recall for A-B, A-D items was higher when change was recollected than when it was not. Furthermore, the conditionalized recall data showed proactive facilitation when change was recollected, and proactive interference when change was not recollected. In addition, conditionalized intrusion rates were lower change was remembered than when it was not.

Critically, there were relationships between attention during the study phase and later recall, both between and within participants. There were significant positive between subject correlations between the on-task proportion during study and subsequent recall. Furthermore, within subjects, conditionalized recall performance was significantly higher when an “on-task” report was made compared to when an “off-task” report was made. For A-B, A-D items, change recollection rates were significantly higher when an “on-task” report was made compared to when an “off-task” report was made. The theoretical implications of these findings for the MFC account and the ways in which the results situate with prior mind wandering work is discussed below.

Attentional Fluctuation and Memory for Changes

The findings from this experiment add to the growing body of literature in support of the MFC account. As described in the Introduction, the account proposes that overall recall performance on A-B, A-D items in a paired associate learning paradigm is comprised of both facilitation and interference effects. When change is detected and recollected, then proactive facilitation is observed. In contrast, if change is detected and not recollected, then proactive interference is observed. Although the current study did not directly measure change detection in List 2, the results from this study are consistent with the MFC account in showing that participants experienced proactive facilitation when change was recollected and proactive interference when change was not recollected.

The novel contribution of this study was to examine the role that attention plays in detecting and recollecting change. By manipulating instructions on where to look for changes, Jacoby et al. (2015) showed that attention can be directed towards changes. The current study is the first to build on this idea and empirically examine the influence of natural variations in attention over the course of a study phase on later change processing. Based on the finding that mind wandering disrupts elaborative encoding and leads to poorer memory performance (e.g., Thomson et al., 2014), I expected that being off-task would be negatively associated with participants' ability to recollect change, which would be associated with poorer memory accuracy on the cued recall test. Consistent with this hypothesis, the results showed that when participants indicated being "off-task" in the last block, they had lower change recollection rates and lower correct recall for A-B, A-D items compared to when they indicated being "on-task". It was also found that across participants, those who paid more attention, as indicated by more "on-task" reports during the last block of the study phase showed higher correct recall for A-B, A-D items than participants who paid less attention, as indicated by more "off-task" reports. Part of the benefit to correct recall may have been driven by higher change recollection rates, as it was found that participants who were on-task more often in the last block recollected change more often than those who were off-task more often. This suggests that attention is a necessary component for detection and later recollection of changes.

Given the evidence for the role of attention in change processing from the current study, the MFC account may need to be refined to further articulate a role for natural variations in attention during encoding (which can stimulate study-phase retrievals). As discussed in the Introduction, the subjective experience of encoding a changed word pair is akin to taking a recognition memory test or to categorizing an object based on both shared and distinctive features. One must decipher if the current pair being studied is the same as the one seen before, or if it includes both shared and changed features. Since change detection and recollection begins with the initial reminding that allows for the responses to become integrated, attention must be given to the changed response to allow this reminding to occur. Attention during encoding is critical for the memorial benefits associated with change recollection, as previous research has shown that without the initial reminding that leads to change detection, there is no further memorial benefit associated with change recollection (Wahlheim & Zacks, 2019). Therefore, further work examining the MFC account should incorporate how natural variation in attention contributes to the initial reminding process. This could be used as one tool to better predict whether one may experience proactive interference when response competition is present.

It was also hypothesized that being off-task during encoding when under conditions that could lead to proactive interference would increase the rate of intrusions on the final test, as disruptions to encoding the changed pair could allow the original pair to become a strong competitor. Instead, the results showed that there was no significant correlation between the on-task proportion and the intrusion rates, and intrusion rates did not differ based on whether an "on-task" or "off-task" report had been made. This relationship could depend on how attention was allocated throughout the previous three blocks. The current study included thought probes inserted pseudo-randomly throughout the blocks in order to capture attention lapses more naturally, but this meant that there was not a direct match for probes to appear after the same items in each block. Consequently, the data reported here does not allow me to draw conclusions about the attention allocation during the study on specific items. Despite this limitation, one can imagine that if participants give attention to an A-B item but not the associated A-D item, intrusions rate may increase. In order to more accurately capture this relationship, future research should consider comparing attention for both the A-B item and the associated A-D item in the study phase and then examining the downstream consequences for later memory.

The focus of the current study was on the role of attention under conditions that could lead to interference, with the hypothesis that the difference between performance when one is on or off-task should be the largest for changed information (A-B, A-D items). Although inconclusive, there was evidence in the pairwise comparisons for recall data conditionalized on probe reports in Block 4 to suggest that attention during study was related to recall performance for both C-D and A-B, A-D items, but not for A-B, A-B items. This pattern suggests that future research should design studies that can examine this relationship in a more powerful way. One technique would be to remove the C-D items from the study and test phase and only use A-B, A-B and A-B, A-D items. This would increase the number of observations for each item type and increase the power to examine such interactions. In addition, as mentioned above, it will be important for future studies to map thought probes onto the same items across blocks. By having probes that appear consistently for the same items across blocks, one may be able to better estimate if the consequences for attention on later recall differ as a function of the item type.

It is also important to point out that there was evidence from the between subject correlations showing that attention during encoding was important for memory for all item types, not just for changed items. Conceptually, being engaged in an off-task thought is like a divided attention-task where one is splitting attention between the external task demands and their own internal thoughts. Therefore, the relationship between attention during encoding and overall memory performance is not surprising given the large body of literature on the detrimental effects on memory when attention is divided (e.g., Anderson, Craik & Naveh-Benjamin, 1998; Craik, Govoni, Naveh-Benjamin & Anderson, 1996; Fernandes & Moscovitch, 2000). In tasks used to study divided attention, participants are usually required to study a list of words and concurrently perform a secondary task, such as a continuous reaction-tasks that requires participants to quickly discriminate between sounds or to find a target in a visual array. Results from these studies show that memory performance is hindered when attention is divided during encoding compared to when attention is not divided. This is consistent with the results reported here showing that recall performance was lower when participants were paying attention to something else other than studying the word pair compared to when they indicated paying attention to the word pair.

Relationship to Mind Wandering Literature

In the current experiment, I utilized thought probes as a tool to measure attentional fluctuation during study. By doing so, the present findings can further contribute to the small body of literature examining the consequences of mind wandering on memory performance in paradigms testing basic memory processes. Previous work has found that the type of processing that participants are engaged in can influence how likely they are to pay attention. When participants are asked to engage in a task that requires rating how pleasant they feel a word is (Maillet & Rajah, 2013), or if the word is too easy or too difficult for them to study (Xu & Metcalfe, 2016), they are more likely to mind wander. Furthermore, when the task requires them to engage more elaborately with the word by imaging its size relative to a computer monitor, participants more often let their thoughts drift away (Thomson et al., 2014). In addition to these factors, the results from the current study suggest that the A-B, A-D items acted as a cue for participants to pay attention because there was an increase in attention from the third to the fourth block of the study phase for A-B, A-D items only. This finding cannot be explained by a novelty effect because this increase in on-task reports was not shown for C-D items, which were also novel to the last block.

This finding of increased attention to changes is somewhat consistent with other work suggesting that situational changes can reduce mind wandering. Specifically, Faber, Radvansky and D’Mello (2018) examined the number of self-caught mind wandering episodes while participants watched a narrative film that included a range of changes, such as different locations or characters being shown. The results showed that more situational changes in the narrative and a higher likelihood of an event boundary occurring were associated with less mind wandering. The authors argue that when there was a change in the narrative, it kept people focused on the film rather than on their own internal thoughts. Although the level of change in the current experiment could have been more subtle given that it was word pairs instead of dynamic stimuli, this may be one reason why the rate of on-task reports increased for the A-B, A-D items from the third to the fourth block.

Another possibility is that retrieving the earlier response when encoding the changed response (which was assumed to occur during change detection) acted as a type of test for participants. We know that the benefits of interleaving testing during study can reduce the rates of mind wandering (Szpunar, Khan, & Schacter, 2013). In their study, Szpunar et al. (2013) had participants study lecture notes in the laboratory and then had to report how often they mind wandered after the lecture (Experiment 1) or were probed if they were mind wandering during the lecture (Experiment 2). The results across both experiments showed that the interpolated testing group showed lower mind wandering rates than the group that had not been tested throughout the lecture segments. According to the MFC account, the presentation of the A-D pair can stimulate spontaneous retrieval of the A-B pair, so the presentation of A-D pairs in the current study could have kept participants more focused on encoding those items compared to other items that did not require such retrieval processes. It could be argued that the presentation of a repeated A-B pairs may also stimulate the retrieval of earlier A-B pairs (Wahlheim, Maddox, & Jacoby, 2014), but this retrieval process would likely be different. The retrievals that are triggered by the presentation of A-D pairs may be more indicative of elaborative processing that requires more focus because it is posited that participants are integrating both responses into a configural representation along with the reminding that bound the responses together (e.g., Jacoby et al., 2015; Wahlheim & Jacoby, 2013).

Finally, the results of the current study inform the utility of the ARCES measure, which was used here in an exploratory manner to examine the relationship between attention to the task during encoding and the propensity to experience errors related to lapses in attention. Prior work has found that the ARCES measure was highly correlated with other measures of the propensity to have attentional lapses and on errors in a SART task (Cheyne, Carriere, & Smilek, 2006; Carriere et al., 2008). However, there has not been research examining how the ARCES score relates to performance on other cognitive tasks. The lack of correlation found here suggests that it does not correlate with the lapses of attention that occur during an episodic memory task. The design of the task used in the current experiment differs significantly from sustained attention measures, and so it may be the utility of the ARCES measure is limited to sustained attention-tasks where an error score can be calculated continuously throughout the task, rather than an paired associate learning task where the "errors" do not appear until the final recall test is taken.

Limitations of the Present Experiment

Although the results of the current study support the relationship between attention and the ability to recollect changes, there are several limitations that should be acknowledged. As with earlier studies relying on self-reported mind wandering episodes, it is unknown how accurate the measurement of attentional fluctuation was. Furthermore, it is possible that variations in the experimental design could influence the results. For example, the probe design used in the current study, which asked participants to indicate if they were on or off-task, deviates from other mind wandering work that uses several categorized thought options (e.g., Kane et al., 2007; Kane et al., 2017) or explicitly gives participants the option to indicate that they are "mind wandering" (e.g., Metcalfe & Xu, 2016; Xu & Metcalfe, 2016). It is unknown exactly how the wording of the probes may impact the proportion of "on-task" reports in the current experiment, but previous work has suggested that mind wandering rates can vary as a function of probe framing (e.g., Weinstein, De Lima, & van der Zee, 2018). Furthermore, due to the constraints of the study design, probes appeared for participants 62 seconds apart on average. This could also impact the proportion of “on-task” reports because prior work found that mind wandering rates increase as the time between probes increases (Seli, Carriere, Levene, & Smilek, 2013). Given these limitations, it is important for future work to examine if the relationship between attention and change processing remains when the framing and timing of the probes is altered.

Concluding Remarks

In sum, the current experiment is the first to attempt a direct examination of the role that natural fluctuation in attention has in change processing. Results showed that recall performance both between and within participants was significantly higher when participants were “on-task” compared to “off-task”. This suggests that attention does impact how often one will engage in change detection and recollection, and that there are consequences for fluctuations in attention on later recall. By testing the assumption from the MFC account that attention is involved in these processes, this experiment further refined the account by identifying attention as one endogenous factor that influences change detection and recollection. Future work should continue to examine how combinations of attention on both the original and change information can influence the processes posited by the MFC account, and should utilize a variety of methods to further test the boundaries of the findings reported here.

REFERENCES

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: I. Evidence from divided attention costs. *Psychology and Aging, 13*(3), 405-423.

Anderson, M. C, & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), Handbook of perception and cognition (2nd ed.): Memory (pp. 237-313). San Diego, CA: Academic Press.

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills, 23*, 399-417.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchinson, K. A., Kessler, B., Loftis, B., . . . Treiman R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445–459.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Carriere, J. S. A., Cheyne, J. A., & Smilek, D. (2008). Everyday attention and memory failures: the affective consequences of mindlessness. *Consciousness and Cognition, 17*, 835-847.

Cheyne, J. A., Carriere, J. S. A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition, 15*, 578-592.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*(2), 159-180.

Faber, M., Radvansky, G., & D'Mello, S.K. (2018). Driven to distraction: A lack of change gives rise to mind wandering. *Cognition, 173*, 133-137.

Farley, J., Risko, E., & Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Frontiers in psychology*, *4*, 619.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General, 129*(2), 155-176.

Fox, J., & Weisburg, S. (2011). *An {R} companion to applied regression, second edition.* Thousand Oaks, CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Giambra, L. M. (1989). Task unrelated thought frequency as a function of age: a laboratory study. *Psychology and Aging, 4*, 136-143.

Hintzman (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition, 32*(2), 336-350.

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. Memory & Cognition, 38, 102–115.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the “coordinates of truth.” Perspectives on Psychological Science, 6, 253–271.

Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language, 35,* 32-52.

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1282-1297.

Jacoby, L. L., Wahlheim, C. N., & Yonelinas, A. P. (2013). The role of detection and recollection of change in list discrimination. *Memory & Cognition, 41*, 638-649.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life.

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348-354.

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General*, *145*(8), 1017-1048.

Kane, M. A., Smeekens, B. A., von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General, 146*(11), 1649-1674.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932.

Lenth, R. (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.2.3. https://CRAN.R-project.org/package=emmeans.

Loh, K. K., Tan, B. Z. H., & Lim, S. W. H. (2016). Media multitasking predicts video-recorded lecture learning performance through mind wandering tendencies. *Computers in Human Behavior*, *63*, 943–947.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavioral Research Methods, Instruments & Computers, 28*(2), 203-208.

Maillet, D., & Rajah, M. N. (2013). Age-related changes in frequency of mind-wandering and task-related interferences during memory encoding and their impact on retrieval. *Memory*, *21*(7), 818-831.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, *35*(1), 196-204.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*(2), 188–197.

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “d’oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 525–549.

Metcalfe, J. & Xu, J. (2016). People mind wander more during massed than spaced inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(6), 978-984.

Negley, J. H., Kelley, C. M., & Jacoby, L. L. (2018). The importance of time to think back: The role of reminding in retroactive effects of memory. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, *44*(9), 1352-1364.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, 1, 19–40.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Riby, L., Smallwood, J. & Gunn, V. P. (2008). Mind wandering and retrieval from episodic memory: A pilot event-related potential study. *Psychological Reports, 102*, 805-818.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*(2), 234-242.

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*, *68*, 275-283.

Seli, P., Carriere, J. S., Levene, M., & Smilek, D. (2013). How few and far between? Examining the effects of probe rate on self-reported mind wandering. *Frontiers in Psychology*, *4*, 430.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. Dialogue: The Official Newsletter for the Society of Personality and Social Psychology, 26, 4-7.

Smallwood, J. M. (2011). Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistics Compass, 5*(2), 63-77.

Smallwood J. M. (2013). Distinguishing how from why the mind wanders: a process-occurrence framework for self generated thought. *Psychological Bulletin, 139,* 519-535.

Smallwood, J. M., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*, 452-484.

Smallwood, J. M., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*(6), 1144-1150.

Smallwood, J. M., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness & Cognition, 18*(1), 118-125.

Smallwood, J. M., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. Consciousness and *Cognition, 15*, 218-231.

Smallwood, J. M., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132,* 946-958.

Smallwood, J. M., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*, 487-518.

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

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, *23*(5), 551-559.

Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition, 27*, 14-26.

Thomson, D. R., Smilek, D., & Besner, D. (2014). On the asymmetric effects of mind-wandering on levels of processing at encoding and retrieval. *Psychonomic Bulletin & Review*, *21*(3), 728-733.

Wahlheim, C. N. (2014). Proactive effects of memory in older and younger adults: The role of change recollection. Memory & Cognition, *42*, 950-964.

Wahlheim, C. N. (2015). Testing can counteract proactive interference by integrating competing information. *Memory & Cognition, 43*, 27-38.

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory & Cognition, 41*, 1–15.

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 94-105.

Wahlheim, C. N., & Zacks, J. M. (2018). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General,* doi:10.1037/xge0000458.

Wammes, J. D., Seli, P., Cheyne, J. A., Boucher, P. O., & Smilek, D. (2016). Mind wandering during lectures II: Relation to academic performance. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 33–48.

Weinstein, Y. (2018). Mind wandering, how do I measure thee with probes? Let me count the ways. *Behavioral Research, 50*, 642-661.

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on-task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review*, *25*(2), 754-760.

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*(5), 681–695.

APPENDIX A

TABLES AND FIGURES

Table 1

Probability of Change Classifications as a Function of Item Type

Item Type		
A-B, A-B	C-D	A-B, A-D
.06 [.05, .07]	.05 [.04, .06]	.39 [.37, .41]

Note. Bootstrap 95% confidence intervals are displayed in brackets.

Table 2

Probability of Change Classification for A-B, A-D Items

Change Classification		
Changed + Recall Earlier Response	Changed + No Recall Earlier Response	Not Changed
.28 [.26, .30]	.12 [.11, .13]	.61 [.59, .63]

Note. Bootstrap 95% confidence intervals are displayed in brackets

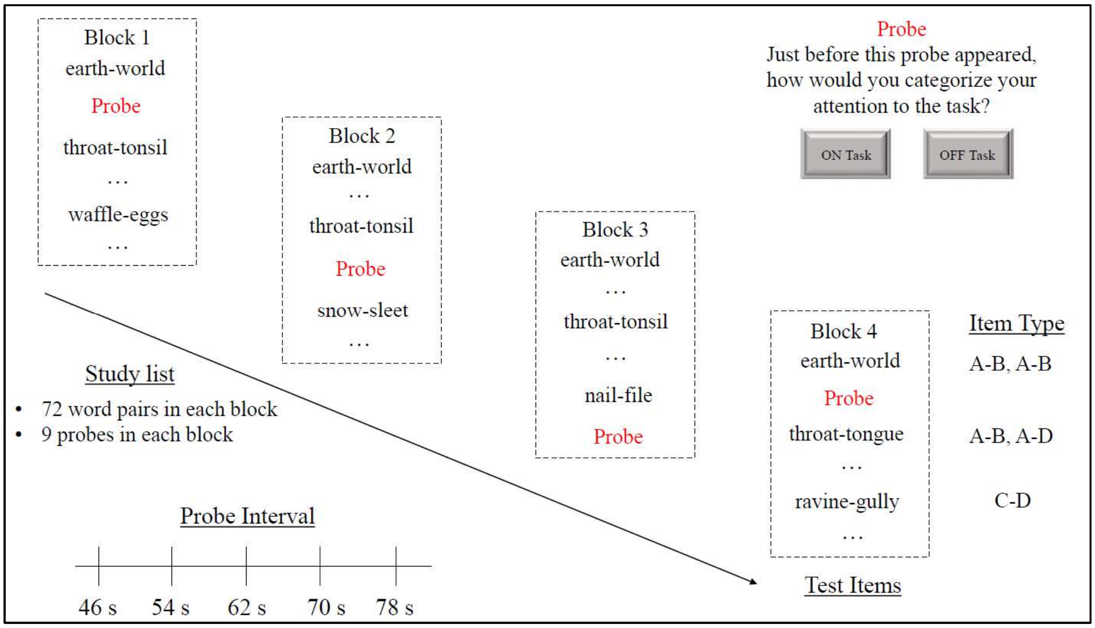

Figure 1. Experimental Design. Participants studied a list that contained four seamless blocks. Each block contained word pairs that repeated across each block (A-B, A-B), repeated in the first three blocks and then had the same cue with a changed response in the fourth block (A-B, A-D), or were new to each block (C-D). Thought probes were dispersed pseudo-randomly such that three probes came after each item type in each block, and the probes were varied to appear 6-10 word pairs apart from each other. The probe appeared immediately after the previous word pair, and asked participants to indicate if they were On-task or Off-task. After making their indication, participants studied the next pair. Participants were tested on all pairs presented in Block 4.

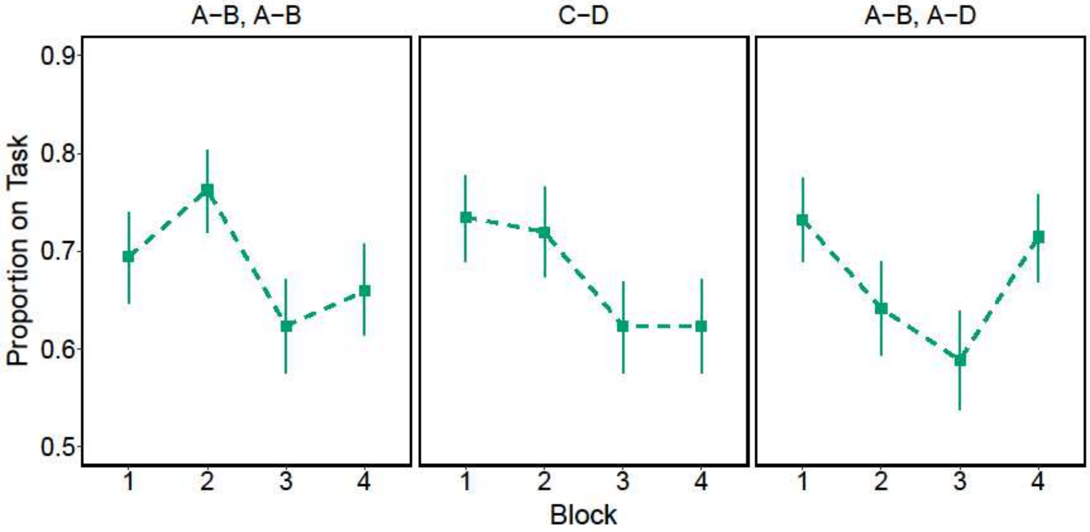

Figure 2. Proportion On-Task as a Function of Item Type and Block. Error bars are bootstrap 95% confidence intervals.

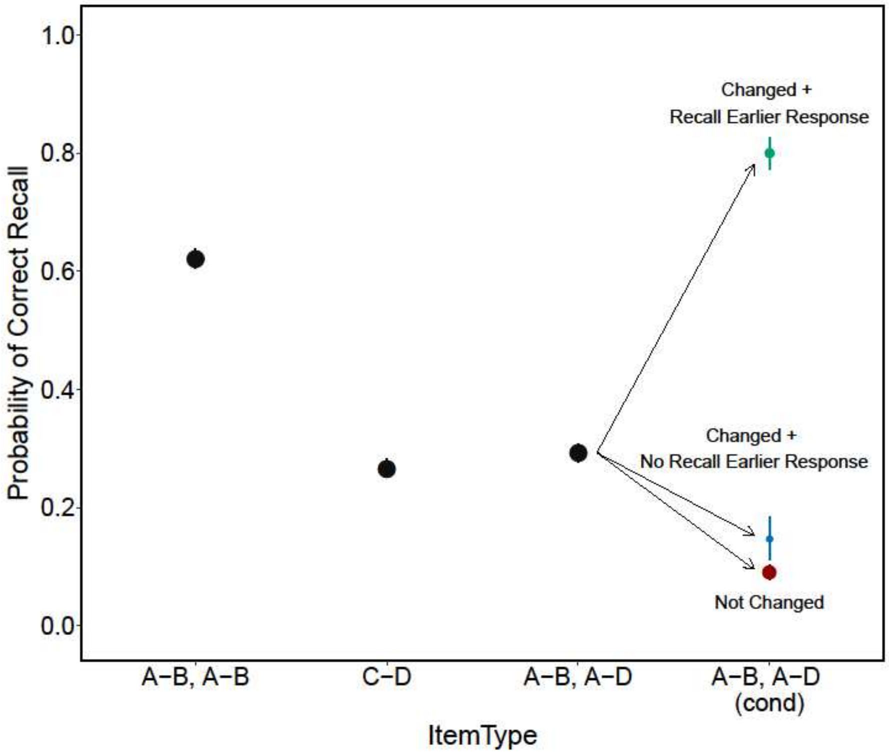

Figure 3. Correct Recall as a Function of Item Type. Black points represent overall correct recall for each item type. The green dot represents conditionalized correct recall for A-B, A-D items given that participants indicated change and were able to recall the earlier response (change recollection). The blue point represents conditionalized correct recall for A-B, A-D items given that participants indicated change and did not correctly recall the earlier response. The red point indicates conditionalized correct recall for A-B, A-D items given that participants did not indicate change. The size of the colored dots represents the proportion of all responses that fall into each of those cells. Error bars are bootstrap 95% confidence intervals.

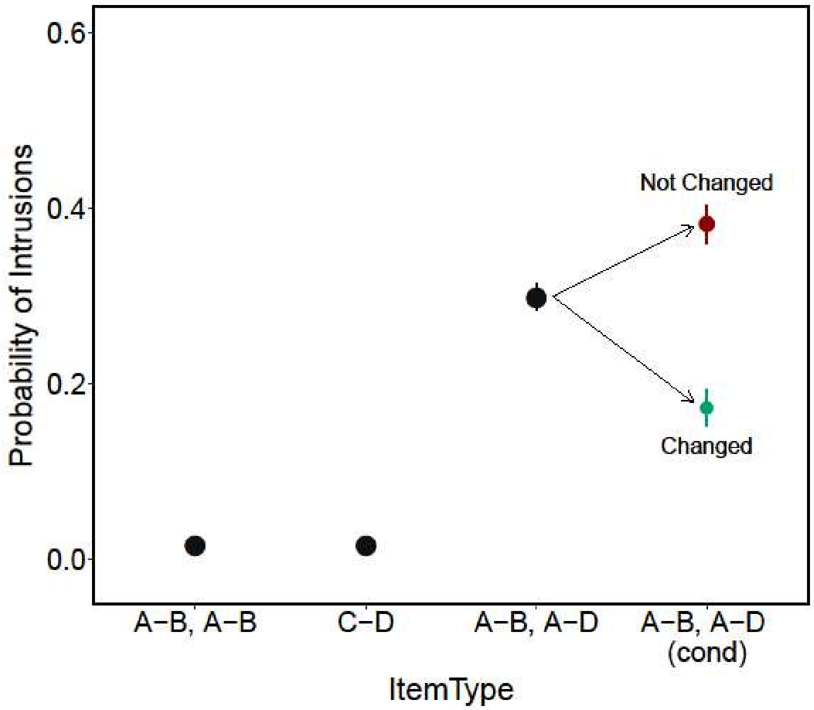

Figure 4. Intrusions as a Function of Item Type. The black points represent the intrusion rates for each item type. The green dot represents the conditionalized intrusion rate given that participants indicated that the word pair changed. The red dot represents the conditionalized intrusion rate given that participants did not indicate that the word pair changed. The size of the dot represents the proportion of responses that fall into each of those cells. Error bars are bootstrap 95% confidence intervals.

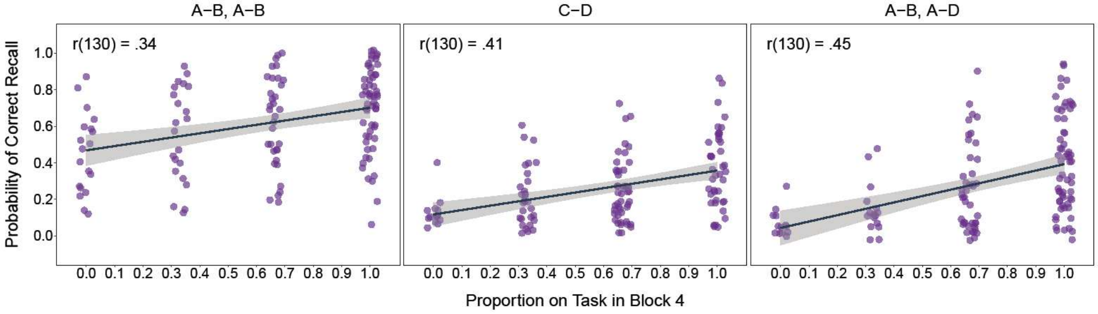

Figure 5. Between-Subjects Correlations Between Block 4 Proportion On-Task and Correct Recall. Given that the on-task proportion is from block 4 only, the on-task proportion was calculated based on 3 probes per participant for each item type. Shaded region shows the 95% confidence interval. The correlation coefficient and degrees of freedom are shown in the upper left-hand corner of each panel.

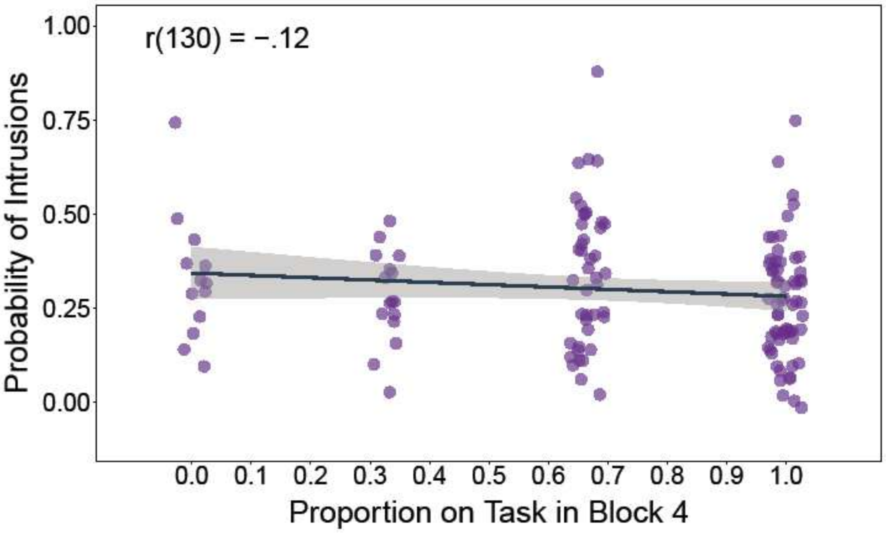

Figure 6. Between-Subjects Correlations Between Block 4 Proportion On-Task and Intrusions for A-B, A-D Items. Given that the on-task proportion is from block 4 only, the proportion on-task was calculated based on 3 probes per participant. Shaded region shows the 95% confidence interval.

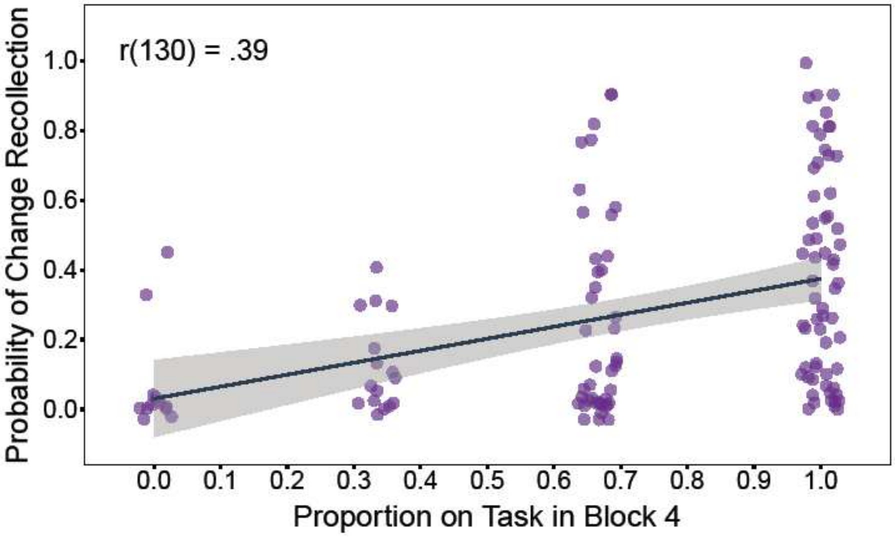

Figure 7. Between-Subjects Correlations Between Block 4 Proportion On-Task and Change Recollection for A-B, A-D Items. Given that the on-task proportion is from block 4 only, the proportion on-task was calculated based on 3 probes per participant. Shaded region shows the 95% confidence interval.

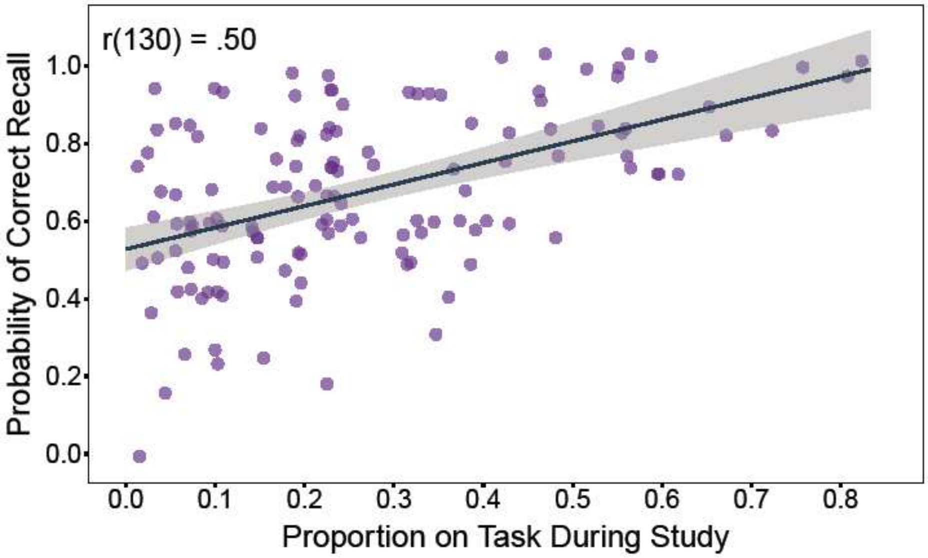

Figure 8. Between-Subjects Correlations Between Proportion On-Task During Study and Correct Recall for C-D Items. The proportion on-task was calculated based on all 36 probes. Shaded region shows the 95% confidence interval.

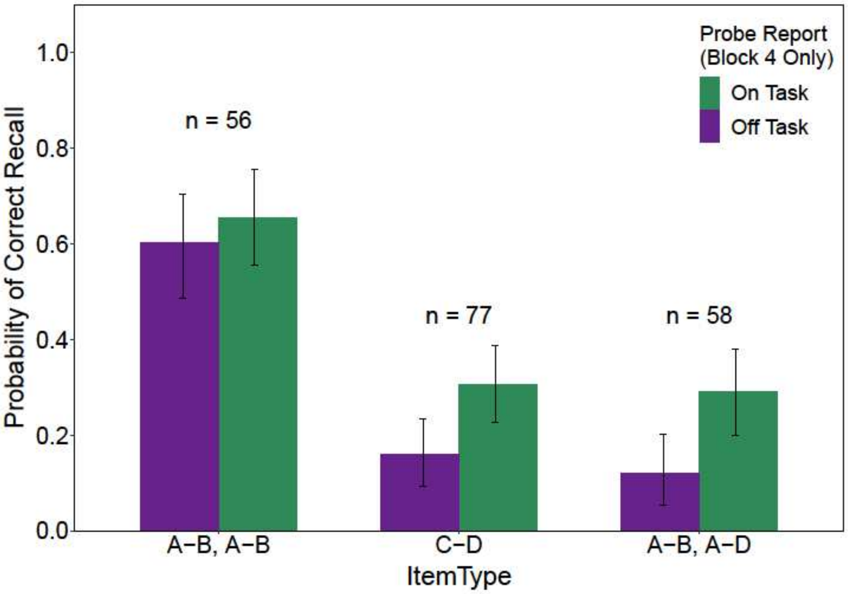

Figure 9. Probability of Correct Recall as a Function of Probe Reports in Block 4. The number of participants that contributed to each on and off-task comparison are labeled above the set of bars for each item type. Error bars are bootstrap 95% confidence intervals.

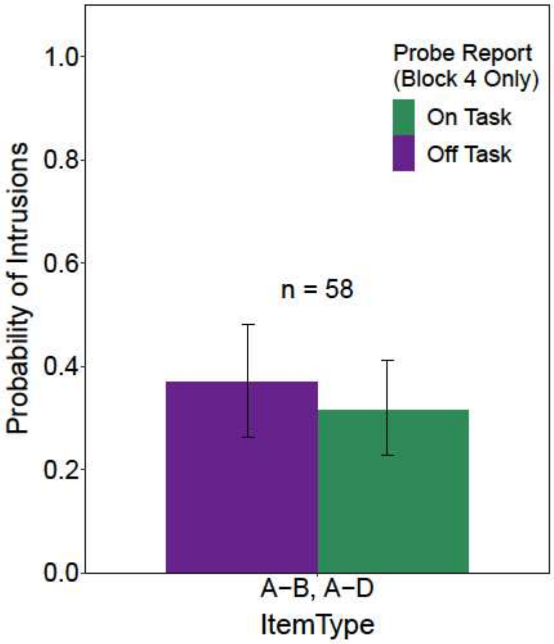

Figure 10. Probability of Intrusions as a Function of Probe Reports in Block 4 for A-B, A-D Items. The number of participants that contributed to the on and off-task comparison is listed above the bars. Error bars are bootstrap 95% confidence intervals.

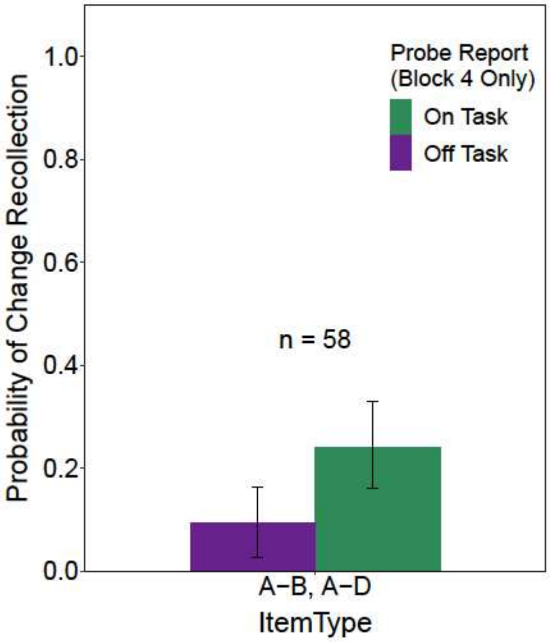

Figure 11. Probability of Change Recollection as a Function of Probe Reports in Block 4 for A-B, A-D Items. The number of participants that contributed to the on and off-task comparison is listed above the bars. Error bars are bootstrap 95% confidence intervals.

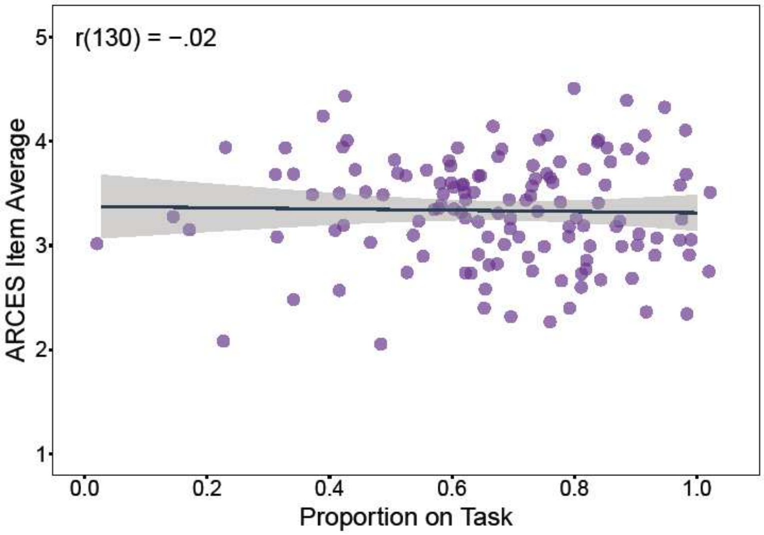

Figure 12. Between-Subjects Correlations Between ARCES and Proportion On-Task During Study. Shaded region shows the 95% confidence interval.

APPENDIX B

STUDY PHASE INSTRUCTIONS

In this part of the experiment, you will be asked to study word pairs for an upcoming test. Some of the word pairs will repeat, some will appear once, and some will change at a later point in the study phase. For changed word pairs, the left-hand member of a pair will be presented later with a different right-hand member (e.g., silly-clown; silly-giggle). Each pair will appear on the screen for 6 seconds. Please learn each pair as best as you can for a test that you will be given at the end of the experiment.

Do you have any questions?

Press the SPACE BAR for more instructions.

While you are studying the word pairs, you may notice that your ability to focus your attention on the task waxes and wanes throughout this period. It is normal for people to experience various levels of attentional engagement. We are interested in the extent to which you experience these variations in task engagement. Every now and then, we will ask you to indicate your current level of engagement during the upcoming study phase. To measure this, we will randomly present a screen that asks you to indicate whether you are on-task or off-task. If your attention just before the probe was firmly directed at learning the word pairs, then indicate that you are On-task. In contrast, if your attention was on something else other than studying the word pairs, then indicate that you are Off-task. You will indicate this by clicking the appropriately labeled button on the screen.

Do you have any questions?

Press the SPACE BAR when you are ready to begin studying.

APPENDIX C

TEST PHASE INSTRUCTIONS

Test Phase

In this part of the experiment, you will be tested on your memory for the word pairs that you studied. You will be presented with the left member of a word pair (e.g., silly - ?), and your task will be to type the word that it was most recently paired with during the study phase.

Do you have any questions?

Press the SPACE BAR for more instructions.

After you have made your response, you will be asked whether the right word changed during the study phase. The question, “Did the right word change during the study phase?” will appear in the middle of the screen with boxes labeled “Yes (1)” and “No (0)” displayed below.

If you think that the right word that was presented with the left-hand member of a pair changed in the study phase (e.g., silly-clown; silly-giggle), then press the "1" on the keyboard. When you indicate that a pair has changed (silly-giggle), you will next be asked to recall what the cue was paired with earlier in the study phase (clown). If you cannot remember the earlier pairing, then it is fine to either guess or pass. Please type your response into the box below and check your spelling carefully. In contrast, if you do not think that the right word of a pair changed during the study phase, then you should press the “0” on the keyboard.

Do you have any questions?

Press the SPACE BAR to start with some practice trials.

APPENDIX D

ANOVAS AND EFFECT SIZES

Attention to task. A two-way ANOVA was performed to compare the effects of Item Type and Block on the proportion of on-task reports. There was a significant effect of Block, $F(3, 4740) = 6.43$, $p < .001$, $\eta_p^2 = .004$, which was qualified by a significant Block × Item Type interaction, $F(6,4740) = 4.00$, $p < .001$, $\eta_p^2 = .005$. As shown in Figure 2, the proportion of on-task thoughts increased from Block 3 to Block 4 for A-B, A-D items but was not significantly different for A-B, A-B or C-D items. The effect of Item Type was not significant, $F(2, 4740) = .94$, $p = .39$, $\eta_p^2 < .001$.

Recall performance. A one-way ANOVA was performed to compare the effects of Item Type on recall performance. The effect of Item Type was significant, $F(2, 9501) = 582.79$, $p < .001$, $\eta_p^2 = .11$. As can be seen in Figure 3 (right panel), recall for A-B, A-B items was significantly higher than the other two item types, but recall did not differ between A-B, A-D and C-D items. A separate one-way ANOVA was performed to compare the effects of Item Type on intrusion rates. The effect of Item Type was significant, $F(2, 9501) = 1051.18$, $p < .001$, $\eta_p^2 = .18$. As shown in Figure 3 (right panel), intrusions were highest for A-B, A-D items, but did not differ between A-B, A-B and C-D items.

Change classifications. A one-way ANOVA was performed to compare the effects of Item Type on change classifications made during on the cued recall test. There was a significant effect of Item Type, $F(2, 9501) = 1048.1$, $p < .001$, $\eta_p^2 = .18$. As shown in Table 1, change classifications were highest for the A-B, A-D items, and change classifications did not differ between A-B, A-B and C-D items.

Recall performance conditionalized on change classifications. A one-way ANOVA was performed to compare the effects of the three levels of change classification outcomes at test on conditionalized correct recall. The levels include indicating change and reporting the earlier response (change recollection), indicating change without reporting the earlier response, and not indicating change. Recall performance on the C-D items was also included as a level in the ANOVA so that proactive effects of memory could be examined. There was a significant effect of Change Classification, $F(2, 9501) = 582.79$, $p < .001$, $\eta_p^2 = .25$. As shown in Figure 3 (left panel), recall performance was higher when participants recollected change compared to when they did not recollect change.

To compare the effect of change classifications on the rate of intrusions, a t-test was conducted that included whether participants indicated change at test or not. There was not a level to include indicating change and reporting the earlier response because doing so would negate producing it as an intrusion. The pairwise comparison showed that intrusions were significantly lower when participants indicated that there was a change at test compared to when they did not indicate change, $t(3061) = 13.56$, $p < .001$, $d = .48$.

Recall performance conditionalized on attention to task. Since there were an unequal number of participants that contributed to recall performance for each Item Type based on the task reports made in Block 4, I fit separate one-way ANOVAs to each Item Type to examine the effects of Task Report on correct recall. The ANOVA for A-B, A-B items showed no significant effect of Task Report, $F(1, 110) = .01$, $p = .92$, $\eta_p^2 < .001$. The ANOVA for C-D items showed a significant effect of Task Report, $F(1, 152) = 5.47$, $p = .02$, $\eta_p^2 = .03$. The ANOVA for A-B, A-D items also showed a significant effect of Task Report, $F(1, 114) = 5.52$, $p = .02$, $\eta_p^2 = .05$.