2022

Divided Attention and its Effect on Forward Testing

Nicholas H. Garcia
University of North Florida, n00898291@unf.edu

Divided Attention and its Effect on Forward Testing

Nicholas Garcia

July 25, 2022

Abstract

The testing effect is a well-studied and robust phenomenon. The forward testing effect is a relatively new phenomenon that has been observed in robust settings with a diverse population. The testing effect (also coined the backwards testing effect) and the forward testing effect share similar benefits and are applicable in similar settings. Research on the forward testing effect has demonstrated underlying mechanisms that differ from the backwards testing effect, illuminating the differences between these two phenomena. Dividing attention during study periods has been revealed to negatively affect the backwards testing effect, significantly reducing its efficacy. The forward testing effect, operating under different underlying mechanisms, may be less susceptible to the negative impact of dividing attention. The current experiment consisted of 135 participants divided into one of four conditions; divided attention/test, divided attention/restudy, full attention/test, and full attention restudy. I hypothesized that dividing attention during encoding will not have the same detrimental effects on the forward testing effect as it does on the backwards testing effect. I hypothesized a main effect of attention, such that the full attention condition will have higher recall on a criterial test. I hypothesized a main effect of testing, such that those in the test condition will have higher recall on a criterial test. A factorial ANOVA analysis revealed a significant forward testing effect. The results of the experiment help to elaborate a resource theory explanation for the FTE. Limitations to be considered include online participants and the inclusion of only an auditory attention manipulation.

Introduction

Testing has been identified to be an integral part of the learning process (Roediger and Karpicke, 2006). Within the past few decades, testing has been associated with robust benefits, including reduced proactive interference (interference that occurs due to prior learning) and increased memorization of studied materials. Testing is utilized most often in the classroom, though testing has benefits to those outside of the classroom (i.e., traumatic brain injury patients in clinical settings).

With testing being a catalyst for more efficient learning, it is essential that the extent of testing and its numerous benefits be investigated. Recently, a new phenomenon within testing has been observed, the forward testing effect (FTE). The FTE refers to a phenomenon in which individuals that utilize testing perform better than their peers on *new* information presented after the tested material (Pastötter and Bäuml, 2014). This FTE appears to be as robust as the testing effect that typically takes place in the classroom. Understanding the FTE, its underlying mechanisms, and how it interacts with other cognitive processes is essential for creating an environment that is optimal for learners. Distractions, such as technological distractions, other individuals, and multi-tasking, often occur in educational settings. The cognitive toll that divided attention takes on individuals utilizing forward testing is important to investigate and understand to utilize the FTE to its fullest potential. Whether the FTE is sustained when faced with divided attention will determine how valuable the FTE is to learners in environments rife with distractions.

The Testing Effect

The testing effect refers to interim testing of learned material leading to greater retention of studied material (Roediger and Karpicke, 2006). The testing effect has been found to be a

robust phenomenon applicable in various settings and with diverse populations. Settings in which the testing effect has been observed include clinical settings in which subjects have experienced traumatic brain injuries (Pastötter et al., 2013) and laboratory settings. Testing effect experiments typically involve two conditions: a tested condition and an untested condition (Rowland, 2014). Participants in the test condition perform interim tests of previously studied material, while other participants either restudy or do not perform any task. Participants in the restudy condition are given the opportunity to study the material a second time for the same duration as the initial study phase. The testing effect has been observed in multiple different test formats, including: word pairs, word lists, prose passages, and nonverbal materials (Rowland, 2014). Within experimental settings, it is typical that those subjects within the test condition (i.e., taking interim tests in between learning segments) outperform subjects in the restudy or no task condition. An additional benefit of the testing effect is the reduction of proactive interference (Pierce and Hawthorne, 2015). Proactive interference refers to the interference of old information in the subsequent learning of new information (Szpunar, et al., 2008).

The Forward Testing Effect

The aforementioned testing effect has also been coined as the backwards testing effect (Yang et al., 2019) in order to differentiate between learning of previously studied materials and learning of new information. The term “backwards” refers to the participants' enhanced retrieval capabilities of already studied materials due to the intermittent testing of material. The FTE (also referred to as: test-potentiated learning and test-potentiated new learning) refers to the finding that participants that partake in intermittent testing of material typically have higher levels of memorization when tested on new material (Chan et al., 2018; Yang et al., 2018; Pastötter and Bäuml, 2014). For example, Szpunar et al., (2008) had participants study five word lists. Half of

the participants received free-recall tests after each new list, while the other half only completed a free-recall test after list 5. All participants performed a filler task after the encoding phase. The critical comparison between the two groups was the performance on list 5. Participants in the test condition had fewer prior list intrusions (i.e., recalling items from lists previously studied that are not from the currently tested material), showing a reduction in proactive interference (i.e., interference with the current test that occurs due to prior learning), and performed better on the critical list 5 test (see Figure 1). In a similar study, Aslan and Bäuml (2016) conducted an experiment where a condition of age group was included. Participants were placed in one of three categories: younger children ($M = 6.7$), older children ($M = 8.8$), and adults ($M = 23.3$). The older children and adults performed as expected and had higher recall than their restudy peers, though the younger children performed roughly the same as their restudy peers. This suggests that the backward testing effect differs from the FTE in their respective underlying mechanics, as the backwards testing effect has been demonstrated in children as young as 3 to 5 years old (Aslan and Bäuml, 2016).

The FTE, similarly to the backwards testing effect, has been demonstrated to be a robust phenomenon, applicable to diverse populations (i.e., traumatic brain injury patients, older children, and adults) and amongst different test formats, including free-recall, prose passages, word lists, and art identification (Szpunar et al., 2008). Forward testing is a particularly impressive finding as the new information does not necessarily need to be related to the previously tested material (Pastötter and Bäuml, 2014; Divis and Benjamin, 2014). The FTE, similar to the backwards testing effect, has been shown to reduce the amount of proactive interference from previous lists (Yang et al., 2018). For example, in the Szpunar et al., (2008) experiment the researchers results demonstrated that participants in the interim test group

recalled a greater number of list 5 words than the no interim test group and also suffered less proactive interference.

Forward testing has also been demonstrated to impact the way in which learners interact with studied material. Szpunar et al., (2014) demonstrated the FTE and its effect on how participants engage with material they are trying to learn. Reduced mind wandering and increased note taking were seen in participants that utilized forward testing. However, interpolated testing did not have as much of an impact on high-school students as it did with college-students.

Underlying Mechanisms

Currently, the FTE has no unifying theories that give a complete account of the underlying mechanisms involved (Chan et al., 2018). A multitude of theories have been proposed that account for the many aspects of the FTE (Pastötter and Bäuml, 2014; Szpunar et al., 2013; Wissman et al., 2011; Chan et al., 2018; Yang et al., 2018). Similar to the backwards testing effect, these theories are not mutually exclusive (Chan et al., 2018). The varying underlying mechanisms for the FTE have been compiled and arranged into four distinct categories. The categories are as follows: resource theories, metacognitive theories, context theories, and integration theories (Chan et al., 2018).

Resource theories consist of theoretical frameworks in which testing increases the cognitive resources available for encoding new information (Chan et al., 2018). Explanations within this category operate under the premise that a reduction of proactive interference occurs during the encoding of new materials and increases attentional resources available for postretrieval encoding operations. The reduction of proactive interference has been demonstrated in many studies (Aslan and Bäuml, 2016; Yang et al., 2019; Nunes and Weinstein,

2012). Pastötter et al., (2011) demonstrated this reduction in proactive interference using electroencephalogram (EEG) recordings of brain activity to measure alpha power, which is hypothesized to indicate failure of attention, across encoding phases of multiple lists (Pastötter et al., 2011). Alpha power was not shown to increase within a testing condition, suggesting that testing can help immunize against proactive interference. This finding is contrasted with the finding that alpha power increased across lists within the comparison condition of restudy. Pastötter and Bäuml (2014) conclude that the finding of lower levels of alpha power suggest that testing creates a reset within the encoding process, effectively making encoding of later lists as effective as the encoding of the previously studied list.

In addition to the cognitive resources renewed through the reduction of proactive interference and the encoding "reset" brought about by testing, Szpunar et al., (2013) suggest interim testing can reduce mind-wandering. In an experiment conducted by Szpunar et al., (2013), students were asked to listen to a video lecture and were tasked with learning the contents of the lecture. Students were informed that the lecture would be divided into four sections with a break between each section. Half of the students were placed in a test condition and the other half were placed in a restudy condition. Students in the test condition reported fewer instances of mind-wandering during the lecture than their peers in the restudy condition. In contrast, Jing et al., (2016) found similar rates of mind wandering between those in the tested and untested condition.

Metacognitive theories encompass theoretical frameworks that suggest that testing enhances later learning due to learners optimizing their encoding strategies, which is mediated by metacognitive knowledge regarding participants' own ability (Chan, et al., 2018). For example, Szpunar et al., (2014) measured participants' judgement of learning while viewing a video

lecture. Participants either took interpolated tests during each section, took a test following the fourth section, or were not tested. Participants were asked to predict how well they would perform on a final cumulative test. Participants were overconfident in their assessments of their ability to perform well on the cumulative test. However, those in the interpolated tests condition were able to bridge the gap and perform closer to their assessments than their peers. The FTE can also enhance students' self-paced learning. Yang et al., (2017) conducted an experiment wherein the participants' encoding phase was self-paced, as opposed to experimenter-paced. Participants were given five study lists of Euskara-English word pairs in anticipation of a cumulative test. Participants were placed in either an interim test condition or a no interim test condition. Those in the no test condition decreased their encoding time across lists, while those in the test condition did not decrease their encoding time across the lists. The results of this experiment suggest that interim testing can help maintain the motivation of individuals across study lists.

The role of test expectancy also provides a viable theoretical framework within the category of metacognitive theories. Weinstein et al., (2014) investigated the role of test expectancy as it related to groups who were either untested and warned about an upcoming test, untested and not warned, tested and warned, or tested and not warned. Participants in the untested/warned group performed better than those in the untested/unwarned group. This suggests that those who expected a test applied better encoding strategies than those who had no warning. Those participants that were warned and/or tested may focus their attention on learning as opposed to their unwarned counterparts. In addition to the direct benefit that test expectancy grants, interpolating tests within material appears to encourage task-relevant activities, such as note taking (Szpunar et al., 2013) and the aforementioned increase in encoding time within learners self-allocating study time (Yang et al., 2017).

Context theories involve theoretical frameworks suggesting that testing enhances new learning via isolating prior learning episodes from new learning episodes (Chan et al., 2018). The basis for these theories is that individuals encode information and this content is stored with the study context. When individuals participate in retrieval of said information, the contextual information that is accessible to the learner impacts the likelihood of retrieving the target information (Chan et al., 2018). For example, Szpunar et al., (2008) discusses the effect of testing on proactive interference and suggests that this reduction of proactive interference occurs due to participants segregation of tested materials. When materials are tested, a reduction of cue overload occurs during retrieval. That is, when the participants were tested, list segregation occurred, reducing the functional search set (i.e., the material retained in the individuals' memory used to recall the answer) to the recently presented words. When restudy occurs, as opposed to testing, a larger functional search set is available, due to cue overload, resulting in proactive interference from previously studied materials

Similarities occur between context theories and resource theories. For example, Yang et al., (2018) describe a finding from Pastötter et al., (2011) proposing that interim testing creates an internal context change. This internal context change is what induces the resetting of encoding. This explanation can be found within the domain of resource theories; however, the difference is found in how this isolation of contexts influences new learning (Chan et al., 2018). According to Chan et al., (2018), resource theories involving isolation increase attentional resources, while isolation within context theories creates an advantage when participants attempt to retrieve the new-learning items during their critical test.

Integration theories function under the premise that testing enhances integration of the new-learning materials either with the retrieval cue or with the original-learning material (Chan

et al., 2018). Theories within the theoretical framework suggest that attempting to answer questions causes an elaborative memory search, which can help in the integration and retention of the question (Chan et al., 2018). The answer to the question then becomes integrated with the question itself, facilitating the learning of the answer. Integration theories also help to explain the presence of an FTE when the material is new, though interconnected with old material. According to integration theories, when new information builds upon earlier information, testing can facilitate in integrating the new information with old information. For example, when learning the parts of a cell and being tested on this material, new information, such as the functions of the cell, may be integrated through testing, as the new information builds upon earlier information. This integration helps to facilitate the learning of new information.

Detriments

The FTE is associated with a wide-range of benefits, including enhanced task-related abilities, such as note taking (Szpunar et al., 2013), a reduction of proactive interference (Pastötter et al., 2011), a reduction of mind-wandering (Szpunar et al., 2013), an increase in self-allocated study time (Yang et al., 2018), and the enhancement of inductive learning, which consists of learning from specific examples that can be generalized to different, but related materials (Yang and Shanks, 2017). However, there are also detriments that are associated with the phenomenon.

Switching between encoding and retrieval tasks appears to have a detrimental effect on new learning (Davis et al., 2017). It has been observed that when encoding and retrieval task frequency are manipulated, higher levels of task-switching result in a greater impairment of new learning. Further, Finn and Roediger (2013), observed that when individuals practice retrieval and immediately learn new information, the learning of the new information is impaired. In their

study, complementary information (a profession associated with a name and face) was to be learned either after a retrieval or a restudy phase. Those in the retrieval phase experienced impairment in their learning of the complementary information. Regardless of whether feedback was provided or not and whether or not there was an immediate final test or a 24-hour delayed final test, new learning was impaired. According to Davis and Chan (2015), this appears to occur due to participants focusing on encoding of old information (names and faces) as opposed to the new information (profession). This suggests that participants' metacognitive judgments about their inability to learn the old information or its perceived importance prioritized encoding of this information over the new information.

Testing has also been shown to increase susceptibility to misinformation (Pastötter and Bäuml, 2014). For example, Chan et al., (2009) gave participants a video to watch, afterwards participants either took a test on the details of the video or completed an unrelated task. Following the test or task, an audio clip containing incorrect details of the video was played. All participants took a final test regarding the original details of the video and those in the test condition had enhanced recollection of the misinformation. This result is to be expected with the FTE and the current plausible underlying mechanisms that enable it. If an encoding reset takes place after testing, misinformation that is presented after a testing period would receive the same attentional resources as the original encoding of the material. Due to the reduction of proactive interference granted from the FTE (Yang et al., 2018), participants should be more likely to recall the most recent information, even if it contradicts earlier learning.

Testing and Divided Attention

With the technological state of the world, online learning is becoming more prevalent amongst learners. In 2018, approximately 35 percent of students attending degree-granting

postsecondary institutions enrolled in some form of distance learning (nces.ed.gov). With the increasing numbers of learners that are learning online, the opportunity for multitasking while trying to learn is ubiquitous. Learners that are more prone to divided attention and multitasking may suffer consequences of decreased encoding of materials during study sessions. Not only do students have more opportunities for distractions online, mind-wandering is also a common occurrence in and outside of the classroom (Jing et al., 2016). To understand how individuals learn most effectively, understanding of divided attention must be a priority.

The attentional state of the learner has consequences on the encoding of material (Mulligan and Picklesimer, 2015). Experiments in which participants have their attention manipulated in such a way that there is a divided attention condition, in comparison to a full attention condition, reveal that divided attention substantially reduces memory performance when attention is divided during the encoding phase of learning (Mulligan and Picklesimer, 2015; Anderson et al., 1998; Buchin and Mulligan, 2017). Mulligan and Picklesimer found that divided attention greatly impairs encoding, while retrieval is minimally impacted. They also found an increase in the backwards testing effect under divided attention, when dividing attention during a retrieval phase, in relation to those in a full attention condition. However, the secondary task was more disrupted by the retrieval than restudy condition. This finding suggests that the encoding effects of retrieval (the effect that retrieving material has on the retention of said material) are not as susceptible to dividing attention as typical encoding. To test this, they had participants learn 60 word pairs, with participants divided into one of four conditions; full attention/restudy, full attention/test, divided attention/restudy, divided attention/test. Attention was divided with a secondary task of participants hearing a sequence of digits playing over headphones and categorizing each digit as even or odd as quickly and accurately as possible.

Participants then took a final cued-recall test of the word pairs. The results of those in the full attention/restudy condition were higher than those in the divided attention/restudy condition, demonstrating the detrimental effects of dividing attention on encoding processes. However, dividing attention had a very modest effect on those in the divided attention/test condition. This resulted in a significantly larger testing effect in the divided attention condition compared to the full attention condition.

While the degradation of memory when attention is manipulated has been demonstrated in the lab (Buchin and Mulligan, 2017), under certain conditions this does not appear to be the case. MacDonald and MacLeod (1998) conducted an experiment which suggests that subtle attentional manipulations impact performance on direct tests of memory (i.e., recognition tests), though indirect tests of memory (i.e., rapid reading tests in which studied words should have a lower latency than new words) had no such impact. It should be noted that with further attentional manipulation, indirect tests of memory also suffer an impairment. This may be the case due to the elaborative account of the testing effect. The elaborative account suggests that the testing effect occurs because retrieval enhances the elaborative processing of the cue and to-be-retrieved target (Buchin and Mulligan, 2017). Due to retrieval enhancing the elaborative process, subtle attentional manipulations may impair the elaborative processes that are happening within direct tests.

The manipulation of attention has not been investigated within the FTE. It appears that different mechanisms underlie the backwards testing effect compared to the FTE. One possible explanation of the FTE is the aforementioned encoding reset due to retrieval (Pastötter et al., 2011). With dividing attention having a negative impact on encoding processes (Mulligan and Picklesimer, 2015), dividing attention during encoding may result in attentional

resources typically used in encoding being preoccupied even with the reset of encoding after testing. This could result in a diminished FTE, compared to those in a full attention/tested condition. Though it could also be argued that the reset of encoding could still preserve an FTE. Compared to a restudy condition, attentional resources would be reset in a tested condition, allowing for less detrimental effects of divided attention on those in a tested condition than the restudy condition. Though recall on a criterial test may be lower in a divided attention condition than a full attention condition, the FTE may appear to be stronger, as the gap in recall between divided attention/restudy and divided attention/test conditions enlarge. Considering the FTE within a context theory also lends support to an FTE resiliency towards divided attention. Context theories suggest that the FTE appears due to a segregation of information into different contexts. Testing helps create unique contexts for the learner such that information is separated and the search set for answers is diminished. Context theories operate under the assumption that retrieval practice is what creates a unique context for information. Because retrieval, and not encoding, creates a specific context, dividing attention during an encoding period may not impact the FTE, as retrieval is not being disrupted.

As attentional manipulation has never been studied in relation to the FTE, this experiment seeks to understand the extent of the detrimental effects, if any are present, that dividing attention presents to the FTE. Should the FTE present itself as resilient to dividing attention, the implications are far-reaching. Students in the classroom often participate in classes that utilize interim tests during the lesson or clicker questions throughout the lecture. Should the FTE be resilient to the effects of divided attention, educators will be able to use these methods effectively to enhance their students' learning.

In the current experiment, attention was divided during the encoding period, as opposed to the retrieval period within the Mulligan and Picklesimer (2015) study. The decision to divide attention during encoding comes from the different underlying mechanisms present in the FTE (i.e., resource theories and context theories). Attention was manipulated auditorily during the encoding process to those participants in the divided attention condition. I hypothesize that there will be a main effect of attention, such that dividing attention will reduce final list recall. I hypothesize that there will be a main effect of testing, such that individuals in the testing condition will have greater final list recall than individuals in a restudy condition. I also hypothesize that a lack of interaction between attention and testing will allow for an FTE to be seen within a divided attention/tested condition, as prior literature suggests that encoding processes are reinvigorated after testing.

Method

Participants

One hundred and fifty-seven participants were gathered using Mechanical Turk for this experiment. Participants were compensated $2.50 for their participation in the experiment. The sample size for this experiment was determined running a power analysis within G*power. With a power of .80 and an alpha error probability of .05, 32 participants per group would be required to detect an effect size of .50 (Cohen's *d*). Failure to follow instructions, as evidenced by a lack of responses or attempts to respond in the divided attention condition, resulted in exclusion from the experiment. All Mechanical Turk participants had a 95% hit rate or better. 23 participants were removed from the experiment due to failure to follow instructions for the divided attention task, resulting in the inclusion of 135 participants. It took participants approximately 30 minutes to complete the experiment. Due to using Mechanical Turk workers as participants, the sample

consisted of adults aged 18 and up. Because the experiment was not restricted by any other demographic information, the results are generalizable to adults in early adulthood through old age.

Design

The experiment used a 2 (Test vs. Restudy) X 2 (Divided Attention vs. Full Attention) between-subjects design. The dependent variable being measured was final test performance measured as a proportion of correctly recalled list 4 items.

Materials

A total of 4 lists with 20 weakly related word pairs per list were used during the encoding phase of the experiment for each condition, totaling four word pair lists. The fourth list of weakly related word pairs was used as a final test to measure the forward testing effect across conditions. The lists of word pairs were derived from Davis and Peterson (2021) who used the Kornell et al., (2009) set. All word pairs (e.g., *Achieve - Gain)* have weak forward and backward associations (i.e., their associations are less than .1), according to Nelson et al., (2004) norms. The word pair lists were counterbalanced. Lists were tested with cued recall tests using the word stems (e.g., *Achieve -?*). In between encoding and test/restudy, participants completed a filler task consisting of word puzzles.

Attention was divided utilizing voice recordings of a male and female voice reciting numbers that were randomly generated. The pacing of the recording had one digit announced every 2.5 seconds, with 25 digits per list. The range of the numbers that were recited was 1-9 and the order of male and female voices was randomized.

Procedure

All materials were presented in Qualtrics survey software. Participants in all conditions were told that they will study a series of word lists and be randomly tested after each encoding phase, with a final test being given after the fourth list. In reality, participants were grouped so they were either in a test or a restudy condition during the initial three lists. Participants in all conditions were shown a single word pair (e.g., *Achieve* - *Gain*) on a computer screen for two-and-a-half seconds. This continued until all 20 word pairs had been shown to the participant. During this encoding phase, participants in the divided attention condition also heard alternating male and female voices reciting a number ranging from 1-9. Participants were given instruction to type in all odd numbers recited by either the male or female voice in a response box beneath the encoding screen. The voice that participants responded to was counterbalanced. Participants in the divided attention condition received feedback on their performance of the secondary task. Following each section of the secondary task, participants were shown their answers, as well as the correct answer. After the encoding phase, participants in all conditions performed a filler task consisting of word puzzles; for example, the word "physiographical" was presented on the computer screen and participants tried to construct as many words from the word on screen as possible within 40 seconds. Participants in the full attention condition received an extra 10 seconds on the word puzzle to make up for the feedback time given to the divided attention condition. Following the filler task, those participants in the test condition were given a cued recall test of the word pairs studied during the encoding phase, such that the cue was provided for the participant, and they had to determine the correct word target. Participants in the restudy condition had the word pairs appear to them a second time for the same duration as the initial encoding period. This procedure continued for a total of three phases. Participants in all

conditions then studied the fourth list, completed the filler task, and were given a cued recall test on list four (see Figure 2).

IRB Approval

The current experiment was submitted and approved by an IRB. The experiment was pre-registered to ensure that hypotheses were not manipulated after the data was analyzed.

Results

In the current experiment, a 2 (Test versus Restudy) × 2 (Full Attention versus Divided Attention) between-subjects factorial ANOVA was ran to measure the effect of testing and attention on list four accuracy (i.e., new learning). A 2 × 2 between-subjects factorial ANOVA was chosen as the appropriate statistical analysis to analyze the interactions and main effects of the current experiment for significance (Allen, 2017). This test was chosen as participants were partitioned into one of four conditions, either a test or a restudy condition paired with a full attention or divided attention condition, creating two between-subjects conditions. The dependent variable being measured was the proportion of list four items that were correctly recalled, which is used to measure the forward testing effect. The assumptions of a factorial ANOVA include interval or ratio data for the dependent variable, no multicollinearity, homoscedasticity, and normality of the data. The assumption of interval or ratio data for the dependent variable has been met, as the dependent variable uses ratio data. The assumption of no multicollinearity has been met as list four accuracy is the only variable being measured. The assumption of homoscedasticity has been violated. This is demonstrated with a significant result of Levene's Test of Equality of Error Variances $F(4, 132) = 7.14$, $p < .001$. The assumption of normality of data has been violated, as demonstrated by a positive skew in the data. While the assumptions of homoscedasticity and normality have been violated, the factorial ANOVA test

has been demonstrated to be a robust test that holds validity when assumptions have been violated (Ziegler et al., 2010). Due to the robustness of the factorial ANOVA test, the data collected will be analyzed using the uncorrected data set.

Descriptive statistics for all conditions are provided in Table 1. As shown in Figure 3, the 2 × 2 factorial ANOVA revealed a significant main effect in the test versus restudy condition, such that participants who utilized testing (M = .32) had higher accuracy on list four materials than individuals who utilized restudy (M = .12), $F(1, 131) = 20.02$, $p < .001$. As shown in Figure 3, a significant main effect was found in the full attention versus divided attention condition, such that participants in the full attention condition had higher accuracy (M = .29) on list four materials than individuals in the divided attention condition (M = .16), $F(1, 131) = 9.36$, $p < .001$. A significant interaction was not found between testing conditions and attention conditions $F(1, 131) = .10$, $p = .75$.

An independent samples t-test was used to compare the mean list four recall between the test and restudy groups within both the divided attention and full attention condition. When attention was divided, a significant effect of testing was found $t(68) = -3.46$, $SE = .054$, $p = .001$. Similarly, when attention was not divided, a significant effect of testing was found $t(63) = -2.95$, $SE = .073$, $p = .004$.

A 2 × 2 between-subjects factorial ANOVA was also used to measure the effect of testing and attention on extralist intrusions (i.e., proactive interference) within list four. The dependent variable being measured was the proportion of extralist intrusions that participants recalled. No significant main effect was found in the test versus restudy condition, indicating there was no difference in the amount of proactive interference between test and restudy groups, $F(1, 131) = 3.47$, $p = .07$. Similarly, no significant main effect was found in the divided attention versus full

attention condition $F(1, 131) = 1.73$, $p = .19$, indicating there was no difference in the amount of proactive interference between attention conditions. A significant interaction was not found between testing conditions and attention conditions $F(1, 131) = .12$, $p = .73$.

Discussion

In the current experiment, I demonstrated that dividing attention during an encoding phase does not eliminate a forward testing effect. Importantly, there are four findings within this experiment that are worth mentioning. The first finding is the main effect that testing had on new learning, as evidenced by higher recall on list four. The results of this experiment replicate previous research findings (Pastötter and Bäuml, 2014; Szpunar et al., 2013; Wissman et al., 2011; Yang et al., 2018) in the forward testing effect literature; that is, participants that utilized a testing paradigm, in comparison to restudy, had higher list four recall than individuals in the restudy condition. The second finding in this experiment worth mentioning is the main effect of attention. Specifically, individuals in a divided attention condition performed worse than individuals in a full attention condition on the criterial test. This finding, while it has not been investigated within a forward testing effect protocol, is expected, as the attention and testing literature have reported the detrimental effects of divided attention on encoding processes (Mulligan and Picklesimer, 2016).

The lack of an interaction between attention and testing is also notable. The results of the experiment support the hypothesis that testing will still confer benefits to participants regardless of their membership in either attention condition. Although participants in the full attention condition had higher recall on list four items, participants in the testing condition had significantly higher recall than participants in the restudy condition on the list four test. These

results are noteworthy as they contribute to the theoretical and practical understanding of the forward testing effect.

Surprisingly, there was no significant effect of attention or testing on the amount of proactive interference within the different groups. This finding is contrary to previous findings in the forward testing effect literature (Aslan and Bäuml, 2016; Yang et al., 2019; Nunes and Weinstein, 2012). It should be noted, however, that although the finding was not significant, the ANOVA analysis reported a finding that approached significance (p =.07) within the testing condition. One possible explanation for the lack of reduction in proactive interference is the material used for the study phases. The word lists used were not designed in a way to test for proactive interference. For example, one word pair that may be presented in a word list would be *Skyscraper - Tower*, while a separate list may have the word pair *Achieve – Gain*. The word pairs themselves have weak forward and backward associations, according to Nelson et al., (2004) norms, and the word lists were not designed to create strong associations between interlist word pairs. This explanation is further supported when examining the mean values of the proactive interference within each group. As seen in Figure 4, the mean values of proactive interference for all four conditions are marginal. Proactive interference, as evidenced by extralist intrusions, was not substantial within any condition, suggesting that the materials used during the encoding phase did not lend themselves to testing for proactive interference.

Theoretical Interpretation

The results of this experiment contribute to the forward testing effect literature in two important ways. First, the results of this experiment contribute to our understanding of the underlying mechanisms behind the forward testing effect. The results of the current experiment contribute to an explanation from resource theories of the FTE. Resource theories state that the

FTE occurs due to an increase in cognitive and attentional resources available to individuals utilizing testing via an "encoding reset", which resets the available cognitive and attentional resources available to individuals following the testing phase (Pastötter et al., 2012). Participants in the test condition demonstrated an FTE, although there was a significant difference in list four recall between participants in the divided attention and full attention conditions. The persistence of an FTE may result from the benefits conferred from the encoding reset that resulted from intermittent testing. These benefits were demonstrated regardless of the attention condition, i.e., the lack of an interaction between attention and testing. However, due to the additional cognitive and attentional resources utilized during the divided attention task, the available cognitive and attentional resources available to the divided attention condition participants may have substantially decreased when encoding list four items. This explanation would account for the difference in list four recall between attention conditions, while still accounting for an FTE demonstrated in the testing condition.

While an explanation from the resource theory family of ideas can account for the present results, context theories also provide a viable option. According to context theories, when items are recalled, a context change occurs, allowing different study and test phases to become isolated. A context change results from a decrease in proactive interference, enabling the separation of study sessions into distinct contexts (Chan et al., 2018). Assuming a context theory, an FTE would be preserved regardless of the attention condition due to the context change that is initiated when testing is utilized. The context of word pairs from list four would be separated from previous word lists, which would reduce proactive interference and result in better encoding of list four items. The difference in recall between the attention conditions may be the result of less effective context shifts within the divided attention/test condition. The detrimental

effects of divided attention would result in a weakened context change, resulting in a decrease in recall performance within the list four test for the divided attention condition. This reduced context shift, according to context theories, would result in a weakened, but still significant, FTE. Given a context theory explanation, we may expect to see more proactive interference in the divided attention/test condition if a weakened context shift accounted for the decline in the divided attention/test condition. However, the results of the extralist intrusion analysis (i.e., proactive interference) were not significant. The results reveal no difference between the proactive interference between the conditions. When further examining the amount of extralist intrusions, as shown in Figure 4, the mean values of extralist intrusions were not high for any condition. This suggests that proactive interference was not contributing to the decline demonstrated in list four performance. Due to the low numbers of proactive interference and the reduced performance in list four recall in the divided attention/test condition, a resource theory explanation is more appropriate for explaining the results of the current experiment.

Second, this experiment is the first to examine the relationship between attention and testing within the parameters of a forward testing effect protocol. This is important as study sessions often take place in environments that give ample opportunity for distraction during encoding phases, such as online environments. An understanding of attentional manipulations and their impact on forward testing in such environments can contribute to educational applications that seek to reap the most benefits from testing. The findings from the current experiment suggest that while the ideal situation for retaining information is study in the absence of distractors, individuals will still reap the benefits of forward testing in the presence of distracting stimuli during study sessions.

Practical Implications

The implications of the current experiment are important as they have not only theoretical applications, but practical applications as well. It is well replicated that testing improves not only learning of the studied materials (Roediger and Karpicke, 2006), but also future learning (Wissman et al., 201; Pastötter et al., 2012; Jing et al., 2016; Szpunar et al., 2013). The results of the current experiment demonstrate that an FTE is present when attention is divided between two different stimuli. Although optimal performance is seen when individuals can devote their full attention to study materials and utilize testing, benefits still occur when attention is divided during the study phase. In a world ubiquitous with distractions, attention is easily divided. Though the learning of new material may be stunted by these distractions, the results of the current experiment suggest that testing is still a viable option to improve the learning of new material. Should individuals seek to study in environments where the elimination of distraction is not possible, they will still receive the benefits of forward testing. In environments with stimuli that compete for attention, such as in public libraries, homes, or classrooms, the filtering out of distracting stimuli is not always possible. Utilizing testing, as opposed to restudy, may allow learners to enhance their learning of new materials in these attention-demanding environments.

Limitations

Within the context of the generalizability of these findings, there are certain limitations to be aware of. First, the sample for the current experiment consisted of Amazon Mechanical Turk workers. These participants were not in a controlled setting, and it is not known whether or not distractors outside the scope of the experiment were present during the completion of the experiment. Due to participants being completely online, it is possible multitasking occurred outside of the parameters of the experiment via technological distractions or interactions with other individuals. Additionally, the divided attention task consisted of a task that included only

an auditory stimulus. Auditory stimuli have been used in other experiments (Buchin and Mulligan, 2019; Mulligan and Picklesimer, 2016), though it is not known whether the results would replicate using a visual stimulus to divided attention. Taking these limitations into consideration, the results of the current experiment are generalizable to adults in early adulthood through an older population, as the sample consisted of participants in this age range. All participants were treated ethically according to the principles of the Belmont report and no risk of harm was involved within this experiment.

Future Directions

Future researchers may consider having participants complete the experiment in person and in a controlled setting. In a controlled setting, researchers will be able to more accurately determine if participants' attention is divided solely from the divided attention task, or if stimuli outside of the experiment are confounding the results. Completing the experiment within a controlled setting may also add to the ecological validity of the experiment. Although learning is increasingly taking place in online environments (Caprara & Caprara, 2022), much learning also takes place outside of a virtual context. Students in the classroom are more likely to be in a controlled setting with limited access to distractors. However, distractors may still be present. These distractors may range from noises within the classroom to visual distractors within the classroom. Researchers may also look at incorporating different stimuli within the divided attention task, specifically visual stimuli. In order to further inclusivity within neurodivergent populations, future researchers may also investigate how populations that suffer from attention deficit disorders are impacted by divided attention within an FTE paradigm. Because the FTE has been discovered in older children and traumatic brain injury patients (Szpunar et al., 2008),

future researchers should investigate whether the FTE is susceptible to the detrimental impact of divided attention within these populations.

In conclusion, the present experiment investigated the robustness of the FTE when attention is divided during studying. The results point to a persistent FTE when attention is divided. This experiment is the first to examine the relationship between attention and the FTE. This helps to illuminate underlying mechanisms of the FTE, which the data support a resource theory class of explanations.

References

Allen, M. (2017). Factorial analysis of variance. *The SAGE Encyclopedia of Communication Research Methods*. https://doi.org/10.4135/9781483381411.n192

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (n.d.). The Attentional Demands of Encoding and Retrieval in Younger and Older Adults: 1. Evidence From Divided Attention Costs. 19.

Aslan, A., & Bäuml, K.-H. T. (2016). Testing Enhances Subsequent Learning in Older but Not in Younger Elementary School Children. *Developmental Science*, *19*(6), 992–998. https://doi.org/10.1111/desc.12340

Buchin, Z. L., & Mulligan, N. W. (2017). The Testing Effect Under Divided Attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(12), 1934–1947. https://doi.org/10.1037/xlm0000427

Caprara, L., Caprara, C. Effects of virtual learning environments: A scoping review of literature. *Educ Inf Technol* **27,** 3683–3722 (2022). https://doi.org/10.1007/s10639-021-10768-w

Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval Potentiates New Learning: A Theoretical and Meta-Analytic Review. *Psychological Bulletin*, *144*(11), 1111–1146. https://doi.org/10.1037/bul0000166

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a Witnessed Event Increases Eyewitness Suggestibility: The Reversed Testing Effect. *Psychological Science*, *20*(1), 66–73. https://doi.org/10.1111/j.1467-9280.2008.02245.x

Davis, S. D., & Chan, J. C. K. (2015). Studying on Borrowed Time: How Does Testing Impair New Learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1741–1754. https://doi.org/10.1037/xlm0000126

Davis, S. D., Chan, J. C. K., & Wilford, M. M. (2017). The Dark Side of Interpolated Testing: Frequent Switching Between Retrieval and Encoding Impairs New Learning. *Journal of Applied Research in Memory and Cognition*, *6*(4), 434–441. https://doi.org/10.1016/j.jarmac.2017.07.002

Davis, S. D., & Peterson, (under review) *Journal for Applied Research in Memory and Cognition.*

Divis, K. M., & Benjamin, A. S. (2014). Retrieval Speeds Context Fluctuation: Why Semantic Generation Enhances Later Learning but Hinders Prior Learning. *Memory & Cognition*, *42*(7), 1049–1062. https://doi.org/10.3758/s13421-014-0425-y

Finn, B., & Roediger, H. L. (2013). Interfering Effects of Retrieval in Learning New Information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1665–1681. https://doi.org/10.1037/a0032377

Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated Testing Influences Focused Attention and Improves Integration of Information During a Video-Recorded Lecture. *Journal of Experimental Psychology: Applied*, *22*(3), 305–318. https://doi.org/10.1037/xap0000087

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful Retrieval Attempts Enhance Subsequent Learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 35*, 989-998.

MacDonald, P. A., & MacLeod, C. M. (1998). The Influence of Attention at Encoding on Direct and Indirect Remembering. *Acta Psychologica*, 20.

Mulligan, N. W., & Picklesimer, M. (2016). Attention and the Testing Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(6), 938–950. https://doi.org/10.1037/xlm0000227

NCES.gov. *The NCES fast Facts tool provides quick answers to many education QUESTIONS (National Center for EDUCATION STATISTICS)*. National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. (n.d.). https://nces.ed.gov/fastfacts/display.asp?id=80.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Nunes, L. D., & Weinstein, Y. (2012). Testing Improves True Recall and Protects Against the Build-Up of Proactive Interference Without Increasing False Recall. *Memory*, *20*(2), 138–154. https://doi.org/10.1080/09658211.2011.648198

Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval Practice Enhances New Learning: The Forward Effect of Testing. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00286

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval During Learning Facilitates Subsequent Memory Encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 287–297. https://doi.org/10.1037/a0021801

Pastötter, B., Weber, J., & Bäuml, K.-H. T. (2013). Using Testing to Improve Learning After Severe Traumatic Brain Injury. *Neuropsychology*, *27*(2), 280–285. https://doi.org/10.1037/a0031797

Pierce, B. H., & Hawthorne, M. J. (2016). Does the Testing Effect Depend on Presentation Modality? *Journal of Applied Research in Memory and Cognition*, *5*(1), 52–58. https://doi.org/10.1016/j.jarmac.2016.01.001

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming Overconfidence in Learning from Video-Recorded Lectures: Implications of Interpolated Testing for Online Education. *Journal of Applied Research in Memory and Cognition*, *3*(3), 161–164. https://doi.org/10.1016/j.jarmac.2014.02.001

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. https://doi.org/10.1073/pnas.1221764110

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing During Study Insulates Against the Buildup of Proactive Interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1392–1399. https://doi.org/10.1037/a0013082

Weinstein, Y., Gilmore, A. W., Szpunar, K. K., & McDermott, K. B. (2014). The Role of Test Expectancy in the Build-Up of Proactive Interference in Long-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1039–1048. https://doi.org/10.1037/a0036164

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The Interim Test Effect: Testing Prior Material can Facilitate the Learning of New Material. *Psychonomic Bulletin & Review*, *18*(6), 1140–1147. https://doi.org/10.3758/s13423-011-0140-7

Yang, C., Chew, S.-J., Sun, B., & Shanks, D. R. (2019). The Forward Effects of Testing Transfer to Different Domains of Learning. *Journal of Educational Psychology*, *111*(5), 809–826. https://doi.org/10.1037/edu0000320

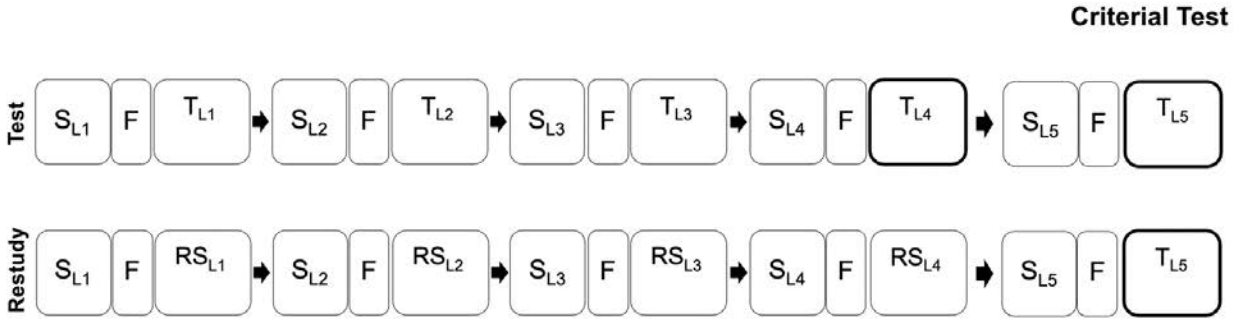
Yang, C., Potts, R., & Shanks, D. R. (2017). The Forward Testing Effect on Self-Regulated Study Time Allocation and Metamemory Monitoring. *Journal of Experimental Psychology: Applied*, *23*(3), 263–277. https://doi.org/10.1037/xap0000122

Yang, C., Potts, R., & Shanks, D. R. (2018). Enhancing Learning and Retrieval of New Information: A Review of the Forward Testing Effect. *Npj Science of Learning*, *3*(1), 8. https://doi.org/10.1038/s41539-018-0024-y

Yang, C., & Shanks, D. R. (2018). The Forward Testing Effect: Interim Testing Enhances Inductive Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(3), 485–492. https://doi.org/10.1037/xlm0000449

Figure 1

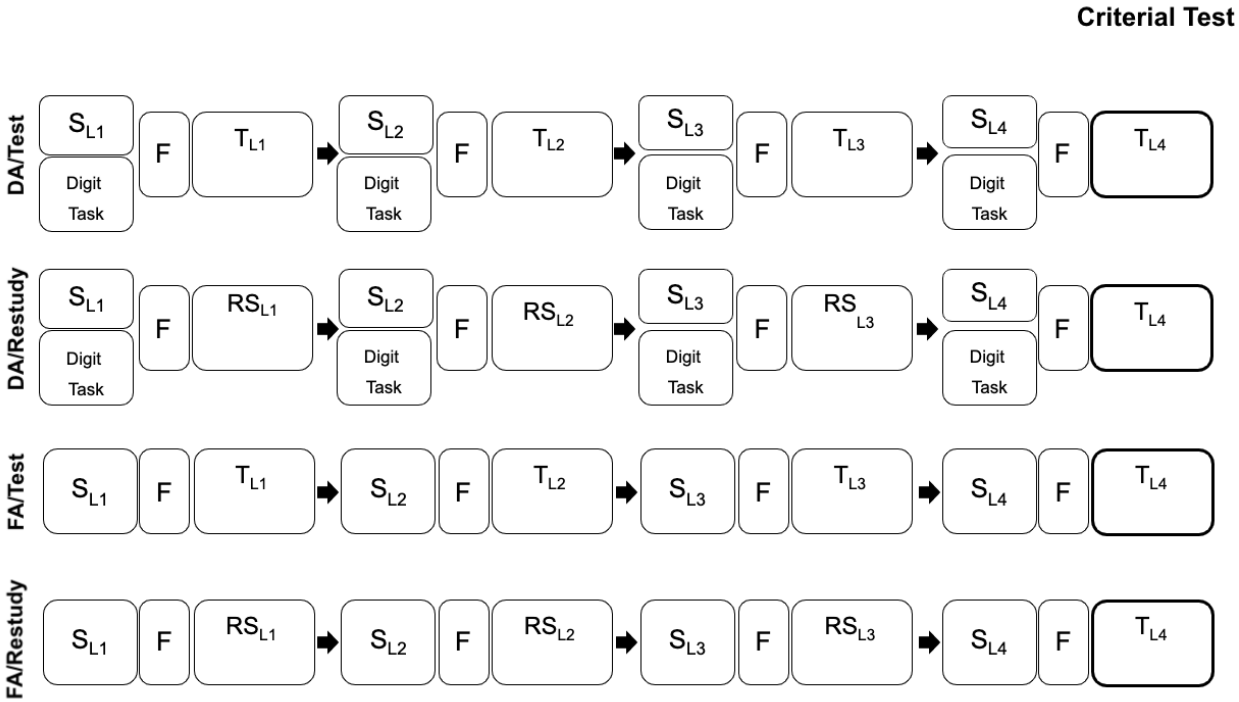
Forward Testing Protocol



Note. S= Study, L1, L2, L3, etc.= List 1, List 2, List 3, etc. F= Filler Task, T= Test, RS= Restudy

Figure 2

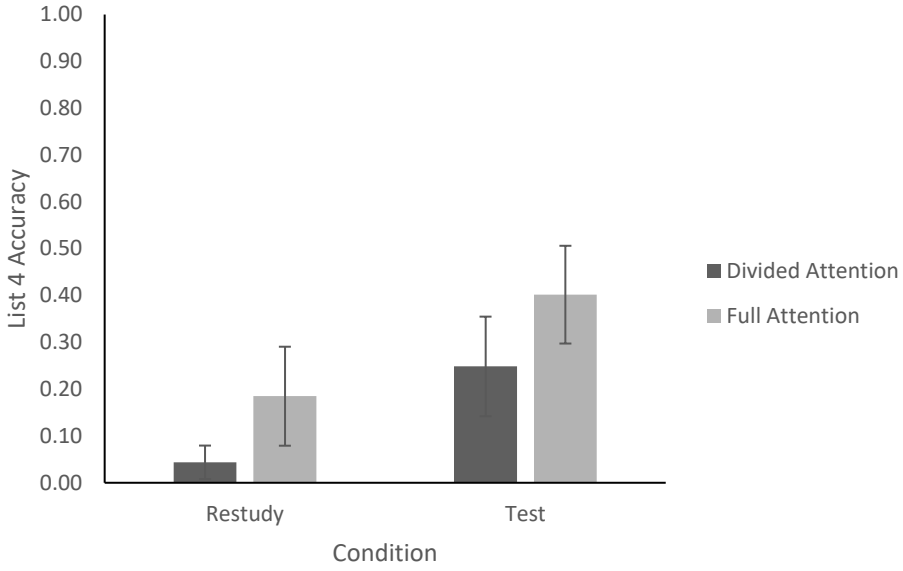
Participant Procedure for Effects of Attention of Restudy vs. Testing



Note. S= Study, L1, L2, L3, etc.= List 1, List 2, List 3, etc. F= Filler Task, T= Test, RS= Restudy

Figure 3

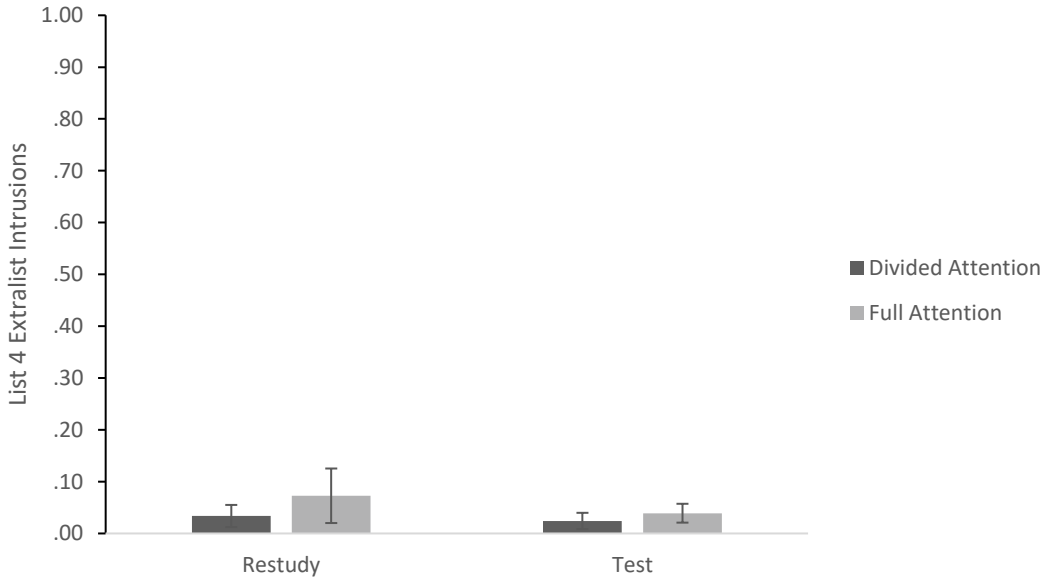
Proportion of List Four Items Correctly Recalled as a Function of Test Condition and Attention Condition



Note: Bars represent 95% Confidence Intervals

Figure 4

Proportion of List Four Extralist Intrusions as a Function of Test Condition and Attention Condition



Note: Bars represent 95% Confidence Intervals

Table 1

Descriptive Statistics for List Four Accuracy

Condition	N	*M*	*SD*
Test/Full Attention	32	.40	.29
Test/Divided Attention	31	.25	.31
Restudy/Full Attention	33	.18	.30
Restudy/Divided Attention	39	.06	.11