Working memory capacity differences in working memory offloading

by

Trevor J. Bell

B.A., Christian Brothers University, 2018
M.S., Kansas State University, 2022

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Abstract

Effective memory performance relies on the dynamic exchange of information between working memory (WM) and long-term memory (LTM). Past research has shown that both semantic and episodic LTM can enhance WM performance, as prior knowledge facilitates recall through mechanisms such as chunking, boosting, and offloading. Recent work by Bartsch and Shepherdson (2023) suggests that offloading—storing previously learned information in LTM to reduce WM demands—may be a key strategy for minimizing the impact of distractions on information maintained in WM. However, the extent to which individual differences in WM capacity (WMC) influence offloading remains unclear. This dissertation had three main goals: to (1) replicate prior findings showing that LTM enhances WM performance, (2) replicate prior findings supporting the mechanism of offloading, and (3) examine whether offloading depends on individual differences in WMC. We hypothesized that (H1a) WM recall performance would be higher for lists containing pre-learned items in LTM, demonstrating that episodic LTM benefits WM. We expected (H2a) that WM performance would be poorer following a distractor task, but more so for lists with only new items, as previously learned items should be offloaded to LTM and thus less affected. Alternatively, (H2b) if pre-learned items are not offloaded to LTM but instead remain in WM, we expected that performance would decline across all lists, regardless of whether they contained only new items or a mix of new and pre-learned items supporting a boosting account. Finally, we hypothesized (H3a) that higher WMC would be associated with more offloading. To test these hypotheses, participants completed a multi-phase memory study. In the LTM learning phase, they first studied word pairs. Next, during the WM phase, they encoded and retrieved lists of word pairs with varying amounts of pre-learned information (0%, 25%, 50%, or 75% Old). WM trials were followed by either a distractor task

designed to disrupt WM maintenance or a blank screen, which allowed us to assess the extent to which distractions disrupt memory performance. Finally, participants completed an LTM retrieval phase to test their memory for previously learned items and reported which strategies they used to remember each word pair. Results indicated that pre-existing episodic LTM representations benefited WM performance, supporting H1a. However, we found mixed evidence for offloading (H2a; H2b). Although WM performance was worse on trials containing a distractor task, this effect did not interact with the presence of previously learned items, which provides some support for the boosting hypothesis (H2b). At the same time, certain list types (e.g., those containing 25% Old and 75% Old items) showed no significant performance difference between blank and distractor conditions, providing some support for the offloading hypothesis (H2b). Thus, we are left with an imperfect account of whether offloading or boosting account better fit our results. Lastly, WMC was not significantly associated with offloading, failing to support H3a and instead aligning with H3b. This finding may suggest that individual differences in WMC stem from other cognitive abilities, such as attentional control or domain-specific skills, rather than from differences in WM offloading. These findings highlight the complex interplay between WM and LTM and suggest that future research should explore offloading under different experimental conditions to further disentangle these accounts.

Working memory capacity differences in working memory offloading

by

Trevor J. Bell

B.A., Christian Brothers University, 2018
M.S., Kansas State University, 2022

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
Dr. Heather R. Bailey

Copyright

Abstract

Effective memory performance relies on the dynamic exchange of information between working memory (WM) and long-term memory (LTM). Past research has shown that both semantic and episodic LTM can enhance WM performance, as prior knowledge facilitates recall through mechanisms such as chunking, boosting, and offloading. Recent work by Bartsch and Shepherdson (2023) suggests that offloading—storing previously learned information in LTM to reduce WM demands—may be a key strategy for minimizing the impact of distractions on information maintained in WM. However, the extent to which individual differences in WM capacity (WMC) influence offloading remains unclear. This dissertation had three main goals: to (1) replicate prior findings showing that LTM enhances WM performance, (2) replicate prior findings supporting the mechanism of offloading, and (3) examine whether offloading depends on individual differences in WMC. We hypothesized that (H1a) WM recall performance would be higher for lists containing pre-learned items in LTM, demonstrating that episodic LTM benefits WM. We expected (H2a) that WM performance would be poorer following a distractor task, but more so for lists with only new items, as previously learned items should be offloaded to LTM and thus less affected. Alternatively, (H2b) if pre-learned items are not offloaded to LTM but instead remain in WM, we expected that performance would decline across all lists, regardless of whether they contained only new items or a mix of new and pre-learned items supporting a boosting account. Finally, we hypothesized (H3a) that higher WMC would be associated with more offloading. To test these hypotheses, participants completed a multi-phase memory study. In the LTM learning phase, they first studied word pairs. Next, during the WM phase, they encoded and retrieved lists of word pairs with varying amounts of pre-learned information (0%, 25%, 50%, or 75% Old). WM trials were followed by either a distractor task

designed to disrupt WM maintenance or a blank screen, which allowed us to assess the extent to which distractions disrupt memory performance. Finally, participants completed an LTM retrieval phase to test their memory for previously learned items and reported which strategies they used to remember each word pair. Results indicated that pre-existing episodic LTM representations benefited WM performance, supporting H1a. However, we found mixed evidence for offloading (H2a; H2b). Although WM performance was worse on trials containing a distractor task, this effect did not interact with the presence of previously learned items, which provides some support for the boosting hypothesis (H2b). At the same time, certain list types (e.g., those containing 25% Old and 75% Old items) showed no significant performance difference between blank and distractor conditions, providing some support for the offloading hypothesis (H2b). Thus, we are left with an imperfect account of whether offloading or boosting account better fit our results. Lastly, WMC was not significantly associated with offloading, failing to support H3a and instead aligning with H3b. This finding may suggest that individual differences in WMC stem from other cognitive abilities, such as attentional control or domain-specific skills, rather than from differences in WM offloading. These findings highlight the complex interplay between WM and LTM and suggest that future research should explore offloading under different experimental conditions to further disentangle these accounts.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to extend a heartfelt thanks to all the undergraduates who helped me with research subjects: Olivia Edwards, Elayna Robinson, and Cora Malone. Their research support and communication while running participants made it easy to collect data quickly and efficiently. I also appreciate their willingness to listen to initial drafts of my research ideas and presentations.

I also want to give a special shoutout to my cohort throughout my years here – Nora Balboa, Troy Fort, and Zach McKinnell. Their constant support and willingness to give advice has helped me over the years. I truly appreciate their kindness that made graduate school, not only manageable, but also enjoyable.

I am especially grateful to Dr. Lea Bartsch for her generosity in providing valuable insights into her own experimental design, analyses, and data. She also provided direct guidance on my dissertation design and provided recommendations to further the research. Her support has been a vital foundation for this project.

Additionally, I want to thank Dr. Heather Bailey for her mentorship throughout my time here. She has significantly influenced both my academic and personal growth. And although I might be biased, I do consider her the best mentor I've had. She was always willing to take time to help me understand issues I've come across in experiments and provided guidance for tough situations.

Finally, to everyone who has been part of my graduate school journal, rather through direct help or providing encouragement, I am genuinely thankful. This dissertation would not have been possible without each of you.

Dedication

I dedicate this dissertation to my family, and especially to my late great grandmother, Vera. My family has been a constant source of support throughout my life. I am honored to be the first in our family to pursue a PhD.

Chapter 1 - Literature Review

Individual Differences in Transfer of Information Between Memory Systems

One major focus in research on memory systems is how well one can memorize necessary information and shift information from one temporary and immediate memory system (working memory; WM) to a more permanent memory system (long term memory; LTM). We know from prior research that these two systems exchange information in both beneficial (e.g., proactive facilitation; Cermak, 1970; Arzi et al., 1985) and harmful ways (e.g., proactive interference; Keppel & Underwood, 1962). One recent account for how LTM enhances WM processing is *offloading*. Offloading posits that when an LTM representation exists for information on a WM task, then it is unnecessary to maintain that information directly in WM. Essentially, an individual can increase WM efficiency by maintaining a representation of the to-be-remembered information outside of WM (in LTM), which frees up resources to store new information in immediate memory (Bartsch & Shepherdson, 2023).

Individuals differ in their ability to maintain information in immediate memory (typically new information; Unsworth & Engle, 2007; Daneman & Carpenter, 1980). If individuals are already limited in their ability to maintain new information, it is possible that their ability to offload old information could determine how well they are able to learn and remember new information. However, only a few studies have specifically evaluated whether offloading occurs and how it affects memory performance (these will be discussed in more detail below in "Information Exchange Between Memory Systems"). Thus, there is a need to understand when offloading occurs, how it affects LTM, and whether people of all abilities offload in a similar manner.

Organization of Memory Systems

Many theories exist regarding the relationship between WM and LTM. One camp of theorists proposes that memory is a unitary system (e.g., Nairne, 2002). According to this camp, all information is part of LTM, and the selection of information relies on the specificity of retrieval cues: The greater the cue specificity, the higher the probability of activating the relevant memory. More specifically, distinctive cues will activate a narrow, more targeted group of representations in LTM. With fewer options to select from, a person is more likely to accurately retrieve the correct memory. For instance, if someone is tasked with learning the word pair, “TABLE-DOG,” using their own dog as a cue for remembering that word pair will help in later recall. The distinct cue of their dog will aid the person during recall by increasing the likelihood of retrieving the word “DOG” when presented with the word “TABLE” and not another distractor word, such as “CAT.”

Conversely, indistinct or general cues might lack the necessary information to activate relevant information in LTM. Because of this, activation may spread to a large amount of information with a low association to the to-be-remembered memory, which will decrease the likelihood of correct retrieval. Using the same previous example, if the person uses the cue “ANIMAL” to search LTM when presented with the word “TABLE”, then many words may be activated (e.g., “DOG”, “CAT”, “FISH”, “GIRAFFE”, etc.) and the likelihood of accurately selecting the correct answer is decreased. Thus, the primary retrieval factor within unitary models is the uniqueness of the cues to differentiate relevant information from irrelevant information. However, unitary models do not distinguish between differences in WM and LTM’s duration (Ebbinghaus, 1875/1913), capacity limits (Miller, 1956; Atkinson & Shiffrin, 1971), and rate of decay (Atkinson & Shiffrin, 1971; Bahrick, Bahrick, & Wittinger, 1975).

Modular theories of memory, on the other hand, separate long-term memory storage from immediate memory usage (i.e., modular systems; see also Atkinson & Shiffrin, 1971). One popular modular theory was proposed by Baddeley and Hitch (1974); see Figure 1. Within this model, WM is separate from LTM with domain-specific storage systems that represent different sensory modalities (i.e., phonological loop for auditory & verbal information; visuospatial sketchpad for visual & spatial information). These two short-term storage systems are modulated by the central executive, which acts as a switchbox alternating between the two systems, depending on a given task. Despite its popularity, this theory does not explain how information is bound together across multi-modal domains (e.g., such as students seeing a PowerPoint slide while hearing the instructor lecture) and, thus, was later updated to include an episodic buffer component. Baddeley (2000) claimed that the episodic buffer integrates information from both sensory storages and transfers this information to and from long-term storage.

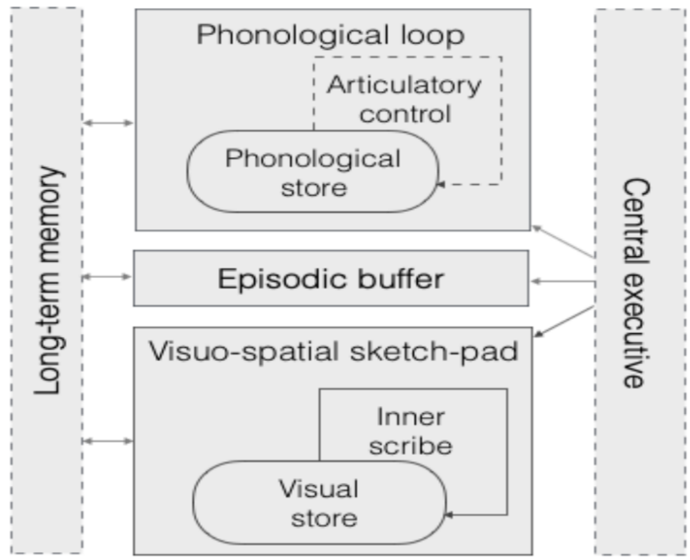

Figure 1. Baddeley's Multicomponent Model of Working Memory. From "Working Memory from the Perspective of the Multicomponent Model of Embedded-Processes Model," by Ozimič, 2020, doi: 10.7906/indecs.18.4.2.

However, there are issues with Baddeley's model of memory, such as there is very little neural data supporting the idea that these different subsystems truly exist in separate brain areas. For example, neural evidence cannot define an exact location for the central executive and episodic buffer, and if they do, their function overlaps with other proposed mechanisms of cognition. More critically, Baddeley's model cannot account for neural studies that suggest there are different levels of activation for given working memory stimuli (for review, see LaRocque, Lewis-Peacock, & Postle, 2014). Specifically, studies have demonstrated that WM items that are not currently in the focus of attention can still be actively maintained in a "activity-silent" state (LaRocque, Lewis-Peacock, & Postle, 2014; Wolff et al., 2015; Sprague, Ester, & Serences, 2016). For instance, LaRocque et al. (2013) used multivariate pattern analysis (MVPA) to distinguish between attended and unattended items in visual WM. They found that although unattended items had reduced neural signals, they were still above baseline brain activity. Furthering extending this idea, Rose et al. (2016) demonstrated that transcranial magnetic stimulation (TMS) could be used to successfully reactivate unattended (but relevant) WM items. Their results provided causal evidence that WM items can be shifted from an inactivated (i.e., lower level of activation) to the focus of attention (i.e., higher level of activation). WM items appear to exist along a continuum of activation that can dynamically shift depending on what is necessary for the task. Baddeley's current multicomponent model lacks a mechanism for how this continuum exists.

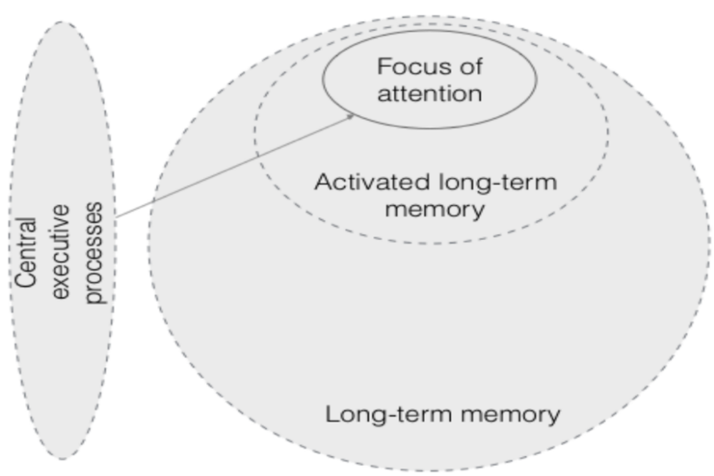

Figure 2. Cowan's Embedded Processes Theory (EPIC) of Working memory. From "Working Memory from the Perspective of the Multicomponent Model and Embedded-Processes Model," by Ozimič, 2020, doi: 10.7906/indecs.18.4.2.

To address the fact that unconscious processes occur in WM and that there is a distinction between WM and LTM in capacity and duration, more recent theories incorporate ideas from both modular and unitary system theories, providing more of a hybrid approach. One such theory by Cowan (1999, 2008; Cowan et al., 2021), the Embedded Processes Theory (EPIC), proposes that information is represented in LTM at varying levels of activation or consciousness (see Figure 2). Information that is currently being attended to is in the focus of attention (e.g., this sentence), whereas information that was recently relevant but is not currently attended to (e.g., information from the beginning of this paragraph) is in activated LTM. All other information is proposed to be in LTM (e.g., grammatical and vocabulary knowledge). The important distinction between these three levels is the amount of activation (guided by attention). Thus, Cowan's theory directly addresses how consciousness plays a role in WM processes, but importantly, he believed that the focus of attention is capacity-limited and does not experience interference (Cowan, 2008; Cowan et al., 2021; Oberauer et al., 2012). His theory also proposes a central executive to guide information from activated LTM into the focus of attention. However, he did not explicitly explain *how* this mechanism might select relevant (or irrelevant) items to place into the focus of attention.

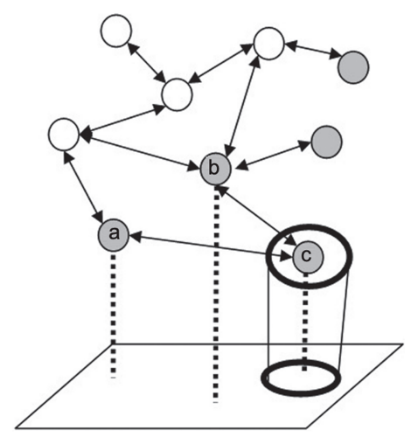

Figure 3. Oberauer's Embedded Component Model of Memory. The letters "a," "b," and "c" represent items that the participant needs to remember for a given task. The dotted lines from "a," "b," and "c" are items that are in the region of direct access. The bold lines around "c" represent the focus of attention. The gray circles represent items that exist in activated LTM. From "Attention to Information in Working Memory," by Oberauer and Hein, 2012, doi: 10.1177/0963721412444727.

Another hybrid model by Oberauer (2002, 2021), Embedded Component Model (EC) asserts that memory representations exist in LTM, which contain information specific to the item's context of encoding. The item's representation activates other parts of LTM, which creates a region of direct access. This region represents information that is directly related to the item's context (e.g., other related stimuli). For example, using Oberauer's memory model from Figure 3, imagine a participant in a memory experiment is asked to learn three word pairs, like “APPLE-READ”, “AIRPORT-COUSIN”, and “ROLE-HOUSE”. Let us say a participant is later shown, “APPLE-____” and asked to recall the other word in the pair. According to Oberauer’s model, various words are activated above a certain baseline, such as “ROLE,” “AIRPORT,” and “READ”, and held in the region of direct access (represented by the gray circles). Then, the focus of attention, represented in Fig. 3 as the bolded outline around item “c,” will select one of the activated representations in the region of direct access to attend to, and this selection will be the participant’s chosen response (e.g., “READ”).

To reconcile differences between memory models, researchers often look beyond behavioral evidence to underlying neural mechanisms. Neural activity during tasks involving memory have three main features: 1) Hippocampal activity is commonly observed when completing LTM tasks, 2) PFC activity is shown when information is represented in WM, and 3) Activity-silent neural activity occurs during memories tasks which can be further decoded to represent WM items and contexts (Rose et al., 2016). Both unitary and modular theories fail to support the above-mentioned neural results. Unitary models cannot account for separate brain activity associated with WM vs. LTM functions (Lewis-Peacock et al., 2012) or the interaction between supposedly separate systems. Baddeley’s modular theory can account for these findings if one argues that the PFC acts as the subsystem responsible for storing working memory representations while the hippocampus is a subsystem response for storing long term memory representations. However, Baddeley’s modular theory fails to explain why activity-silent neural information may exist when participants are not consciously attending to the information. Modular models also fail to fully explore how contextual information associated with memories (e.g., information condensing multiple subsystems) can be encoded into memory.

Hybrid models appear to be the most robust for explaining neural results. Both Cowan's and Oberauer's model support the idea that activity in the PFC could reflect information represented within the central executive processes (Cowan) or region of direct access (Oberauer), whereas activity in the hippocampus could reflect information represented in activated LTM or LTM. When it comes to activity-silent neural activity, memory representations are held in activated, but dormant, LTM until they are brought to conscious awareness (e.g., Oberauer's region of direct access and Cowan's activated long-term memory) and then attended to (e.g., focus of attention). It is important to note that both Cowan's EPIC theory and Oberauer's EC theory can account for behavioral and neural data mentioned above. However, this proposal will test various research questions that rely upon assumptions made by Oberauer's model of memory (Oberauer, 2002, 2009, 2021).

Information Exchange Between Memory Systems

For memories to be effectively encoded and retrieved, information must be exchanged between WM and LTM. Evidence for this comes from findings that both semantic LTM (e.g., memories for general facts and knowledge; Tulving, 1972) and episodic LTM (e.g., memories someone has that include specific temporal and spatial details; Tulving, 1972) influence performance on WM tasks. For instance, semantic LTM aids in performance on WM tests when materials included in the task are familiar to participants. Some well-known examples of semantic LTM benefitting WM includes the concreteness effect (e.g., participants recall words at a higher rate when they are concrete, imageable words with one clear representation compared to abstract words that contain multiple representations; Walker & Hulme, 1999; Tse & Altarriba, 2022; Chubala et al, 2019), the word frequency effect (e.g., participants recall words at a higher rate when they are words that appear more often in one's native language; Monsell, Doyle, &

Haggard, 1989), and the chunking effect (e.g., participants recall words at a higher rate when they match new information to previously learned semantic chunks, like PDF or CIA; Chase & Simon, 1973; Portrat et al., 2016; Thalmann et al., 2019; Lörch, Lemaire, & Portrat, 2023).

Further, WM performance is influenced not only by semantic LTM, but also by episodic LTM representations. Episodic LTM information can either increase WM performance (e.g., *proactive facilitation*; Cermak, 1970; Arzi et al., 1985) or decrease WM performance (e.g., *proactive interference;* Keppel & Underwood, 1962). Chen and Cowan (2005, 2009) showed that adding previously learned information (i.e., items in which participants have already created an episodic LTM representation) to WM tasks increased participants' overall performance. Further, Hoskin et al. (2016) found the reinstating episodic memories during WM tasks can disrupt memory performance if the items are too like one another.

Numerous studies above provide evidence that shows both episodic and semantic LTM benefit WM, but few have attempted to connect these to theoretical mechanisms within memory models. There are three potential explanations for why LTM benefits WM: e.g., chunking, offloading and boosting accounts. For instance, a chunking account claims that several items are combined into one singular item, which is thought to primarily occur in WM. Some researchers theorize that chunking allows people to organize information into meaningful groups, thereby maximizing the amount of information that can be held in WM (Miller, 1956). Others contend that information can be "offloaded" to LTM. In other words, previously learned information can be represented in LTM without needing to be maintained in WM (Thalmann et al., 2019; Bartsch & Shepherdson, 2023; Kowialiewski et al., 2021). Lastly, the boosting account states that the strength of an episodic memory will facilitate recall and reduce potential confusion with irrelevant information (Unsworth & Engle, 2007). Importantly, according to the boosting account, the representation of the episodic memory still exists in WM to guide information, in contrast to the offloading account that claims WM is bypassed altogether.

Bartsch and Shepherdson (2023) aimed to differentiate between these three accounts (chunking, offloading & boosting) to fully understand the process behind how episodic LTM influences WM. A modified schematic of their Experiment 2 methods, which is most relevant to the proposed study, is shown in Figure 4. Participants completed four phases of memory testing: 1) LTM learning phase, 2) WM encoding task, 3) WM retrieval test and 4) a LTM retrieval test. In the LTM learning phase, participants saw 42 pairs of words for 4000 ms each followed by 1000 ms inter-stimulus-interval (ISI). Immediately after the LTM learning phase, participants completed the WM task. During the WM encoding task, participants viewed four pairs of words for 1000 ms with a 500 ms ISI and afterwards either saw a blank screen for 10 sec or completed a distractor task for 10 sec. All WM encoding trials were capped at four total word pairs with either 0 or 2 previously learned word pairs (from the LTM learning phase). After four WM trials were encoded (e.g., four word pairs were shown), they completed a four-alternative forced choice task (4-AFC) with the following set up for the responses: 1) target word, 2) another item paired with a different word in the same trial (within-trial protrusion), 3) an item presented in the LTM learning phase but not the current WM trial (LTM lure), and 4) a completely new item.

Following all the WM retrieval trials, participants completed a final LTM retrieval test using a 4-AFC recall paradigm.

Bartsch and Shepherdson (2023) found that performance was higher when LTM pairs were included in the WM task than when they were not included. Importantly, though, the effects of the distractor task could help differentiate between the boosting and offloading accounts. According to the boosting account, LTM exerts its benefits on information being maintained in WM. If this account were supported, then disrupting WM maintenance using a distractor task should impair WM test performance on all four conditions (lists containing LTM-available word pairs and lists containing only new word pairs). By contrast, the offloading account proposes that previously learned information can be offloaded to (i.e., maintained in) LTM without the use of WM. If this accounted were supported, then the distractor task should impair performance on lists with new word pairs more so than performance on lists with LTM-available pairs because the distractor task should have a minimal effect on the pre-learnt/offloaded information maintained in LTM. Bartsch and Shepherdson (2023) found evidence in favor of the offloading

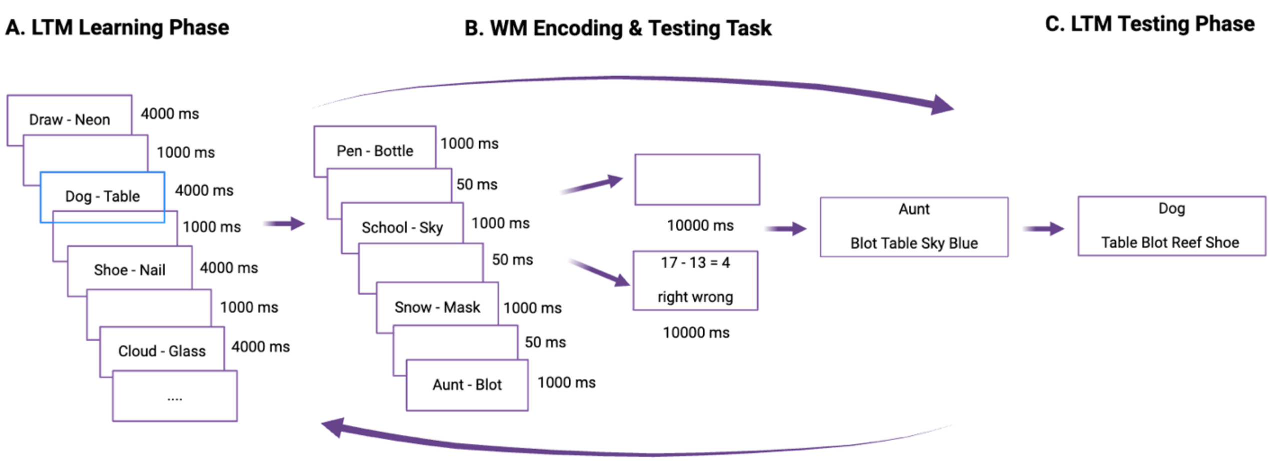

Figure 4. Experiment 2 paradigm from "Chunking, boosting, or offloading? Using serial position to investigate long-term memory's enhancement of verbal working memory performance," from Bartsch. L. M. & Shepherdson, P., 2023, *Attention, Perception, and Psychophysics, 85*, 1570. Reprinted.

account: Performance was disproportionately impaired by the distractor task on lists with only new pairs. This result provided support for the offloading account because disrupting WM maintenance did not have as large of an impact memory when participants had a pre-existing episodic LTM representation for the to-be-remembered pairs. Such evidence suggests that, when people offload information to LTM, they can focus their WM resources on new information rather than on both the new information and maintaining previously learned information. In other words, the LTM information remains in LTM until participants need to draw from it for subsequent recall, similar to a gating mechanism between WM and LTM (Oberauer et al., 2016). Thus, Bartsch and Shepherdson (2023) provide empirical evidence for a theoretical mechanism by which information is exchanged between WM and LTM.

Individual Differences in Working Memory Capacity (WMC)

While Bartsch and Shepherdson (2023) provided support for an offloading mechanism, they did not consider individual differences in this process. Differences in memory performance could be influenced by how much information certain individuals can retain in WM versus how much they "offload" to LTM. Individual differences in WMC are well documented (Oberauer et al., 2016; Engle, 2001; Daneman & Carpenter, 1980; Barrett, Tugade, & Engle, 2004; Cantor & Engle, 1993; Unsworth & Engle, 2005; Just & Carpenter, 1985; La Pointe & Engle, 1990), but the mechanism underlying these differences is still debated.

One potential explanation for individual differences in WM that have already been described in the literature is differences in attentional control, with high WMC individuals having better attentional control than low WMC individuals (e.g., Conway, 1996; Engle et al., 1992). In other words, high WMC individuals perform better on tasks that require attentional flexibility because they are better able to direct attention to relevant stimuli and ignore distracting information. Attentional control may influence the extent to which people access previously learned LTM representation (i.e., offload). That is, poor attentional control may prevent people from selectively activating/accessing LTM representations, which would force them to represent the incoming information in WM.

Another source of individual differences in WMC could be domain-specific skills. Daneman and Carpenter (1980) claim that high WMC individuals have stronger task-specific skills than do low WMC individuals, which essentially frees up resources for remembering information. For example, when performing a verbal working memory task called Reading Span in which the individual must determine whether a sentence is grammatically correct and then remember a word, an efficient reader needs to devote less attention to comprehending the sentence and, thus, has more attention available to encode the word memoranda compared to a less efficient reader who must expend more effort while reading the stimuli. Alternatively, Unsworth and Engle (2007) argue that the main difference between high and low span individuals is their ability to discriminate between target and distracting information. Specifically, they claim that high WMC individuals use more diagnostic retrieval cues that guide their search for relevant information. In turn, this allows high WMC individuals to discriminate relevant memory stimuli from irrelevant or interfering information. For offloading, this could lead to differences because of how much information high and low WMC individuals are able to successfully offload. For example, high WMC individuals could be more efficient at offloading memory representations to LTM. Thus, when participants are tasked with remembering four pieces of information, their region of direct access will retain a smaller number of items than individuals with low WMC. As such, high WMC individuals will have less items to cause potential interference compared to low WMC individuals.

Finally, another potential source of individual differences in WMC is in their ability to utilize strategies to remember key information (Cokely et al., 2006; Bailey et al., 2008; Bartsch, Singmann, Oberauer, 2018; Bartsch et al., 2019). The efficacy of memory strategies may depend on the level of processing each strategy requires (Craik & Lockhart, 1972). When information involves more elaborative processing of information, such as making connections between stimuli or between the stimulus and information in LTM, it can be recalled more readily. Elaborative strategies (e.g., sentence-linking and imagery), often lead to a better memory performance for verbal information. These strategies typically require the participant to consider the meaning of the to-be-remembered. For example, sentence-linking allows the participant to connect the to-be-remembered information to other to-be-remembered words (e.g., linking the words “DOG” and “BREAD” together as “my dog, Miller, likes to steal bread from the counter”), which increases the association in long-term memory.

On the other hand, less elaborative strategies (e.g., rehearsing a word repeatedly), are associated with lower performance compared to when people report using more elaborative strategies. Rehearsal, for instance, only requires the participant to simply repeat the information internally to keep it active in memory but does not require the participant to make associations between the to-be-remembered information and long-term memory (Craik & Watkins, 1973; Richardson, 1998). In the context of offloading, individuals who report using more elaborative strategies (often high WMC capacity individuals) create well-formed memory representations, potentially allowing them to offload information to LTM and free up resources in WM. On the other hand, individuals who report using less elaborative strategies may not be creating stable LTM representations (i.e., not offloading), and thus, relying more upon WM.

It is important to note that these theories are not mutually exclusive. For instance, increased attentional control could allow for both the use of better strategies and more constrained search of LTM (Astle et al., 2012; Poole & Kane, 2009; Kane, Bleckley, & Conway, 2001; Unsworth, Brewer, & Spillers, 2012). This proposal's aim is not to distinguish between these theories, but instead, to propose offloading as a potential mechanism to explain individual differences in WMC.

Current Study

To summarize, although researchers may disagree about the mechanism underlying individual differences in WM, theories agree on: 1) attention is limited and must be directed to focus on certain stimuli, 2) distracting information must be inhibited, and 3) episodic LTM can aid WM. The first goal of the proposed study is to replicate Bartsch and Shepherdson's (2023) work showing that pre-existing episodic representations can be offloaded and maintained outside of WM. Offloading information to LTM provides more attentional resources to encode new items by minimizing the amount of interfering information in WM. The second goal of the proposed study is to evaluate whether high and low WMC individuals differ in this offloading process, given their differences in both attentional resources and increased susceptibility to interference. To do so, we will modify the procedures used in Bartsch and Shepherdson (2023). Specifically, we will manipulate the amount of pre-learned items (0-3 items) on a given WM trial and will assess WMC.

Research Questions (RQ) and Hypotheses (H)

RQ1: Do pre-existing episodic LTM representations benefit WM performance? (Replication of Bartsch & Shepherdson, 2023; Main effect of List Condition)

H1a: WM recall performance will be higher on lists with pre-learned items compared to lists with only new items. Such a result would demonstrate that episodic LTM representations benefit WM.

H1b: WM recall performance will not differ between lists with and without pre-learned items. Such a result would demonstrate that LTM representations do not benefit WM.

RQ2: Do individuals offload information to LTM? (Replication of Bartsch & Shepherdson, 2023; List x Task Interaction)

H2a: WM recall performance will be lower when participants are faced with a distractor task compared to when they are shown a blank screen. However, the distractor task will have a greater effect on lists with new items only (all presumably represented in WM) compared to lists containing previously learned items (some information presumably represented in LTM). Such a result would support an offloading account because items represented in LTM would be less affected by a WM distractor task.

H2b: WM recall performance will be lower when participants are faced with a distractor task compared to a blank screen, and this will be true for lists with new items only and lists containing previously learned items. This result would support a boosting account, which assumes that all items are represented in WM and, thus, would be affected by a distractor task.

RQ3: Do individual differences in WMC predict the likelihood of offloading? (Extension; WMC x List x Task interaction)

It is possible that not all individuals offload information to the same extent: Low spans may not offload information to LTM and, thus, unnecessarily “clutter" WM. I hypothesize that the offloading effect (i.e., List x Task interaction) will depend on an individual’s WMC.

H3a. Individual differences will influence the extent to which individuals offload. If this is true, we expect to find a WMC x List x Task interaction, such that higher WMC will be associated with a stronger List x Task interaction. Presumably, this effect will occur because individuals with higher WMC are able to offload information to LTM (and thus, those representations in LTM will be less affected by the distractor task), whereas lower WMC will not offload as readily (i.e., all items will be represented in WM) and be affected by the distractor task. (It is possible we observe other forms of interactions, but this is the most logical one that fits with the literature review).

H3b: All individuals offload similarly, and there is no 3-way interaction, only a two way interaction for List Type x Task (i.e., offloading). Such a result would indicate that individuals can use previously information to offload resources from WM, regardless of their WMC. Thus, individual differences in WMC are due to other explanations other than the extent to which they offload.

Chapter 2 - Methods and Materials

Power Analysis

To estimate the appropriate sample size for the study, I used the *simr* package (Green & MacLeod, 2016) in R to conduct simulations of different sample sizes using informed parameter estimates from Bartsch and Shepherdson's (2023) dataset on simulated data. This process had four main steps: 1) create a reference model to inform known parameter estimates, 2) create an artificial dataset that houses all the variables, so the simulation knows what variables are included, 3) create a model with both known and unknown parameters estimates, 4) run the model on simulations of the artificial dataset with different sample sizes.

Prior to the power analysis, I used *lme4* package (Bates et al., 2015) to conduct a mixed effects model to investigate the fixed effects of List x Filler Task with the random intercept and slope for participants only on data from Bartsch and Shepherdson (2023). The output from the model was used to inform known parameter estimates (List, Filler Task, all interactions, and intercept). Next, I created an artificial dataset with the added variable, WMC (continuous), using the packages and libraries, *dplyr* (Wickham et al., 2023) and *tidyverse* (Wickham et al., 2019). From there, I used the gathered parameter estimates for the known effects and substituted Cohen's $d = 0.3$, which is a typical effect size in individual differences work, to inform the unknown parameters that included WMC as a variable. Lastly, I used the library *mixedpower* to run iterations of simulations that used the sample sizes of 50, 100, 150, 200, and 250 participants. These simulations indicated that, to detect the three-way interaction (List x Filler Task x WMC) at 80% power, I would need to run 50 participants. However, this sample size is not large enough to detect any of the two-way interactions (which recommended 100+ participants). Given this, I decided to collect 150 usable participants, which is a typical sample size for individual differences studies in working memory capacity (Miller & Unsworth, 2018; Robinson et al., 2018; Aslan & Bäuml, 2010; Unsworth & Miller, 2024).

Participants

One hundred and eighty-seven participants were recruited from a public, midwestern university's psychology courses and earned either course credit or extra credit for participating. Participants ranged in age from 18-47 years of age (*M* = 19; *SD* = 2) and were primarily White/Caucasian (79.7%; *N* = 149) and female biased (77.2%; *N* = 142). Participants were excluded from individual tasks based on their performance. The exclusion criteria for each of the tasks was chance performance: 25% or 32 words correct (n = 2 were excluded) for WM recognition; 25% or 12 words correct (n = 8 were excluded) for LTM recognition. Based on past literature, we excluded participants from WMC analyses if they scored lower than 80% correct on the processing component of the Symmetry Span (n = 22; see below for more details; Shah & Miyake, 1996). The total sample size for each of the individual tasks after exclusions was: 186 participants for WM recognition analyses, 179 participants for LTM recognition, and 165 participants for WMC.

Materials

Working Memory Capacity Assessment

Participants completed a version of the Symmetry Span Task (SST; Kane et al., 2004) on OpenSesame (Mathôt et al., 2012) reported in Monteiro et al. (2024). Prior to all memory trials, participants completed a calibration task to determine how long they would be given in the processing task (e.g., deciding symmetrical or not) for the SST testing phase. For the calibration portion of the task, participants were shown a fixation cross for 500 ms following by an 8 x 8 matrix of squares. This matrix had squares that are filled in black color. The pattern of the filled-in squares was either be symmetrical or asymmetrical for half the trials. Participants had unlimited time to respond with "1" for symmetrical or "2" for asymmetrical. Participants completed two practice trials in which they were asked to solely make decisions on whether a square was symmetrical for 20 trials.

Immediately following the calibration task, participants were shown a square matrix and asked to determine whether the pattern in the matrix was symmetrical (see Figure 5). Then, they were shown another 4 x 4 matrix of squares with one of the squares colored red for 1000 ms, and they were instructed to remember the position of the red square. This procedure (symmetrical/asymmetrical processing trial; red square memory trial) repeated until participants finished their designated set size. Set sizes ranged from 2-6 to-be-remembered red square locations. Once the entire set size was shown, participants were given a blank 4 x 4 matrix and asked to click the locations of the squares in the order in which they appeared. Their previously-selected square remained on screen for 500 ms and disappeared before they continued to choose locations for the subsequent squares. They were not given a time limit to recall the squares. They received feedback on their performance for the designated set before moving to the next set. Participants completed 12 trials per set size presented in ascending order (N = 60 trials).

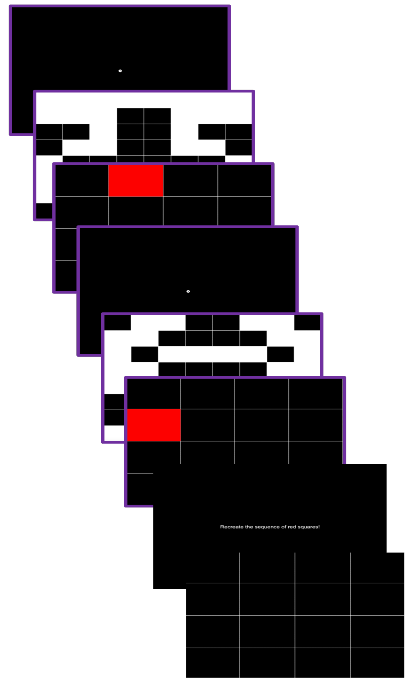

Figure 5. Example of a Symmetry Span Task used on OpenSesame from Monteiro et al. (2024).

Working memory capacity was scored using procedures and scripts from Monteiro et al. (2024). This procedure used partial scoring for symmetry span (e.g., participants received credit for each correctly recalled square, even if the sequence in which squares were clicked was incorrect). Raw scores were calculated as the total the number of correct responses. Normalized scores were calculated as the sum of the number of correct responses divided by the number of total trials. Normalized symmetry span scores were used for the WMC variable in the study.

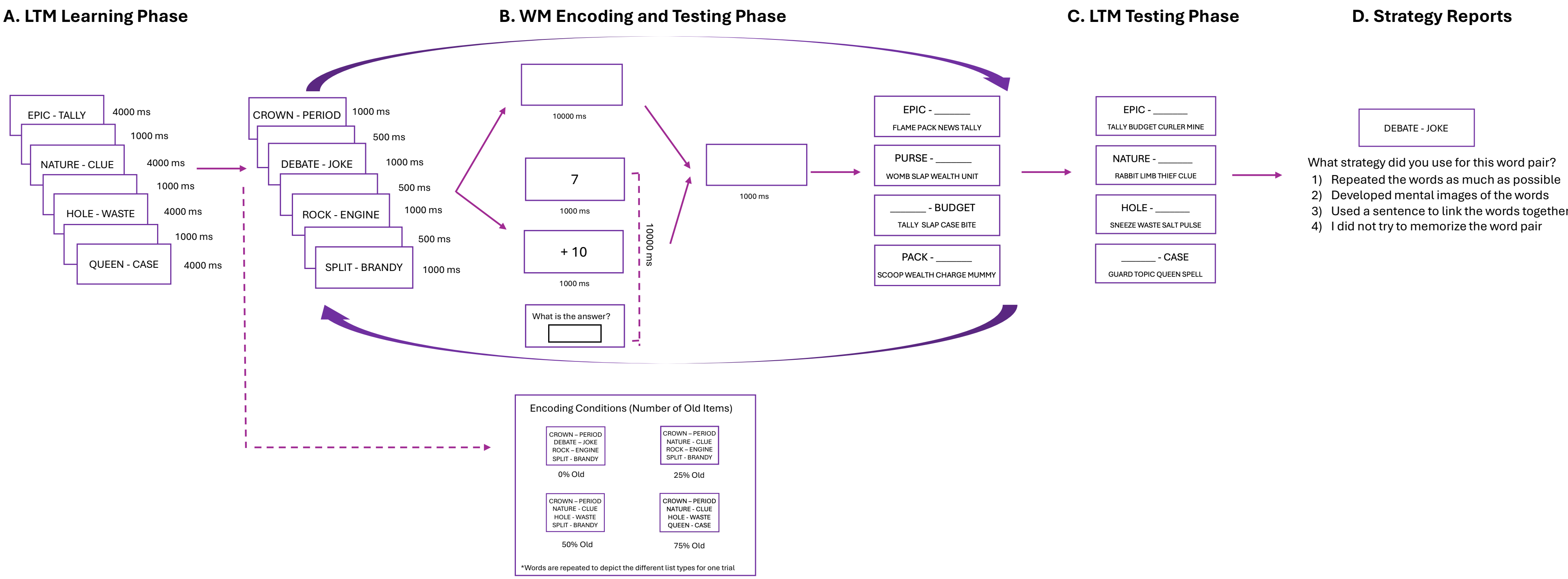

Figure 6. A graphical depiction of the experimental design for the study, showing the order of the tasks as A) a LTM learning phase, B) WM Encoding Task, which after one set of four pairs is shown, participants will either see a blank screen or math equation. Not pictured is the WMC assessment and filler task prior to the LTM learning phase. Following presentation of all four word pairs in one set, the participants will complete the 4 AFC for each word pair shown in a set until all sixteen sets (8 followed by a blank, 8 followed by a math equation) are completed. Lastly, participants will be asked to complete the C) LTM 4-AFC retrieval task. The bottom of the figure shows the four encoding conditions in the experiment. Thus, each condition will be repeated four times.

Long-Term Memory and Working Memory Tasks

Word Stimuli. The Medical Research Council (MRC) Psycholinguistic Database (Coltheart, 1981) were used to select stimuli in the experiment, controlling for number of letters, number of syllables, concreteness, imageability, word frequency, and familiarity. These word stimuli were used before in previous studies in our laboratory. A total of 512 words were used to create randomly generated word pairs for the study (e.g., 256 words for the entire task, 128 new, never assigned words were used for option choices within recognition test trials for both the WM trials, and an additional 128 new, never assigned words used for option choice within the LTM recognition trials).

Long-Term Memory Learning Phase

The LTM phase was identical to that of Bartsch and Shepherdson (2023). Participants were shown one word pair at a time for 4000 ms followed by a 1000 ms interstimulus interval, for a total of 48 word pairs. The order of the word pairs was randomized within participants. Word pairs were randomly assigned to be long-term stimuli (N = 48 word pairs). Then, they were randomly assigned into list type (e.g., 25%, 50%, and 75%, N = 8-24 word pairs per list). For the response options, the words used as within-trial lures were randomly drawn from one of the four word pairs displayed in the trial.

For the response options, there were four response options: a) working memory intrusion option, b) long term memory intrusion option, c) new, never seen option, and d) the target response option. The working memory intrusion options was randomly drawn from somewhere in the list of four words show in the trial. The long-term memory intrusion option was randomly drawn from words used in the LTM task but not from the current WM trial. The new, never seen options were drawn from a random list of word pairs from a previous study used in our laboratory, but they were not shown at all during the study. Lastly, the response options were also options that were previously used in our laboratory.

Working Memory Encoding and Recognition Task

Word pairs were randomly selected to be working memory stimuli (*N* = 80 word pairs). They were randomly assigned into list type (e.g., 0%, 25%, 50%, and 75%, *N* = 32 word pairs per list). For the response options, the words used as within-trial lures were randomly drawn from one of the four word pairs displayed in the trial. The same word randomization process for LTM lists was repeated for WM lists, meaning, the within-trial intrusion lures and LTM-trial lures were randomized again using the same process described in the previous sentence, and an additional 128 words were used for the non-shown lure choice. All option choices for recognition tests were randomized upon appearance.

I will briefly define the language for the design used prior to describing the task: 1) trial = one word pair, 2) WM set = a set of four word pairs. For the WM task, each word pair was shown for 1000 ms followed by a 500 ms interstimulus interval. Each set included four word pairs. In each set, we manipulated the amount of pre-learned pairs from the LTM learning phase. This word list manipulation included sets in which none (0%; *N* = 8 sets), one pair (25%; *N* = 8 sets), two pairs (50%; *N* = 8 sets), or three pairs (75%; *N* = 8 sets) in the set were pre-learned information. Each of these sets ended in either a blank screen or a filler task. Half of the sets were followed by a blank screen (i.e., 16 WM sets; 4 sets of no pre-learned items with blank, 4 sets of 1 pre-learned items with blank, etc.) and half were followed by a filler task (i.e., 16 WM sets; 4 sets of no pre-learned items with math equations, 4 sets of 1 pre-learned items with math questions, etc.)), for a total of 32 WM sets (see Figure 6).

After the WM encoding and filler tasks (i.e., blank screen or filler) for one WM set concluded, WM recognition was assessed through a four alternative force choice task (4-AFC), which consisted of a cue word and four response options (see Figure 6). The participant's task was to identify the target word that was originally paired with the cue word during the WM encoding phase. In line with Bartsch and Shepherdson (2023), the answers contained four of the following: 1) the target word that was correctly paired with the word, 2) an item paired with another word in the WM task (within-trial intrusions), 3) an item presented in the LTM learning phase that was not in the WM task, and 4) a new item that was not shown in the task. Participants were given unlimited time to make their selection on each recognition trial. They completed the 4-AFC task for each of the 4 word pairs in the WM set, one at a time. Then, participants were given a new set of 4 word pairs to encode and the cycle was repeated (e.g., four word pairs for encoding, filler task, four word pairs for testing). This continued until the participant completed all 32 WM sets.

Working Memory Encoding Filler Task

For half of the trials, participants engaged in a filler task to take up the focus of attention/WM resources. The filler task consisted of a series of math equations, and the participant were asked to solve the equation. In this task, participants were shown half of a math equation (e.g., 5) on one screen for 1000 ms before being given the next half of the equation on a second screen (e.g., + 6) for 1000 ms. This was followed by a screen that asked the participant "What is the answer?" Participants had an unlimited time to answer. Once they had responded, a new math equation appeared. This process continued until 10 seconds elapsed, at which point they saw a blank screen for 1000 ms to let them finish pressing any keys on the keyboard without accidentally clicking through trials before being directed to the recognition test. For the other half of the trials, participants saw a blank screen for 10 seconds with the additional 1000 ms screen for consistency.

Long-Term Memory Recognition Task

Participants completed an LTM task, which consisted of a similar 4-AFC procedure as the WM task. On each trial, participants were presented with the cue word and 4 response options that consisted of a target word, another item paired with a different word in the WM task phase (WM lure), an item presented in the LTM learning phase but with the not current LTM cue word (LTM lure), and a new item. Participants worked through all 48 LTM words at their own pace. (It is also worth noting that 59 participants were shown 128 word pairs, which were all WM and LTM word pairs used in the study, at their own pace. After further discussion regarding the timing of the experiment and potential fatigue, this was reduced back down to the original 48 word pairs.)

Strategy Reports

Participants were also asked to report the strategies they used to remember all the word pairs that were shown throughout the study (e.g., word pairs in both LTM and WM task). Participants were simply asked, "What strategy did you use to remember this word pair?" and were shown a list with four, force-choice options: 1) I repeated the words in my head, 2) I created an image of the word, 3) I used a sentence to link the words together, 4) I didn't try to remember the words.

Procedure

Participants were recruited via SONA Systems, which is an online recruitment platform for students in psychology courses. Participants were run in groups of 1-3 people. Upon entering the laboratory, they read and signed a consent form outlining the risks and benefits of the study.

Participants also completed a demographics questionnaire. Then, a research assistant gave verbal instructions for how to navigate through the computer tasks and an overview of the goals in each part of the study. First, participants completed the Symmetry Span task, which consisted of two phases (e.g., calibration phase and memory testing phase). They completed two practice trials for the calibration phase and the memory testing phase before proceeding to task. Next, the researcher would come over to the participant's station and switch the computer program to PsychoPy (Peirce et al., 2019), which contained the rest of the experiment. Then, participants were instructed to begin the LTM learning phase. Then, they completed the WM encoding phase, the filler task, and WM testing phase. After the WM task, participants completed the LTM testing task. After the LTM testing, participants were asked to complete the strategy report. Following completion of the study, participants were thanked for their participation in the study, received a debriefing statement explaining the importance of the study, and given credit for their time. The experiment took anywhere from about 45 minutes to 75 minutes to complete.

Chapter 3 - Results

Statistical Approach

All analyses were conducted in R (R Core Team, 2024, version 4.4.2). Mean values reported in the analyses were obtained using the 'emmeans' library (Lenth, 2020). Tukey's post-hoc analyses for all categorical independent variables were also performed using 'emmeans.' To evaluate the significance of fixed effects, the 'car' library (Fox & Weisburg, 2019) to conduct an analysis of deviance (Type III Wald Chi-Square Test), facilitating ease of interpretation.

A generalized multilevel mixed-effects (MLM) model was implemented using the lme4 package (Bates et al., 2015). This model investigated the fixed effects of working memory capacity (WMC), Filler Task, and List Type, as well as random effects, including the random intercept (representing initial performance within trials) and random slope (rate of performance change across trials) for participants across Filler Task and List Type. The outcome variable was modeled as a binomial variable (0/1), with 1 indicating correct responses. Both Filler Task (two levels: Math or Blank) and List Type (four levels: 0% Old, 25% Old, 50% Old, and 75% Old) were treated as categorical variables and effect coded for all analyses. WMC was treated as a continuous predictor and was mean-centered. This model was used to test the following effects:

1) Main effect of List Type (RQ1).
2) List Type × Filler Task interaction (RQ2).
3) WMC × List Type × Filler Task interaction (RQ3).

The random effect structure was adapted from Bartsch et al. (2023). The base optimizer in the lme4 package uses the BOBYQA algorithm to constrain optimization parameters (Powell, 2009). However, this optimizer may not be appropriate for some datasets, potentially leading to model convergence failures. To address this issue, I used allFit() to test a range of optimizers on the model to identify the most suitable option for my dataset. Of the optimizers tested, only *nlminbwrap*, *L-BFGS-B*, and *nlopt_ln_neldermead* successfully converged. The *nlminwrap* optimizer was selected for the final model due to its ability to handle boundary constraints (e.g., data constrained to 0 and 1) and its capacity to manage complex nonlinear interaction models (Gay, 1990).

General Descriptives

Performance across all three tasks (e.g., symmetry span, WM recognition, and LTM recognition) was negatively skewed. Performance on the processing component of the Symmetry Span ranged from 80% – 100% (*M* = 91.57%, *SD* = 4.72; see Figure 7A). Working memory recognition ranged from 34.38% – 95.31% (*M* = 68.09%, *SD* = 14.64; see Figure 7B). Long term memory recognition ranged from 26.56% – 100% (*M* = 63.36%, *SD* = 20.46; see Figure 7C). Working memory capacity ranged from .20 – .98 (*M* = .71, *SD* = .15; see Figure 7D).

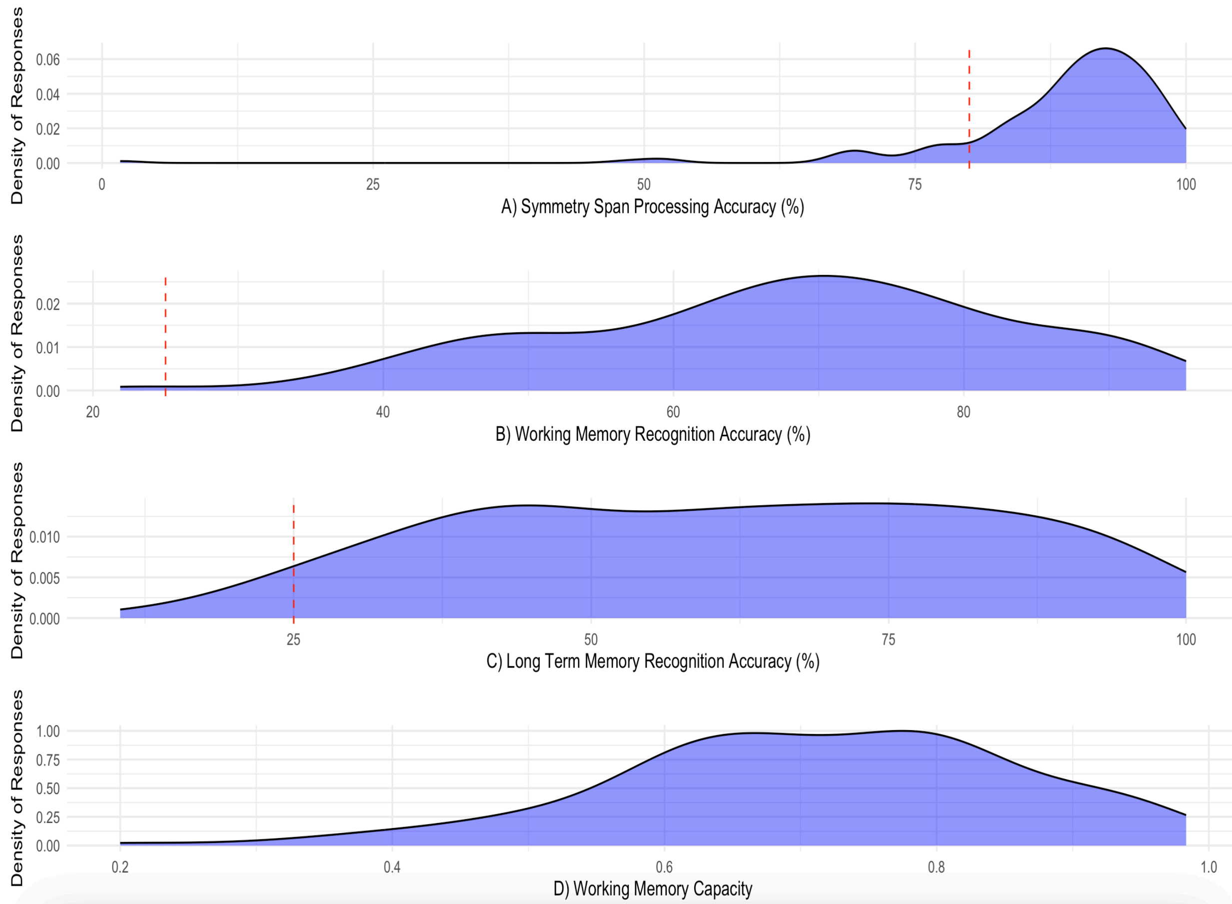

Figure 7. Density plots depicting the negatively skewed distribution of performance (in percentage (%)) for all three tasks: A) Symmetry Span Task (SST), B) Working Memory Recognition Accuracy (WMR), C) Long Term Memory Recognition Accuracy (LTMR), and D) Working Memory Capacity (WMC). The y-axis represents the density of the response (e.g., the higher the value, the more responses fell there). The red line in each graph shows the cutoff criteria used in each task (e.g., 80% in SST, 25% in WMR and LTMR).

To evaluate the residuals of the multilevel logistic regression I used the "DHARMa" package (Hartig, 2022, version 0.4.6; Figure 8). This approach assesses the model's ability to predict error in the data. The DHARMa package also provides tests for dispersion, outliers, and the Kolmogorov-Smirnov (KS) test. None of these tests indicated significant issues with the model residuals. The residuals are well-distributed and accurately predicted by the model.

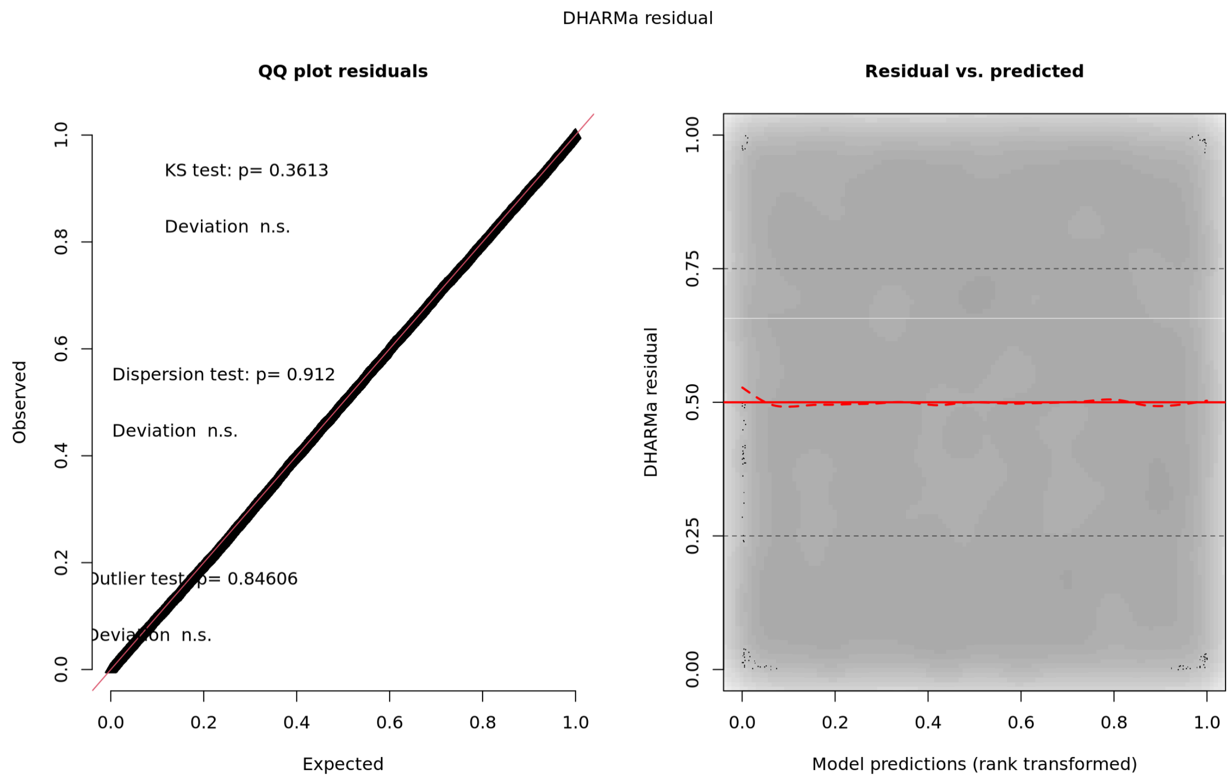

Figure 8. The left panel, "QQ Plot Residuals," compares the expected residuals from the model to the observed residuals. Ideally, the black dots should align closely with the red line, indicating minimal deviation. As shown in the figure, there is very little divergence between the black dots and the red line, suggesting good model fit. The right panel, "Residual vs. Predicted," assesses the alignment of the model residuals with predicted residuals. Here, the dotted line should align closely with the solid red line. Once again, minimal deviation is observed, further supporting the adequacy of the model.

Hypothesis 1: Main effect of List Type

My first set of hypotheses aimed to understand whether pre-existing episodic LTM representations benefit WM performance. My two competing hypotheses were: Pre-existing episodic LTM representation H1a) would or H1b) would not benefit WM performance. To evaluate these hypotheses, I analyzed the main effect of list type (e.g., 0% old, 25% old, 50% old, and 75% old), which varied the number of pre-existing episodic LTM representations.

There was a significant main effect of list type ($\chi^2(3) = 123.89$, $p < .001$; see Figure 9). Participants had a higher probability of responding correctly when presented with lists containing 75% pre-learned information ($M = 0.77$, $SE = 0.01$) compared to 0% pre-learned lists ($z = 11.00$, $p < .001$; $M = 0.66$, $SE = 0.01$), 25% pre-learned lists ($z = 6.93$, $p < .001$; $M = 0.70$, $SE = 0.01$), and 50% pre-learned lists ($z = 5.05$, $p < .001$; $M = 0.72$, $SE = 0.01$). Moreover, lists with 0% pre-learned information had a significantly lower probability of responding correctly compared to all list types (see Table 1 for a list of all post hoc comparisons). These findings demonstrate that increasing pre-learned information significantly enhances the likelihood of correct responses, highlighting the influence of episodic LTM on WM task performance.

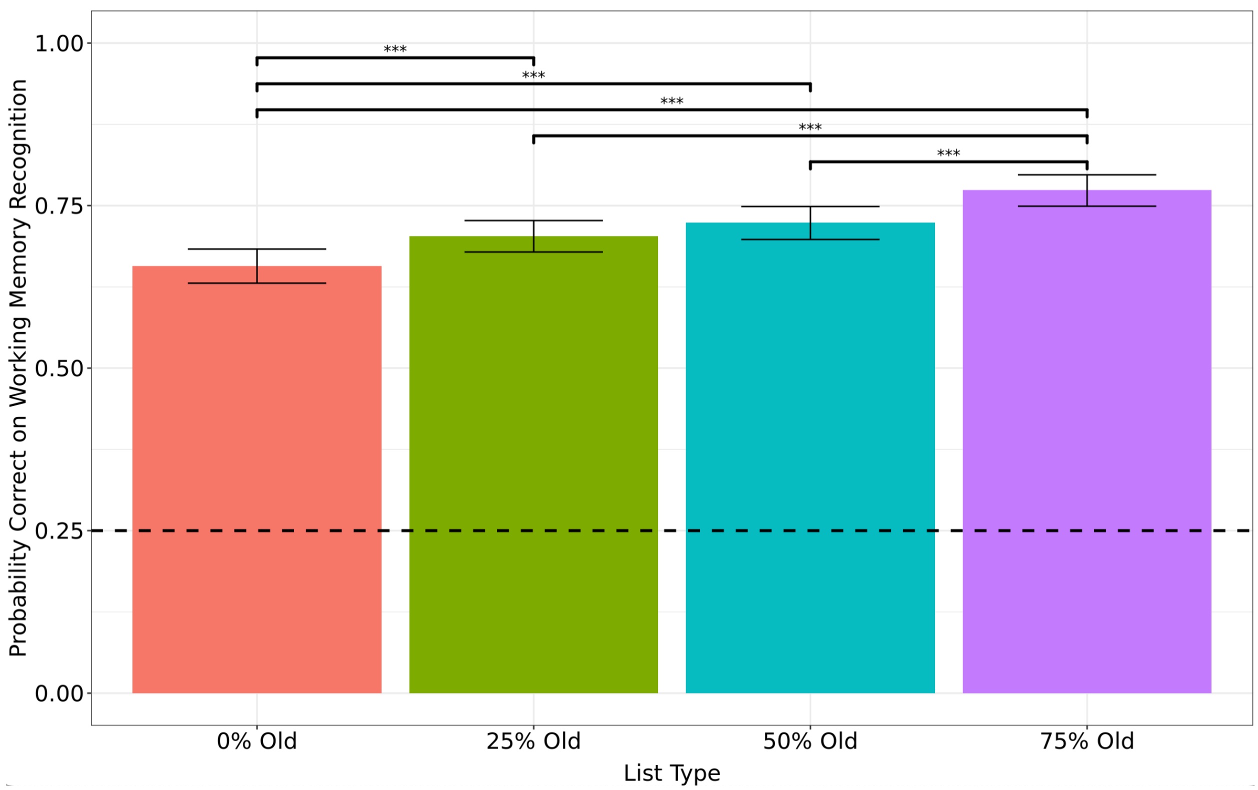

Figure 9. Predicted probability correct on the working memory recognition test by list types. The dotted black line represents 25% chance on the task. Error bars represented 95% confidence intervals.

Table 1. Post hoc analyses for main effect of List Type

Contrast	Mean Odds Ratio	*SE*	*z*-ratio	*p*-value
0% Old – 25% Old	.81	.04	-4.67	< .001
0% Old – 50% Old	.73	.04	-6.42	< .001
0% Old – 75% Old	.56	.03	-11.00	< .001
25% Old – 50% Old	.90	.04	-2.06	.17
25% Old – 75% Old	.69	.04	-6.93	< .001
50% Old – 75% Old	.77	.04	-5.05	< .001

Hypothesis 2: Interaction between List Type and Filler Task

My second set of hypotheses aimed to understand whether individuals offload information represented in WM to LTM to free up resources for subsequent information. My two competing hypotheses were the offloading hypothesis and the boosting hypothesis. The offloading hypothesis is that previously learned information can be offloaded to (i.e., maintained in) LTM without the use of WM. If this hypothesis is supported, then the math filler task should impair WM performance on lists with new word pairs only more so than performance on lists with pre-learned word pairs because the math filler task should have minimal effect on the offloaded information maintained in LTM. The boosting hypothesis states that LTM affects information *within* WM. If this is correct, then information is not offloaded to LTM but rather is maintained within WM and I should observe that WM performance across all tasks is impaired by the math filler task compared to the blank screen, regardless of whether the lists are comprised of new or pre-learned word pairs.

To evaluate these hypotheses, I analyzed the interaction between list type (e.g., 0%, 25%, 50%, and 75% old items) and Filler Task (e.g., math or blank) on the probability of correctly responding on the WM recognition task, and it was not significant, $\chi^2(3) = 5.44$, $p = .14$ (see

Figure 10). This indicates that the effect of Filler Task (e.g., blank or math equation) on probability correct did not depend on which List Type participants were currently completing. Interestingly, though, Bartsch and Shepherdson (2023) also did not find a significant List Type x Filler Task interaction for offloading, but instead, found that probing the group differences offered conflicting evidence to fully support an offloading account. Because of this, I conducted a Tukey's post-hoc analysis to investigate differences in performance between different levels of my list type and filler task variables.

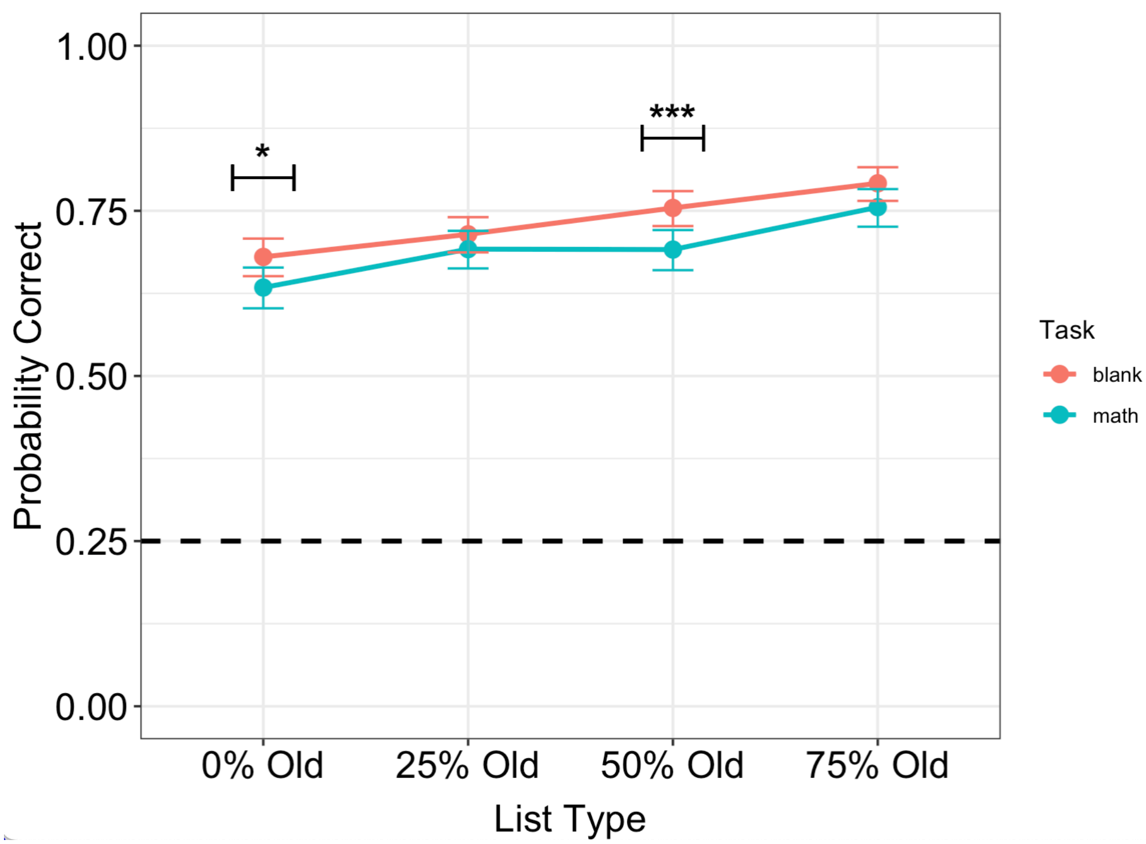

Figure 10. Predicted probability correct on the working memory recognition test by list types when trials ended in either a blank screen (pink, salmon line) or math task (blue, teal line). The dotted black line represents 25% chance on the task. Error bars represented 95% confidence intervals.

In line with the boosting hypothesis, we found that WM performance was significantly impaired on math filler task trials compared to blank trials for lists containing 0% Old items (M_{blank} = 0.68, SE_{blank} = .01 vs. M_{math} = 0.63, SE_{math} = .02) and lists containing 50% Old items (M_{blank} = 0.75, SE_{blank} = .01 vs. M_{math} = 0.69, SE_{math} = .02; see Figure 10). However, the boosting account cannot explain the fact that we observed no differences in WM performance between math filler task and blank trials on lists containing 25% (M_{blank} = 0.72, SE_{blank} = .01 vs. M_{math} = 0.69, SE_{math} = .01) and 75% Old items (M_{blank} = 0.79, SE_{blank} = .01 vs. M_{math} = 0.76, SE_{math} = .01). In other words, if all information is always maintained in WM (and not offloaded to LTM), then why was performance unaffected by the math filler task on these lists?

Table 2. Post hoc analyses for interaction between List Type and Filler Task

Contrast	**Mean Odds Ratio**	***SE***	***z*-ratio**	***p*-value**
0% Old Blank – 25% Old Blank	.85	.05	-2.56	.17
0% Old Blank – 50% Old Blank	.69	.05	-5.47	< .001
0% Old Blank – 75% Old Blank	.56	.04	-8.17	< .001
0% Old Blank – 0% Old Math	1.23	.08	3.26	.02
0% Old Blank – 25% Old Math	.95	.06	-0.84	.99
0% Old Blank – 50% Old Math	.95	.06	-0.77	.99
0% Old Blank – 75% Old Math	.69	.05	-5.23	< .001
25% Old Blank – 50% Old Blank	.82	.06	-3.03	.05
25% Old Blank – 75% Old Blank	.66	.05	-5.83	< .001
25% Old Blank – 0% Old Math	1.45	.09	5.67	< .001
25% Old Blank – 25% Old Math	1.11	.07	1.66	.71
25% Old Blank– 50% Old Math	1.12	.08	1.64	.73
25% Old Blank– 75% Old Math	.81	.06	-2.93	.07
50% Old Blank – 75% Old Blank	.81	.06	-2.94	.07
50% Old Blank – 0% Old Math	1.77	.12	8.35	< .001
50% Old Blank – 25% Old Math	1.37	.10	4.49	< .001
50% Old Blank – 50% Old Math	1.37	.09	4.74	< .001
50% Old Blank – 75% Old Math	.99	.07	-0.09	1.00
75% Old Blank – 0% Old Math	2.20	.16	10.84	< .001
75% Old Blank – 25% Old Math	1.69	.12	7.14	< .001
75% Old Blank – 50% Old Math	1.70	.12	7.24	< .001
75% Old Blank – 75% Old Math	1.23	.09	2.93	.07
0% Old Math – 25% Old Math	.77	.05	-4.21	< .001

0% Old Math – 50% Old Math	.77	.05	-4.02	.002
0% Old Math – 75% Old Math	.56	.04	-8.51	< .001
25% Old Math – 50% Old Math	1.00	.06	0.05	1.00
25% Old Math – 75% Old Math	.73	.05	-4.63	< .001
50% Old Math – 75% Old Math	.73	.05	-4.71	< .001

Next, I conducted a model fit analysis to investigate whether the null model (with no interaction) was preferred over the interaction model. This approach allows researchers to understand how well different models fit a particular dataset. If model fit indices are higher for the interaction model, then the added (offloading) interaction better fits the current dataset compared to the model without the interaction. Following this approach, I created an interaction model that included the main effects of Filler Task and List Type, their interaction and the random effects and a null model that included only the main effects and the same random effect structure. I then compared the Akaike Information Criterion (AIC) values, in which lower values indicate better model fit (Vrieze, 2012): the null model AIC = 27719.91 and the interaction model AIC = 27720.82, resulting in a difference (ΔAIC) of 0.91. Using the formula for the likelihood ratio between models,

L(Mi|Mni)=exp(−12Δi(AIC)),L(M_i|M_{ni}) = \exp\left(-\frac{1}{2}

\Delta_i(\text{AIC})\right),L(Mi|Mni)=exp(−21Δi(AIC)), the null model was 0.63 times as likely to produce the observed data as the interaction model. This is a negligible difference between models, indicating that both models are equally likely to produce the observed data. Thus, while we replicated the overall pattern of results reported by Bartsch and Shepherdson (2023), our findings also indicate that the model containing the offloading effect (i.e., List Type x Filler Task) does not improve model fit over the null model for this dataset.

Evaluating Offloading with Word Type in Lieu of List Type (Exploratory)

To more directly compare our results to Bartsch and Shepherdson (2023), I reran this analysis with word type (e.g., old vs. new items) instead of List Type. The analysis kept the same random effect structure and contrast coding discussed in the Statistical Approach section above. I will evaluate the two-way interaction (Word Type x Filler Task), post-hoc analyses between our conditions, and comparisons between an interaction and no-interaction model.

First, I analyzed the interaction between word type (e.g., new and old word pairs) and Filler Task (e.g., math or blank) on the probability of correctly responding on the WM recognition task, and it was not significant, $\chi^2(1) = 0.35$, $p = .55$ (see Figure 11). The effect of Filler Task (e.g., blank or math equation) on probability correct did not depend on which Word Type participants studied. Again, we chose to continue with post-hoc analyses to evaluate if they align more with a boosting account or offloading account. Both theories would predict recognition performance of new items will be worse on the math filler trials than the blank trials; however, these accounts have different predictions for the old items. According to the boosting account, performance should still be worse on math vs. blank trials but, according to the offloading account, old items would be less affected by the math filler task.

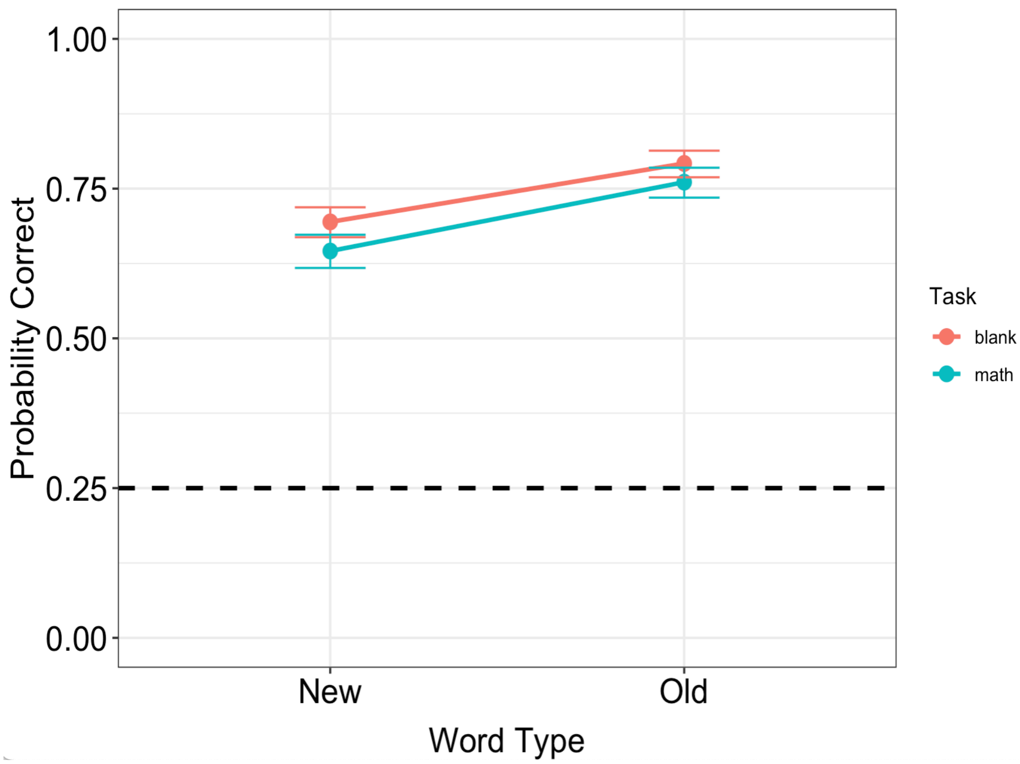

Figure 11. Predicted probability correct on the working memory recognition test by word type when trials ended in either a blank screen (pink, salmon line) or math task (blue, teal line). The dotted black line represents 25% chance on the task. Error bars represent 95% confidence intervals.

We found that WM performance was significantly impaired on math filler task trials compared to blank trials for both new items (M_{blank} = 0.69, SD_{blank} = .02 vs. M_{math} = 0.65, SD_{math} = .01) and old items (M_{blank} = 0.79, SD_{blank} = .01 vs. M_{math} = 0.76, SD_{math}= .01; see Table 3). Again, the fact that participants have a decrease in performance while studying new items in the math filler task is not definitive support for either theory given that they can both account for this. The primary difference lies in whether performance on old items is different between task type. Given that we did not find support for an interaction, this result aligns more with a boosting account as it predicts performance is harmed in the distractor task condition, regardless of word type. However, like the above post-hoc analysis, mean difference between conditions is relatively small (~3-4% difference) and should be interpreted with caution.

Table 3. Post hoc analyses between Word Type and Filler Task

Contrast	Mean Odds Ratio	*SE*	*z*-ratio	*p*-value
New Blank – Old Blank	0.60	.03	-10.22	< .001
New Blank – New Math	1.25	.05	5.15	< .001
New Blank – Old Math	.72	.04	-6.39	< .001
New Math – Old Math	0.57	.03	-11.45	< .001
Old Blank – New Math	2.09	.11	13.77	< .001
Old Blank – Old Math	1.20	.07	3.05	.01

Lastly, like the model comparison in the above analysis, the null model was 2.29 times likely to produce the observed data compared to the interaction model. This is weak evidence in favorite of the no interaction model, and thus, does not provide evidence for the offloading account.

Hypothesis 3: Three-Way Interaction of WMC x List Type x Filler Task

Although there was no offloading effect (the List Type x Filler Task interaction was not significant), it is possible that offloading may depend on working memory capacity. My third set of hypotheses were related to whether working memory capacity affects offloading. My two competing hypotheses were: 1) WMC will be predict offloading such that higher WMC will be associated with offloading; 2) WMC will not predict offloading. To evaluate these hypotheses, I analyzed the three-way interaction between WMC (continuous predictor), List Type (e.g., 0% old, 25% old, 50% old, and 75% old) and Filler Task (e.g., math or blank) on WM recognition. The three-way was not significant, $\chi^2(3) = 0.84$, $p = .84$ (see Figure 12). This null effect indicates that WMC was not associated with the likelihood of offloading. I also ran this analysis with the exploratory model with Word Type in lieu of List Type. Like the main model, this three-way was not significant, $\chi^2(1) = 0.02$, $p = .88$. Implications of this will be discussed later.

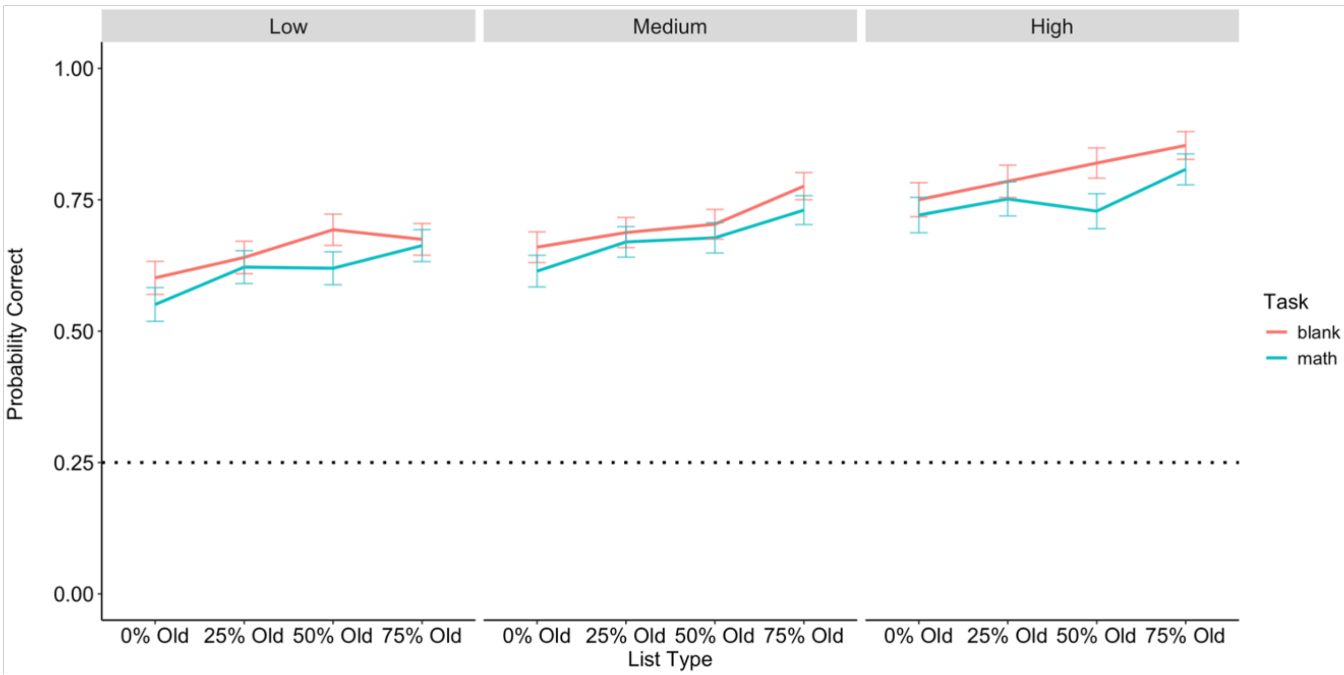

Figure 12. Raw mean probability correct on the working memory recognition test by list types when they ended in either a blank screen (pink, salmon bar) or math task (blue, teal bar). The x-axis is List Type and the panels represent low (bottom 33rd quartile), medium (between 33rd to 66th quartile), and high (above 66th quartile) WMC. The dotted black line represents 25% chance on the task. There was no three-way interaction, but there was a main effect of task, list type, WMC. There was one two-way interaction between WMC and task. Error bars represented 95% confidence intervals.

Additional Analyses

The main crux of this dissertation proposal focuses on the planned analyses for the specified hypotheses; however, there were other effects that we found in the study when analyzing the full model. I will describe the rest of the three-way interaction for full transparency of results.

Unsurprisingly, I observed a main effect of WMC, $\chi^2(1) = 26.34$, $p < .001$, such that those with higher WMC ability had a higher probability of correctly recalling words (Figure 13; Wilhelm et al., 2013; Rosen & Engle, 1997; Cantor & Engle, 1993; Just & Carpenter, 1992).

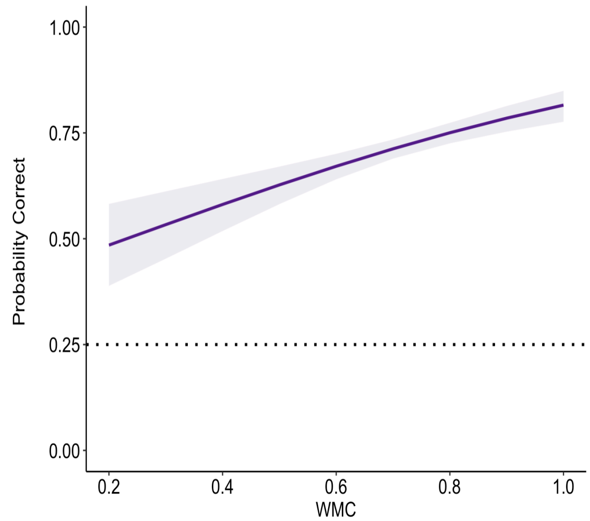

Figure 13. Graphical depiction of the main effect of WMC, plotted on the x-axis, on probability correct for the WM recognition task. As WMC increases, the probability of correctly recognizing the word that belongs in a word pair increases. The dotted line represents chance performance (25%). Error bars represented 95% confidence intervals.

Further, there was a main effect of Filler Task, $\chi^2(1) = 32.46$, $p < .001$, with the probability of correctly recalling words being higher when the filler task was a blank screen ($M = .74$, $SD = .01$) compared to when it involved solving the filler math task ($M = .70$, $SD = .01$; Figure 14). Again, this finding replicates past work on task interference in which distractors task that engage WM decrease overall performance because of the need for competing cognitive resources, like

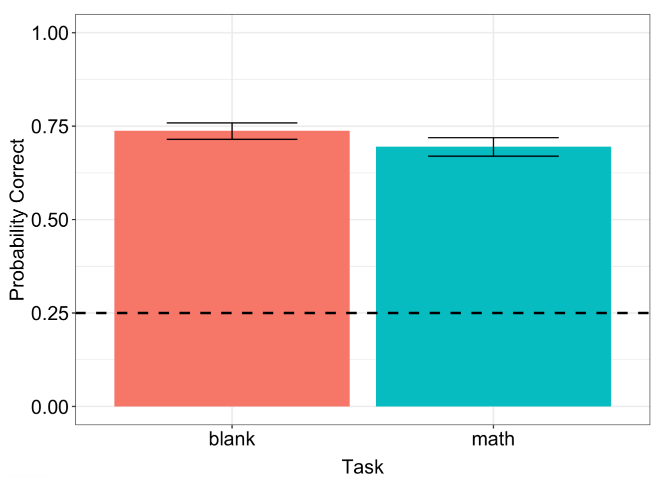

Figure 14. The probability of correctly recognizing the word that belongs in a word pair increases when the word pairs end in a blank screen (pink bar) compared to the math filler task (blue bar). The dotted line represents chance performance (25%). Error bars represented 95% confidence intervals.

attention (Geffen et al., 1997; Kane & Engle, 2000; Cowan et al., 2005; Unsworth & Engle, 2007).

Next, I observed a significant WMC x Filler Task interaction, $\chi^2(1) = 4.88$, $p = .03$ (see Figure 15). At lower levels of WMC, there is no difference in performance between list types that end with a blank screen compared to those that end with math equations. However, as WMC increases, the difference in performance between task increases, such that those with higher WMC were more likely to correctly remember words when there was a blank screen versus math problems as the filler task.

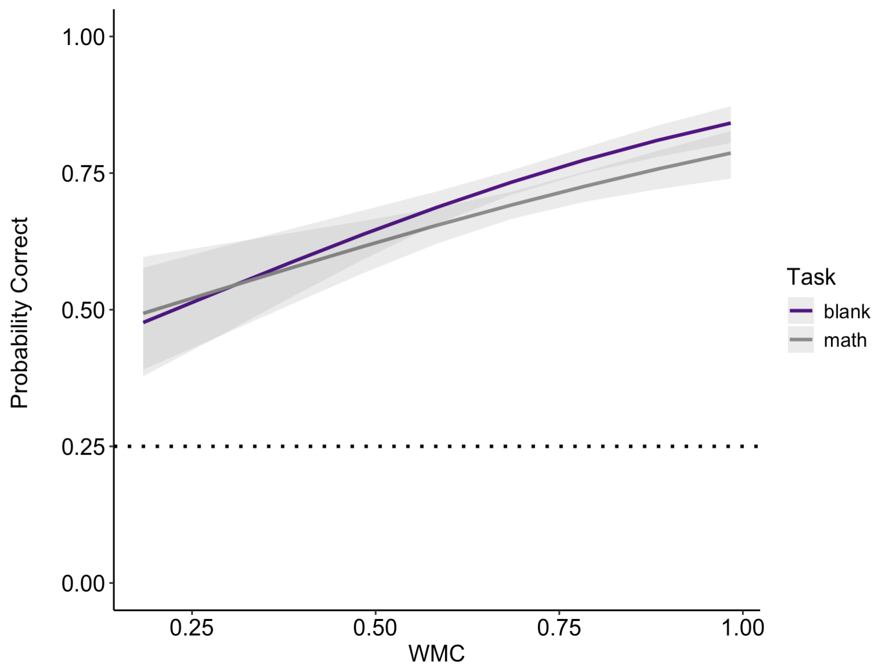

Figure 15. At low levels of WMC, the probability of correctly recognizing the word that belongs in a word pair is not moderated by task; however, as WMC increases, performance for word pairs that end in blank task increases compared to when they end in math distractor task. The dotted line represents chance performance (25%).

Given that high WMC individual generally do better on cognitive tasks, we wanted to follow up this analysis with a brief probe of how many problems low, medium, and high WMC individuals completed. WMC was split into three groups with those in the upper 66th quartile classified as "high WMC," below 33rd percentile classified as "low WMC," and scores in between as "middle WMC". High WMC individuals (M = 26.7, SD = 6.7, range = 15 - 44) completed more average math problems than low WMC individuals (M = 22.6, SD = 5.6, range = 8 - 34). A quick one-way ANOVA showed that there were differences in performance on problems, $F(2,162) = 5.01$, $p = .01$. Specifically, high WMC individuals completed more math problems than low WMC individuals ($M_{diff} = 4.15$, $p = .01$). Middle WMC individuals completed similar math problems to both low WMC individuals ($M_{diff} = 1.70$, $p = .33$) and high WMC individuals ($M_{diff} = 2.45$, $p = .14$).

Self-Reported Strategies

It is also possible that this effect could relate to the strong link between strategies and working memory capacity. Past research has found that high WMC individuals are more likely to use more elaborative strategies, such as sentence generation and imagery (Bailey et al., 2008; Dunlosky & Kane, 2007; Unsworth & Spillers, 2010). It is possible that one reason for this interaction effect is that strategies are driving the performance difference for high spans. In other words, when WM is not disrupted by a math filler task, high spans may be making use of the extra time to implement an elaborative strategy that, in turns, helps increase memory performance more during the blank filler task.

I calculated the proportion of trials that each strategy was reported as well as WM performance as a function of the strategy the participant reported using on a given trial. First, participants report using repetition more often than the other strategies (see Table 4).

Table 4. Proportion of trials participants reported using each strategy

Less Effective Strategies	Effective Strategies		
Repeat	Sentence	Imagery	Didn't Try
0.31 (0.01)	0.23 (0.01)	0.23 (0.01)	0.23 (0.01)

Note: Parentheses = standard error of the mean

To investigate whether strategies had a role in the WMC x Filler Type interaction, I ran a generalized multilevel logistic regression evaluating the effects of WMC, Filler Task, Strategy and their interactions on WM recognition performance. Strategy choice was treated a categorical variable with four levels: Did Not Try, Repetition, Sentence Generation, and Imagery. Give that strategy choice and task were within-subjects variables, I randomized the intercept and slope for participants across strategy choice and task. The original model failed to converge, so I used the allFit() command to assess an appropriate control. All controllers failed to lead to any convergence, so I reverted to a basic RE structure (e.g., randomized the intercept and slope for participants across list type and task). This model converged with the standard BOBYQA method with no additional controllers needed.

The final model showed support for the three main effects of WMC ($\chi^2(1) = 8.41$, $p < .001$), Filler Task ($\chi^2(1) = 14.49$, $p < .001$), and Strategy Choice ($\chi^2(3) = 969.11$, $p < .001$). I used a Tukey's HSD to test significant differences in performance between the strategy choice levels (main effect of Strategy Type; see Figure 16). I found that when participants reported they Did Not Try, they had significantly lower performance ($M = .53$, $SE = .02$) than when they reported doing repetition ($M = .69$, $SE = .01$; $z = 12.91$, $p < .001$), sentence generation ($M = .82$, $SE = 0.01$; $z = 24.5$, $p < .001$), and imagery ($M = .84$, $SE = 0.01$; $z = 27.03$, $p < .001$). Further, repetition had significantly lower performance compared to both sentence generation ($z = 13.96$,

$p < .001$) and imagery ($z = 16.58$, $p < .001$), but there was no significant difference in performance between elaborative strategies ($z = 2.35$ $p = .09$). These effects replicate past work

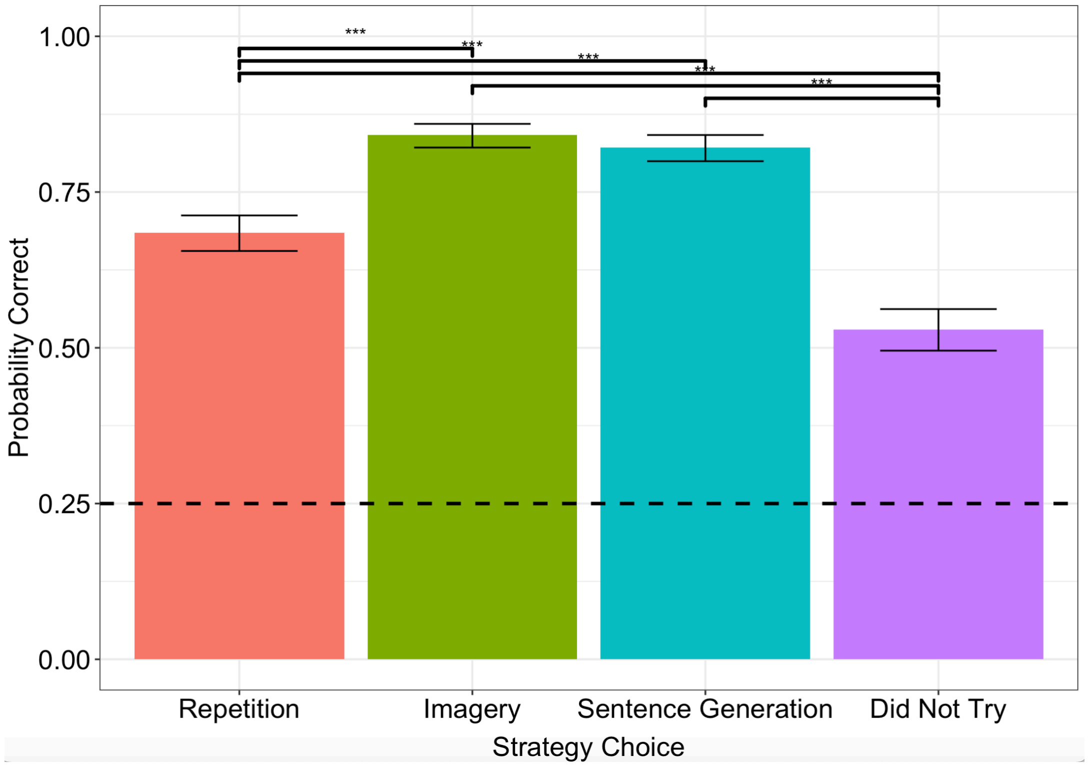

Figure 16. Predicted probability correct on the WM recognition task between strategy choices. The dashed, black line represents chance performance. Elaborative strategies (e.g., sentence generation and imagery) had higher probability correct compared to both non-elaborative strategies (e.g., repetition) and not trying. Non-elaborative strategies had significantly higher performance compared to not trying. Elaborative strategies were not different in performance from one another.

showing that people remember more information when they report using an elaborative compared to a non-elaborative strategy (Bailey et al., 2008; Bailey et al., 2011; Dunlosky & Kane, 2007).

There was a significant interaction between WMC and Strategy Choice, $\chi^2(3) = 10.31$, $p = .02$ (see Figure 17). Using emtrends(), I tested whether the slope of WMC differed depending on the strategy used. We found that WMC significantly predicted performance while using various strategies. The slopes for each of the strategies were the following: Imagery ($b = 2.36$, *SE* = 0.45, 95% CI [1.49, 3.24]), Did Not Try (*b* = 1.82, *SE* = 0.41, 95% CI [1.01, 2.63]), Sentence Generation (*b* = 1.59, *SE* = 0.45, 95% CI [0.72, 2.47]), and Repetition (*b* = 1.17, *SE* = 0.40, 95% CI [0.38, 1.96]). Elaborative strategies (e.g., Imagery and Sentence Generation) had steeper slopes than non-elaborative strategies (e.g., Repetition). Pairwise comparisons of each strategy's slope showed that the effect of WMC was significantly greater when participants used Imagery than Repetition (b_{diff} = -1.20, *SE* = 0.38, *z* = -3.15, *p* = .01). The slope analysis suggests

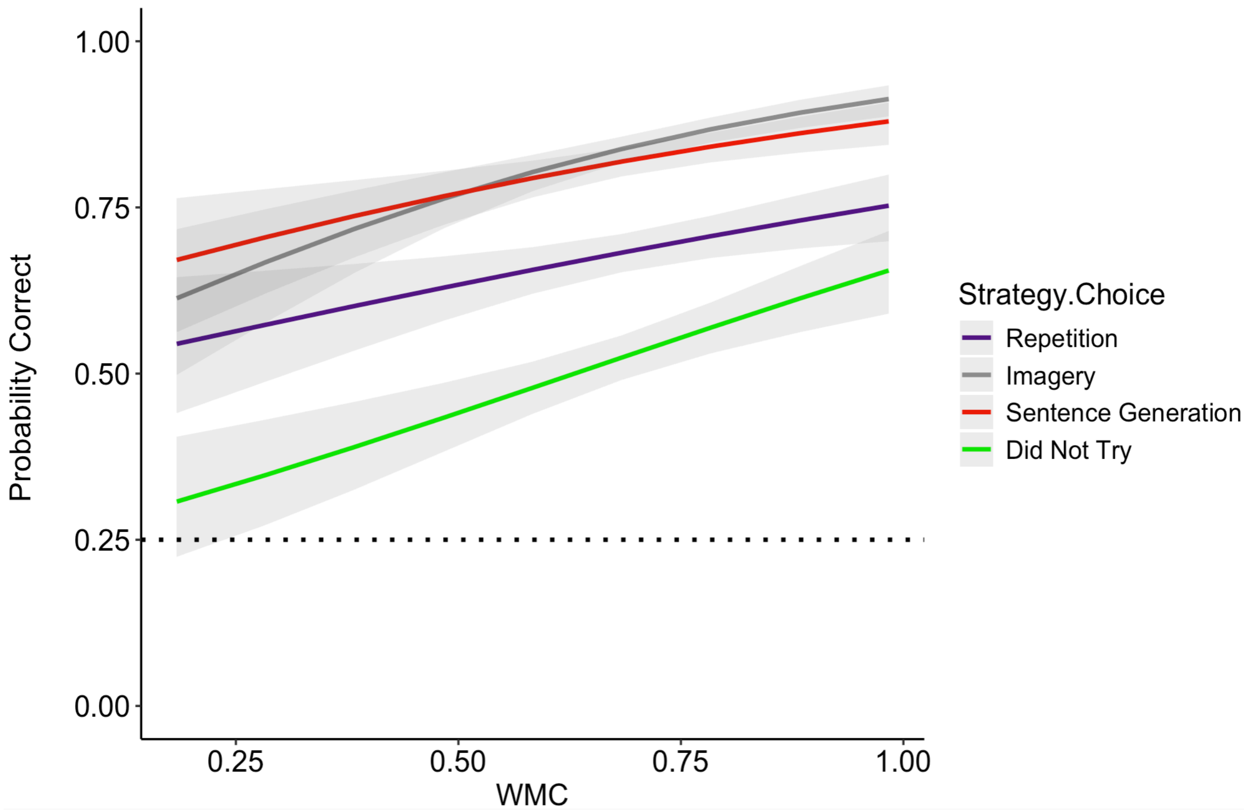

Figure 17. Predicted probability correct on the WMC recognition task by WMC and Strategy Choice. The dotted, black line represents chance performance.

that working memory was strongly associated with performance when participants used elaborative strategies. This further suggests that elaborative strategies resulted in better performance compared to non-elaborative strategies overall, but this difference in performance was more apparent for those with lower WMC compared to those with higher WMC. Meaning, higher WMC individuals benefited more in performance from effective strategies (e.g., imagery and sentence generation) while those with lower WMC saw less of a performance boost.

No other interaction, including the WMC x Filler Task interaction, the Filler Task x Strategy Choice interaction, nor the three-way interaction for WMC, Task, and Strategy Choice were significant. Thus, it seems unlikely that strategies are the explanation for why those with higher WMC remember more on blank trials compared with math filler task trials, whereas those with lower WMC do not show a difference.

Chapter 4 - Discussion

The purpose of this proposal was two-fold: 1) to replicate the newly proposed mechanism, offloading, and 2) to extend this line of research to evaluate whether offloading depends on individual differences in WMC. Offloading is proposed to be an explanation for how we can increase immediate memory by relying on LTM representations and thus freeing up WM resources (Bartsch & Shepherdson, 2022; Bartsch & Shepherdson, 2023; Bartsch, Frischkorn, & Shepherdson, 2024). This dissertation explored offloading as a potential explanation for low spans' limited capabilities (e.g., lower attentional control and interference management; Unsworth, 2019; Heibel & Zimmer, 2015; Kane et al., 2001; Mecklinger et al., 2003; Kane & Engle, 2000). It is possible that low spans do not offload information to LTM and instead overburden WM resources. If this is the case, then high spans may outperform low spans on tasks that require more WM resources (e.g., math filler task) because they alleviate the demand on WM by using episodic memory representations in LTM. To address these goals, the current study was built around three main research questions described below.

RQ1: Do pre-existing episodic LTM representations benefit WM performance?

We aimed to answer evaluate whether pre-existing episodic LTM representations benefit WM performance. I hypothesized two competing hypotheses to answer this question: H1a) WM recall performance would be higher for lists with pre-learned items compared to lists with new items; H1b) WM recall performance would not differ between lists with and without pre-learned items. We supported **H1a** and replicated prior studies that demonstrating that LTM aids WM (Cermak, 1970; Arzi et al., 1985; Shiffrin & Atkinson, 1969; Unsworth & Engle, 2006; Beukers et al., 2021; Bartsch & Shepherdson, 2023). Our research supports the idea that prior exposure to information does increase the likelihood of later recalling it.

Theoretical Implications for Hybrid Models

The results of our study align well with assumptions made by hybrid models of memory, like the Embedded Process Theory (EPIC; Cowan, 1999, 2008; Cowan et al., 2021) and Embedded Component Model (EC; Oberauer, 2002, 2021). For Cowan's EPIC model, the process would occur because the central executive transfers information from the focus of attention directly to activated long-term memory. Previously learned items would exist within activated long-term memory (or even long-term memory, depending on one's familiarity with word pairs or strategy for remembering the words). When the participant sees the cue word, participants can narrow their search down to the correct cue more accurately for previously learned words than newly learned items. This same logic applies to Oberauer's EC model of memory. When participants see a cue word, their focus of attention will select one of the activated representations in the region of direct access. Previously learned items are more likely to be in the region of direct access, presumably because of prior exposure, and are more readily selected as an answer. In either framework, this prior exposure facilitates more efficient retrieval when participants are exposed to previously learned information.

RQ2: Do individuals offload information to LTM?

An important part of the research process is replication of prior research studies. For this proposal, I sought to replicate effects from Bartsch and Shepherdson (2023) that supported the idea that individuals offload information to LTM. To evaluate this question, I have two competing hypotheses to explain potential effects. The *offloading hypothesis* (H2a) states that previously learned information can be stored in LTM without the need for active maintenance in WM. If this account were supported, then performance on lists containing old pairs should be relatively unaffected by the math filler task since these pairs can be retrieved from LTM, whereas new word pairs would rely upon WM maintenance. Thus, we would not expect to see impairments on math filler trials compared to blank trials on lists that contain new word pairs. In

By contrast, according to the *boosting hypothesis* (H2b), all information must be maintained in WM; therefore, a math filler task should prevent information from being maintained in WM, which in turn, will impair performance compared to when that information is maintained in WM with no math filler task. Importantly, this means we would expect that WM performance on blank trials should always be higher than performance on math filler trials, regardless of whether the word pairs are old or new, because the attentional resources needed to maintain information in WM will be allocated to the math filler task rather than the to-be-remembered word pairs.

In the current study, the overall analysis (List Type x Filler Task) did not provide support for the offloading hypothesis. It is important to note that Bartsch and Shepherdson (2023) also found weak evidence in favor of a model including an offloading effect. However, they did find evidence in support of offloading when looking at how word type (e.g., new vs. old items) moderated performance during task. Their study found that WM performance was more affected by the math filler task on lists with new pairs only compared to lists with old pairs (i.e., "offloading" effect). Thus, we conducted post-hoc analysis to parallel those of Bartsch and Shepherdson (2023). The results of our post-hoc analyses of the Filler Task x List Type interaction do not fully support either hypothesis.

We found evidence that supports the boosting account under some conditions. The offloading hypothesis claims that the old word pairs should be represented in LTM and, thus, not affected by a WM math filler task; however, we found that WM performance was significantly lower on math filler trials compared to blank trials on lists with 50% old word pairs, providing support for a boosting account. We also found that evidence that supports the offloading hypothesis under other conditions. Specifically, we did not find a significant difference in performance between the math filler task and the blank screen on lists containing 25% and 75% older words. Further, both hypotheses predict that there should be a significant difference (blank > math filler task) on lists containing all new pairs (i.e., 0% old), which is what we found in the current study. Thus, we are unable to fully support either hypothesis. It should be noted, though, that even on the lists in which significantly more words were remembered on blank versus math filler task trials, the actual difference in performance was quite small (see Figure 10; approximately 5% difference in performance in the 0% and 50% old lists). This effect will be discussed further in the Limitations and Future Directions section below.

Again, to parallel Bartsch and Shepherdson (2023), we chose to conduct an additional analysis that collapsed across list type and looked at a three-way interaction with item type (e.g., old vs. new items) in hopes to clarify our results. Similar to the List Type analysis, we found no overall two-way interaction. When I ran post-hoc analyses, the results showed that both new items and old items had lower performance when they ended in a math filler task compared to the blank filler task. This indicates that, regardless of word type, performance was generally lower when participants were required to remember interfering stimuli (i.e., numbers) than when they were given time to maintain information freely, and the information that was presumably represented in LTM was still affected by the interfering task. This result aligns more with the boosting account, which would predict a decrease in performance for the math filler task compared to the blank filler task, regardless of word type. Much like the differences in the List Type x Task interaction, the actual difference in performance were small (approximately 3-4% between Task).

To conclude, it is hard to reconcile our conflicting results to fully support evidence of offloading. Our model comparisons do not support interaction models (List Type/Word Type x Task) over no interaction models, indicating that adding the offloading effect did not produce better model fit indices for our analyses. The two-way interaction analyses also do not support evidence of a significant interaction in support of offloading. Adopting Bartsch and Shepherdson (2023) original approach to offloading (e.g., Word Type x Task) also did not show support for offloading. It is only when we probe the post-hoc analyses in our List Type x Task interaction that we see evidence in support of offloading (albeit, small performance differences) that cannot be explained fully by boosting. Given our original hypotheses and analysis (List Type x Task), I will conclude that we do not support either H2A or H2B as neither theory can fully account for our conflicting results.

Theoretical Implications for Hybrid Models

Once again, the two hybrid models of memory discussed above can account for these results. In both Cowan's and Oberauer's model, information is exchanged between WM and LTM between the focus of attention (Cowan and Oberauer) to activated LTM (Cowan) or the region of direct access (Oberauer). When the focus of attention is at or beyond capacity (e.g., 1 item or 4 items), information may not be transferred to LTM. According to the boosting account, since both new and old items are maintained in WM, then introducing a math filler task would exceed the capacity of WM and require participants to strategically reallocate WM resources from the focus of attention to complete the math filler task. Doing so would significantly decrease the likelihood of successfully transferring the to-be-remembered words to activated LTM or the region of direct access for later retrieval.

On the other hand, the offloading account can be supported with a similar, but nuanced, line of thinking. Specifically, if previously learned items are already stored in activated long term memory or the region of direct access, there is no need for those items to be maintained solely in the focus of attention. Instead, they can be accessed directly from activated LTM or region of direct access without allocating attentional resources to maintain them in WM. Rather attentional resources can be allocated to maintaining novel information as well as solving equations in the math filler task.

RQ3: Do individual differences in WMC predict offloading?

The crux of this proposal was evaluating whether individual differences in WMC is associated with offloading. Thus, I evaluated two competing hypotheses: H3a) WMC will be positively associated with offloading (i.e. those with higher WMC will offload more); H3b) WMC will not be associated with offloading (i.e., individuals with varying levels of WMC will offload similarly). This research question was evaluated through the three-way interaction between WMC, List Type and Filler Task. This three-way interaction was not statistically significant, supporting neither H3a nor **H3b**, which suggests that WMC does not moderate offloading in our study. This result suggests that WMC differences may not be due to the extent to which people exchange information between WM and LTM. Instead, individual differences in WMC could be due to alternative explanations, such as attentional capacity differences, domain-specific skills, and strategies. Although we cannot differentiate between these alternative explanations in the current study, the current study adds to the growing literature for individual differences in WM performance that shows there are individual differences in WMC that are associated with high and low WM performance (Conway, 1996; Engle et al., 1999; Unsworth, 2010; Miller & Unsworth, 2018) and that our study failed to find that people were offloading differently based on these differences.

Although the 3-way interaction was not significant, there was a significant WMC x List Type interaction such that, for those with higher WMC, memory performance was higher on the trials followed by a blank screen than trials followed by a math filler task. On the other hand, those with lower WMC showed no difference in performance on trials with a blank screen compared to a math filler task. We did find that high WMC completed significantly more math problems than low WMC; therefore, the math filler task may have created more interference for the high WMC participants than it did for the low WMC participants.

However, I also evaluated whether this interaction was explained by self-reported strategy use, but strategy choice did not interact with either list type or the WMC and List Type interaction. The only interaction we did find was between WMC and strategy choice. Specifically, elaborative strategies (e.g., imagery and sentence generation) led to improved performance compared to non-elaborative strategies (e.g., repetition), but this performance benefit was apparent in those with higher WMC. So, although this was not specifically the point of the proposal, we did replicate past research that showed higher WMC individuals are more likely to use elaborative strategies efficiently (Daneman & Carpenter, 1980; Just & Carpenter, 1992; McNamara & Scott, 2001; Engle et al., 1999).

Theoretical Implications for Hybrid Models

Although we did not find evidence that supported WMC-related differences in offloading, hybrid memory models can still provide guidelines for how and why this might happen. The primary focus of my proposal was geared toward understanding how differences in WMC affect encoding processes, specifically, offloading. Instead, I found that participants in my study were offloading similarly to one another, regardless of their WMC. It is possible that WMC differences influence processes involved in retrieval of information more than encoding of information. One feature that both hybrid models specify is that LTM representations of recently encoded information have a special status (e.g., in activated LTM or region of direct access). Perhaps, instead of individuals differing in how they handle attentional resources in the focus of attention, they differ in their ability to use cues to retrieve information from activated LTM and/or region of direct access, with high spans using retrieval cues that activate less (but more precise) information in LTM compared to low spans (Unsworth et al., 2012; Unsworth, 2019). Further, if activated LTM and region of direct access is constrained by the limited capacity of the focus of attention (e.g., 1 item or 4 items), individual differences in WMC may systematically influence the efficiency of attentional control and/or resource allocation. High WMC individuals, who can maintain more information or perhaps more efficiently utilize their focus of attention, would be more likely to handle the interference of the math filler task while transferring newly learned information to activated long term memory or the region of direct access.

These ideas discussed above that high WMC may better handle interference better conflicts with the results we found that high WMC were more affected by the math filler task.

However, it is possible that high WMC participants might invest more cognitive resources during encoding, making their strategies more susceptible to interruption by secondary tasks (like the math filler task). However, once encoding is complete, high-WMC individuals are also more adept at using precise retrieval cues, which helps them recover from interference more effectively during retrieval.

Limitations and Future Directions

One potential limitation for our study exists within the assumption for why offloading exists. Presumably, the concept of offloading depends on exceeding one's limitation capacity. Meaning, it is necessary to "offload" information to LTM rather than keeping all information in WM and exceeding one's capacity limitations. Both Oberauer's region of direct access and Cowan's focus of attention has a has a capacity limit of 4 items (Oberauer, 2002, 2013, 2021; Cowan, 1999, 2009, 201l; Cowan et al., 2021). Based on the limitations put forth by these theories, the current study and that conducted by Bartsch and Shepherdson (2023) used trials capped at 4 word pairs. With that said, the number of word pairs is, theoretically, near the top of everyone's capacity limit but not necessarily exceeding it. Thus, it is possible that our participants may have been able to successfully maintain all four items within WM without the need to offload to LTM. Thus, offloading may still occur, but this possibility could explain the limited evidence we found in support of offloading. Further, having list lengths that do not exceed WMC may also explain why WMC did not interact with offloading.

In line with this possibility, a follow up study by Bartsch, Frischkorn, and Shepherdson (2024) explored the extent to which LTM influences WM when exposed to varying amounts of cognitive load. Across a set of four studies, they varied to-be-remembered items from 2-, 3-, 4- or 6-word pairs in a set. They found that when memory load was low (e.g., 2 word pairs), pre-learned items in episodic LTM did not benefit WM performance. However, as memory load increased, previously learned LTM representations were helpful for WM performance. This lends support to the need for future research to replicate these effects (e.g., with lists of to-be-remembered items exceeding the typical WMC), especially when evaluating the influence of individual differences in WMC. Doing so could strengthen the effect of offloading by forcing participants into situations in which they either offload (successfully) or suffer a decrease in performance.

A second potential limitation of our current study, and of Bartsch and Shepherdson (2023), that may also underestimate the effect of offloading that occurs in the real world is the amount and type of exposure to the previously learned items. Prior to the WM task, participants were only exposed to pre-learned items once (e.g., during the LTM learning phase). We assume that participants are creating a strong episodic LTM representation; however, past research has shown that LTM improves when people practice retrieving information as compared to when they only study information (the *testing effect*; e.g., Roediger & Karpicke, 2006; McDaniel et al., 2007) and when information is accurately retrieved at least three times (*criterion learning, diminishing returns;* Pyc & Rawson, 2009; Bjork & Bjork, 2013). Our study did not ask participants to retrieve any of their pre-learned items until the final testing phase. Thus, adding one or two immediate "retrieval phases" would likely strengthen the episodic LTM representation of the pre-learned items and potentially increase one's ability to engage in offloading.

Along this same vein, a key assumption to offloading is that for items to be successfully offloaded to LTM, they must be effectively encoded. In both our study (*M* = 63.36%, SD = 20.46) and Bartsch and Shepherdson (*M* = 70.7%, *SD* = 18.5), participants recalled a substantial portion of word pairs, which suggests both samples were successfully encoding the to-be-remembered items to some degree. However, the final analyses in both studies examined overall WM performance for all word pairs, regardless of whether those word pairs were successfully retrieved during the LTM testing phase. This raises an important question: To what extent are we measuring offloading if items we assume to be offloaded were never stored in LTM? A potential solution to this issue would involve only analyzing word pairs that were successfully recalled on the LTM test. By isolating these items, we can gain a more precise understanding of how offloading contributes to WM performance. In other words, we may have underestimated the size of the offloading effect because the analyses included both information that was later successfully and unsuccessfully retrieved from episodic LTM. Unfortunately, the current study was only powered to detect the offloading effect on the total number of trials (not remembered trials only). Thus, future work could better evaluate this possibility by including more trials or potentially having participants learn word pairs to a criterion.

The final potential limitation to our study is that our measure of WMC may not be sensitive enough to detect differences in offloading. For time's sake, we could only reliably collect data using one WMC assessment in addition to the main tasks. We chose to use a pre-programed assessment of the Symmetry Span task (Monteiro et al., 2024), which only measures one facet of WMC (i.e., visuospatial skills). It is better to use multiple assessments of WM using a variety of span tasks (e.g., spatial, verbal, simple, complex) to create a composite score for WMC. This more reliably captures WM ability and minimizes the influence of domain-specific abilities associated with processing tasks (e.g., math/OSPAN; reading comprehension/RSPAN; etc.) on overall performance.

Conclusion

The primary purpose of this study was to explore the memory mechanism of "offloading." Offloading refers to a process in which information that is already stored in episodic LTM is not actively held in WM, allowing more space for new information that is not yet in LTM. Our main goals were to replicate previous findings that: 1) pre-existing LTM representations benefit WM performance, 2) individuals offload pre-existing items to LTM, freeing up WM resources, and 3) individual differences in WMC influence individuals' ability to offload. We replicated past work showing that pre-existing LTM representations benefit WM performance. However, the results provided mixed support for the idea that participants offload these representations to episodic LTM, as some evidence suggested that they may continue to occupy WM resources. Additionally, we found no evidence that offloading was related to an individuals' WMC. Future research should explore methods to increase the likelihood that participants engage in offloading, which will help clarify the extent to which this mechanism occurs. Overall, our findings underscore the benefit of LTM in enhancing WM performance, while highlighting areas for further investigation.

References

Arzi, H. J., Ben-Zvi, R., & Ganiel, U. (1985). Proactive and retroactive facilitation of long-term retention by curriculum continuity. *American Educational Research Journal*, *22*(3), 369-388.

Aslan, A., & Bäuml, K. H. T. (2011). Individual differences in working memory capacity predict retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 264.

Astle, D. E., Nobre, A. C., & Scerif, G. (2012). Attentional control constrains visual short-term memory: Insights from developmental and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(2), 277-294.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*(2), 82-91.

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition*, *36*, 1383-1390.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*(4), 553.

Bartsch, L. M., Frischkorn, G. T., & Shepherdson, P. (2024). When Load is Low, Working Memory is Shielded From Long-Term Memory's Influence. *Journal of Cognition*, *7*(1), 44.

Bartsch, L. M., Loaiza, V. M., Jäncke, L., Oberauer, K., & Lewis-Peacock, J. A. (2019). Dissociating refreshing and elaboration and their impacts on memory. *NeuroImage*, *199*, 585-597.

Bartsch, L. M., & Shepherdson, P. (2022). Freeing capacity in working memory (WM) through the use of long-term memory (LTM) representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(4), 465.

Bartsch, L. M., & Shepherdson, P. (2023). Chunking, boosting, or offloading? Using serial position to investigate long-term memory's enhancement of verbal working memory performance. *Attention, Perception, & Psychophysics*, *85*(5), 1566-1581.

Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory & Cognition*, *46*, 796-808.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences*, *4*(11), 417-423.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1),54–75.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Usinglme4. *Journal of Statistical Software*, *67*(1).

Bjork, R. A., & Bjork, E. L. (2013). Optimizing treatment and instruction: Implications of a new theory of disuse. In *Memory and Society* (pp. 109-133). Psychology Press.

Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: an individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1101.

Cermak, L. S. (1970). Proactive facilitation in short-term memory. *Journal of Experimental Psychology*, *85*(2), 305.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81.

Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: a reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1235.

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: The limit in words retained without covert articulation. *Quarterly Journal of Experimental Psychology*, *62*(7), 1420-1429.

Chubala, C. M., Neath, I., & Surprenant, A. M. (2019). A comparison of immediate serial recall and immediate serial recognition. *Canadian Journal of Experimental Psychology*, *73*(1), 5.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*, 991-997.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497-505.

Conway, A. R. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*(6), 577-590.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control*, *20*, 506.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in Brain Research*, *169*, 323-338.

Cowan, N, Morey, C. C., & Naveh-Benjamin, M. (2021). An embedded-process approach to working memory: How is it distinct from other approaches, and to what ends?. In Logie, R., Camos, V., and Cowan, N (Eds.), *Working Memory: State of the Science.* (pp. 116-149). Oxford University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *12*, 588-607

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *Quarterly Journal of Experimental Psychology*, *60*(9), 1227-1245.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The Nature of Remembering: Essays in honor of Robert G. Crowder* (pp. 297- 314). Washington, DC: American Psychological Association Press.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 972.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309.

Fox, J., & Weisberg, S. (2019). Using car functions in other functions. *CRAN R*.

Gay, D. M. (1990). Usage summary for selected optimization routines. *Computing Science Technical Report*, *153*(153), 1-21.

Geffen, G. M., Wright, M. J., Green, H. J., Gillespie, N. A., Smyth, D. C., Evans, D. M., & Geffen, L. B. (1997). Effects of memory load and distraction on performance and event-related slow potentials in a visuospatial working memory task. *Journal of Cognitive Neuroscience*, *9*(6), 743-757.

Green, P., & MacLeod, C.J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493-498.

Hiebel, N., & Zimmer, H. D. (2015). Individual differences in working memory capacity and attentional control. *Canadian Journal of Experimental Psychology*, *69*(1), 17.

Hoskin, A. N., Bornstein, A. M., Norman, K. A., & Cohen, J. D. (2019). Refresh my memory: Episodic memory reinstatements intrude on working memory maintenance. *Cognitive, Affective, & Behavioral Neuroscience*, *19*, 338-354.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, *92*(2), 137.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*(3), 153-161.

Kowialiewski, B., Gorin, S., & Majerus, S. (2021). Semantic knowledge constrains the processing of serial order information in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(12), 1958.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1118.

LaRocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It'sa matter of attention. *Frontiers in Human Neuroscience*, *8*, 5.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61-79.

Lörch, L., Lemaire, B., & Portrat, S. (2023). A Hebbian Model to Account for Musical Expertise Differences in a Working Memory Task. *Cognitive Computation*, 1-20.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*(4-5), 494-513.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*, 10-17.

Mecklinger, A., Weber, K., Gunter, T. C., & Engle, R. W. (2003). Dissociable brain mechanisms for inhibitory control: effects of interference content and working memory capacity. *Cognitive Brain Research*, *18*(1), 26-38.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Miller, A. L., & Unsworth, N. (2018). Individual differences in working memory capacity and search efficiency. *Memory & Cognition*, *46*(7), 1149-1163.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they?.*Journal of Experimental Psychology: General*, *118*(1), 43.

Monteiro, F., Nascimento, L. B., Leitão, J., Santos, E. J., Rodrigues, P., Santos, I. M., ... & Nascimento, C. S. (2024). OpenWMB: An open-source and automated working memory task battery for OpenSesame. *Behavior Research Methods*, *56*(7), 6880-6903.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53-81.

Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 948.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411.

Oberauer, K. (2005). Control of the contents of working memory--a comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 714.

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45-100.

Oberauer, K. (2013). The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in Human Neuroscience*, *7*, 673.

Oberauer, K. (2021). Towards a theory of working memory: From metaphors to mechanisms. In Logie, R., Camos, V., and Cowan, N (Eds.), *Working Memory: State of the Science.* (pp. 116-149). Oxford University Press.

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity?. *Psychological Bulletin*, *142*(7), 758.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*, 779-819.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195-203.

Poole, B. J., & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, *62*(7), 1430-1454.

Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, *26*, 26-46.

Portrat, S., Guida, A., Phénix, T., & Lemaire, B. (2016). Promoting the experimental dialogue between working memory and chunking: Behavioral data and simulation. *Memory & Cognition*, *44*, 420-434.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, *60*(4), 437-447.

Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin & Review*, *5*(4), 597-614.

Robison, M. K., Miller, A. L., & Unsworth, N. (2018). Individual differences in working memory capacity and filtering. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(7), 1038.

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

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136-1139.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*, 1839-1860.

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, *91*(3), 694-707.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 37.

Tse, C. S., & Altarriba, J. (2022). Independent effects of word concreteness and word valence on immediate serial recall. *British Journal of Psychology*, *113*(3), 820-834.

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence*, *38*(2), 255-267.

Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin*, *145*(1), 79.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language*, *67*(1), 1-16.

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, *33*(2), 213-220.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104.

Unsworth, N., McMillan, B. D., Brewer, G. A., & Spillers, G. J. (2012). Everyday attention failures: an individual differences investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1765.

Unsworth, N., & Miller, A. L. (2021). Individual differences in the intensity and consistency of attention. *Current Directions in Psychological Science*, *30*(5), 391-400.

Unsworth, N., & Miller, A. L. (2024). An examination of individual differences in levels of processing. *Memory*, *32*(5), 615-626.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*, 200-205.

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*(2), 228.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pederson, T. L., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), Article 1686.

Wickham H, François R, Henry L, Müller K, Vaughan D (2023). dplyr: A Grammar of Data Manipulation. R package version 1.1.4, https://github.com/tidyverse/dplyr.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it?. *Frontiers in Psychology*, *4*, 433.

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*, 123.