Neural oscillations in memory strategies

by

Trevor J. Bell

B.A., Christian Brothers University, 2018

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Dr. Heather Bailey

Abstract

We rely on our long-term memory (LTM) system to remember a wide range of information on a daily basis. However, we know from experience, whether forgetting a necessary item on a grocery list or answers to a test question, we are not always able to accurately recall that information. What determines how well information is stored in LTM? One factor is how memory stimuli are encoded. Memory strategies are designed to help counteract failures by providing memory cues to associate with necessary memory stimuli. Past literature has evaluated the effects of strategies on memory accuracy, but less is known about the neural mechanisms underlying such strategies. The current proposal aims to alleviate this gap by using electroencephalography (EEG) while participants completed a modified version of a Paired-Associates Learning task and reported which strategy they used to encode the paired words. Overall, the proposal answered three key research questions: 1) What neural patterns are active in the brain when we experience a memory failure?, 2) Do effective memory strategies differ neurophysiologically from less effective memory strategies?, and 3) Do neural signatures validate strategy reports and memory performance? The current proposal used time-frequency analyses to look at multiple processes occurring in the brain at once. It is proposed that the theta band will help participants actively encode and refresh information in short-term memory (i.e., attention or processing), gamma band will help participants store memory representations (i.e., storage), and alpha will work to suppress irrelevant information during the task. More importantly, the current proposal is interested in how theta and gamma work together to shift information to LTM. We found neural evidence that memory failures exhibit lower theta power and differing gamma power depending on how much information is required to remember (e.g., 3 sets of word pairs vs. 10 sets of word pairs). We also found evidence that effective strategies recruit higher gamma power, which could be related to storing or tying current information in LTM. Lastly, we found that brain patterns based on “effective” and “less effective” behavioral data is hard to dissociate, but we suggest that more information is needed to explore this question. Specifically, the unexpected alpha power enhancement and gamma suppression indicates that there is more to these bands than previously thought. Exploratory analyses also found an increase in beta power for Less Effective – Correct strategies compared to all other conditions. It is possible that beta power could be an internal rehearsal loop for strategy types. Future research is needed to understand what alpha and beta enhancement is doing and replicate the pattern of gamma suppression found in the current study.

Key words: memory failures, time frequency, short-term memory, long-term memory

Table of Contents

List of Figures....................v
List of Tables.................... vii
Acknowledgements viii
Neural Oscillations in Memory Strategies 1
  Theories About the Relationship Between Short- and Long-Term Memory 1
  Memory Encoding Strategies 4
  Electroencephalographic Correlates of Memory....................6
  Frequency Bands Involved in Memory 8
  Research Summary....................14
Research Questions (RQ) and Hypotheses (H) 15
Pilot Experiment....................17
  Methods 17
  Results and Discussion 21
Main Experiment....................21
  Methods 21
Behavioral Results....................26
  Set Size 28
  Strategy....................29
EEG Preprocessing & Analysis Procedure....................31
  Electrodes of Interest....................32
  Event-Related Spectral Perturbations (ERSPs) Procedure....................33
EEG Results....................35
  Memory Performance....................37
  Strategy Type....................42
  Strategy Type x Performance 46
  Set Size 52
  Exploratory EEG Analyses....................57
General Discussion 60
  Neural Markers of Memory Failures....................60

Neural Markers of Encoding Strategies 63
Neural Markers of Effectiveness of Encoding Strategies..65
Theoretical Relevance 66
Practical Relevance 70
Limitations..71
Future Directions 72
References 74
Appendix A - Pilot Experiment Word Stimuli 86
Appendix B - Multilevel Modeling Replication of ANOVA Results 97
Appendix C - ERSPs 102

List of Figures

Figure 1 *Frequency bands of interest in the proposal* 8

Figure 2 *Theta and gamma coupling dynamics* 13

Figure 3 *Modified version of the paired associates learning task (A) and the cued recall task (B)* 19

Figure 4 *Example of a strategy report* 20

Figure 5 *Distribution of the proportion of correctly recalled words. The distribution is roughly normal with a negative skew* 27

Figure 6 *The mean proportion of correctly recalled words by set size* 28

Figure 7 *The mean proportion of correctly recalled words by strategy type* 30

Figure 8 *Topography of the electrode distribution for the EGI system from Smith (2021)* 32

Figure 9 *Power values for all frequency bands averaged across all participants, channels, and conditions* 35

Figure 10 *The mean theta power value for when participants correctly recalled or incorrectly recalled words* 37

Figure 11 *(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for performance only* 39

Figure 12 *Mean gamma power value for correctly and incorrectly recalled words* 40

Figure 13 *Mean alpha power value for correctly and incorrectly recalled words* 40

Figure 14 *Topography scalp map for distribution of correlation strength between theta and gamma power across the scalp for A) when participants recalled words correctly and B) when participants recalled words incorrectly* 41

Figure 15 *Violin plot showing the spread of the theta-gamma coupling power values for performance* 41

Figure 16 *Mean theta power value for effective and less effective strategies* 42

Figure 17 *Mean gamma power values for effective and less effective strategies* 42

Figure 18 *(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for strategies only* 44

Figure 19 *Mean alpha power values for effective and less effective strategies* 45

Figure 20 *Scalp distribution of theta-gamma correlation strength across all electrodes for participants when they were A) using effective strategies or B) using less effective strategies* 45
Figure 21 *Violin plot showing the spread of the theta-gamma coupling power values for strategy type* 46
Figure 22 *Theta power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words* 47
Figure 23 *(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for strategies, depending on performance* .48
Figure 24 *Gamma power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words* 49
Figure 25 *Alpha power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words* 50
Figure 26 *Scalp distribution of correlation strength between theta and gamma values for when participants were using different strategies and whether they correctly recalled words* 51
Figure 27 *Violin plot depicting the spread of the datapoints for strategy types by performance* 52
Figure 28 *Theta power values for each of the set sizes* 53
Figure 29 *(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for set sizes* 54
Figure 30 *Gamma power values for set sizes* 55
Figure 31 *Topography distributions of the correlation in theta-gamma values across set sizes* .55
Figure 32 *Violin plot depicting the spread of the theta-gamma correlation values for set sizes* .56
Figure 33 *Mean beta power values for performance* 57
Figure 34 *Mean beta power values between types of strategies* 58
Figure 35 *Mean beta power values between strategies depending on whether participants later correctly recalled words* 59
Figure 36 *Mean beta power recruitment between set sizes* 60
Figure 37 *Distribution of scores on the paired-recall task* 97
Figure 38 *Set size was predictive of recall performance* 99
Figure 39 *Type of strategy to remember words was predictive of recall performance* 100

List of Tables

Table 1 *Summary statistics of word characteristics used in the pilot experiment* 18

Table 2 *Summary statistics of word characteristics to be used in experiment* 23

Table 3 *Proportion of trials reported using each strategy while studying word pairs* 29

Table 4 *Total number of trials for each condition within participants in the EEG analyses* 34

Table 5 *Proportion of trials reported using each strategy while studying word pairs* 98

Acknowledgements

I would like to acknowledge Dr. Matt Wisniewski and Dr. Alexandria Zakrzewski for their expertise and guidance on this project. I would also like to acknowledge the many undergraduates who helped with data collection: Kelly Krehbiel, Alexys Anguiano, and Annie Harper. Without their help, data collection and cleaning would not have gone as smoothly. Lastly, I would like to acknowledge the Cognitive and Neurobiological Approaches to Plasticity (CNAP) Center of Biomedical Research Excellence (COBRE) of the National Institutes of Health under grant number P20GM113109 at Kansas State University. CNAP provided participant compensation as well as protective gear for running subjects during the COVID-19 pandemic. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

Neural Oscillations in Memory Strategies

Short-term memory (STM) and long-term memory (LTM) are necessary for everyday use, such as taking notes in the classroom, grocery shopping, cooking favorite recipes, and often, are used in tandem with one another. STM is the ability to store information while simultaneously processing other information, whereas LTM is the extended storage of information, such as memory for events, actions, or factual information (Atkinson & Shiffrin, 1968). Yet, there are failures in both types of memory where important information (e.g., answers to an exam question) is lost, which can have detrimental effects on subsequent test performance, for instance. Fortunately, some of these failures can be counteracted using strategies, which promote successful encoding and retrieval (Richardson, 1998; McNamara & Scott, 2001; Dunlosky & Kane, 2007; Turley-Ames & Whitfield, 2003). Giving rise to a need to understand how these strategies are implemented and effective for both LTM and STM memory retrieval.

Theories About the Relationship Between Short- and Long-Term Memory

Theories about how working memory/short-term and long-term memory interact disagree about a few issues. One of the main issues is the extent to which STM and LTM are from distinct memory systems: one class of theories claims that STM and LTM are unitary (e.g., there is no differentiation – all STM is LTM) whereas the other class claims that STM and LTM are non-unitary (e.g., separate processes from one another).

Importantly, unitary theories do not differentiate between STM and LTM (e.g., Nairne, 2002); instead, these theories propose that there are differing levels of activation such that the information currently being attended to (e.g., what others consider STM) is highly activated, whereas information that is not in the focus of attention (e.g., what others consider LTM) is less activated. In other words, information currently being worked with is not stored in a separate storage buffer, but instead, is used as a retrieval cue to search through memory representations. These cues allow the participant to reconstruct the context in which a memory occurred, thus, boosting performance on the task (Nairne, 1988, 1990). By this account, failures in memory constitute a poor retrieval cue to narrow the search of information in memory. In addition, STM and LTM also share similar temporal functions (i.e., you can extinguish primacy effects while still having recency effects and vice versa) and do not require separation from one another (Crowder & Neath, 1991).

On the other side, proponents for non-unitary models argue that these two concepts are completely distinct based upon their duration (Ebbinghaus, 1875/1913), capacity (Miller, 1956; Atkinson & Shiffrin, 1971), and rate of decay (Atkinson & Shriffin, 1971; Bahrick, Bahrick, & Wittinger, 1975). Further evidence comes from brain lesion data demonstrating a double dissociation between STM and LTM. Patients who suffered damage to their medial temporal lobe (MTL) had deficits in LTM performance, but not in STM performance (e.g., patient "H. M."; Scoville & Milner, 1957). Conversely, a patient, K. F., who suffered a different pattern of brain damage (to perisylvanian cortex, not MTL) had deficits in STM performance, but not LTM performance (Shallice & Warrington, 1970). The results of brain lesioning seem to imply that disconnection of one area responsible for either STM or LTM does not impact the other area responsible for the opposing memory system. If the assumption is that these concepts are not separate, damage to one area should severely impact the other.

The current proposal operates under the assumption that STM is analogous to internal attention, which guides information to and from LTM (Cowan, 1999). Specifically, 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 paper) is in activated LTM. The important distinction is the level of activation (guided by attention). This view is supported by behavioral evidence (Cowan et al., 1990) but also evidence from neuroimaging and studies. These brain-related data suggest that both STM and LTM rely upon frontal regions (Cabeza, Dolcos, Graham & Nyberg, 2002) during retrieval, and that STM can be impaired by MTL damage (Ranganath & Blumenfeld, 2005). When new information is encoded, frontal processes are required to maintain attention (Wager & Smith, 2003); however, as the process starts to shift towards storing that updated information, the hippocampal area is necessary for consolidating the new information into memory (Lisman & Grace, 2005; Backus et al., 2016; Kluen, Dandolo, Jocham, & Schwabe, 2019). This theory would explain neuroimaging evidence showing that prefrontal processes are involved in LTM, not just STM, (Simons & Spiers, 2003; Meeuwissen, Takashima, Fernández, & Jenson, 2011; Melrose et al., 2020), and hippocampus involvement in STM, not just LTM, (Bergmann et al., 2012; Melrose et al., 2020).

Despite the differences amongst various theories of memory, the one thing they all agree on is that participants use a variety of different strategies that affect performance on any given task (Nairne, 2002; Cowan, 1999; Waugh & Norman, 1968; Baddeley & Hitch, 1974, Oberauer, 2002). However, all these models failed to test or even acknowledge the role of memory strategies within their own model. When strategy use is withheld, memory performance can fail altogether or fall well below the projected number seven plus or minus two items (Miller, 1956). Strategies provide a way for participants to connect information from STM (or what they are attending to now) to information in LTM or provide a strong and unique cue to the correct answer to later guide someone's LTM, depending on the given model of memory one subscribes to. These strategies are an important part of the encoding and retrieval processes and provide an interesting line of potential research.

Memory Encoding Strategies

Performance on memory tasks depends on the successful encoding and retrieval of information over both short and long intervals. Participants rely on STM to remember a recent trial or to maintain the current goal for the task (Unsworth & Engle, 2007) while also relying on LTM to maintain information from earlier trials. One factor that often influences memory is the extent to which individuals are engaging in strategic behavior when encoding information. Encoding strategies are used to help a participant learn information for later recall (Richardson, 1980). These strategies differ from retrieval strategies, which are strategies employed at the time of retrieval to help narrow down the search through LTM. The extent to which people use encoding strategies can vary across individuals and within individuals depending on the to-be-remembered item or task (Dunlosky & Kane, 2007), meaning some participants choose to stick to one particular strategy throughout a task whereas other participants will shift from one strategy to another to figure out what works best for the task.

The efficacy of these chosen strategies could 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. *Normatively effective strategies*, or more 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, *normatively less effective strategies*, or less elaborative strategies (e.g., rehearsing a word repeatedly), are associated with lower performance compared to when people report using the normatively effective 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). However, it is important to mention that some have criticized this theory for circular reasoning (Olson, 1980). It is possible that “depth” is conflated with “accuracy,” meaning researchers should look towards measures of strategy quality that are not reliant on memory performance, and therefore, should evaluate and use more objective measures, such as measures of brain response.

To summarize, encoding strategies are implemented to improve the quality of the memory trace and to decrease chances of forgetting information. But individuals sometimes fail to properly employ effective strategies. However, some issues have been raised with current measures of encoding strategy use. Current studies often rely on self-report data, which has several issues, including demand characteristics. In the case of self-reported strategies, there is also the possibility that participants may not be metacognitively aware of the strategies they used or able to accurately remember which strategies they used during the task (Richardson, 1980). It is also possible that the strategy reports may have reactive effects on participants’ future strategic behavior (Ericsson & Simon, 1984). Self-reports are practical for larger samples, but what if there was a more objective way to determine whether individuals were using effective or less effective strategies using neurophysiological measures?

The current study aims to use electroencephalography (EEG) to measure brain activity when strategies are successfully or unsuccessfully implemented. Behavioral reports of strategies may be corroborated through reliable brain signals, and thus, would address issues in previous studies. In combination with behavioral data, this will provide a full view of the cognitive processes involved when a participant uses a given strategy to counteract memory failures. However, the underlying neural mechanisms of both memory failures and the strategies used to counteract them are not well understood.

Electroencephalographic Correlates of Memory

Event-related potentials (ERPs) are electrophysiological brain activity paired with a certain event or stimuli (Luck, 2005). It is locked to the onset of a stimulus and thought to be activity driven by the shown stimuli. This activity differs from fMRI because it does not localize activity, but instead, measures precise timing of cognitive functions. In general, ERPs are averaged across trials, and the increase (or decrease) in amplitude is compared between several conditions. Studies focusing on working memory or STM generally look at an ERP component known as the contralateral delay activity (CDA). The CDA has possible implications for memory, specifically fluctuating with increased memory load (Vogel & Machizawa, 2004; Rajsic, Burton, & Woodman, 2018; Adam, Robison, & Vogel, 2018). However, one key downfall to ERPs is that they represent a small portion of activity in the brain (Luck, 2005). They focus on voltage fluctuations in the time domain, which ignores the fact that the brain operates at different phases at these points in time when completing memory tasks (Cohen, 2014). The phase of the frequency cycle (e.g., a trough or a peak) alters the voltage measure for ERPs, whereas these phases can be looked at with a time-frequency approach.

Given that much more than voltage fluctuations occur across the scalp when people are encoding and retrieving information, more recent research has begun to evaluate brain oscillations. Oscillations represent the frequency at which the brain operates, and importantly, the brain can oscillate at *multiple* frequencies at one timepoint (Cohen, 2014). Time-frequency analyses break moments in time down into the frequencies at which brain waves oscillate. Frequency bands are often referenced by their hertz (Hz), power, and phase. For instance, the theta frequency band oscillates at a rate of four to seven cycles per minute (e.g., 4-7 Hz), whereas the gamma frequency band oscillations at a rate of 30 to 200 cycles per minute. Power represents the strength of the signal that is present (e.g., if a signal is strong, power will increase), whereas the phase refers to the point at which the frequency band is in its cycle, such as a peak (e.g., an increase) or a trough (e.g., a decrease).

In comparison to ERPs, time-frequency analyses look at a variety of cognitive mechanisms occurring across the brain at the same point in time. However, while this approach provides a broader view of brain activity, one disadvantage of time-frequency analysis is that time precision can be hindered by parameters (e.g., the Morlet wavelet transformation) used to calculate time-frequency series (Cohen, 2014). Despite this limitation, the time-frequency approach is well-suited for studies looking for neural signatures of cognitive processes rather than looking for precise timing of experimental manipulations.

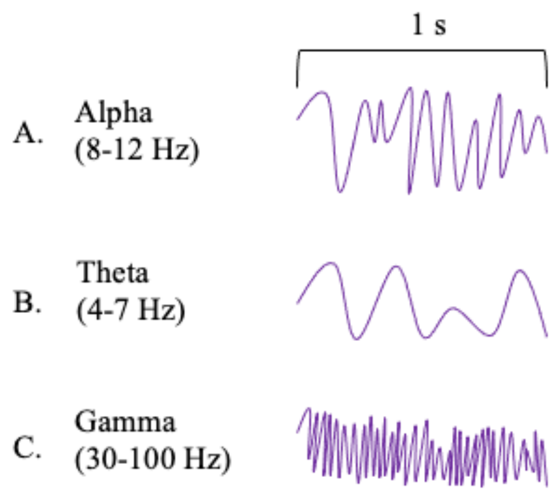
The current proposal focuses on the neural signatures of memory successes and memory failures—the timing of when they occur is of less interest. Rather, the current study is focused on the brain signatures that predict memory failures and using memory strategies to evaluate how memory failures can be counteracted. Thus, time-frequency bands, specifically alpha (8-12 Hz), theta (4-7 Hz), gamma (30-100 Hz), will be evaluated in this proposal.

Frequency Bands Involved in Memory

Alpha

Recent studies have shown that the alpha band (Figure 1A) reflects a suppression mechanism in cognitive tasks (for review, Jensen & Mazaheri, 2010). Theoretically meaning, alpha power could provide a cortical state of rest (Pfurtscheller et al., 1996), such that unnecessary regions are offline and, thus, cannot provide irrelevant information. Cognitive task performance is better when participants exhibit high alpha power in task-unnecessary regions compared to task-necessary regions (Jokisch & Jensen, 2007; Sauseng et al., 2009). Thus, alpha as inhibition reflects a shift cognitive resources from focusing on task-irrelevant information to task-relevant information, resulting in less alpha power in task-relevant areas.

Figure 1
Frequency bands of interest in the proposal



Additionally, Sauseng et al. (2009) found that individuals high in working memory capacity were able to inhibit irrelevant information (e.g., more alpha suppression) and recall more words than individuals low in working memory capacity. Prior research has shown that these individuals report using effective strategies at higher rates (Bailey, Dunlosky, & Kane, 2011), which may indicate that effective strategies help participants shift attention to relevant information more efficiently. Therefore, if alpha power is used to inhibit information and shift cognitive resources to relevant task-related information, alpha power should be the lowest when participants are using strategies that help them focus because they are more efficient at shifting their cognitive resources.

Theta.

Patterns of theta (see Figure 1B) in working memory paradigms vary. One finding asserts that theta power varies with set sizes such that theta power increases as the number of items stored in memory increases (Jensen & Tesche, 2002; Lee et al., 2005), reflecting increasing mental effort, or attention, to keep information in memory. However, a competing finding showed that theta power did not vary with set sizes (Raghavachari et al., 2001), rather they found that theta power increased during the retention interval for trials that were later successfully recalled. These results indicate that theta power may increase and remain stable (or continue to increase) as more memory items are added since they are actively being processed for later storage. Because of this, theta power remains stable until the end of the encoding interval when processing is no longer necessary. Raghavarchari et al. (2001) also proposed that theta has a gated nature, and subsequently, resets its phase with every new item added to the to-be-remembered words, presumably because attention is needed to actively engage the item.

There are important factors that could explain these conflicting results between these studies. One such factor could be which brain regions and/or electrodes were analyzed. Onton, Delorme, and Makeig (2005) only found memory load differences after conducting an independent-component cluster analysis (ICA). By isolating the theta activity independent of channels, the study found an increase in theta power as memory load increased. Further demonstrating that Raghavachari et al. (2001) might not have found theta differences because they chose a channel-analysis approach. However, it is important to note that Jensen and Tesche (2002) also took a channel analysis approach and still found memory load effects in theta power.

As such, a more plausible explanation for the study difference is that Raghavachari et al. (2001) used fewer words than Jensen and Tesche (2002). Theories disagree as to the total capacity limit of working memory, with some models postulating four items (Cowan, 2002) and others postulating up to 10 items (Atkinson & Shiffrin, 1968). Cowan's primary argument for this difference of opinion is the use of strategies to connect information to activated long term memory. Strategy-free memory would reveal only a four-item capacity. However, Raghavachari et al. (2001) did not prevent participants from using any type of rehearsal strategy during their experiment. As such, the lack of power differences between set sizes could be the task tapping into something different from Jensen and Tesche (2002) or Onton et al. (2005).

It is equally possible that the task did not tax memory load enough because participants engaged in strategy use during the task. Raghavachari et al. (2001) only required participants to remember four items. Connecting incoming information with information stored in LTM via elaborative encoding strategies would be much easier to do with fewer word stimuli than with the larger set sizes (up to 7 items) and the unpredictable, varying set sizes (e.g., set sizes ranging from 1-7 items) used in Jensen and Tesche (2002) and Onton et al. (2005). In addition, Onton et al. (2005) varied the maintenance period of the letter Sternberg task from 2 – 4 s and Jensen and Tesche (2002) kept the maintenance period of the digit Sternberg task held constant at 3 s, whereas Raghavachari et al. (2001) varied the maintenance period of the letter Sternberg task from 0.9 s – 2 s. Given Raghavachari et al.'s sample (e.g., epileptic participants), the short maintenance period is expected, but with minimal time between trials and fewer word stimuli, their task could either be 1) easier to recall because of a small maintenance period or 2) easier to connect the word stimuli with strategies.

Given the above information, the current proposal assumes that the paired-associates task will encourage the use of strategies for the to-be-remembered words. It is predicted that theta power should increase among set sizes in the current task. If theta power is associated with increased mental effort and attention, its function is akin to the central executive (Baddeley & Hitch, 1974) or the focus of attention (Cowan, 1999; Oberauer, 2002) in previous memory models. When set sizes increase the amount of information to-be-remembered, participants will be required to focus their attention and their mental effort will increase. The current proposal also plans to take a channel analysis approach. If null results are found, and theta does not increase with set size, it would be better to reanalyze this data in the future using the ICA approach in Onton et al. (2005) for a more sensitive measure of the theta frequency band.

Gamma (see Figure 1C).

When looking at memory tasks differentiating between recollection and familiarity, it has been proposed that the gamma band is related to binding features into memory (Burgess and Ali, 2002; Gruber et al., 2008). Recognition is seen as objectively easier because it provides response options and only requires identification of the target information. When responding to a recognition trial, a person can attempt to recollect the information or they can simply rely on feelings of familiarity for the provided response options (e.g., Mandler, 1980; Jacoby, 1991; Yonelinas, 1994). On the other hand, recall tasks provide no memory cue and force the person to recollect the information to generate one's own answer. On recall tasks, gamma relates, not only to binding features into memory, but also holding specific items for memory storage (Fell et al., 2003; Gruber et al., 2001).

Given that the current study will utilize a modified paired-associates learning (PAL) paradigm, a cued-recall task that requires recollection of information, we expect that gamma will correlated with the maintenance of memory representations (i.e., storing information into LTM). This prediction is in line with Fell et al. (2003) and Gruber et al. (2001): We expect higher gamma power at higher set sizes (Howard et al., 2003) and gamma power may correlated with memory performance (Roux et al., 2012). This is also in line with research on storage in working memory studies in which more items require more storage of information and the efficacy in storage can affect how well someone does on a working memory task.

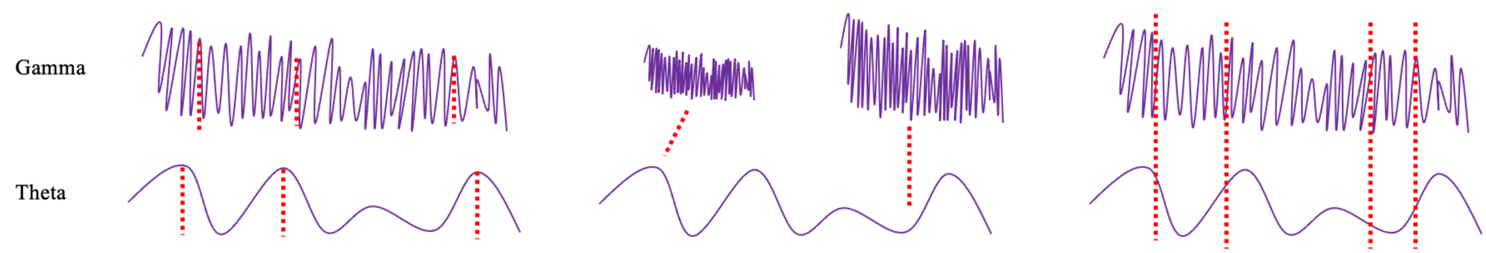
Theta-Gamma Coupling.

Though the frequency bands on their own may each serve an important, independent function for memory, the focus of the proposal will be on how *both* the theta and gamma band work in tandem to support memory processing. Gamma often works in tandem with theta, such that the two bands are *coupled* together functionally. One potential interpretation of this coupling is that information is being shifted from short-term memory (or focus of attention) to long-term memory (Lisman & Jensen, 2013; Baddeley & Hitch, 1974; Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). Research studies have shown that theta-gamma coupling occurs primarily in the hippocampus (Justras, Fries, & Buffalo, 2009; Belluscio et al., 2012). When gamma oscillations are found in the prefrontal region, they are modulated by theta oscillations in the hippocampus, possibly increasing communication between cortical areas involved in maintaining attention and storing memory information (Sirota et al., 2008; Spellman et al., 2015).

There are many different types of theta-gamma coupling, but they all correlated with enhanced memory performance. The first type of coupling is power-to-power (see Figure 2A). This type of coupling occurs when power in both bands increase simultaneously. When this occurs, subsequent memory performance is better (Sederburg et al., 2003). One important thing to note about power-power coupling is that the coupling must exist in the same area for conclusions to be valid (e.g., theta and gamma power in the frontal region). There are situations in which frequency bands can work independently of one another in different areas (e.g., theta in the frontal region and gamma in the temporal regions). For instance, the gamma band can represent an independent cognitive process that happened to be activated at the same time that the theta band was active in another cognitive process. Therefore, analyses will primarily focus on coupling only in the frontal region.

Figure 2

Theta and gamma coupling dynamics



A. Gamma power is coupled with the equivalent amplitude of the peaks and troughs of the theta waves. The length of the dotted red line in both gamma and theta waves are the same length, so they have locked power

B. Gamma power is phase-locked to the peaks and troughs of the theta wave. The dotted red line on the right is an example of low gamma power paired with the peaks of theta. The dotted red line on the left is an example of high gamma power paired with the troughs of theta.

C. Gamma phase is locked in similar phase pattern as the theta wave. The waves between the dotted red lines are points in time where gamma phase is aligned with theta phase.

Another type of coupling is known as power-to-phase coupling. This type of coupling occurs when the power of one band is phase-locked to the oscillations of another band (see Figure 2B). When participants are asked to remember information, the power at which the gamma frequency works is coupled to the phase of the theta band (Canolty et al., 2006). More specifically, when theta is in peak phase, gamma power increases; whereas, when theta is in the trough phase, gamma power decreases. When gamma power-theta phase coupling is stronger, it was associated with participants maintaining multiple items in memory (Axmacher et al., 2010) and remembering items more accurately (Friese et al., 2012), ultimately showing that communication between the two bands are a key component for subsequent memory of items. It is possible gamma power works to store information in certain cycles of theta activity (Fell et al., 2003). It is possible that theta phase activity is associated with actively maintain or focusing attention on items that gamma power is cycling through. In essence, the attention on these memory items increases their ability to be retrieved at a later time.

The last type of coupling is known as phase-to-phase coupling. For example, when theta peaks, gamma will also peak in alignment (see Figure 2C). The alignment of phase is correlated with an increase in memory performance, such that if both gamma and theta are well-aligned, memory for items increases (for review, see Fell & Axmacher, 2011; Schack et al., 2002). While phase-to-phase coupling is important to mention, analyses for this type coupling are beyond the scope of this project. As such, this proposal will focus primarily on power-to-power for ease of analyses. Phase-to-power coupling and Phase-to-phase coupling will be examined but these analyses will be purely exploratory.

To conclude, the coupling of theta and gamma frequency bands is correlated with facilitating memories by updating short-term information (or focusing on a given piece of information) and storing that information into long-term memory (Lisman & Jensen, 2013). Different types of coupling (e.g., power-to-phase, phase-to-phase, power-to-power) can occur in the same study (for review, Jensen & Mazaheri, 2010). The proposed study will not attempt to disentangle the different possible types of coupling, but rather, considers them from a holistic view of memory. Therefore, it is important to study coupling to see how it plays a role in shifting short-term memories into long-term memory.

Research Summary

It is important to note that very little work has evaluated the neural oscillations of encoding strategies. These frequency bands are linked with task difficulty or encoding effort, but not to specific encoding strategies. However, memory theories have proposed that elaborative encoding strategies are controlled processes that allow people to make connections between the to-be-remembered information and information stored in long-term memory (Craik & Watkins, 1973; Richardson, 1998). In other words, elaborative encoding strategies take more mental effort (i.e., attention and storage) to connect word pairs than do less elaborative encodings strategies. The current proposal operates under the assumption that effective strategies should require more mental resources to use; thus, brain dynamics (e.g., power and power coupling for theta, gamma, and alpha) will be stronger for these strategies than less effective strategies.

Research Questions (RQ) and Hypotheses (H)

RQ1: What neural patterns are active in the brain when we experience memory failures? (replication of Weiss and Rappelsberger, 2000)

H1: Theta and gamma will positively correlate with memory performance, such that power in both bands will be weaker when a participant experiences a memory failure as compared to a memory success.

H1a: Set size will moderate memory outcome, which in turn, will influence the neural mechanisms. In line with Fell et al. (2003), Gruber et al. (2001), and Onton et al. (2005), both theta power and gamma power will increasing as set size increases.

H2: Theta-gamma power coupling will positively correlate with when a participant experiences a memory failure as compared to a memory success. If power-to-power coupling is found, theta power will increase as gamma power increases during the encoding intervals.

H2a: Set size will moderate memory outcome, which in turn, will influence coupling of the neural mechanisms. We expect that theta-gamma power coupling will show increasing as set size increases.

RQ2: Do effective strategies differ neurologically from less effective strategies? (extension)

H3: Prior research has shown behavioral differences in memory performance between effective strategies (e.g., imagery or sentence generation) and less effective strategies (e.g., rehearsal or merely reading a word). As an extension to behavioral analyses, theta and gamma power will be stronger for effective strategies than less effective strategies, regardless of memory performance.

H3a: Similar to H3, theta-gamma power coupling dynamics will be stronger when participants report using effective strategies than when they report using less effective strategies, regardless of memory performance.

H3b: Because alpha inhibits information, alpha power will be lower when participants report using effective strategies than when they report using less effective strategies because it correlates with when participants must put forth more cognitive effort to suppress irrelevant information (e.g., words they were exposed to from other sets that do not belong with the word that needs to be recalled).

RQ3: Do neural signatures validate strategy reports and memory performance? (extension)

H4: Theta, gamma, and alpha power will positively correlate when participants effectively use elaborate (and more effortful) strategies. Thus, theta and gamma will be stronger while alpha power will be the lowest on correct-recall trials in which participants reported using effective strategies compared to all other trials.

H4a. Theta-gamma power coupling will be positively correlation with participant's recall accuracy and strategy type (i.e., effective or less effective). Thus, theta-gamma power coupling will be stronger on correct-recall trials in which participants reported using effective strategies compared to all other trials. In other words, when participants put effort into elaborately encoding information, theta-gamma coupling will be stronger.

H4b. The strength of alpha inhibition will correlate with a participant's recall accuracy and strategy type. Thus, alpha power will be lower on correct-recall trials in which participants reported using effective strategies compared to all other trials.

Pilot Experiment

Methods

In a paired-associates task, some items are more likely to be remembered whereas others are not, depending on many factors (e.g., word frequency, word length, association strength). The purpose of the pilot experiment was to develop a set of word pairs that generated a 25% error rate in participants, meaning only words that, on average, were recalled 75% (or less) correctly will be included in the proposed experiment.

Participants

One hundred and eighty-eight participants were recruited from a public, midwestern university. Participants were students from psychology courses and completed the study as course or extra credit. They were primarily female (58%; $N = 108$) and Caucasian (79%; $N = 145$). Participants ranged in age from 18 – 45 years ($M = 20$ years, $SD = 3$). No participants were excluded from the analyses.

Tasks

Word Stimuli. Stimuli used in the experiment were screened to ensure that difficulty of the words would not play a significant role in the memorization of words. The Medical Research Council (MRC) Psycholinguistic Database (Coltheart, 1981) was used to choose stimuli in the experiment. Stimuli were controlled for based on the following: number of letters, number of syllables, concreteness, imageability, word frequency, and familiarity (see Table 1). These parameters controlled the level of difficulty within the stimuli. A total of 624 words were placed into 312 randomly generated word pairs to use in the pilot experiment.

Table 1
Summary statistics of word characteristics used in the pilot experiment

	M	*SD*	*IQR*
Concreteness	481.65	108.77	205
Familiarity	539.63	46.03	49.75
Imageability	502.28	90.23	152
Number of Letters	4.94	0.81	2
Paivio's Meaningfulness	648.44	99.50	144
Number of Syllables	1.39	0.51	1
Brown's Word Frequency	8.16	13.23	8

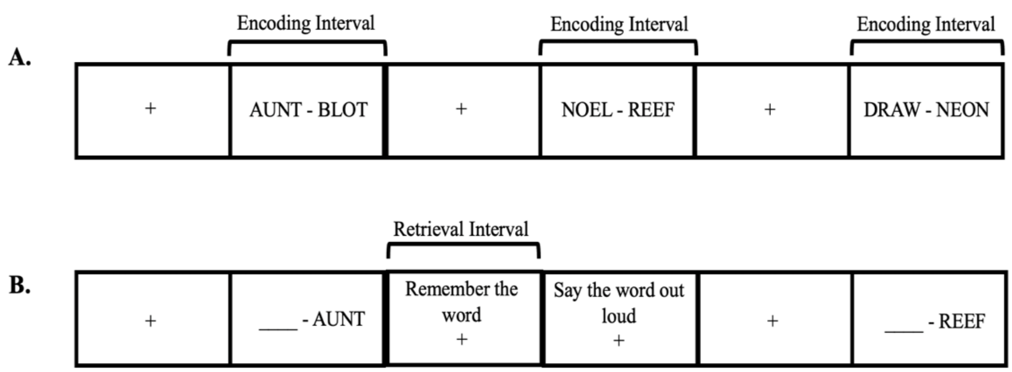
Note: M = mean, SD = standard deviation, IQR = interquartile range

Set sizes varied from 3-10 pairs to vary task difficulty. Each block consisted of eight sets, one from each set size (3, 4, 5, 6, 7, 8, 9, and 10 pairs). This study was conducted online; therefore, to reduce the reduce participant fatigue and increase retention throughout the session, participants did not see all 312 word pairs. Instead, each participant completed five blocks of words pairs for a total of 34-45 word pairs (i.e., 68-90 words). Thus, memory performance data *for each item* is based on data from between 22-25 participants rather than the full number of participants. See Appendix A for the individual word pairings and characteristics.

Modified Paired Associates Task. The encoding task was a modified version of the Paired Associates Learning Task. Participants studied a word pair (e.g., AUNT – BLOT; see Figure 3A) for 4 seconds. Each word pair was followed by a 4 second interstimulus encoding interval, indicated by a fixation cross. This process repeated for each word pair until the set was complete. After the final word pair, participants completed a cued-recall test. In this test, participants saw the word pair again with one of the words omitted (e.g., AUNT – ________; see Figure 3B). The position of the target word (e.g., whether the first or second word) was counterbalanced to prevent participants from learning which word to memorize for subsequent trials. The cue was shown for five seconds, which allowed participants time to type the missing word to the pair into a text entry box. Each word pair was followed by a 4 second interstimulus interval consisting of a fixation cross. This process was repeated for each word pair until the set is complete. After the last cued-recall trial, the participants started the encoding task for the next set. Set sizes were pseudo-randomized so that participants only saw 6 blocks of words (e.g., one participant would get Block 1, which had set size 3, 4, 5, 7, 8, 10,

Figure 3

Modified version of the paired associates learning task (A) and the cued recall task (B)



while another participant would get Block 2, which had set size 3, 4, 6, 8, 9, 10). Set sizes within the block were randomized upon participants starting the block (e.g., 3, 4, 5, 7, 8, 10 versus 5, 4, 10, 7, 3, 9).

Strategy Reports. After the cued-recall task for each set of words in a particular block, participants saw the word pairs one at a time. Participants chose the strategy they used to encode each word pair from a list of options with one of the options indicating that they did not try to memorize the words (see Figure 4). They reported a strategy for each word pair. Afterwards, they were shown new word pairs to memorize until the completion of the entire task.

Figure 4
Example of a strategy report

GHOST – PIZZA

1. Repeated the words as much as possible
2. Developed mental images of the words
3. Used a sentence to link the words together
4. Grouped the words in a meaningful way
5. I did not try to memorize the word pair

Procedure

Participants were recruited via SONA Systems, which is an online recruitment platform for students in psychology courses. Participants signed up for the study and received a link to a Qualtrics survey. Upon clicking the link, participants were required to read the consent elements, which outlined that participation was volunteering and could be terminated at any time. If participants did not agree to the consent form, they were not allowed to complete the study. If participants did agree to the consent form, participants were guided to the instruction page. The instruction page listed how to complete the tasks, including the procedure for encoding words, retrieving words, and answering strategy questions about how they memorized the words.

After the instructions were read, participants participated in a brief practice task. The practice task consisted of sample instructions and an example. After ensuring the participants read the instructions, participants then completed a practice round. The practice round consisted of two word pairs and was not scored.

Once participants finished the practice, they were sent to an instruction screen that allowed them to review the procedures once more. Then, participants were told whether to "encode" or "remember" words during their respective tasks. After the cued recall task, participants were, once again, instructed to describe the strategy they used to encode the words on the previous screen. Once all ratings were completed for the entire task (e.g., 4 blocks of word pairs ranging from 3-10 pairs), participants were asked to disclose their demographics and shown a debriefing statement that explained the importance of the study and thanked them for their participation.

Results and Discussion

Each target word was coded as correctly recalled (1) or incorrectly recalled (0). To get an overall accuracy score for each item, we calculated the proportion of participants that correctly recalled each word pair. To do so, the total number of participants that correctly recalled the word was summed and divided by the total number of participants who saw that word pair. Appendix A shows performance on each word pair. The purpose of this study was to identify relatively difficult stimuli in which 75% or less of the participants correctly remembered a word pair. This criterion would increase the likelihood that participants in the main study would experience memory failures. Of the 312 word pairs, 85 word pairs were excluded, resulting in 227 word pairs whose overall accuracy scores were below 75%.

Main Experiment

Methods

The purpose of the proposed main experiment was to evaluate the neural signatures of encoding strategies and subsequent memory failures. Prior research suggested that brain signatures differ when words are correctly or incorrectly recalled. The proposed experiment extended this literature by evaluating different underlying brain patterns in memory processes while using strategies. Strategy reports were collected at the end of each trial. Data analyses were conducted to see if the individual strategies that participants use yield conclusive, and different, results in brain patterns.

Participants

Thirty-one participants were recruited from a public, midwestern university. Participants were students from psychology courses as well as recruited, paid subjects from advertisements around campus or online. They completed this study as course credit, extra credit, or for $15/hour for a total of $45. Participants were primarily female (61%; *N* = 19) and Caucasian (71%; *N* = 22). Participants ranged in age from 18 – 38 years (*M* = 21 years, *SD* = 4). Two participants were excluded from the behavioral and EEG analyses because of audio recording failures and one other participant was excluded from the behavioral and EEG analyses due to extenuating circumstances. Four participants were excluded from the EEG analyses for excessive movements. All participants completed 208 trials, except for one participant who chose to withdraw consent after 104 trials. The final behavioral data analysis included 28 participants and the final EEG data analysis included 21 participants.

Word Stimuli.

Stimuli used in the experiment were chosen from the above pilot study. After controlling for performance (i.e., selecting words that produced errors in at least 25% of the pilot sample), 208 word pairs were chosen for the experiment along with seven word pairs for the practice blocks. Characteristics of the 208 chosen words to be used in the experiment can be seen in Table 2. Words were randomly divided into four blocks of 52 words. All words remained in the same block for each participant, but (1) the blocks were randomized for each participant (e.g., Participant 1 might get 1, 2, 3, 4 while Participant 2 gets 3, 2, 4, 1); (2) the order of the word pairs within the block was randomized (e.g., Participant 1 might see word 1, 2, 3 while Participant 2 might see word 3, 2, 1); and (3) the words within the word pair were randomized to be on the left or right side for each participant (e.g., Participant 1 might get AUNT – BLOT while Participant 2 gets BLOT – AUNT).

Table 2
Summary statistics of word characteristics to be used in experiment

	M	*SD*	*IQR*
Concreteness	481.37	110.02	209
Familiarity	537.77	47.53	45.25
Imageability	500.43	90.72	160.50
Number of Letters	5.00	0.80	2
Paivio's Meaningfulness	643.68	104.87	157.75
Number of Syllables	1.40	0.52	1
Brown's Word Frequency	8.03	13.47	8

Note: M = mean, SD = standard deviation, IQR = interquartile range

Modified Paired Associates Learning Task. Participants completed the modified Paired Associates Learning Task from the Pilot Experiment. The encoding task was the exact same except that EEG data was recorded. The cued verbal recall task procedure was the same as in the

Pilot Experiment, but with the following exceptions. Participants were shown a fixation cross for 4 s prior to the start of each block of words to provide a baseline for EEG analyses.

Participants first saw a 4-second fixation cross before each set of word pairs that indicated that they should prepare for a new set of trials. Then a word pair was presented onscreen for 2 s. After the presentation of the word pair, participants were shown another fixation cross and given 4 seconds to encode the words. This repeated for each pair in the entire set. Set sizes varied from 3-10 word pairs.

After the final word pair in each set, participants completed the "recall" period. They were shown a 4 s fixation cross to prepare the participant for the trials and then the cue word was presented onscreen for 2 s with the target word missing (the cue and the target word were randomly selected). Then, participants were prompted with the instructions "start remembering the word", in which they were given 4 s to try to recall the target word in their head. Finally, they were shown another fixation cross with the words "please say the missing word" for 4 s to orally recall the target word from that pair. Voice recordings were collected during the cued-recall task. Prior to the experiment trials, participants completed two practice blocks. The practice block consisted of two set sizes: set size three and set size four. This allowed participants to get used to the task set-up and ask any questions.

Audio Recordings. For recording voice responses, Focusrite Scarlett audio interface (Focusrite, UK) triggered a microphone to begin recording as soon as the fixation cross appeared with the words, "Please say the missing word", and recorded their oral responses for to 4 seconds. These were saved as ".wav" audio recordings for scoring. Scoring was done by two trained research assistants (interrater reliability = 98.7%). Both assistants scored the same recordings for 24 subjects. Given the high interrater reliability, one assistant continued to score the rest of the participants solo (N = 5 participants). Discrepancies were solved by a third-party listener, who scored only the trials where the discrepancies were (N = 60 trials).

Strategy Reports. Participants used the same strategy reports described in the Pilot Experiment. The only change between the Pilot Experiment and Main Experiment strategy reports was the deletion of the "grouping" and the addition of a "Read the words as they appeared" strategy. Given that the paired-associates task presents words in a grouped manner, this strategy was dropped because participants reported using this strategy often since they were already grouped with one another. Reading was added to further replicate previous studies by Bailey et al. (2011) and Dunlosky and Kane (2007) and provide a way to determine whether less effective strategies were better than participants not trying to remember words. Further, we added another response option: "I did not try to remember the words." Any trial where this option is chosen was excluded from further analyses.

Apparatus. A 64-channel Geodesic system (EGI: Electrical Geodesics Inc., Eugene, OR, USA) measured neural oscillations during task completion. Participants responses were recorded by Focusrite Scarlett audio (Focusrite, UK) and their oral responses were recorded using a dynamic vocal microphone (Shure SM58), which was mounted on a stand ~6 inches from the participant. The participant was seated ~12 inches from an LCD monitor. The program was coded using Experiment Builder (SR Research Experiment Builder, Canada).

Procedure

Participants signed a consent form which that stated the experiment's purpose and risks. They filled out a short demographics form, including their age, ethnicity, and sex. Once participants were fitted with an EEG cap, participants were seated in a quiet room to complete the task. They completed one practice block and four experimental blocks. For each block, participants always completed three tasks: encoding the paired-associates, completing the cued-recall trials and then completing the strategy reports. After each block, participants were offered a short break to allow the research assistant to rewet electrodes in the cap. This break was ~2 minutes. Once all tasks are finished, participants were given a debriefing study, thanked for their participation, and received compensation if they were paid research subjects.

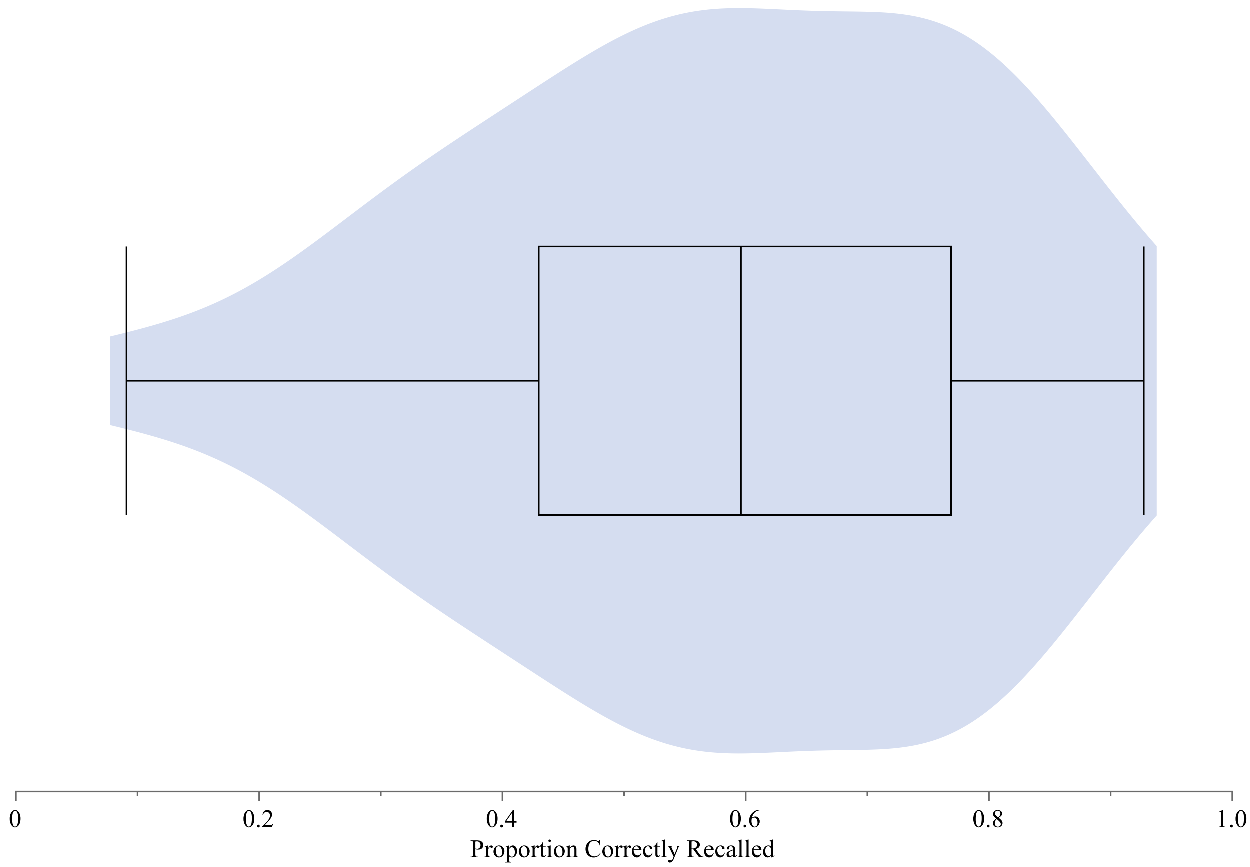
Behavioral Results

The purpose of the current analysis was to compare behavioral effects of the study to prior behavioral effects from previous studies (Weiss and Rappelsberger, 2000). In general, it is expected that both set size and strategy type will predict performance. Performance should be higher (e.g., more words recalled) on lower set sizes than on higher set sizes. Performance should also be higher when participants report using effective strategies compared to when they report using less effective strategies or no strategy at all. First, the variables used in the model will be briefly described, and then all main effects of set size and strategy on performance will be probed. All analyses were done in JMP Pro 16 (SAS Institute Inc., 2021). To look at results derived from a generalized linear modeling approach, see Appendix B.

Model Description. We used two repeated-measures ANOVA to investigate the main effect of strategy use and set size (separately) on proportion of correctly recalled items (i.e., "Proportion Correctly Recalled"). Proportion Correctly Recalled was scored as the participant's total word count divided by the total number of trials completed. Proportion of correctly recalled words could range from 0% to 100% (*M* = 0.59, *SD* = 0.20, *SE* = 0.003; see Figure 5 for a distribution of the raw scores on the task).

Figure 5

Distribution of the proportion of correctly recalled words. The distribution is roughly normal with a negative skew



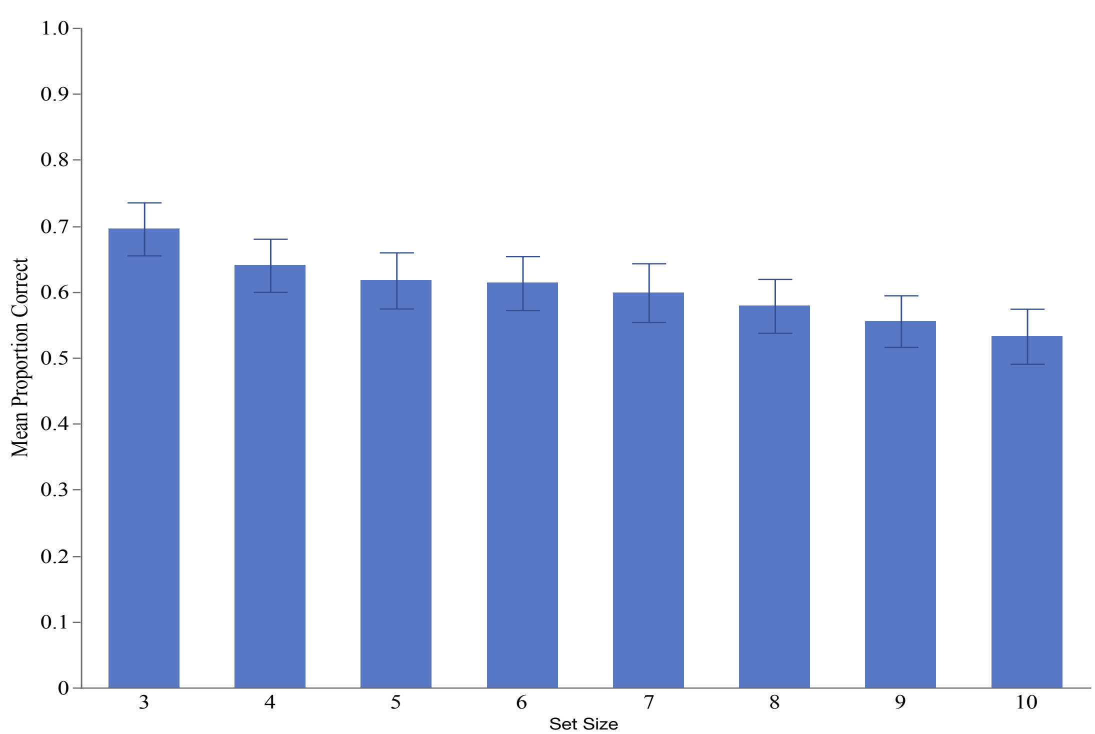
Note. The plotted box represents the range of the data with a mean performance of 0.59.

Variable Description. The outcome variable for the current analyses was a continuous variable (e.g., 0-100%) with two categorical predictors, set size (e.g., number of word pairs in each set; this ranged from 3-10 pairs per set) and strategy type (e.g., Didn't Try, Effective or Less Effective). Trials in which participants reported using with imagery or sentence generation were categorized as "effective," trials in which they reported using rehearsal or reading a word were categorized as "less effective," and trials in which participants reported that they didn't try were categorized as "Didn't Try" (see Table 3 for proportion of trials that participants report using effective strategies and less effective strategies).

Set Size

Results of the repeated-measures ANOVA indicated that there was a significant effect of set size, $F(7, 196) = 5.95$, $p = .001$. A Tukey HSD post-hoc analyses indicated that correctly recalled words for set size 3 was significantly different from that of set size 7 ($M_{diff} = 0.10$, $SE_{diff} = 0.03$), set size 8 ($M_{diff} = 0.12$, $SE_{diff} = 0.03$), set size 9 ($M_{diff} = 0.14$, $SE_{diff} = 0.03$), and set size 10 ($M_{diff} = 0.16$, $SE_{diff} = 0.03$). Correctly recalled words for set size 4 was also significantly different from that of set size 10 ($M_{diff} = -0.11$, $SE_{diff} = 0.03$). No other paired comparison differed significantly (see Figure 6 for the plotted proportion correctly recalled words for each set size).

Figure 6
The mean proportion of correctly recalled words by set size



Note. The error bars represent standard error.

Strategy

The proportion of reported strategies are similar those reported in Dunlosky and Hertzog (2001), who compared rates of strategy use between younger and older adults on a similar paired-associates task.

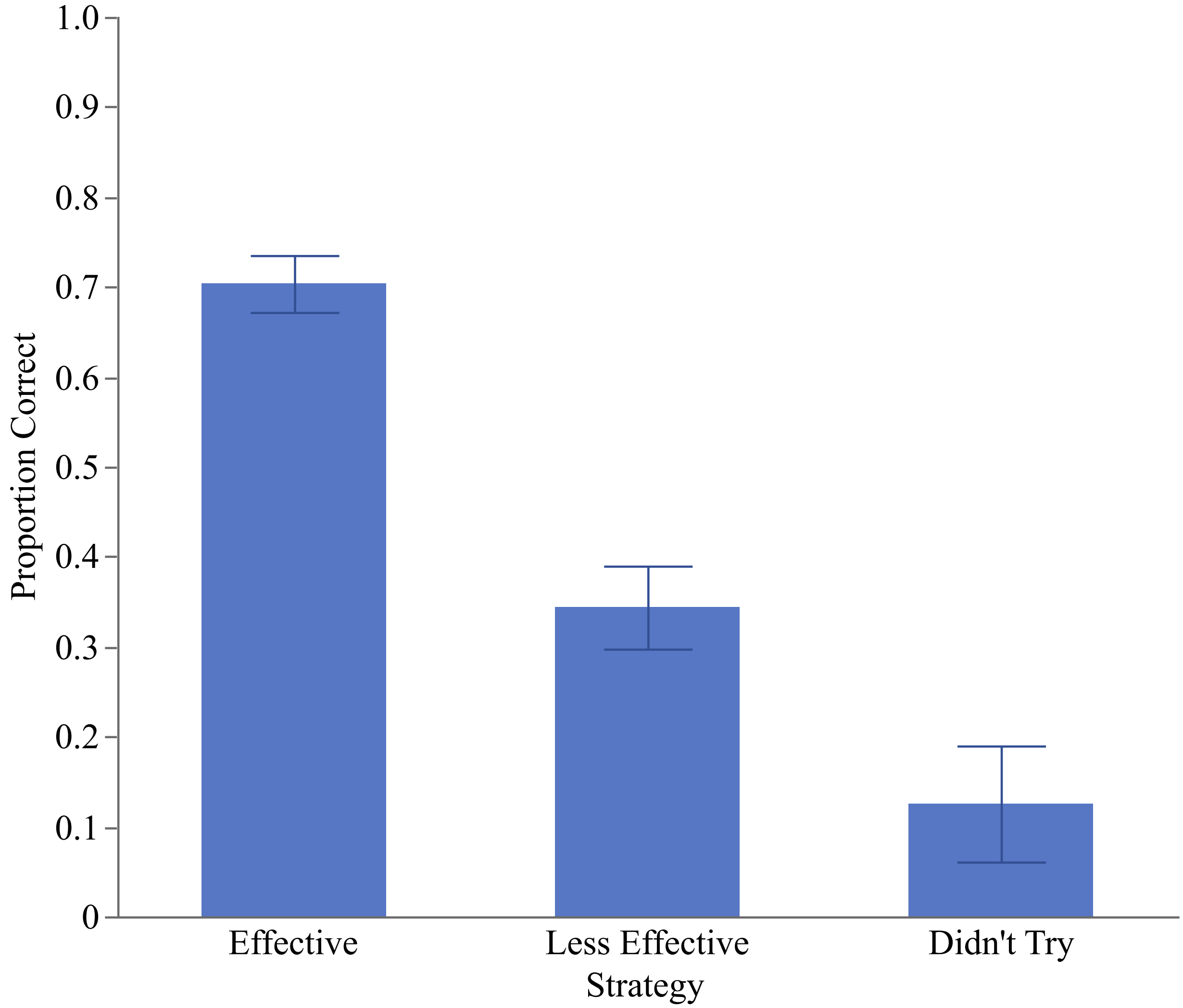
Table 3
Proportion of trials reported using each strategy while studying word pairs

Less Effective Strategies		Effective Strategies		
Repeat	Read	Sentence	Imagery	Didn't Try
0.13 (0.03)	0.12 (0.02)	0.48 (0.06)	0.23 (0.05)	0.04 (0.01)

Note: Parentheses = standard error of the mean

As expected, the omnibus test showed a significant effect of strategy on recall performance, $F(2, 99.15) = 44.84$, $p = .001$ (see Figure 7 for the average proportion correctly recalled words by strategy)[2]. A Tukey HSD post-hoc analyses indicated that effective strategies yielded significantly higher proportion correctly recalled words than less effective strategies ($M = 0.35$, $SE = 0.05$) and when participants did not try ($M = 0.58$, $SE = 0.07$). Less effective strategies also yielded higher proportion correctly recalled words than when participants did not try ($M = 0.23$, $SE = 0.07$).

Figure 7
The mean proportion of correctly recalled words by strategy type



Note. Effective strategies had a higher proportion correct compared to both less effective strategies and when participants did not try.

The observed patterns in the behavioral data replicated previous studies (Dunlosky & Kane, 2001; Dunlosky & Hertzog, 2001; Bailey et al., 2011) such that memory performance was better on smaller set sizes and also when participants reported using effective strategies. Further, our participants reported using both normatively effective and less effective strategies. Next, we will analyze the EEG data to investigate theta, gamma, and alpha band power differences between performance (e.g., correct vs. incorrect), strategy type (e.g., effective vs. less effective), the interaction between (e.g., correct vs. incorrect effective & correct vs. incorrect effective), and set size (e.g., sets 3-10).

EEG Preprocessing & Analysis Procedure

Prior to analysis, the data was subjected to an independent component analysis (ICA). Noise-related components, such as blinks and horizontal eye movements, were identified using their distinctive topography distribution before being removed from the overall electrode recordings. ICLabel was also used to identify other noise-related components, such as heart noise, line noise, muscle movement, and channel noise. If a component was labelled as 75% or more probability of being one of these, it was removed from the ICA for cleaning. All other components remained in the dataset unless their topography uniquely resembled noise-related potentials. Offline analyses were performed using EEGLAB (Delorme & Makeig, 2004). The data was referenced to the average and resampled at 256 Hz. In all participants, eye electrodes were removed because this data was collected during the COVID-19 pandemic, and participants were required to wear masks that interfered with these electrodes.

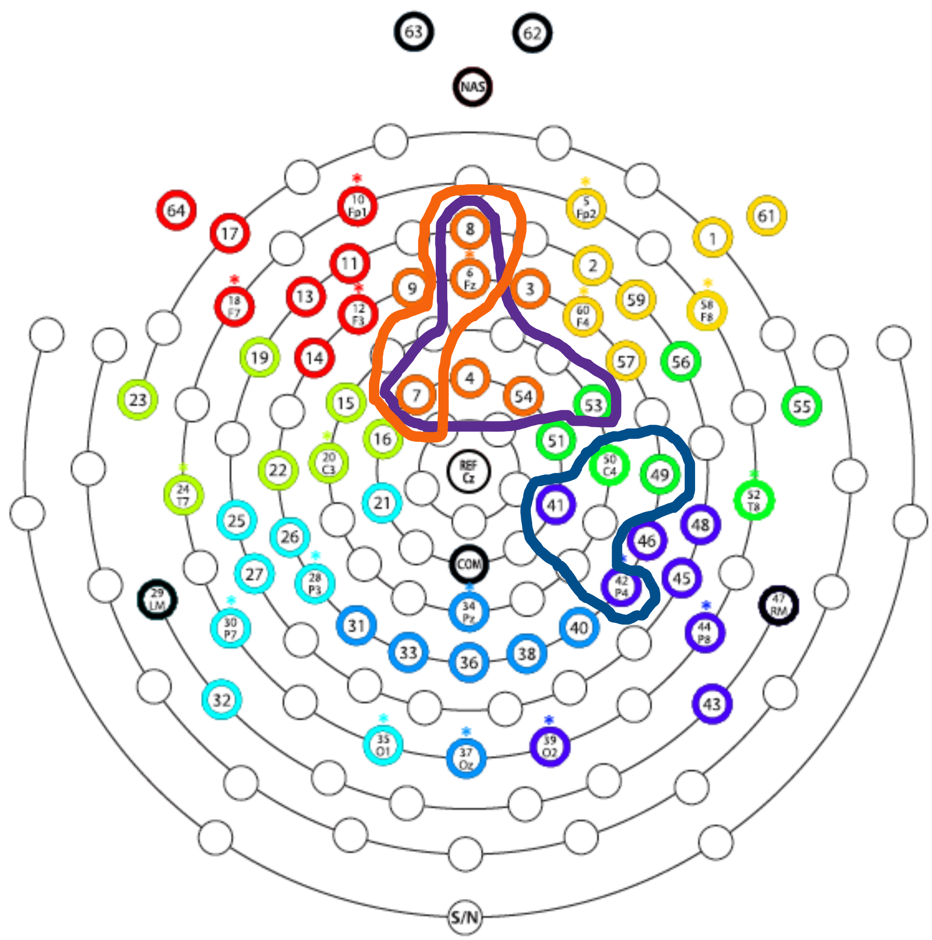
All the EEG analyses were epoched from the encoding interval. Each epoch was -2 s before word pair onset to 4 s after word pair onset. Prior to the start of each full block, participants were show an additional fixation cross for 4 s before encoding the words and after being instructed on how to recall each word (but prior to the initial recall period). This resulted in eight total appearances of these fixation crosses (i.e., four fixation crosses for encode periods & four fixation crosses for recall periods), which together served as the baseline periods for the encode and recall analyses, respectively. This chosen baseline period contained minimally overlapping brain activity, and therefore, ensured comparisons were made against a period where no brain activity was caused by experimental stimuli.

Electrodes of Interest

The decision for which electrodes were analyzed for key patterns rested on the overall topography of the ERSPs across all participants, regardless of the condition (Appendix C). Theta and gamma power values were stronger around a cluster of frontal-midline electrodes. Theta power values were strongest for E4, E6, E7, E8, E54, E53 (purple outline in Figure 8). Gamma power values were strongest for E6, E7, and E8 (orange outline in Figure 8). Alpha values were stronger for electrodes on the right temporal area of the head (E41, P4, E49, C4; blue color in Figure 8). All values were chosen based on clustering of high power values observed in the ERSPs in Appendix C.

Figure 8

Topography of the electrode distribution for the EGI system from Smith (2021)



Note. The orange outline represents clusters of electrodes (Fz. E7, E8) that showed higher gamma values, the purple outline is electrodes (E4, Fz, E7, E8, E53, E54) that showed higher theta values, and the blue outline is electrodes that showed higher alpha values (E41, P4, 59, C4).

Onton et al. (2005) showed that key differences in theta and gamma power depend on electrodes chosen for analyses. Their results were strongest at Fz. Given the above figure, power values for both theta and gamma were clustered around the frontal electrodes, so all power values for theta and gamma will be collected and analyzed from channel Fz (E6), including power coupling analyses. Further analyses could investigate these clusters of electrodes or analyze independent component clusters to focus on brain activity rather than potential electrode power value variations (e.g., due to variations in cortical folding, which in turn, influences where power shows up for each individual participant).

Event-Related Spectral Perturbations (ERSPs) Procedure

Prior to conducting the ERSP analyses and graphing, adjustments were made to account for missing data. Because participants self-reported their strategies, the trial count for each strategy type, recall accuracy type, and set size could differ within participants which would produce noise in the ERSPs. As such, the lowest trial count for participants for each condition was identified (see Table 4). The time-frequency data was randomly permutated and reduced to the lowest number of trials per participant. Afterwards, the event-related spectrum perturbations (ERSPs) were calculated from this randomly permuted data. Based on this reduce trial count, the total number of trials for each condition for all twenty participants was 171 trials.

All time-frequency analyses were logarithmically scaled so the lower frequency changes were easier to see (Cohen, 2014). A complex Morlet wave was used to transform the raw spectral data with a minimum frequency of 4 Hz and a maximum frequency of 100 Hz. Seven cycles were used since lower cycles (e.g., under seven) are better for temporal precision (Cohen, 2014). The total number of frequencies used for the analyses was 96 (Cohen, 2014).

Table 4
Total number of trials for each condition within participants in the EEG analyses

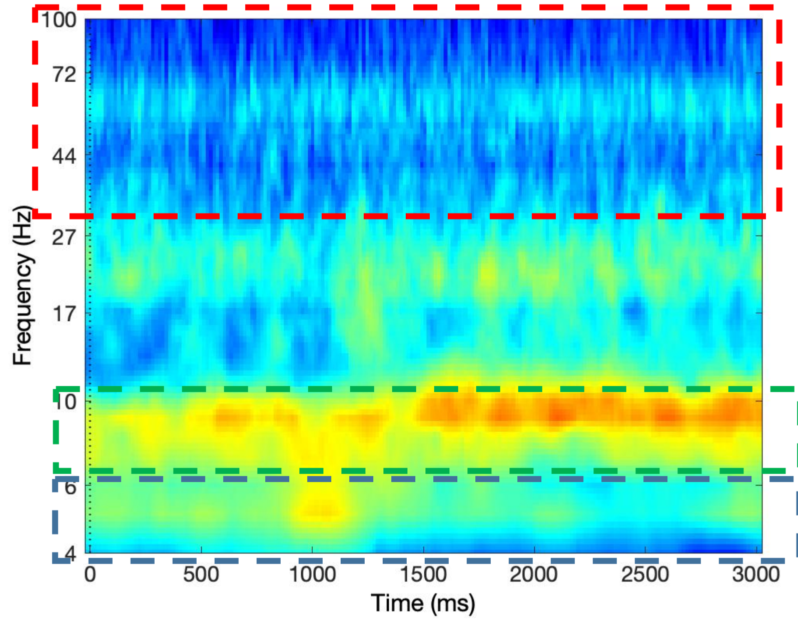
Subject	Baseline	Strategy Type: Effective	Strategy Type: Less Effective	Performance: Correct	Performance: Incorrect	Set Size: Low	Set Size: Med	Set Size: High	Correct x Strategy: Effective	Correct x Strategy: Less Effective	Incorrect x Strategy: Effective	Incorrect x Strategy: Less Effective
1	EXCLUDED BEFORE BEHAVIORAL + EEG ANALYSES – PROGRAM ISSUE											
2	2	73	130	53	125	39	42	97	39	**14***	26	99
3	EXCLUDED BEFORE BEHAVIORAL + EEG ANALYSES – PROGRAM ISSUE											
4	3	75	80	77	78	36	38	81	54	**23***	21	57
5	3	154	43	135	62	42	51	104	122	**13***	32	30
6	3	156	7	108	55	37	43	83	107	**1***	49	6
7	EXCLUDED BEFORE BEHAVIORAL + EEG ANALYSES – PROGRAM ISSUE											
8	4	106	48	152	45	48	43	6	149	**3***	42	3
9	1	174	18	163	29	44	47	101	155	**8***	19	10
10	4	194	3	144	53	46	48	103	143	**1***	51	2
11	4	111	47	79	83	37	39	82	67	**12***	44	35
12	4	70	122	121	71	46	47	99	65	56	**5***	66
13	4	132	67	157	42	42	50	107	130	27	**2***	40
14	3	130	42	137	35	43	47	82	120	17	**10***	25
15	EXCLUDED BEFORE EEG ANALYSES – NO BASELINE											
16	EXCLUDED DURING EEG ANALYSES -- MISSING TRIALS IN ONE CONDITION											
17	EXCLUDED DURING EEG ANALYSES -- MISSING TRIALS IN ONE CONDITION											
18	3	132	44	80	96	35	40	101	72	**8***	60	36
19	4	144	56	154	46	48	51	101	131	23	**13***	33
20	4	117	68	99	86	43	48	94	78	**21***	39	47
21	2	33	141	34	140	39	45	90	23	11	**10***	130
22	EXCLUDED BEFORE EEG ANALYSES -- EXCESSIVE NOISE											
23	EXCLUDED BEFORE EEG ANALYSES -- EXCESSIVE NOISE											
24	4	193	3	163	33	44	49	103	162	**1***	31	2
25	EXCLUDED BEFORE EEG ANALYSES -- EXCESSIVE NOISE											
26	2	96	70	56	110	36	39	91	48	**8***	48	62
27	2	67	13	54	26	22	17	41	52	**2***	15	11
28	4	66	105	47	124	37	43	91	40	**7***	26	98
29	EXCLUDED DURING EEG ANALYSES -- MISSING TRIALS IN ONE CONDITION											
30	4	95	86	70	111	41	47	93	61	**9***	34	77
31	EXCLUDED DURING EEG ANALYSES -- MISSING TRIALS IN ONE CONDITION											

Note: The boldfaced font and asterisk denote the lowest trial number for a given participant. This number represents the reduced trial number across all conditions for the given participant (e.g., Subject 1 had 14 trials for all conditions)

Lastly, the time window for each of the frequency bands of interest was determined by evaluating the mean power values for all frequency bands averaged across all participants and conditions for channel Fz. This allowed us to identify key areas during which theta, gamma, and alpha power are strongest in the data, regardless of condition (see Figure 9). Overall, within Fz, theta power (4-7 Hz) was strongest from 800 ms – 1100 ms, and alpha power (8-10 Hz) was stronger from 1500 ms – 3000 ms. Gamma power (30 – 100 Hz) did not differ across this time window, so to best probe the theta-gamma coupling analyses, the same time window will be used for all analyses, including power coupling analyses for theta-gamma power values.

Figure 9

Power values for all frequency bands averaged across all participants, channels, and conditions



Note. Warmer colors (i.e., dark orange, red, yellow) indicate more power while cooler colors (i.e., dark blue) indicate less power**.**

EEG Results

For all results, the mean power value per frequency band for each condition (e.g., a mean theta power value for correct, a mean power value for incorrect, etc.) were calculated for each individual participant in the time window of interest (e.g., 800 – 1100 ms for theta and gamma and 1500 ms – 3000 ms for alpha). Then, these mean power values were averaged across all participants. From there, three separate 2 (performance: correct vs. incorrect) x 2 (strategy:

effective vs. less effective) repeated-measures ANOVAs were run to analyze significant differences in power for theta, gamma and alpha. For each frequency band, the main effect of performance for each frequency band will be discussed first, followed by the main effect of strategy and then the interaction between strategy and performance will be probed for theta, gamma, and alpha frequency band.

For coupling analyses, the average theta power value and the average gamma power value from the time window of interest was correlated for each participant across all electrode locations. Specifically, a correlation between theta power and gamma power was calculated across all trials for each condition, resulting in 4 correlation values for each participant. Another 2 (performance: correct vs. incorrect) x 2 (strategy: effective vs. less effective) repeated-measures ANOVA was run to analyze significant differences in mean theta-gamma coupling power values for each of these conditions. This analysis will be at the end of each of the main effects and interaction section.

Next, I will discuss each of the frequency bands and how their power values changed across different set sizes. Theta and gamma power values were assessed at the time window of interest (e.g., 0 to 4s) for all participants for each set size (e.g., low, medium, and high). Based on the behavioral analyses for set sizes, low set sizes were collapsed across sets of 3, 4, and 5 because none of these set sizes differed from one another. Medium set sizes were collapsed across sets of 6 and 7. Lastly, high set sizes were collapsed across sets of 8, 9, and 10 because these were significantly different from set sizes of 3. Additionally, these set cutoffs align with Dunlosky and Kane (2007) and Bailey, Dunlosky, and Kane (2008), who considered smaller set sizes to be 2 and 3 or 3 and 4, respectively, and larger set sizes to be 4 and 5 and 6 and 7, respectively. Both studies did not go past sets of 7; however, behavioral analyses showed a difference in sets of 7 compared to 10, thus, "medium" sets were created to accommodate that separation.

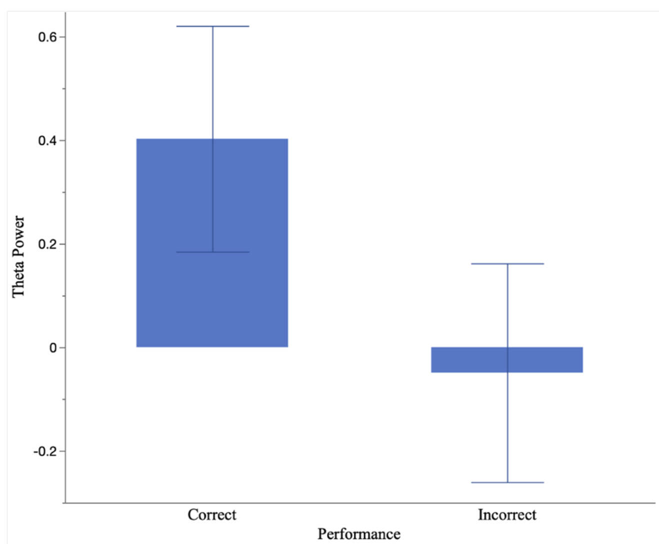
Lastly, exploratory EEG analyses will investigate alpha power at a parietal electrode as well as beta power at Fz for each of the hypotheses. The beta band was chosen at 17-27 Hz since Figure 9 shows a clear enhancement in this portion of the beta band. The exploratory analysis for alpha power will allow me to assess whether alpha power enhancement was only in the frontal electrodes as this result was unexpected. Additionally, beta power was not anticipated to differ between conditions, but it has a strong appearance in each of the ERSPs. As such, analyses at electrode Fz will be conducted to investigate this potential effect. Based on Figure 9, the beta band will be analyzed at 1100 – 2000 ms as this is where it is the strongest.

Memory Performance

Theta. It was hypothesized that power in the theta band would be weaker when participants experienced a memory failure (**H1**). When looking at the ERSP for theta power patterns (see Figure 11, blue dotted line), it visually appears more prevalent for when participants got a problem correctly recalled words than when they incorrectly recalled words.

Figure 10

The mean theta power value for when participants correctly recalled or incorrectly recalled words



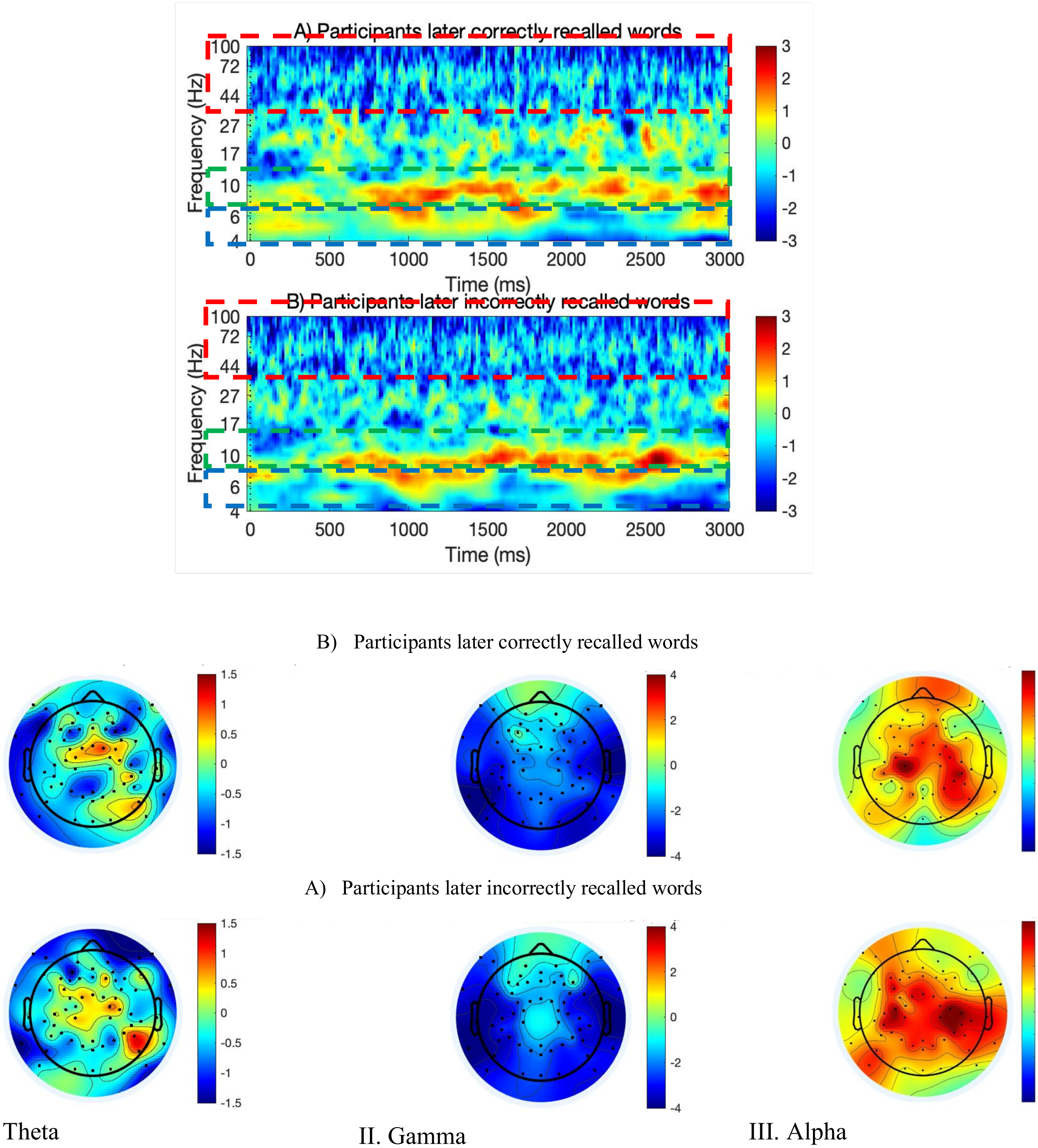
Note. Theta power was significantly higher for correctly recalled words. The error bars represent standard error of the mean.

Interestingly, theta power is more consistent early on (e.g., 0 – 1500 ms) across the encode window for when participants later correctly recalled words.

Importantly, there was a main effect of memory performance for theta power, $F(1, 57) = 4.63$, $p = 0.04$ (see Figure 10), such that mean theta power for correctly recalled words ($M = 0.40$. SE = 0.27) was higher than incorrectly recalled words ($M = -0.05$, $SE = 0.27$).

Figure 11

(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for performance only

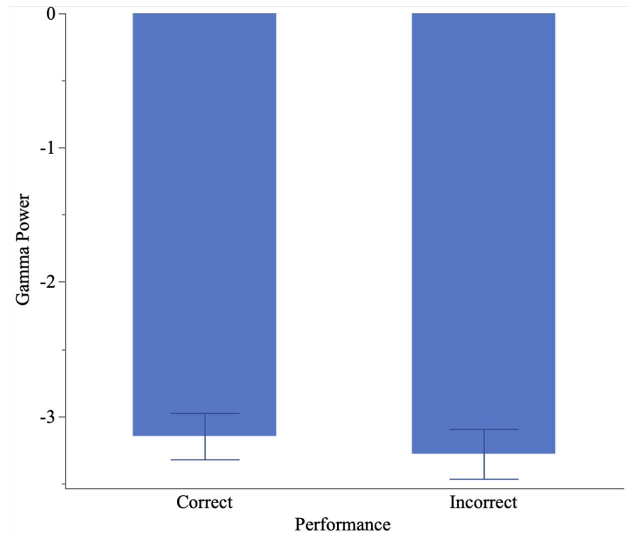


Note. (top) Event-related spectral perturbation activity for A) Later correctly recalled words and B) later incorrectly recalled words for the entire encoding period. The dotted blue line is the theta band, the dotted red line is the gamma band, the dotted green line is the alpha band.; (bottom) Power value across all 62 electrodes for A) Correctly recalled words and B) Incorrectly recalled words for I) Theta power, II) Gamma power, III) Alpha power. Warmer colors (e.g., more red) indicate higher power values than baseline activity; whereas, cooler colors (e.g., blue) represent less power values than baseline activity.

Gamma. It was hypothesized that power in the gamma band would be weaker when participants experienced a memory failure (**H1**). When looking at the ERSP to visually evaluate how the gamma band looked across the entire encode period, there was not a specific time window during which either correctly or incorrectly recalled words had higher values. Power values were consistently the same across the encode window regardless of performance (see Figure 10, red dotted line). This was confirmed by the ANOVA: There was no significant main effect of memory performance for gamma power, $F(1, 57) = 0.73$, $p = 0.39$ (see Figure 12; M correct = -3.15, $SE = 0.22$; M incorrect = -3.28, $SE = 0.22$).

Figure 12

Mean gamma power value for correctly and incorrectly recalled words

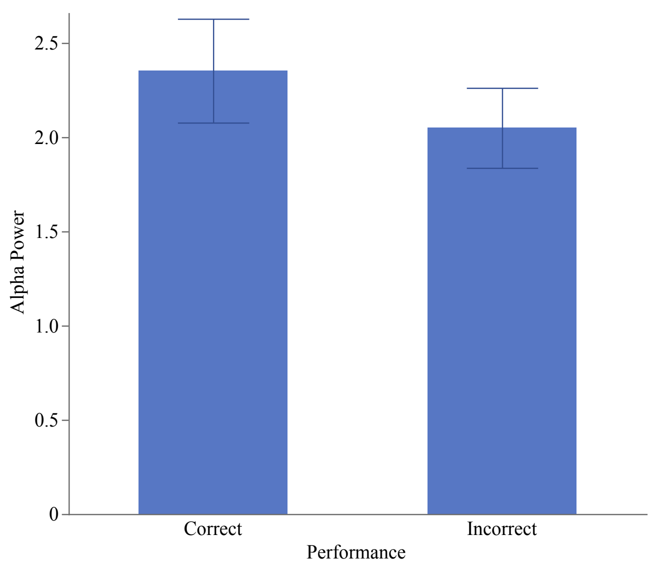


***Note.* There was no difference in gamma power for correctly and incorrectly words. Error bars represent standard error of the mean.**

Alpha. To be complete, the full ANOVA is reported for alpha power; however, there are no aprior hypotheses about alpha power's influence on memory

Figure 13

Mean alpha power value for correctly and incorrectly recalled words



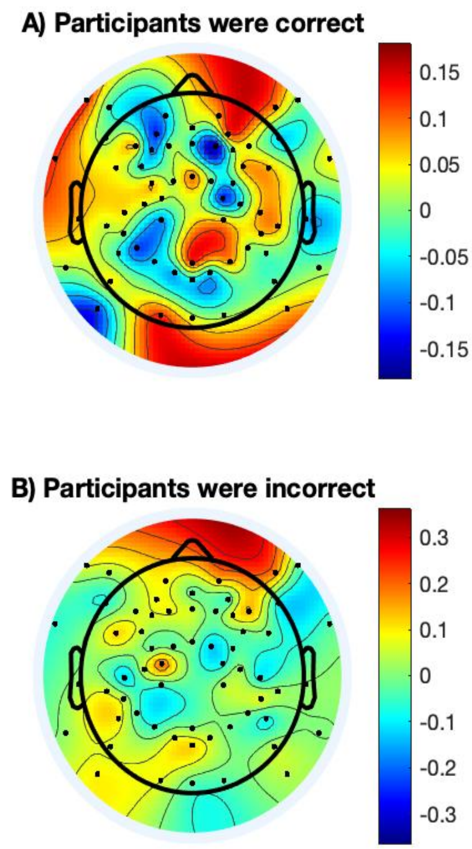
***Note.* Alpha power was significantly higher for correctly recalled words. Error bars represent standard error of the mean.**

performance. When looking at the ERSP for alpha power patterns (see Figure 10, green dotted line), it is visually more consistent and higher across the entire time window for later incorrectly recalled words than later correctly recalled strategies. There was a significant main effect of performance for alpha power, $F(1, 57) = 6.37$, $p = .01$, such that correctly recalled words ($M = 2.35$, $SE = 0.23$) had higher alpha power than incorrectly recalled words ($M = 2.05$, $SE = 0.23$; see Figure 13).

Theta-Gamma Coupling. Lastly, I hypothesized that theta-gamma coupling power values would be weaker when participants experience memory failures (**H2**). Based on the topography scalp distribution of the correlations, it looked like stronger negative correlations between theta and gamma power towards the front of the scalp when participants were correct (see Figure 14), but

Figure 14

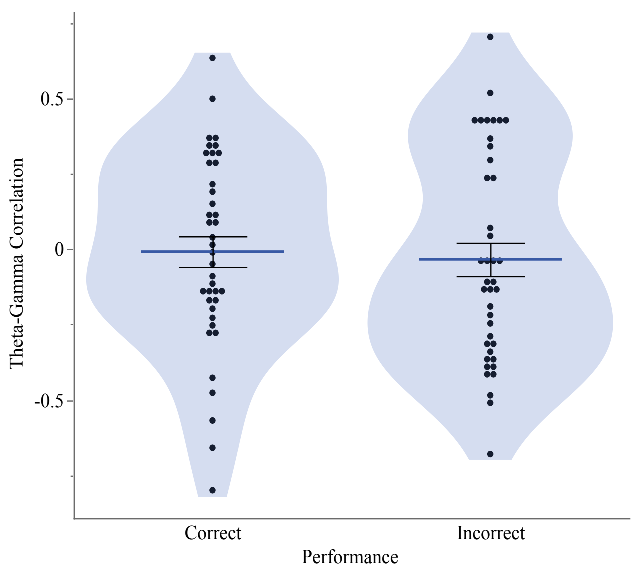
Topography scalp map for distribution of correlation strength between theta and gamma power across the scalp for A) when participants recalled words correctly and B) when participants recalled words incorrectly



Note. Cooler colors indicate higher, negative correlations & warmer colors indicate higher, positive correlations.

Figure 15

Violin plot showing the spread of the theta-gamma coupling power values for performance



Note. There was no difference in correlation strength between correctly and incorrectly recalled words. The blue, horizontal line represents the mean value of the data. The error bars represent standard error of the mean.

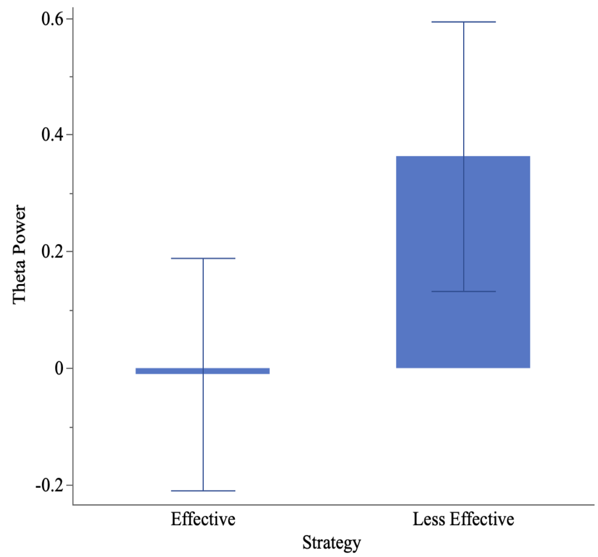
there was no main effect of performance on theta-gamma coupling correlations, $F(1, 57) = 0.10$, $p = 0.75$ (see Figure 15). Regardless of whether participants later remembered words correctly or incorrectly, theta-gamma coupling was recruited similarly.

Strategy Type

Theta. It was hypothesized that theta power would be higher when participants use effective strategies compared to less effective strategies (**H3**). When looking at the ERSP for theta power (see Figure 18, blue dotted line), it is visually more present for less effective strategies than effective strategies, specifically early on in the encoding window. However, there was no significant main effect of strategy type, $F(1, 57) = 3.16$, $p = 0.08$ (see Figure 16), although the difference in mean theta power was marginal between effective ($M = -0.01$, $SE = 0.27$) and less effective strategies ($M = 0.36$, $SE = 0.27$).

Figure 16

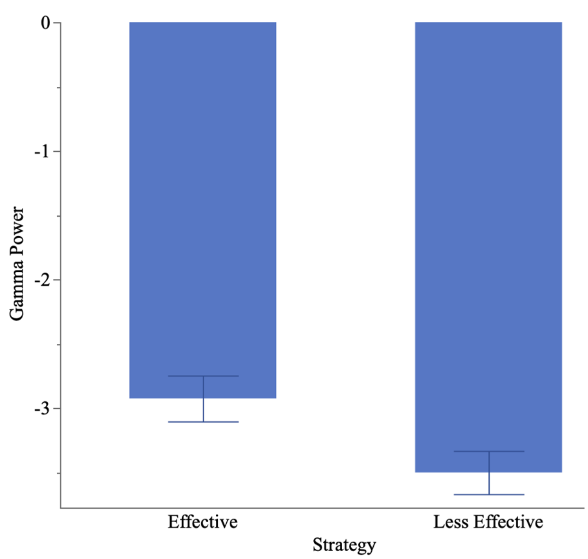
Mean theta power value for effective and less effective strategies



Note. There was no difference between strategies. Error bars represent standard error of the mean.

Figure 17

Mean gamma power values for effective and less effective strategies

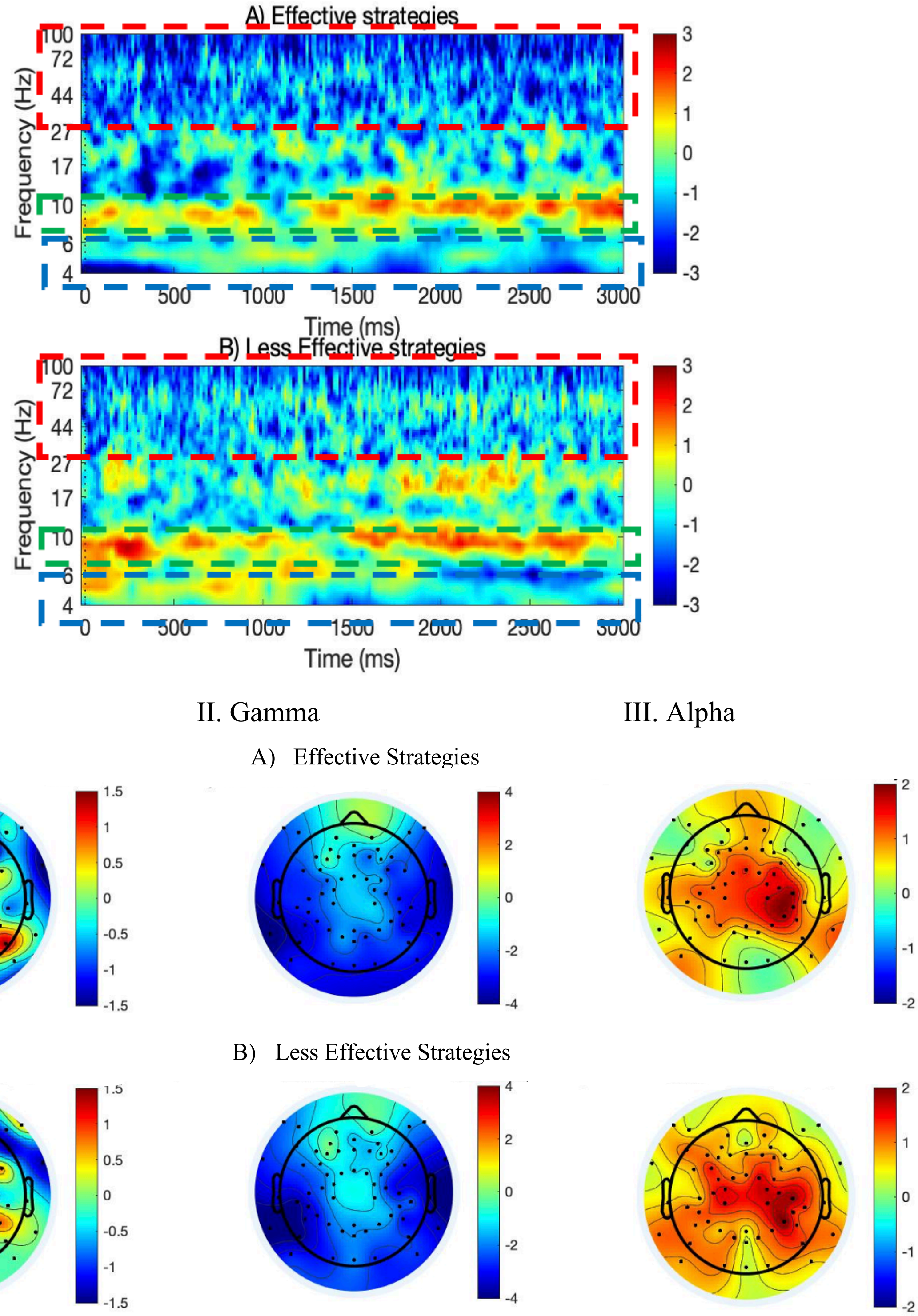


Note. Gamma power was significantly higher for less effective strategies. Error bar represents standard error of the mean.

Gamma. It was hypothesized that gamma power would be higher when participants use effective strategies compared to less effective strategies (**H3**). When looking at the ERSP for gamma power patterns (see Figure 18, red dotted line), it is visually more present for less effective strategies across the entire encoding window. The main effect of strategy type was significant for gamma power, $F(1, 57) = 13.89$, $p < 0.001$ (see Figure 17), with effective strategies ($M = -2.93$, $SE = 0.22$) having higher gamma power values than less effective strategies ($M = -3.50$, $SE = 0.22$).

Figure 18

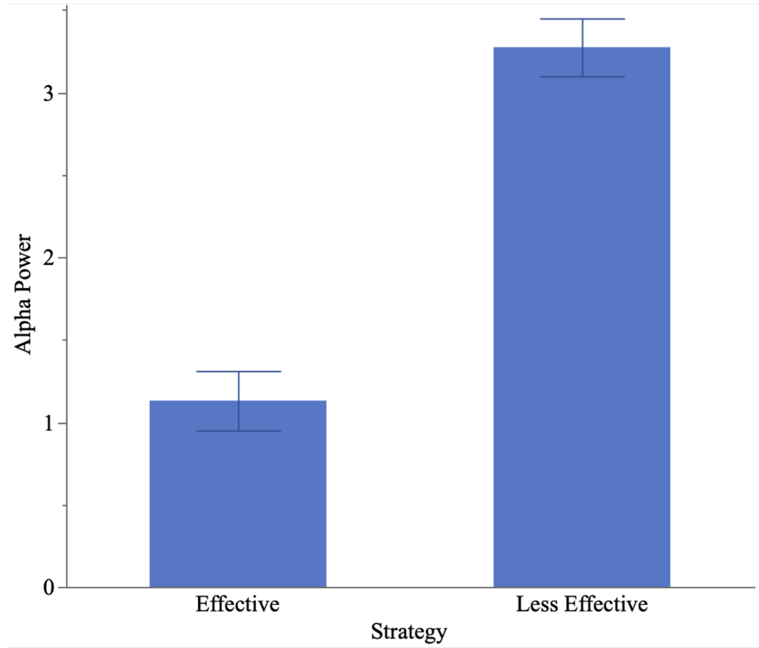
(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for strategies only



Note. (top) Event-related spectral perturbation activity for A) Effective strategies (e.g., sentence generation or imagery) and B) less effective strategies (e.g., reading or rehearsing the word) for the entire encoding period. The dotted blue line is the theta band, the dotted red line is the gamma band, the dotted green line is the alpha band.; (bottom) Power values for A) Effective strategies and B) Less effective Strategies for I) Theta power, II) Gamma power, III) Alpha power. Warmer colors (e.g., more red) indicate higher power values than baseline activity; whereas, cooler colors (e.g., blue) represent less power values than baseline activity.

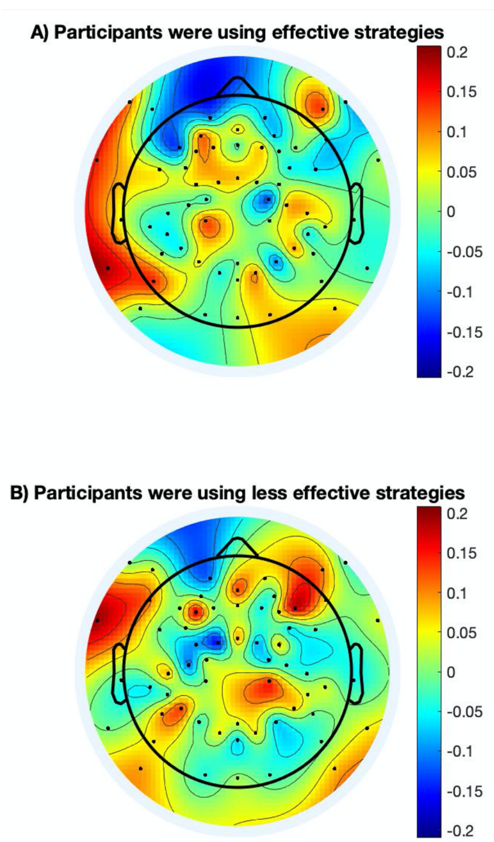
Alpha. It was hypothesized that effective strategies will yield higher power values in the alpha band when compared to the less effective strategies (**H3b**). When looking at the ERSP for alpha power patterns (see Figure 18, green dotted line), it is visually more consistent across the entire time window for less effective strategies than effective strategies. This was confirmed with the ANOVA: There was a significant main effect of strategy type, $F(1, 57) = 317.68$, $p < .001$, such that less effective ($M = 3.27$, $SE = 0.23$) had higher alpha power than effective strategies ($M = 1.13$, $SE = 0.23$; see Figure 19).

Figure 19
Mean alpha power values for effective and less effective strategies



Note. Alpha power was significantly higher for less effective strategies. Error bars represent standard error of the mean.

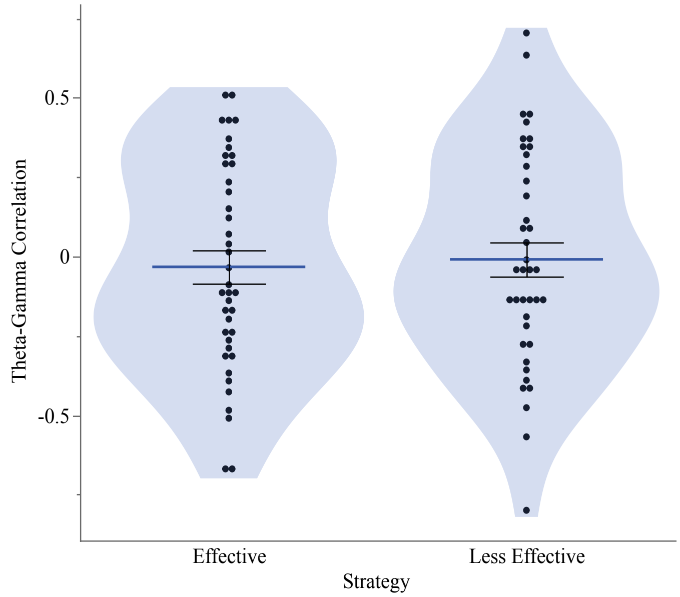
Figure 20
Scalp distribution of theta-gamma correlation strength across all electrodes for participants when they were A) using effective strategies or B) using less effective strategies



Theta-Gamma Coupling. Lastly, it was hypothesized that (**H3a**) that theta-gamma coupling would be strongest for participants using effective strategies. For each participant, I calculated the mean theta and gamma power for each trial in which participants reported using a less effective strategy and the mean theta and gamma power for each trial in which participants reported using an effective strategy. When looking at topography of correlation values across the scalp (see Figure 20), participants had higher correlation values between theta and gamma power on the left side of the scalp and lower correlation values between theta and gamma power on front, left side. Theta-gamma correlation values were higher for frontal electrodes when participants used less effective strategies than when they used more effective strategies (see clustered electrodes in Figure 8); however, there was no significant main effect of strategy type on theta-gamma correlations, $F(1, 57) = 0.09$, $p = 0.77$ (see Figure 21).

Figure 21

Violin plot showing the spread of the theta-gamma coupling power values for strategy type



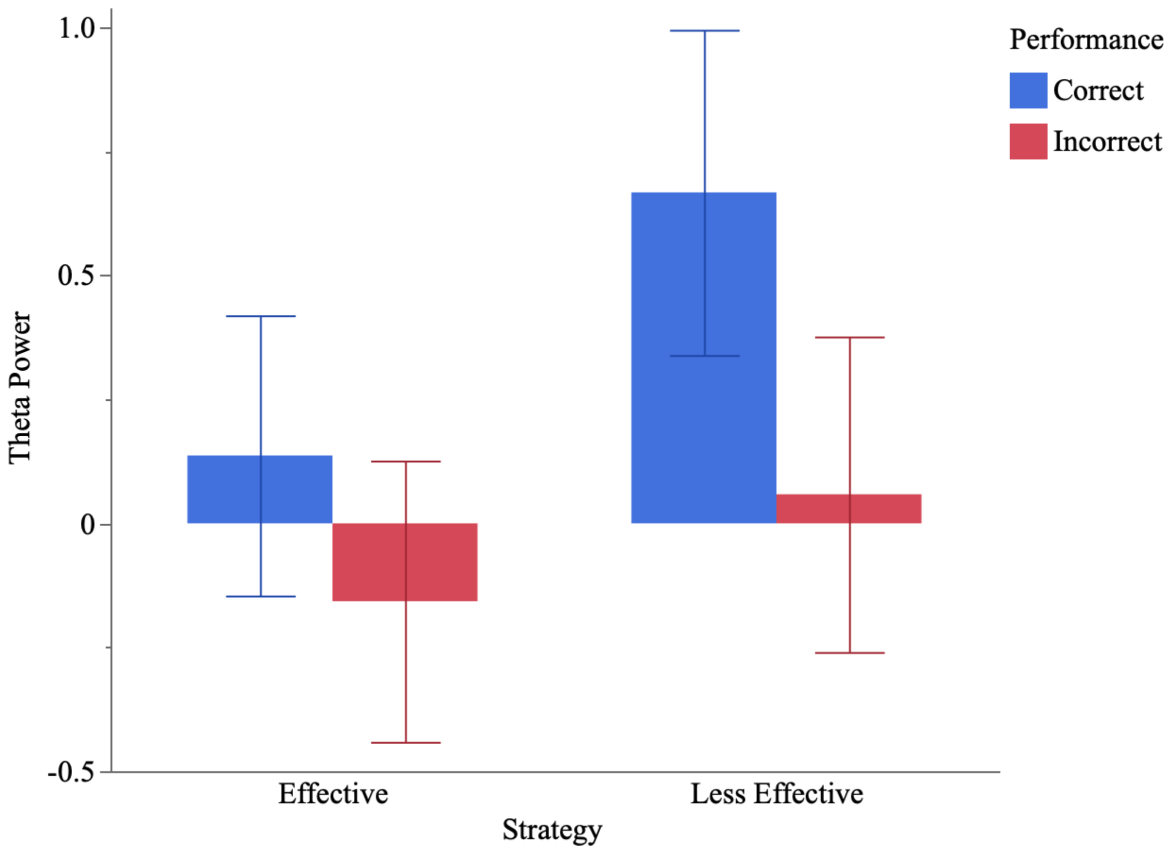
Note. There was no difference in theta-gamma correlations between strategies. The blue, horizontal line represents the mean value of the data. The error bars represent standard error of the mean.

Strategy Type x Performance

Theta. It was hypothesized that theta would increase when participants effectively use elaborate (and more effortful) strategies (**H4**). When looking at the ERSP for these four conditions, a few patterns emerge overall (see Figure 23, see blue dotted line). Specifically, theta power is noticeably higher for the Correctly Recalled Less Effective Strategies condition and the Incorrectly Recalled Effective Strategies than the other two conditions. However, theta power is more consistent across the encode window for correctly recalled words while using effective strategies and incorrectly recalled words while using less effective strategies. However, the interaction was not significant for theta power, $F(1, 57) = 0.56$, $p = 0.47$ (see Figure 22).

Figure 22

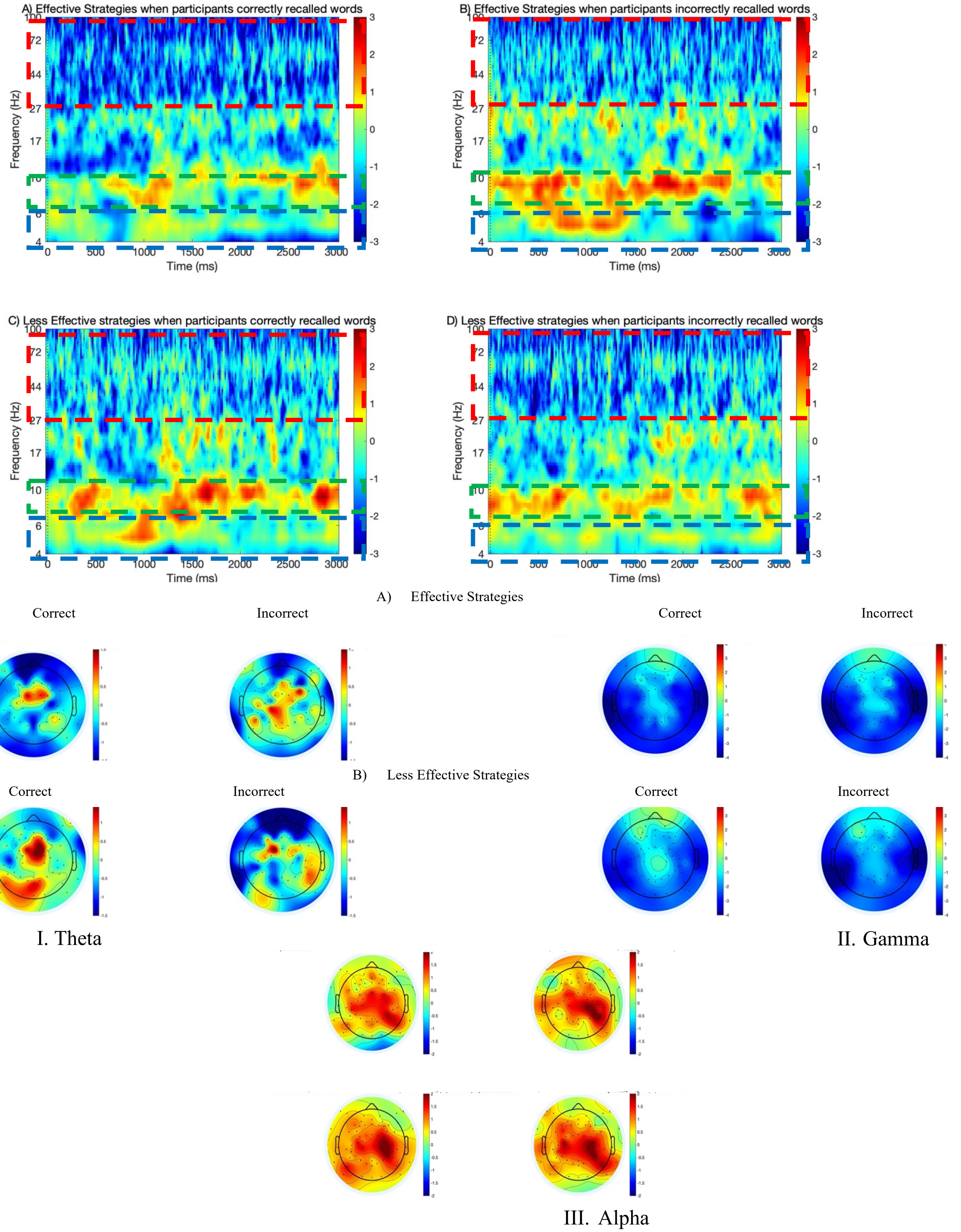
Theta power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words



Note. Error bars represent standard error of the mean. There was no difference between strategies that depended on whether participants were correct.

Figure 23

(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for strategies, depending on performance

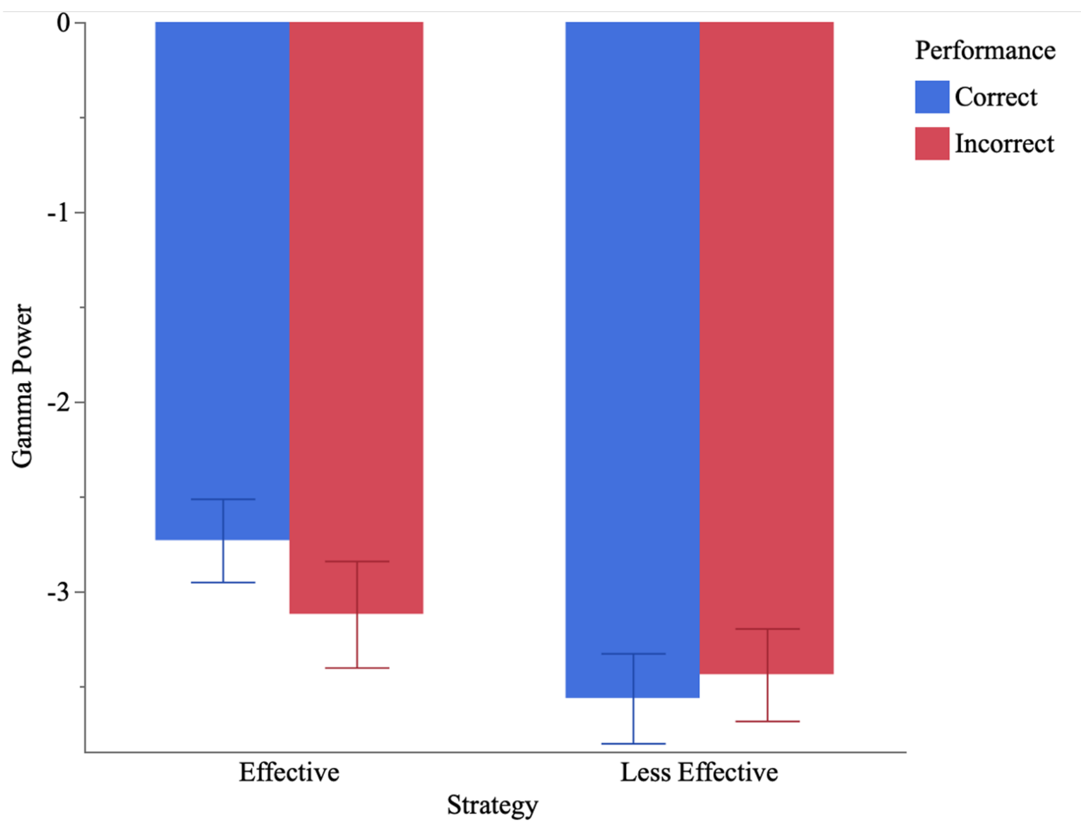


Note. (top) Event-related spectral perturbation activity for A) Effective – Correct, B) Effective – Incorrect, C) Less Effective – Correct, and D) Less Effective – Incorrect for the entire encoding period. The dotted blue line is the theta band, the dotted red line is the gamma band, the dotted green line is the alpha band.; (bottom) Power values for A) Effective strategies and B) Less effective Strategies for I) Theta power, II) Gamma power, III) Alpha power.

Gamma. It was hypothesized that gamma when participants effectively use elaborate (and more effortful) strategies (**H4**). When looking at the ERSP for these four conditions across the entire time window (rather than analysis window), a few patterns emerge overall (see Figure 23, see red dotted line). Similar to the theta band, the gamma band is (visually) more noticeable for incorrectly recalled words while using effective strategies and correctly recalled words while using less effective strategies; however, the interaction was not significant for gamma power, $F(1, 57) = 2.77$, $p = 0.10$ (see Figure 24).

Figure 24

Gamma power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words



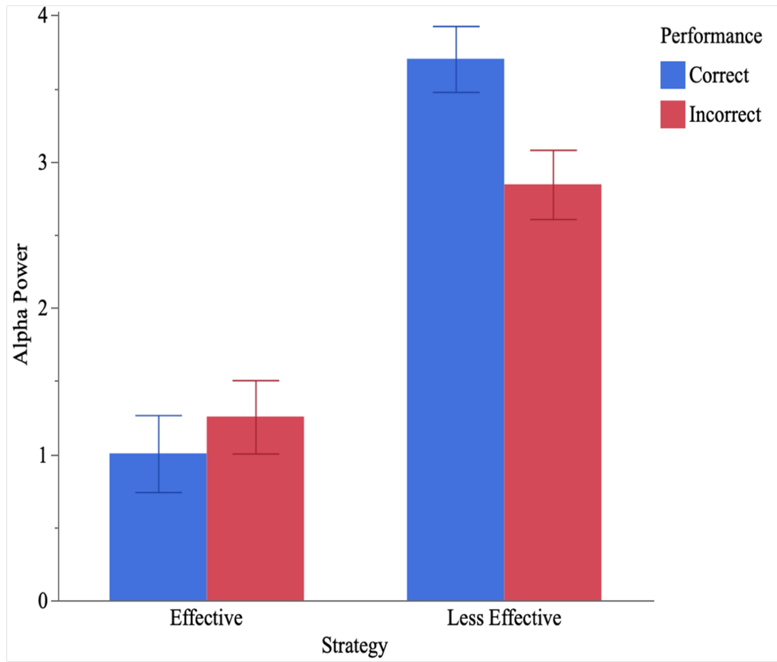
Note. Error bars represent standard error of the mean. There was no difference between strategies that depended on whether participants were correct.

Alpha. It was hypothesized that mean alpha power would be lower when participants used effective strategies and were correctly recalling words than in the other 3 conditions (**H4b**).

When looking at the ERSP for these four conditions, a few patterns emerge overall (see Figure 23, see green dotted line). Alpha is visually more noticeable for incorrectly recalled words while using effective strategies and correctly recalled words while using less effective strategies. The Strategy Type x Performance interaction was significant for alpha power, $F(1, 57) = 21.27$, $p < .001$ (see Figure 25). A Tukey HSD probed the mean difference between the four conditions (e.g., Correct – Effective, Correct – Less Effective, Incorrect – Effective, Incorrect – Less Effective). Alpha power was higher when participants used less effective strategies to correctly recall words ($M = 3.70$) than all other conditions (i.e., Incorrect – Less Effective $M_{diff} = 0.86$, $SE = 0.17$; Correct – Effective $M_{diff} = 2.70$, $SE = 0.17$; Incorrect – Effective $M_{diff} = 2.45$, $SE = 0.17$). Alpha power was significantly higher for Incorrect – Less Effective ($M = 2.84$) than for Correct – Effective ($M_{diff} = 1.84$, $SE = 0.17$) and Incorrect – Effective ($M_{diff} = 1.59$, $SE = 0.17$). No other differences were significant.

Figure 25

Alpha power values for effective and less effective strategies depending on whether participants later correctly (blue) or incorrectly (red) recalled words



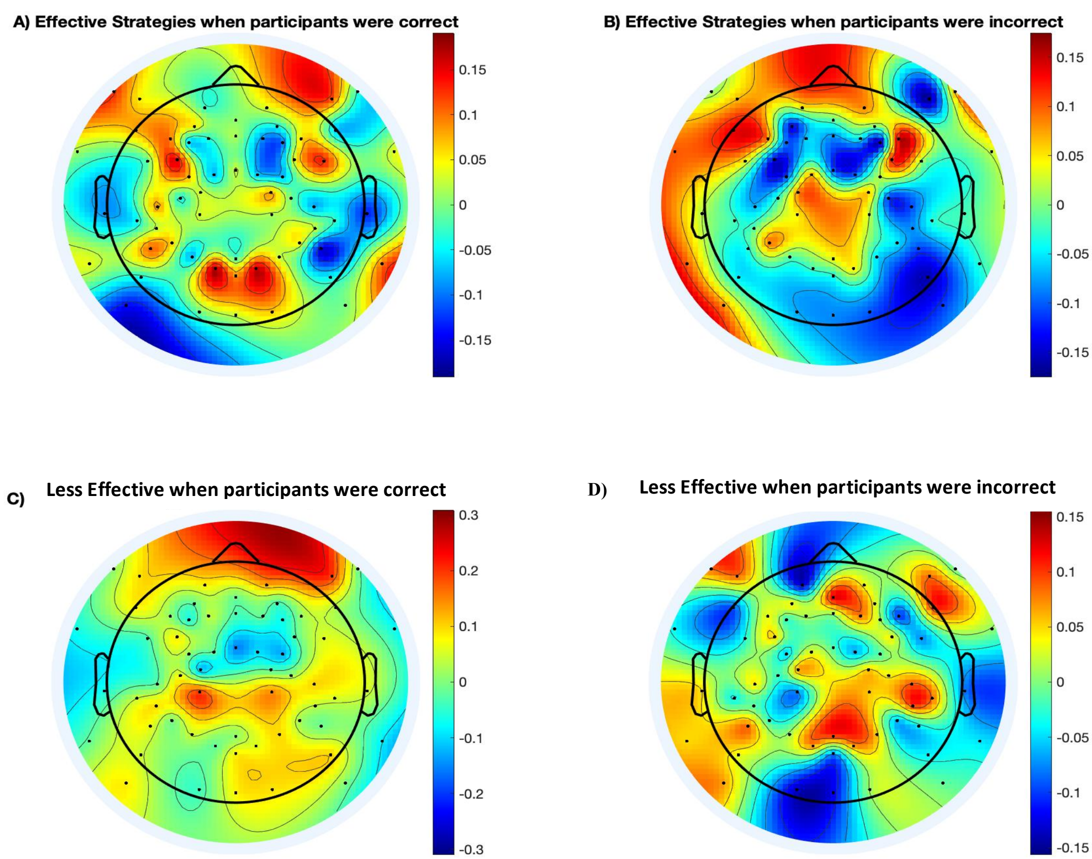
Note. Error bars represent standard error of the mean. Less Effective – Correct had higher alpha power than all other conditions. Less Effective – Incorrect had higher power values than effective strategies, regardless of performance.

Theta-Gamma Coupling.

To test whether theta-gamma coupling power values increases depended on type of strategy participants used (**H4a**), the mean theta and gamma trials for the four conditions were calculated for each participant and then averaged across all participants. The scalp distribution of the correlation strength between theta and gamma values demonstrates higher correlation values in the frontal region of the scalp when participants were using less effective strategies and were later correct (see Figure 26). Interestingly, theta-gamma power values were more coupled together for occipital regions when participants were using effective strategies to correctly recall words or were using less effective strategies to incorrectly recall words. Lastly, occipital regions demonstrated higher, negative correlations in the occipital regions for all conditions except when

Figure 26

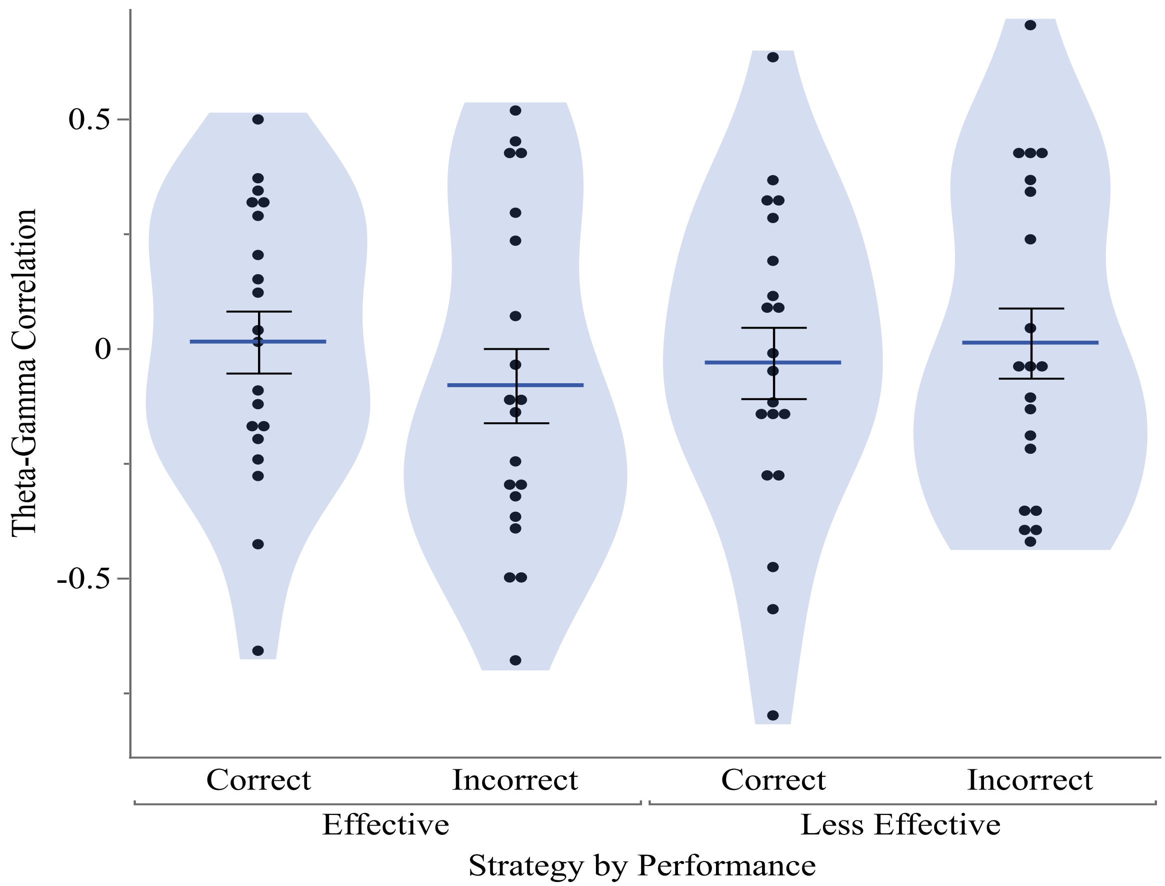
Scalp distribution of correlation strength between theta and gamma values for when participants were using different strategies and whether they correctly recalled words



participants were using less effective strategies incorrectly. However, there was no significant interaction between strategy type and performance on theta-gamma correlations at Fz, $F(1, 57) = 0.74$, $p = 0.39$ (see Figure 27).

Figure 27

Violin plot depicting the spread of the datapoints for strategy types by performance



Note. There was no difference in theta gamma correlation strength between strategies that depended on performance. The blue, horizontal line is the mean value and the error bars represent standard error.

Set Size

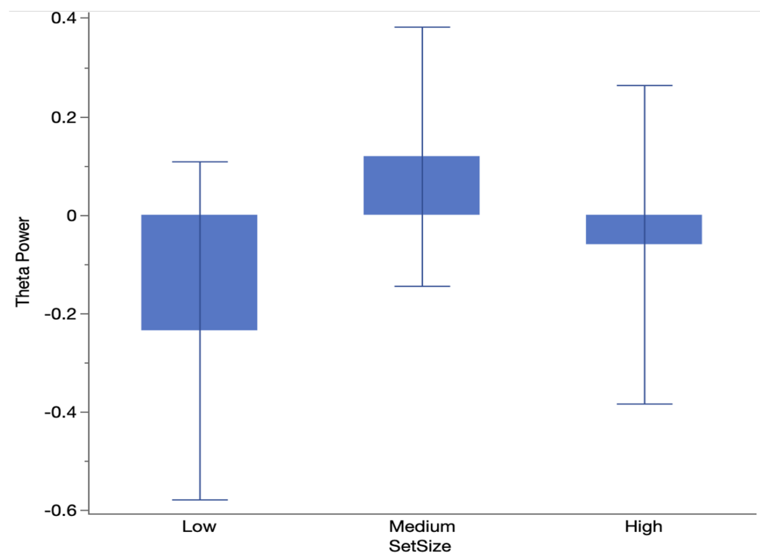
For all set size analyses, the mean power value per frequency band was calculated for each set size for each individual participant in the time window of interest (e.g., 800 – 1100 ms for theta and gamma and 1500 ms – 3000 ms for alpha) and then averaged across participants.

From there, three separate one-way ANOVAs investigated the mean power differences for small, medium, and large set sizes for theta, gamma, and theta-gamma coupling.

Theta. It was hypothesized that theta power would increase as set size increases (**H1a**). When looking at the entire encode window, theta power increased as set size increased (see Figure 29, blue dotted line), but not consistently across the time window for any of the set sizes. Thus, specifically looking at the time window of interest (800-1100 ms), there was no significant difference in theta power between low (M = -0.23, SE = 0.31), medium (M = 0.12, SE = 0.31), and high set sizes (M = -0.06, SE = 0.31), $F(2, 38) = 0.65$, $p = 0.53$ (see Figure 28).

Figure 28

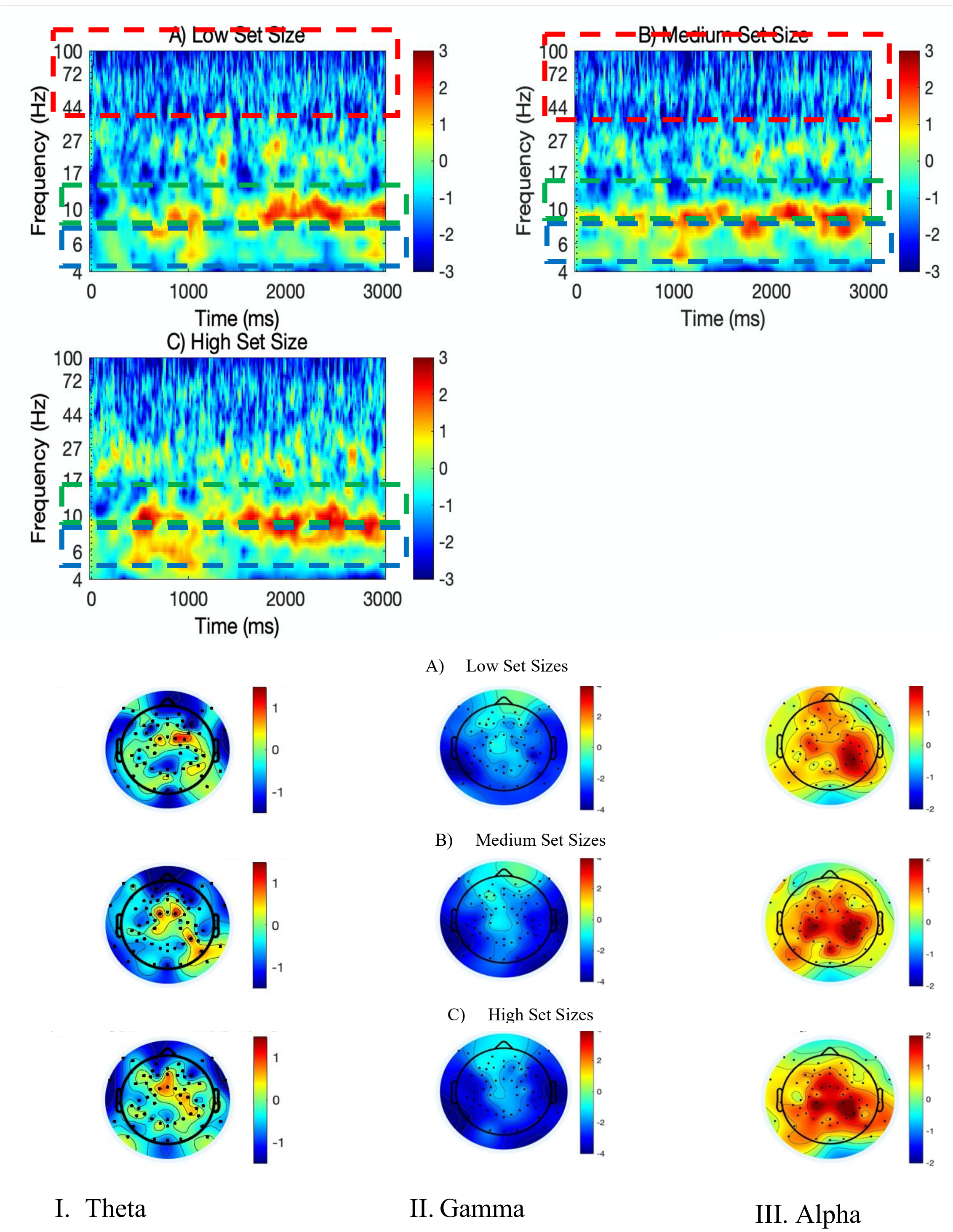
Theta power values for each of the set sizes



Note. The error bars represent standard error of the mean. There was no difference in theta power between set sizes.

Figure 29

(top) Event-related spectral perturbation activity during encoding window and (bottom) power values across all 62 electrodes for set sizes

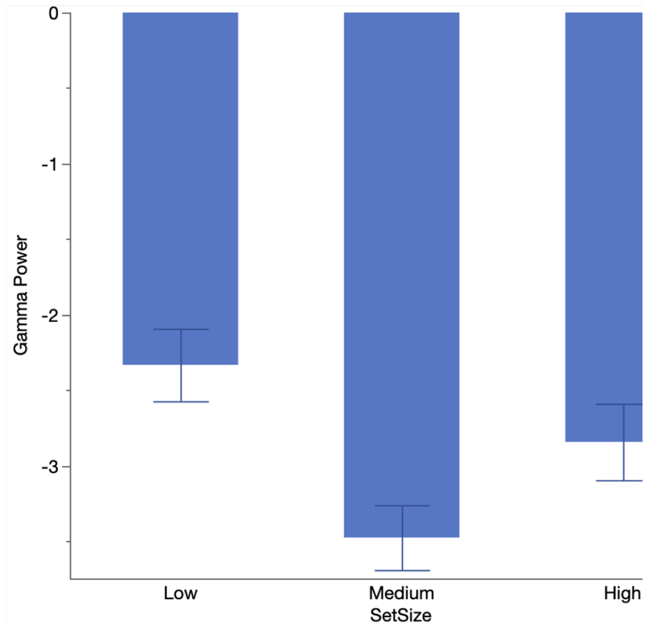


Note. (top) Event-related spectral perturbations for A)) Low set sizes (3-5 word pairs), B) Medium set sizes (6 & 7 word pairs), C) High set sizes (8-10 word pairs) across the entire encoding period; (bottom) Power values across all 62 electrodes for A) low set sizes, B) medium set sizes, C) high set sizes for I) Theta power, II) Gamma power, III) Alpha power. Warmer colors (red) indicated higher power values than baseline. Darker colors (more blue) indicate lower power values than at baseline. The dotted blue line represents the theta band, the dotted red line represents the gamma band, and the dotted green line represents alpha band.

Gamma. It was hypothesized that gamma power would increase as set size increases (**H1a**). When looking at the entire encoding window, gamma power did not increase noticeably as set size increased (see Figure 29, red dotted line); however, there was a significant difference in gamma power between low, medium, and high set sizes, $F(2, 38) = 10.00$, $p < .001$ (see Figure 30). A Tukey's HSD probed the difference in gamma power between set sizes. Medium set sizes ($M = -3.47$) had significantly lower gamma power than low set sizes ($M_{diff} = -1.14$, $SE = 0.26$) and high set sizes ($M_{diff} = -0.63$, $SE = 0.26$), but low set sizes did not differ from high set sizes ($M_{diff} = 0.51$, $SE = 0.26$).

Figure 30

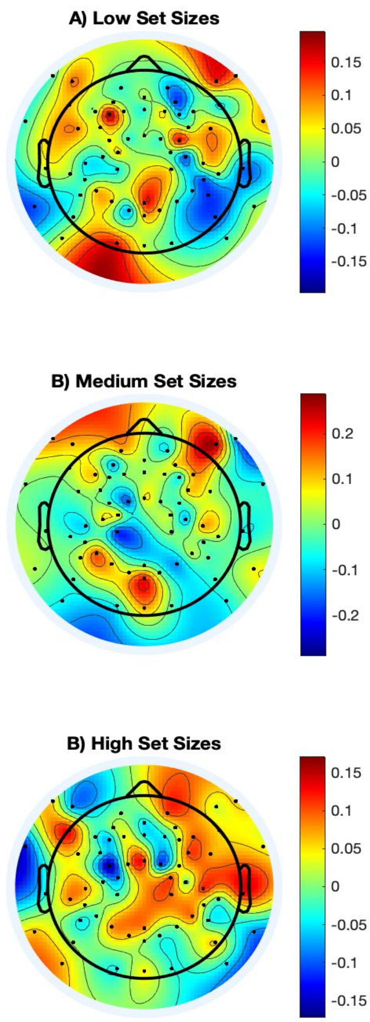
Gamma power values for set sizes



Note. Error bar represents standard error of the mean. Medium set sizes had significantly lower power values than low and high set sizes.

Figure 31

Topography distributions of the correlation in theta-gamma values across set sizes

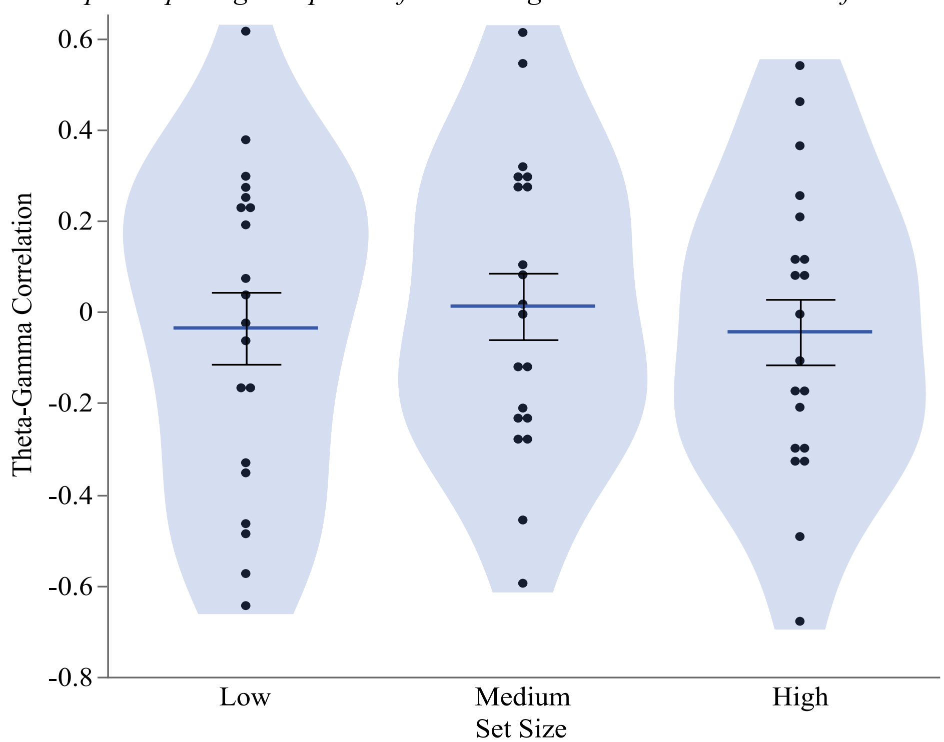


Note. Cooler, blue colors represent lower correlations and warmer, red colors represent higher correlations.

Theta-Gamma Coupling. To test whether theta-gamma coupling power values increased across set sizes (**H2a**), the mean theta and gamma power values for each individual trial within a set size category was averaged across participants (e.g., all trials from low, medium and high set sizes). When looking at the topography of the scalp distribution for the correlation strengths, higher set sizes had higher theta-gamma power correlations around the right temporal region compared to the other two conditions (see Figure 31). Low set sizes also had higher theta-gamma correlations in the left occipital region of the scalp. It was hypothesized that higher set sizes (e.g., set sizes 8, 9, and 10) would yield higher correlations between theta and gamma values than lower set sizes (e.g., set sizes 3, 4, and 5; **H2a**). The one-way repeated measures ANOVA revealed no significant effect of set size on theta-gamma coupling values, $F(2, 38) = 0.17$, $p = 0.85$ (see Figure 32).

Figure 32

Violin plot depicting the spread of the theta-gamma correlation values for set sizes

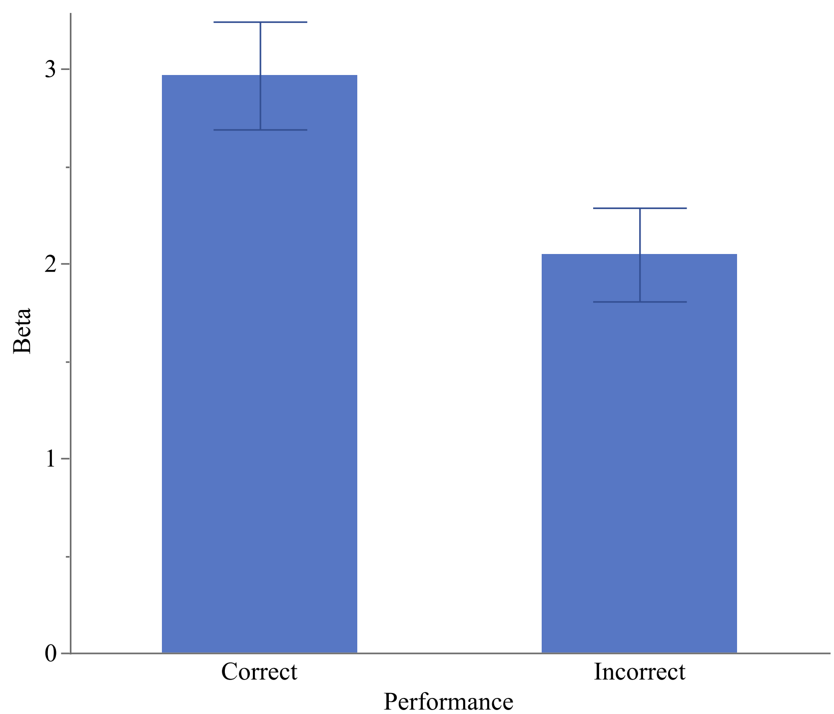


Note. There was no difference in theta-gamma correlation strength between set sizes. The blue horizontal line represents the mean of the data and the error bars represent standard error.

Exploratory EEG Analyses

To explore alpha power more, I used a data-driven approach to evaluate whether Fz was the best electrode for detecting alpha power (e.g., the strongest alpha power). When looking the figure in Appendix B, the electrode with the highest alpha activity was at P4. Exploratory analyses were re-run to evaluate whether effects found at Fz would replicate. When looking at P4 for alpha power values, similar results emerged, where there was a significant difference in alpha power values (M_{diff} = 0.91, *SE* = 0.29), $t(19) = 3.18$, $p = .005$, with incorrectly recalled words (*M* = 2.53) having higher alpha power than correctly recalled words (*M* = 1.62). There was also a significant difference between less effective (*M* = 2.96) and effective strategies (*M* = 1.02), $t(19) = 6.63$, $p < .001$, with less effective strategies contributing to more alpha power. As such, Fz was sensitive to picking up alpha power, despite being a frontal electrode.

Figure 33
Mean beta power values for performance

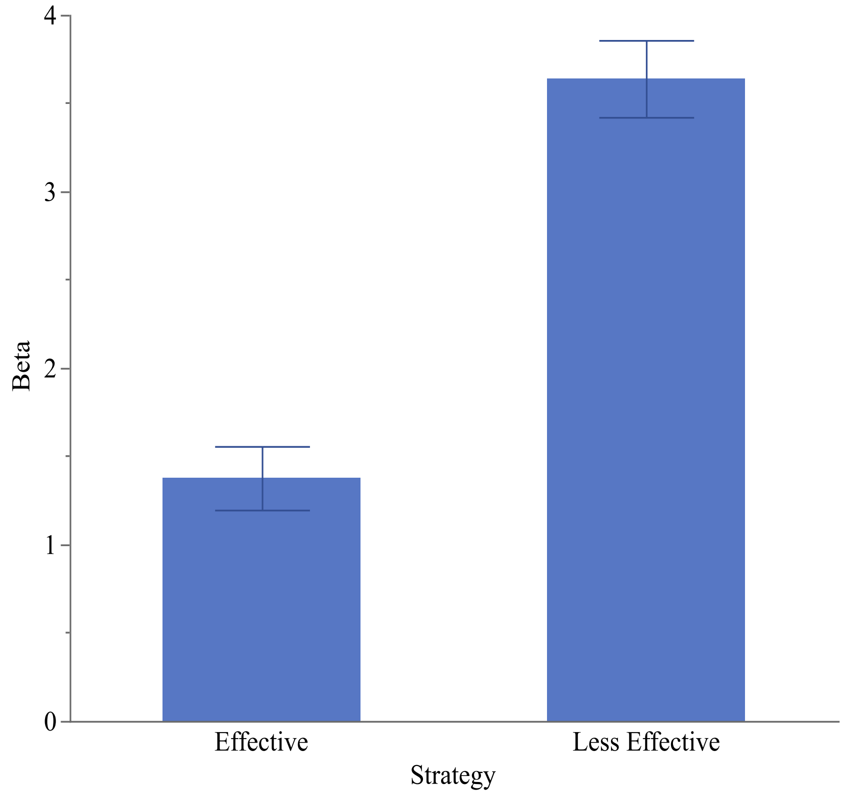


Note. There was a significant difference, such that correctly recalled words recruited higher beta power. Error bars represent standard error of the mean.

More interestingly, one of the most active frequency bands in the ERSPs (Figure 11, 18, 23, and 29) was the beta power band (17 – 27 Hz). Given there was no a-prior predictions for this band, exploratory analyses were run to evaluate the effects of beta activity at Fz. A 2 (performance: correct vs. incorrect) by 2 (strategy type) repeated-measures ANOVA was run to evaluate differences in beta power. There was a main effect of performance for beta power, F(1, 57) = 38.49, $p < .001$ (see Figure 33), such that beta power was higher for correctly recalled words ($M = 2.97$, $SE = 0.24$) than for incorrectly recalled words ($M = 2.05$, $SE = 0.24$). There was a main effect of strategy for beta power, F(1, 57) = 232.52, $p < .001$ (see Figure 34). Less effective strategies ($M = 3.63$, $SE = 0.24$) recruited higher beta power values than effective strategies ($M = 1.38$, $SE = 0.24$). Further, the Strategy x Performance interaction for beta power was also significant, F(1, 57) = 12.43, $p = .001$ (see Figure 35). A Tukey HSD poct-hoc analysis showed that Correct – Less Effective words ($M = 4.26$, $SE = 0.26$) recruited higher beta power than Correct - Effective ($M_{diff} = 2.79$, $SE = 0.21$), Incorrect - Effective ($M_{diff} = 3.18$, $SE = 0.21$), and Incorrect – Less Effective words ($M_{diff} = 1.44$, $SE = 0.21$). Incorrect – Less Effective ($M = 2.92$, $SE = 0.26$) recruited higher beta power than Correct - Effective ($M_{diff} = 1.34$, $SE = 0.21$) and Incorrect - Effective ($M_{diff} = 1.74$, $SE = 0.21$). No other significant differences were found.

Figure 34

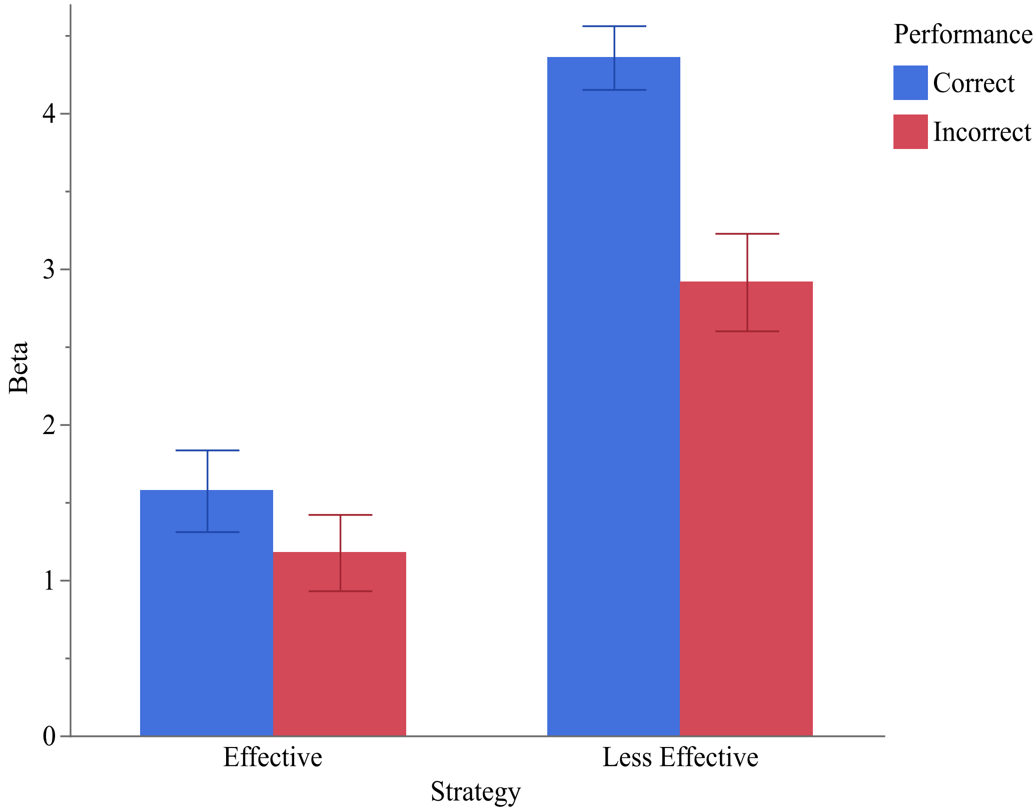
Mean beta power values between types of strategies



Note. There was a significant difference in beta power, such that less effective strategies recruited higher beta power. The error bars represent standard error of the mean.

Figure 35

Mean beta power values between strategies depending on whether participants later correctly recalled words

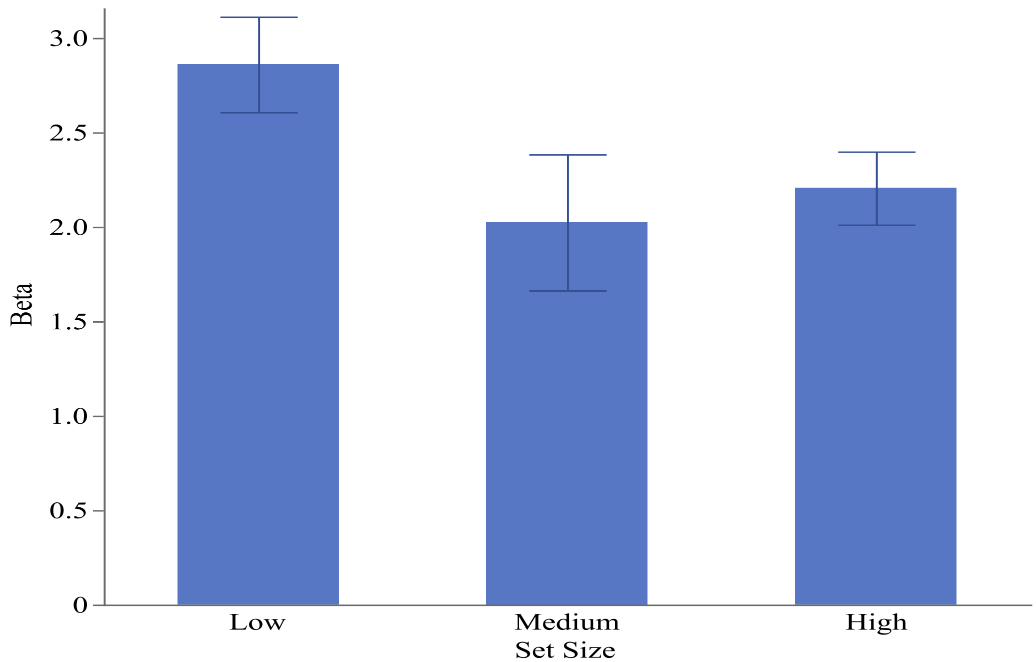


Note. Less effective strategies used to correctly recall words recruited higher beta power than all other conditions. Overall, less effective strategies recruited higher beta power.

Lastly, the difference in beta power recruitment for set sizes was assessed by a repeat-measures one-way ANOVA. There was a significant difference in beta power recruitment between the set sizes, $F(2, 38) = 3.72$, $p = 0.03$ (see Figure 36). A Tukey HSD post-hoc analysis showed that low set sizes ($M = 2.86$, $SE = 0.28$) recruited significantly higher beta power than medium set sizes ($M_{diff} = 0.84$, $SE = 0.32$). No other significant differences were found.

Figure 36

Mean beta power recruitment between set sizes



Note. Low set sizes recruited significantly higher beta power than only medium set sizes.

General Discussion

The primary purpose of the current study was to evaluate neural signatures associated with memories failures and memory strategies. In general, the above study evaluated the following: whether neural patterns differed between successful and unsuccessful memory performance (**RQ1**), whether effective strategies yielded different neural patterns from less effective strategies (**RQ2**) and these neural patterns validated behavioral strategy reports and performance (**RQ3**).

Neural Markers of Memory Failures

Importantly, differences in strategy performance were revealed in the behavioral analyses of the current study. Participants demonstrated that using more effective strategies increases one's ability to recall words on a paired-associates task replicating past work (Bailey et al., 2011; Dunlosky & Kane, 2007; Zakrzewski, Sanders, & Berry, 2021, Dunlosky & Hertzog, 2001), This is presumably because more effort and deeper processing is used to connect word stimuli to one another (i.e., chunking) to increase later recall (Craik & Watkins, 1973; Richardson, 1998). Additionally, participants recalled words in lower set sizes (e.g., set size of 3 and 4) as compared to words in higher set sizes (e.g., set size of 9 and 10; Jensen and Tesche, 2002; Cowan et al., 2012; Waters & Caplan, 1996). Once again, aligning with past studies that suggest it is easier to recall words when there were only a few words required to connect to one another and remember. As such, the behavioral analyses have given validity to the task participants used to remember words in the current study.

More relevant to this proposal, **RQ1** aimed to replicate brain activity behind memory performance differences in past studies. It was hypothesized that both theta and gamma power would be weaker for memory failures (**H1**). This pattern was observed in theta power, which was higher for correctly recalled words than incorrectly recalled words, replicating past studies (Fell et al., 2003; Gruber et al., 2001; Onton et al., 2005). However, this pattern was not observed for gamma power, which fails to replicate past studies (Fell et al., 2003; Gruber et al., 2001; Onton et al., 2005). Again, gamma power values were negative, indicating they had less gamma than in the baseline. This gamma suppression was not predicted and has not been shown in similar prior studies evaluating verbal memory. Similar to the current study, Gruber et al. (2001) used a paired-associates learning task and found gamma enhancement, but their study did not vary set sizes like the current study did. As such, the lack of gamma power differences could be due to differences in task. The current study had varying set sizes rather than a continuous list, which are two conceptually different ways of evaluating memory recall because one introduces periodic breaks and leaves less study time. So, **H1** was partially supported.

Relatedly, it was hypothesized that theta-gamma coupling power values would also be lower for memory failures (**H2**). However, there was no main effect of performance for theta-gamma coupling, indicating that participants recruited similar theta-gamma values in Fz regardless of whether a word was later recalled correctly or not. Thus, we failed to replicate past studies that showed higher theta-gamma power coupling for later remembered words (Axmacher et al., 2010, Canolty et al., 2006, Friese et al., 2012; Schack et al., 2002; Sederburg et al., 2003) and did not support **H2**. However, it is important to note that this study only evaluated power-power coupling, whereas the past work has evaluated mostly power-phase or phase-phase coupling. Sederburg et al. (2003) was one study that looked at power-power coupling, but they evaluated this question using a delayed free-recall task. So, it is possible that our task has enough important differences from those used in past studies, which primarily looked at memory on Sternberg tasks (Axmacher et al, 2010; Schack et al, 2002), auditory working memory (Canolty et al., 2006), or remember/know procedures (Friese et al., 2012). Moreover, we only evaluated theta and gamma activity at Fz, whereas other relevant work has looked at theta-gamma coupling in the hippocampal region (Justras, Fries, & Buffalo, 2009; Belluscio et al., 2012; Sirota et al., 2008; Spellman et al., 2015). This work used a variety of translational research using monkeys and rats to observe patterns of activities in these regions. To evaluate this in humans using EEG, we could utilize dipole modeling and analyze brain activity stemming from this region. This would let us isolate brain activity from only this region to determine patterns of activity.

Lastly, **RQ1** sought to replicate Fell et al. (2003) by finding that theta and gamma power values (**H1a**) and theta-gamma coupling values (**H2a**) would increase as set size increased. The current analysis showed there was no significant difference in theta power between the set sizes, but there was a significant different in gamma power for set size. Interestingly, gamma power increased more for medium set sizes when compared to low set sizes, but low set sizes were not different from higher set sizes. Assuming that gamma power is correlated with binding incoming information into LTM, it is possible that the lack of difference between low and high set sizes could be due to difficulty of integrating information. Behavioral studies have shown it is easier to remember smaller set sizes than higher set sizes. In the current study, it is possible that the lower set sizes were cognitively easy, and therefore, did not require much integration. On the other hand, higher set sizes could be too cognitively taxing, and therefore, tapping out the participant's capacity limit. Thus, it could explain why medium set sizes, which provide more of a cognitive challenge than low set sizes but not tapping out resources quite yet, had higher gamma power values. However, it is important to note that Figure 6 showed that set sizes of 3 were not at ceiling level, and instead, sat at ~0.69% of correctly recalled words. As such, set size 3 presented enough of a cognitive challenge to deplete performance. Another explanation is that the current study was not powered enough to detect these set size differences. Using an IC approach could remedy this issue and allow us to evaluate theta and gamma across different set sizes. Interestingly, although the theta band did increase throughout set sizes, it was not statistically different in any of the conditions. It is possible that theta differences do exist, but the current analysis approach was not sensitive enough to detect it. If we were to use the approach by Onton et al., (2005) and look at the clustered individual components, this difference might be more apparent. We partially supported **H1a**, but the pattern of gamma power increase was not linear as expected. Lastly, the analysis did not show any differences in theta-gamma coupling correlations between set sizes. Set sizes recruited similar levels of theta-gamma coupling, so we failed to support **H2a**.

Neural Markers of Encoding Strategies

RQ2 aimed to extend prior research by evaluating whether effective strategies activate different neural patterns than do less effective strategies. Specifically, the current hypothesis was that effective strategies would yield both higher theta and gamma power as well as theta-gamma coupling (**H3 and H3a**). Gamma power was significantly higher for effective strategies, and theta power was marginally higher when participants reported using less effective strategies compared to normatively effective strategies. However, it is important to note that mean gamma values were lower than baseline, which may indicate that this effect is not theoretically relevant to the current question. Alternatively, gamma suppression can occur when attention is being redirected (Fell et al., 2002) or when participants are ignoring presented stimuli (Bonnefond & Jensen, 2016). So, while we partially supported **H3**, it is worth noting that these values were lower than expected, which could be gamma suppression and is worth investigating further. Lastly, theta-gamma coupling values did not differ by strategy type; thus, we failed to support **H3a**.

Lastly, given that alpha power inhibits irrelevant information, we expected that participants using effective strategies would suppress irrelevant task information better than those using less effective strategies (and therefore, exhibiting lower alpha power) (**H3b**). There was a significant main effect of strategy type; however, once again, the effect was in the opposite direction. Alpha power was statistically lower for effective strategies, but it is unclear whether this is an advantage or disadvantage for these strategies. Specifically, for this hypothesis, I expected there to be statistically significantly more alpha suppression (e.g., statistically LOWER values than the baseline). For alpha in this task, alpha power enhancement was shown across the task and lower for effective strategies. So, even though alpha was statistically lower for effective strategies, participants demonstrated lower alpha power enhancement rather than alpha power suppression. So, we failed to support **H3b**.

Neural Markers of Effectiveness of Encoding Strategies

Lastly, **RQ3** aimed to build upon strategy differences in the literature by validating behavioral strategy reports. Past research assumed that effective are, not only effortful, but also better for participant performance. Behavioral research conflates accuracy with effort, which is not always the case. So, I expected we would find that theta and gamma values would be stronger when participants effectively use elaborate (and more effortful) strategies (**H4**). In other words, power in these bands were expected to be highest when effective strategies are reported and the word is later correctly recalled. Similarly, I expected theta-gamma coupling increase on correct-recall trials in which participants reported using effective strategies compared to all other trials (**H4a**). Lastly, I hypothesized that alpha power inhibition would be stronger on correct-recall trials in which participants reported using effective strategies compared to all other trials (**H4b**). Thus, supporting behavioral research concluding that effective strategies more beneficial for remembering information and require more cognitive effort to, not only remember, but also to suppress irrelevant information (**H4, H4a, H4b**). We failed to support **H4** and **H4a**: Theta power values, gamma power values, or theta-gamma coupling correlations were not highest for correct-recall trials for effective strategies compared to all other trials (e.g., incorrect and effective, correct and less effective, incorrect and less effective). However, there was a significant interaction for alpha power. Alpha power was highest when participants used less effective strategies used to correctly recall words than all other conditions, and it was also higher for less effective strategies used to incorrectly recall words was higher than whenever participants used effective strategies, regardless of performance. In other words, alpha power enhancement was higher for less effective strategies, but more enhanced when they were later correct, than effective strategies. So, although effective strategies have lower alpha power values, they were shown to have lower alpha enhancement, not lower alpha suppression. Thus, we failed to support **H4b**.

Theoretical Relevance

For RQ1, given what hypotheses were supported (**H1 for theta power only and H1a higher gamma power for medium set sizes only**), it seems that correctly recalled words were associated with increase mental effort and attention at encoding. Participants who successfully attended to and encoded the information using their central executive (Baddeley & Hitch, 1974) or focus of attention (Cowan, 1999; Oberauer, 2002) were better able to later recall these words. The lack of gamma power value difference could be explained by one of three things: 1) the modified version of the paired-associates learning task tapped more into the central executive/focus of attention portion of memory than LTM storage (Baddeley & Hitch, 1974) or activated LTM (Cowan, 1999), 2) the current channel analysis approach is less sensitive to capturing high frequency brain activity, or 3) the free-recall memory tasks and/or simple Sternberg procedures are qualitatively different than the PAL task used in the current study. Free-recall memory tasks require participants to rely on internally generated cues, whereas the current task could provide an external cue for the participants. In addition, the simple Sternberg task (e.g., participants only remember one letter or digit at a time for varying sets) would require less associative learning than the current task. Some of the gamma differences could correlate with the associative learning required in this task compared to other tasks traditionally used. Specifically, given that we saw an increase in gamma suppression, it is possible that this neural marker correlated with the subject's ability to ignore incoming external cues that were not related to the currently-maintained information (Bonnefond & Jensen, 2016). It was necessary to ignore this information to avoid interference from external cues; whereas in non-associative learning paradigms, a subject could internally generate a new and unique cue to the stimuli to avoid similar interference. To conclude an answer **RQ1**, the neural patterns active when we experience memory failures appears to be lower theta power and differing gamma values correlated with how much information we are tasked with remembering.

For **RQ2**, only **H3** was partially supported (e.g., higher gamma power for effective strategies) while **H3a** (e.g., higher theta-gamma power coupling for effective strategies) and **H3b** (e.g., higher alpha suppression for effective strategies) were not supported. Like above, effective strategies were hypothesized to offset working memory limitations by activating information into LTM and linking it with the to-be-remembered information. The higher gamma power values could support this conclusion. However, as mentioned in the above paragraph, gamma power values were still lower than baseline. If future work were to investigate this further with more controlled strategy trials (e.g., more trials overall or explicitly instructing participants to use a specific strategy), it would be easier to determine the true relationship between gamma power and strategy type. The lack of support for H3a could be interpreted as either 1) theta-gamma power coupling is not a sensitive coupling measure and should be explored with more complex coupling that has shown more promising results (e.g., phase/power or phase/phase coupling; Gruber et al., 2004; Klimesch, 1996; Hanslmayr, 2011) or 2) the modified version of the paired-associates learning task tapped more into the central executive/focus of attention portion of memory than LTM storage (Baddeley & Hitch, 1974) or activated LTM (Cowan, 1999). Again, the latter explanation seems unlikely given past research showing the involvement of theta-gamma in associative memory paradigms (Gruber et al., 2001; Tort et al., 2009).

For alpha power (H3b), it was expected that alpha power would correlate with suppressing irrelevant information from previous trials. It was predicted that participants would put more mental effort into inhibiting irrelevant information from previous trials and blocks. However, the current study found no evidence of alpha suppression, and instead, found strong alpha enhancement across the entirety of the task. Other work has also shown alpha enhancement in other tasks such as during sentence processing on a verbal working memory task (Meyer, Obleser, & Friederici, 2013). It is plausible that the current task (e.g., modified version of the paired associates learning task) was more focused on processing of sentences and/or words than suppression of irrelevant information (e.g., participants reported using sentence generation on ~49% of trials and rehearsal ~13% of trials), explaining the increase in alpha power rather than suppression of alpha. An additional study found that alpha enhancement positively correlated with later memory performance (Fell et al., 2011). However, alpha enhancement was higher for less effective strategies, despite the poor probability of recall shown in the behavioral analyses. A third possibility is that the alpha activity correlates with item maintenance (Hsieh, Ekstrom, & Ranganath, 2011; Jensen et al., 2002; Schack & Klimesch, 2002). It is unclear how alpha power for item maintenance differs from theta power item maintenance, but Hseih, Ekstrom, and Ranganath (2011) offer a potential explanation: alpha power correlates with item maintenance and theta power correlates with the temporal order of the items within a set. As such, the high effects of alpha power in the current study could be related to actual encoding of individual item information. To conclude an answer for **RQ2**, based on the current data, effective strategies appear to recruit higher gamma power, which could be related to storing information into LTM (either by tying incoming information to representations in LTM or a storage system of its own).

For **RQ3**, we failed to support **H4**, **H4a**, and **H4b**. We did not find differing theta, gamma, or alpha power for strategies depending on whether words were later correctly or incorrectly recalled. Despite this, visually, patterns between conditions trend towards a potential difference. Specifically, theta power increases more when participants used effective strategies and later incorrectly recalled words and when participants used less effective strategies to later correctly recall words. This could indicate that these two conditions recruit similar attentional demands. In other words, participants may put effort into using effective strategies, yet still not recall the word, while putting similar effort into less effective strategies to get it right. It is also possible that participants may be using more than one strategy to help them learn a word pair. In the current study, participants were only allowed to report one strategy per trial. However, participants may be creating a sentence that combines the cue and target words and then mentally rehearse that sentence. Future work should evaluate how often people report using multiple strategies, how it affects memory performance and how it affects brain activity. Alpha enhancement is also increased in these two conditions, possibly because it is correlated with maintenance of item information or potentially with processing sentences (e.g., sentence-generation for effective strategies; Meyer, Obleser, & Friederici, 2013). To conclude **RQ3**, clear brain patterns based on what previous research deemed “effective” versus “less effective” is hard to dissociate in the current study but patterns in frequency bands demonstrate a potential to tease this apart in future studies.

Lastly, and related to the exploratory analyses, beta was correlated with the use of verbal encoding strategies. It could represent an internal auditory loop when people mentally rehearse information (Hwang et al., 2005; Gelastopoulos, Whittington, & Kopell, 2019). However, we also observed beta power recruitment when participants report using effective strategies, which could correlate with an internal rehearsal loop for sentence generation. It is also possible that this same internal rehearsal loop could exist for imagery, but this has not been fully fleshed out. More information is needed to understand what the beta band is doing for the current task, but as of now, it tentatively seems plausible to correlate with brain recruitment needed for the "inner voice" in the phonological loop for Baddeley's model of memory (Baddeley & Hitch, 1974; Gelastopoulos, Whittington, & Kopell, 2019). Interestingly, the fact that lower set sizes had higher beta power failed to replicate prior studies showing an increase in beta power across set sizes (Proskovec, Heinrichs-Graham, & Wilson, 2019), but this does lend credit to beta's correlation with the inner rehearsal loop as it is easier to mentally rehearse 3-5 items continuously than medium and high set sizes. However, Onton et al. (2005) speculate that the beta band could be involved in integrating memory information. This interpretation can fit well with the current results, as it is easier to integrate three items together than multiple items past one's capacity. It would also explain why beta power was higher for correct words (e.g., better integration) than incorrect words (e.g., worse integration due to interference or failure in encoding). This would be an interesting conclusion, as it would suggest less effective strategies are more integrative than effective strategies. As such, the beta band will need to be explored more to determine whether it is directly related to internal rehearsal or the integration of information overall.

Practical Relevance

Neural signatures underlying memory encoding strategies could be used to validate self-report measures in future studies. As of now, research regarding the mental effort behind memory encoding strategies is purely speculative and conflated with accuracy. Instead, brain data in this study gave an unbiased inside look into how and why certain memory encoding strategies are superior to others. The primary proposal was designed to evaluate brain signatures in memory strategies; however, this was the first step to encourage future research to target these signatures for investigation. By identifying why and how memory failures occur, future research studies can use results of the current study to target specific neural bands. Increasing neural signatures, possibly through repetitive transcranial magnetic stimulation (rTMS) or transcranial direct current stimulation (TDCS), could help decrease the knowledge gap between individuals who have low memory capacity and individuals who have high memory capacity.

Limitations

The biggest limitation in the current study was the lack of control for what strategies participants used. Given that different participants use different strategies at different rates, it made comparisons across conditions difficult. However, this study was the first of its kind and is a good starting point for future work, which can utilize different paradigms and designs to evaluate similar research questions. For instance, researchers could control for strategy use, whether through instruction or manipulate of word abstractedness, to increase the probability of certain strategies. Further, because little to no prior work was available, there was no guidance as to which specific frequency bands to focus on and which time windows to use.

A second limitation to the study is the channel analysis approach. Relying on one electrode, rather than an independent component, clouds brain activity with unnecessary channel noise. A cleaner data analysis approach aimed at investigating clusters of independent components appearing in all participants could yield more clear and promising results.

A third limitation to the current study is the type of coupling analysis used. Power coupling between theta and gamma may not be the most robust approach. Most memory studies prefer to use phase coupling to address how these two bands work together. Therefore, phase coupling might be a more sensitive type of coupling to fully understand the relationship between these two bands. It is possible power-power coupling relies too heavily on channel analyses, which can reduce the effects one might find. Further, and related to the second limitation, the coupling analyses may be more informative if specific components are identified for each frequency band as compared to only using data from one electrode.

Future Directions

Future analyses will investigate phase coupling between gamma and theta power during encoding. Phase coupling is theorized to represent cortical communication between different areas of the brain (Fries, 2005). Given that information is shifted from working memory to long-term memory, phase coupling represents a potential mechanism that aids this information shift. Therefore, it is theorized that, phases between gamma and theta power will be more aligned when participants remember information correctly.

Another future direction I could take is to investigate how well aligned neural patterns are during the encoding and recall intervals. In other words, if participants are accessing specific neural patterns during encoding then those same neural patterns should be reactivated to promote healthy recall (Gelbard-Sagiv et al., 2008; for review, Staresina & Wimber, 2019). As such, both power and coupling time-frequency dynamics could be explored to see if the strength of power and phase are more similar between encoding and recall intervals for correctly recalled words than incorrectly recalled words.

The third direction I would like to go is re-analyzing the current data using an ICA-cluster analysis approach. This will allow me to isolate brain activity more clearly than analyzing multiple sources of brain activity mixed together in the raw channel data. By doing so, I can focus my analyses on a cleaner pattern of activity for the encoding window, which could help clarify the conflicting results in the analysis (e.g., lower values where I hypothesized there are higher values) and establish a clearer picture of new patterns of activation I did not anticipate (e.g., alpha enhancement).

A fourth direction I would like to take is analyze the serial position recall for the current study. With how this study was designed, it was easy to collapse across set size for analysis. A more complex way to look at the interaction between working memory and LTM would be to analyze the brain activity for serial position recall. I have not looked at this data to see, but I would expect earlier trials to have less theta, gamma, and theta-gamma coupling than later trials when it is imperative to shift information from WM/STM to LTM. This procedure would be similar to that used by Sederberg et al. (2006). Instead of averaging over data from all items in a set (e.g., across all ten-word pairs in a set of 10), they looked at how brain activity changed depending on the position of the words in the set and found that gamma power was more predictive of success for early serial positions while theta and alpha power were more predictive of success for later serial positions. Therefore, it is possible that averaging across all items in a set washed out some of the frequency band effects in the current study.

The last future direction I would like to take is to control for participant strategy use. To do so, I could either instruct participants to use a given strategy (e.g., rehearsal) for a block of words. Either that or using the same database I used to gather words in Experiment 1, I can manipulate how abstract or related the word pairs are. This would allow participants to readily use imagery, sentence generation, or rehearsal. Then, I can even out the trial number for each strategy to make comparisons between strategies stronger.

References

Adam, K. C., Robison, M. K., & Vogel, E. K. (2018). Contralateral delay activity tracks fluctuations in working memory performance. *Journal of Cognitive Neuroscience*, *30*(9), 1229-1240.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*(4), 89-195.

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

Axmacher, N., Henseler, M. M., Jensen, O., Winreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences, 107*(7), 3228-3233.

Backus, A. R., Schoffelen, J. M., Szebényi, S., Hanslmayr, S., & Doeller, C. F. (2016). Hippocampal-prefrontal theta oscillations support memory integration. *Current Biology*, *26*(4), 450-457.

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.

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition*, *36*(8), 1383-1390.

Bailey, H., Dunlosky, J., & Kane, M. J. (2011). Contribution of strategy use to performance on complex and simple span tasks. *Memory & cognition*, *39*(3), 447-461.

Belluscio, M. A., Mizuseki, K., Schmidt, R., Kempter, R., & Buzsáki, G. (2012). Cross frequency phase–phase coupling between theta and gamma oscillations in the hippocampus. *Journal of Neuroscience*, *32*(2), 423-435.

Bergmann, H. C., Rijpkema, M., Fernández, G., & Kessels, R. P. (2012). Distinct neural correlates of associative working memory and long-term memory encoding in the medial temporal lobe. *Neuroimage*, *63*(2), 989-997.

Bonnefond, M., & Jensen, O. (2013). The role of gamma and alpha oscillations for blocking out distraction. *Communicative & Integrative Biology*, *6*(1), e22702.

Burgess, A. P., & Ali, L. (2002). Functional connectivity of gamma EEG activity is modulated at low frequency during conscious recollection. *International Journal of Psychophysiology*, *46*(2), 91-100.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, *16*(2), 317-330.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in the human neocortex. *Science*, *313*(5793), 1626-1628.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

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

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. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science : A Journal of the American Psychological Society, 19*(1), 51-57.

Cowan, N., Lichty, W., & Grove, T. R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 258.

Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: What does it take to make them work?. *Psychological Review*, *119*(3), 480.

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.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience and Methods*, *134*(1), 9-21.

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition*, *29*(2), 247-253.

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.

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences*, *20*(4), 155.

Electrical Geodesic Systems, Inc., 500 East 4th Ave. Suite 200, Eugene, OR 97401, USA.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. MIT Press.

Fell, J., & Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, *12*(2), 105-118.

Fell, J., Klaver, P., Elfadil, H., Schaller, C., Elger, C. E., & Fernández, G. (2003). Rhinal hippocampal theta coherence during declarative memory formation: interaction with gamma synchronization?. *European Journal of Neuroscience*, *17*(5), 1082-1088.

Fell, J., Klaver, P., Elger, C. E., & Fernández, G. (2002). Suppression of EEG gamma activity may cause the attentional blink. *Consciousness and Cognition*, *11*(1), 114-122.

Fell, J., Ludowig, E., Staresina, B. P., Wagner, T., Kranz, T., Elger, C. E., & Axmacher, N. (2011). Medial temporal theta/alpha power enhancement precedes successful memory encoding: evidence based on intracranial EEG. *Journal of Neuroscience*, *31*(14), 5392-5397.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474-480.

Friese, U., Köster, M., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2013). Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *NeuroImage*, *66*, 642-647.

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. *Science*, *322*(5898), 96-101.

Gelastopoulos, A., Whittington, M. A., & Kopell, N. J. (2019). Parietal low beta rhythm provides a dynamical substrate for a working memory buffer. *Proceedings of the National Academy of Sciences*, *116*(33), 16613-16620.

Gruber, T., Keil, A., & Müller, M. M. (2001). Modulation of induced gamma band responses and phase synchrony in a paired associate learning task in the human EEG. *Neuroscience letters*, *316*(1), 29-32.

Gruber, T., Tsivilis, D., Giabbiconi, C. M., & Müller, M. M. (2008). Induced electroencephalogram oscillations during source memory: Familiarity is reflected in the gamma band, recollection in the theta band. *Journal of Cognitive Neuroscience*, *20*(6), 1043-1053.

Gruber, T., Tsivilis, D., Montaldi, D., & Müller, M. M. (2004). Induced gamma band responses: An early marker of memory encoding and retrieval. *Neuroreport*, *15*(11), 1837-1841.

Hanslmayr, S., Volberg, G., Wimber, M., Raabe, M., Greenlee, M. W., & Bäuml, K. H. T. (2011). The relationship between brain oscillations and BOLD signal during memory formation: a combined EEG–fMRI study. *Journal of Neuroscience*, *31*(44), 15674-15680.

Howard, M. W., Rizzuto, D. S., Caplan, J. B., Madsen, J. R., Lisman, J. E., Aschenbrenner-Scheibe, R., Schultze-Bonhage, A., & Kahana, M. J. (2003). Gamma oscillations correlate with working memory load in humans. *Cerebral Cortex*, *13*(12), 1369-1374.

Hsieh, L. T., Ekstrom, A. D., & Ranganath, C. (2011). Neural oscillations associated with item and temporal order maintenance in working memory. *Journal of Neuroscience*, *31*(30), 10803-10810.

Hwang, G., Jacobs, J., Geller, A., Danker, J., Sekuler, R., & Kahana, M. J. (2005). EEG correlates of verbal and nonverbal working memory. *Behavioral and Brain Functions*, *1*(1), 1-13.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513-541.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*(8), 877-882.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395-1399.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.

JMP®, Version *16*. SAS Institute Inc., Cary, NC, 1989–2021.

Jokisch, D., &. Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *Journal of Neuroscience*, *27*(12), 3244-3251.

Jutras, M. J., Fries, P., & Buffalo, E. A. (2009). Gamma-band synchronization in the macaque hippocampus and memory formation. *Journal of Neuroscience*, *29*(40), 12521-12531.

Klimesch, W., & Doppelmayr, M. (1996). encoding of new. *Neuroreport*, *7*, 1235-1240.

Kluen, L. M., Dandolo, L. C., Jocham, G., & Schwabe, L. (2019). Dorsolateral prefrontal cortex enables updating of established memories. *Cerebral Cortex*, *29*(10), 4154-4168.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703-713.

Lisman, J. E., & Jenson, O. (2013). The theta-gamma neural code. *Neuron*, *77*(6), 1002-1016.

Lee, H., Simpson, G. V., Logothetis, N. K., &. Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, *45*(1), 147-156.

Luck, S. (2014). *An introduction to the event-related potential technique* (Second ed.). MIT Press

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252-271.

McDougall, S. J. P., & Donohoe, R. (2002). Reading ability and memory span: Long-term memory contributions to span for good and poor readers. *Reading and Writing: An Interdisciplinary Journal*, *15*, 359-387.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*(1), 10-17.

Meeuwissen, E. B., Takashima, A., Fernández, G., & Jensen, O. (2011). Increase in posterior alpha activity during rehearsal predicts successful long-term memory formation of word sequences. *Human brain mapping*, *32*(12), 2045-2053.

Melrose, R. J., Zahniser, E., Wilkins, S. S., Veliz, J., Hasratian, A. S., Sultzer, D. L., & Jimenez, A. M. (2020). Prefrontal working memory activity predicts episodic memory performance: A neuroimaging study. *Behavioural Brain Research*, *379*, 112-307.

Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*(3), 711-721.

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.

Sirota, A., Montgomery, S., Fujisawa, S., Isomura, Y., Zugaro, M., & Buzsáki, G. (2008). Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. *Neuron*, *60*(4), 683-697.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition*, *16*(4), 343-352.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*(3), 251-269.

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

Olson, J. C. (1980). Encoding processes: Levels of processing and existing knowledge structures. *ACR North American Advances*.

Onton, J., Deleorme, A. & Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *NeuroImage*, *27*(5), 341-356.

Pfurtscheller, G., Stancak Jr, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*(1-2), 39-46.

Proskovec, A. L., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load modulates the alpha and beta oscillatory dynamics serving verbal working memory. *NeuroImage*, *184*, 256-265.

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., Madsen, J. R., & Lisman, J. E. (2001). Gating of human theta oscillations by a working memory task. *Journal of Neuroscience*, *21*(9), 3175-3183.

Rajsic, J., Burton, J. A., & Woodman, G. F. (2019). Contralateral delay activity tracks the storage of visually presented letters and words. *Psychophysiology*, *56*(1), 1-10.

Ranganath, Charan, & Blumenfeld, Robert S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences, 9*(8), 374-380.

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.

Roux, F., Wibral, M., Mohr, F. M., Singer, W., & Uhlhaas, P. J. (2012). Gamma-band activity in the human prefrontal cortex codes for the number of relevant items maintained in working memory. *Journal of Neuroscience*, *32*(36), 12411-12420.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R, Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., & Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Current Biology*, *19*(21), 1846-1852.

Sederberg, P. B., Gauthier, L. V., Terushkin, V., Miller, J. F., Barnathan, J. A., & Kahana, M. J. (2006). Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage*, *32*(3), 1422-1431.

Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, *23*(34), 10809-10814.

Schack, B., & Klimesch, W. (2002). Frequency characteristics of evoked and oscillatory electroencephalic activity in a human memory scanning task. *Neuroscience Letters*, *331*(2), 107-110.

Schack, B., Vath, N., Petsche, H., Geissler, H. G., & Möller, E. (2002). Phase-coupling of theta -gamma EEG rhythms during short-term memory processing. *International Journal of Psychophysiology*, *44*(2), 143-163.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11.

Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., &. Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, *23*(34), 10809-10814.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long term memory. *Nature Reviews Neuroscience*, *4*(8), 637-648.

Sols, I., DuBrow, S., Davachi, L., Fuentemilla, L. (2017). Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in long-term memory. *Current Biology*, *27*(22), 3499-3504.

Spellman, T., Rigotti, M., Ahmari, S. E., Fusi, S., Gogos, J. A., & Gordon, J. A. (2015). Hippocampal–prefrontal input supports spatial encoding in working memory. *Nature*, *522*(7556), 309-314.

SR Research Experiment Builder 1.10.165 [Computer software]. (2011). Mississauga, Ontario, Canada: SR Research Ltd.

Staresina, B. P., & Wimber, M. (2019). A neural chronometry of memory recall. *Trends in cognitive sciences*, *23*(12), 1071-1085.

Tort, A. B., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009). Theta–gamma coupling increases during the learning of item–context associations. *Proceedings of the National Academy of Sciences*, *106*(49), 20942-20947.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*(4), 446-468.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748-751.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 255-274.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 51-79.

Waugh, N. C., & Norman, D. A. (1968). The measure of interference in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *7*(3), 617-626.

Weiss, S., & Rappelsberger, P. (2000). Long-range EEG synchronization during word encoding correlates with successful memory performance. *Cognitive Brain Research*, *9*(3), 299 -312.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341.

Zakrzewski, A. C., Sanders, E. C., & Berry, J. M. (2021). Evidence for age-equivalent and task-dissociative metacognition in the memory domain. *Frontiers in psychology*, *12*, 57.

Appendix A - Pilot Experiment Word Stimuli

Word Pair	Set Size	Total Times Recalled	Total Participants	Recall %
STROKE-VALLEY	3	20	24	0.83
COLD-WANT	3	20	24	0.83
ESSAY-USAGE	3	19	24	0.79
PRIZE-SHARK	5	22	24	0.92
BELL-SUNSET	5	20	24	0.83
FINE-HALT	5	18	24	0.75
HUMOR-SCENT	5	13	24	0.54
SHEET-UNDER	5	18	24	0.75
ABRUPT-QUIZ	7	20	24	0.83
MOMENT-KETTLE	7	15	24	0.63
CROWN-PERIOD	7	17	24	0.71
DROP-PLUNGE	7	14	24	0.58
NOVEL-VELVET	7	15	24	0.63
ACCENT-BONE	7	12	24	0.50
BLUE-FENCE	7	19	24	0.79
SERMON-DOLL	9	19	24	0.79
SELF-QUICK	9	19	24	0.79
SPLIT-BRANDY	9	13	24	0.54
PARDON-BRAKE	9	9	24	0.38
FARCE-CELL	9	13	24	0.54
PAGE-SEAT	9	10	24	0.42
RESULT-PEAR	9	5	24	0.21
CLAUSE-ANNOY	9	11	24	0.46
TRUTH-HYBRID	9	13	24	0.54
CLOCK-WINDOW	4	23	24	0.96
TUNE-KIDDY	4	22	24	0.92
LINK-FANCY	4	16	24	0.67
SNAIL-DEBUT	4	21	24	0.88

GROAN-LEMON	6	18	24	0.75
FLOOR-BREW	6	19	24	0.79
WEAPON-TURN	6	18	24	0.75
SOUND-WOOL	6	13	24	0.54
ROCK-ENGINE	6	17	24	0.71
FORUM-BLANK	6	16	24	0.67
WRONG-ITEM	8	15	25	0.60
SLAP-BUDGET	8	16	25	0.64
APPLE-IVORY	8	22	25	0.88
YARD-CAREER	8	14	25	0.56
PROMPT-BURN	8	10	25	0.40
PURSE-WEALTH	8	17	25	0.68
ISLAND-DRIP	8	21	25	0.84
MISTER-FOREST	8	15	25	0.60
SLOW-ASIDE	10	12	25	0.48
REMARK-DOUBLE	10	13	25	0.52
SICK-THING	10	13	25	0.52
NARROW-RESCUE	10	11	25	0.44
STRAW-MONEY	10	12	25	0.48
KIND-EASE	10	24	25	0.96
WALLET-SCALE	10	13	25	0.52
CAPE-SCENE	10	12	25	0.48
BODY-ORIGIN	10	16	25	0.64
BUTTER-CALM	10	15	25	0.60
PUBLIC-DELAY	4	18	25	0.72
TINY-SWEET	4	17	25	0.68
TURTLE-BLEND	4	21	25	0.84
NEUTER-BIGGER	4	21	25	0.84
EVENT-BARREL	5	20	25	0.80
ORDER-NOUN	5	15	25	0.60
HOLE-WASTE	5	18	25	0.72

DECK-HEIGHT	5	17	25	0.68
SIMPLE-LEADER	5	17	25	0.68
WELL-MINOR	7	15	25	0.60
NATIVE-TANK	7	24	25	0.96
OUTFIT-BAND	7	17	25	0.68
SENATE-WRAP	7	12	25	0.48
CALL-FLOW	7	12	25	0.48
PEOPLE-TONGUE	7	17	25	0.68
SPLASH-PORK	7	14	25	0.56
GUEST-FAIL	3	14	25	0.56
TROUGH-SPOIL	3	15	25	0.60
HINT-BANKER	3	17	25	0.68
WART-TASTE	8	21	24	0.88
RIGHT-SWIM	8	16	24	0.67
THEORY-LAWYER	8	21	24	0.88
TABLET-RICE	8	16	24	0.67
SITE-YELLOW	8	14	24	0.58
SUMMER-LEAN	8	19	24	0.79
LEFT-FEAR	8	17	24	0.71
SURVEY-INFORM	8	23	24	0.96
ROOF-WOBBLE	10	18	24	0.75
ORANGE-PLANT	10	21	24	0.88
TWANG-SPEECH	10	16	24	0.67
BEAT-SCREW	10	14	24	0.58
WORKER-LADY	10	22	24	0.92
TRUCK-BRIBE	10	16	24	0.67
FRUIT-PIECE	10	15	24	0.63
GRIP-ZERO	10	19	24	0.79
HEART-CHUCK	10	14	24	0.58
SPRING-CHINA	10	20	24	0.83
PRAISE-COPY	6	12	24	0.50

PITY-FISH	6	18	24	0.75
SNOW-DIET	6	19	24	0.79
MEDAL-VARNISH	6	10	24	0.42
BARGE-TRAY	6	11	24	0.46
JERSEY-CHAIN	6	16	24	0.67
SOFA-NUMBER	9	15	24	0.63
MOOSE-SEAM	9	18	24	0.75
EFFORT-SLOPE	9	7	24	0.29
EAST-CLUB	9	12	24	0.50
COFFIN-EDGE	9	13	24	0.54
SHAWL-THEME	9	10	24	0.42
DEFEAT-REST	9	8	24	0.33
BENT-SINGER	9	14	24	0.58
PERMIT-ERROR	9	7	24	0.29
GAIN-POTTY	4	19	24	0.79
MOTOR-SPREAD	4	20	24	0.83
HATE-CARD	4	19	24	0.79
JUMP-SINE	4	15	24	0.63
DESK-SHIRT	3	20	24	0.83
TWICE-SNIDE	3	16	24	0.67
VULGAR-SYMBOL	3	18	24	0.75
LACK-FLICK	5	15	24	0.63
PUDDLE-DEAL	5	15	24	0.63
FLAP-TREND	5	14	24	0.58
BOAT-DOWN	5	19	24	0.79
THROAT-NIGHT	5	15	24	0.63
SPOKE-DUST	7	17	24	0.71
MEET-BILL	7	20	24	0.83
FOLLY-DANGER	7	12	24	0.50
SLIP-FOIL	7	14	24	0.58
CUSTOM-HELMET	7	18	24	0.75

CLOSET-JOIN	7	14	24	0.58
SLIGHT-RETORT	7	7	24	0.29
VIEW-STEAK	8	20	24	0.83
HUNT-FARM	8	21	24	0.88
JAIL-SHIVER	8	20	24	0.83
ATOM-PLATE	8	14	24	0.58
QUID-DEMAND	8	12	24	0.50
NAME-SACK	8	19	24	0.79
RAVE-BOIL	8	14	24	0.58
AREA-CHART	8	11	24	0.46
SAGA-BEAR	9	20	24	0.83
RECESS-CANDLE	9	14	24	0.58
TROT-ANSWER	9	9	24	0.38
CHERRY-MISS	9	16	24	0.67
BASS-DESIGN	9	11	24	0.46
WEIGHT-SOCIAL	9	13	24	0.54
NEEDLE-FIELD	9	14	24	0.58
FRINGE-METHOD	9	11	24	0.46
PLOT-CURRY	9	8	24	0.33
REJECT-MARCH	6	17	24	0.71
TEXT-CYCLE	6	21	24	0.88
LEAVE-PAIR	6	13	24	0.54
SAUCE-ASSENT	6	13	24	0.54
SHIP-MINUTE	6	14	24	0.58
STOCK-BIRTH	6	18	24	0.75
GROUND-ARMY	10	20	24	0.83
TICK-CHEAP	10	18	24	0.75
SCREAM-HUMAN	10	22	24	0.92
APPEAL-ROAD	10	11	24	0.46
SPRAY-FREAK	10	11	24	0.46
SIGHT-BLUFF	10	9	24	0.38

OFFICE-SMILE	10	16	24	0.67
RUSTIC-FAST	10	7	24	0.29
SALT-ALERT	10	11	24	0.46
HEALTH-BREAK	10	13	24	0.54
STEEL-WINE	7	18	22	0.82
FRONT-PROXY	7	17	22	0.77
ROLE-DRINK	7	10	22	0.45
SPEAR-LOOSE	7	11	22	0.50
MIRROR-FRESH	7	10	22	0.45
TRACK-OCEAN	7	13	22	0.59
CLEAN-HILL	7	12	22	0.55
PRIDE-DECREE	8	15	22	0.68
STOLE-MEAN	8	17	22	0.77
REASON-STAND	8	13	22	0.59
PROFIT-SHELL	8	10	22	0.45
FAIR-QUIET	8	13	22	0.59
LENGTH-SOUP	8	14	22	0.64
DAILY-THRILL	8	10	22	0.45
CASTLE-TEST	8	9	22	0.41
GRAPE-YOUTH	9	19	22	0.86
RUSH-BLIND	9	18	22	0.82
JEEP-QUEUE	9	13	22	0.59
LAMP-SKULL	4	20	22	0.91
DEBATE-JOKE	4	17	22	0.77
PIANO-DUCK	4	19	22	0.86
NEST-BALLOT	4	15	22	0.68
EXCUSE-CHARGE	8	16	22	0.73
STATE-COUCH	8	12	22	0.55
HORSE-CLOSE	8	13	22	0.59
YEAR-GHOST	8	9	22	0.41
SIGNAL-CRAWL	8	7	22	0.32

HOLD-MYTH	8	8	22	0.36
QUEEN-CASE	8	10	22	0.45
PAUSE-AWARD	8	9	22	0.41
FIFTH-BITE	9	17	22	0.77
SIZE-ROLL	9	14	22	0.64
CURE-WARM	9	20	22	0.91
TOPIC-POUR	9	9	22	0.41
TWIST-PASTE	9	9	22	0.41
COURSE-FORM	9	7	22	0.32
DRAIN-WASH	9	11	22	0.50
CROWD-SCHOOL	9	16	22	0.73
CROP-UNION	9	14	22	0.64
PARTY-OBJECT	5	17	23	0.74
SCOTCH-ZIGZAG	5	17	23	0.74
GLUT-HOCKEY	5	17	23	0.74
LOCKER-THREAT	5	15	23	0.65
PERSON-GRAVY	5	18	23	0.78
DESERT-WEAR	4	22	23	0.96
UNCLE-CREDIT	4	16	23	0.70
MATCH-FIGHT	4	17	23	0.74
REVISE-SALAD	4	14	23	0.61
CURVE-HERO	6	16	23	0.70
TITLE-CABIN	6	17	23	0.74
CURLER-ECHO	6	9	23	0.39
ROSE-LUMP	6	9	23	0.39
NATURE-CLUE	6	11	23	0.48
COAST-DETAIL	6	10	23	0.43
DIME-LIMIT	3	20	23	0.87
CHILLY-REPORT	3	11	23	0.48
SALARY-BRANCH	3	17	23	0.74
TOMATO-DOZEN	9	17	23	0.74

HUSKY-POOR	9	18	23	0.78
MAIN-INDIAN	9	11	23	0.48
GIANT-MARINE	9	13	23	0.57
TRIBE-RULE	9	10	23	0.43
DOGMA-LOOK	9	7	23	0.30
PARCEL-BUSH	9	10	23	0.43
FILE-TENSE	9	10	23	0.43
BEACH-STAIN	9	14	23	0.61
RAPID-SMART	10	15	23	0.65
SUPPLY-AUTHOR	10	16	23	0.70
SNEEZE-COTTON	10	16	23	0.70
CHOIR-BLOCK	10	10	23	0.43
MAID-PARENT	10	14	23	0.61
WHEAT-SHOT	10	14	23	0.61
TRUST-GRADE	10	11	23	0.48
PACK-SCOOP	10	9	23	0.39
KNEE-URGE	10	12	23	0.52
LIFE-STAFF	10	10	23	0.43
TWELVE-SINK	7	11	22	0.50
BULLET-AUTUMN	7	16	22	0.73
BEAUTY-MINE	7	14	22	0.64
CLASS-SOFT	7	14	22	0.64
COFFEE-INCOME	7	17	22	0.77
NEWS-CHANGE	7	17	22	0.77
PLAY-SALE	7	9	22	0.41
MOOD-RATE	9	15	22	0.68
SPARK-TICKLE	9	15	22	0.68
GUARD-SENSE	9	8	22	0.36
TWIG-CHEESE	9	14	22	0.64
FORCE-CONE	9	6	22	0.27
DEVICE-MOVE	9	7	22	0.32

PLUS-THIEF	9	8	22	0.36
WOMB-GOLD	9	10	22	0.45
BOTTOM-DRAW	9	6	22	0.27
BUCKET-MUMMY	4	12	22	0.55
SHOVEL-PEANUT	4	16	22	0.73
FRAUD-RISE	4	14	22	0.64
BOWL-SELL	4	14	22	0.64
DUTY-EMPTY	6	15	22	0.68
CONVOY-MOLD	6	17	22	0.77
STILL-HEAVEN	6	17	22	0.77
MENU-PULSE	6	14	22	0.64
TREE-SKIN	6	16	22	0.73
FORK-LEGEND	6	15	22	0.68
TOWN-SKUNK	10	16	22	0.73
REPEAT-CHAP	10	11	22	0.50
AMOUNT-HOTEL	10	10	22	0.45
UNIT-SPELL	10	9	22	0.41
SHOP-FROWN	10	11	22	0.50
LEAF-HOUSE	10	7	22	0.32
DONKEY-POSTER	10	11	22	0.50
STAGE-AUNT	10	15	22	0.68
ADULT-STYLE	10	16	22	0.73
LEAP-SILVER	10	11	22	0.50
EPIC-TALLY	5	18	22	0.82
BACK-FETCH	5	16	22	0.73
HOUR-BRICK	5	9	22	0.41
POET-DIMMER	5	12	22	0.55
RIOT-JUGGLE	5	14	22	0.64
OUGHT-FACE	10	16	25	0.64
TRAIN-YAWN	10	23	25	0.92
TUSK-CART	10	13	25	0.52

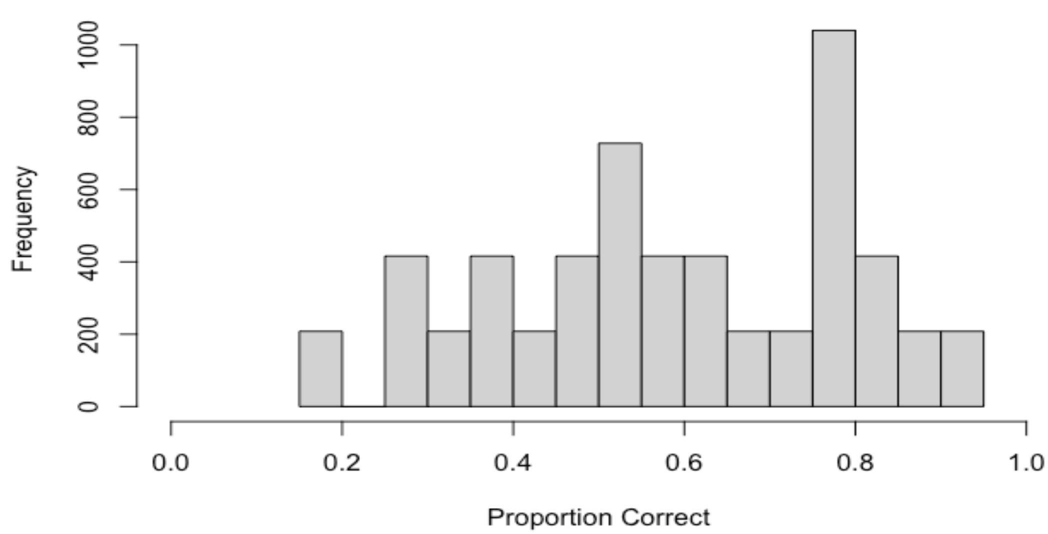
LAUGH-FIFTY	10	17	25	0.68
PASS-COLLAR	10	11	25	0.44
PUPIL-SURF	10	16	25	0.64
COUNT-BRIDGE	10	13	25	0.52
OPTION-VOTE	10	17	25	0.68
BELT-WHITE	10	16	25	0.64
KNIFE-SCARE	10	21	25	0.84
BLOOD-TOOL	3	23	25	0.92
SECRET-TREATY	3	22	25	0.88
FLAME-OUNCE	3	15	25	0.60
BROOM-STOVE	6	22	25	0.88
FEEL-STORE	6	20	25	0.80
BATH-RABBIT	6	19	25	0.76
SMACK-DANCE	6	19	25	0.76
ROUTE-COURT	6	15	25	0.60
SPILL-NECK	6	16	25	0.64
CHIEF-POCKET	5	19	25	0.76
FOLD-SORROW	5	14	25	0.56
FLOOD-BLUSH	5	18	25	0.72
SOCK-BABE	5	15	25	0.60
RECOIL-TRACE	5	16	25	0.64
CIRCLE-ISSUE	7	18	25	0.72
TEAM-GOSSIP	7	18	25	0.72
TOAST-BABY	7	18	25	0.72
ACCORD-MARKET	7	13	25	0.52
POISON-LIMB	7	19	25	0.76
VERBAL-ENEMY	7	13	25	0.52
SHED-MOON	7	17	25	0.68
SAVE-TALK	8	15	25	0.60
NINETY-SISTER	8	21	25	0.84
KISS-PICK	8	18	25	0.72

REACH-PHONEY	8	13	25	0.52
FALL-HOBBY	8	19	25	0.76
SPEED-TEETH	8	17	25	0.68
SONG-MAPLE	8	16	25	0.64
COLUMN-GRANT	8	15	25	0.60

Appendix B - Multilevel Modeling Replication of ANOVA Results

Model Description. We used a generalized multilevel mixed effects model to investigate the main effect of strategy use and set size on proportion of correctly recalled items (i.e., proportion correct). A multilevel mixed effects approach allows flexibility to account for variations in performance within a participant (e.g., a participant might score an average 75% for one block of the study but score an average 85% on the next block). By modeling the variations within participants, the model is more sensitive to differences within and between participants. A generalized approach, specifically modeling binomial outcomes, allowed the model to give accurate and more realistic predictions for recall percentages as it restricts the range to 0 to 1. Traditional general linear modeling assumes a range outside of 0 and 1, which is not correct for the performance in the current study. For purposes of data visualization, aggregated data will be represented as the proportion of correctly recalled words (total number of words recalled divided by the total number of trials completed). Proportion Correct could range from 0.00 to 1.00 (M = 0.59, SD = 0.20, SE = 0.002) (see Figure 37 for a distribution of raw scores on the task).

Figure 37
Distribution of scores on the paired-recall task



Note. Performance is normally distributed, ranging from 18% to 93% correct.

Variable Description. The outcome variable for the current analyses was a binomial variable (e.g., 1 = correct, 0 = incorrect for each trial) with two categorical predictors, set size (e.g., number of word pairs in each set; this ranged from 3-10 pairs per set) and strategy type (e.g., Didn't Try, Effective or Less Effective). Trials in which participants reported using with imagery or sentence generation were categorized as "effective," trials in which they reported using rehearsal or reading a word were categorized as "less effective," and trials in which participants reported that they didn't try were categorized as "Didn't Try" (see Table 5 for proportion of trials that participants report using effective strategies and less effective strategies). The proportion of reported strategies are similar those reported in Dunlosky and Hertzog (2001), who compared rates of strategy use between younger and older adults on a similar paired-associates task. Within the current analysis, only the intercept (starting performance within trials) and slope (rate of performance within trials) for participant performance was allowed to vary within participants as a random effect. All categorical variables were dummy coded. Set Size was dummy coded such that set size of 3 was the comparison group. Strategy was dummy coded such that "Less Effective" was the comparison group.

Table 5
Proportion of trials reported using each strategy while studying word pairs

Less Effective Strategies		Effective Strategies		
Repeat	Read	Sentence	Imagery	Didn't Try
0.13 (0.03)	0.12 (0.02)	0.49 (0.06)	0.21 (0.04)	0.04 (0.01)

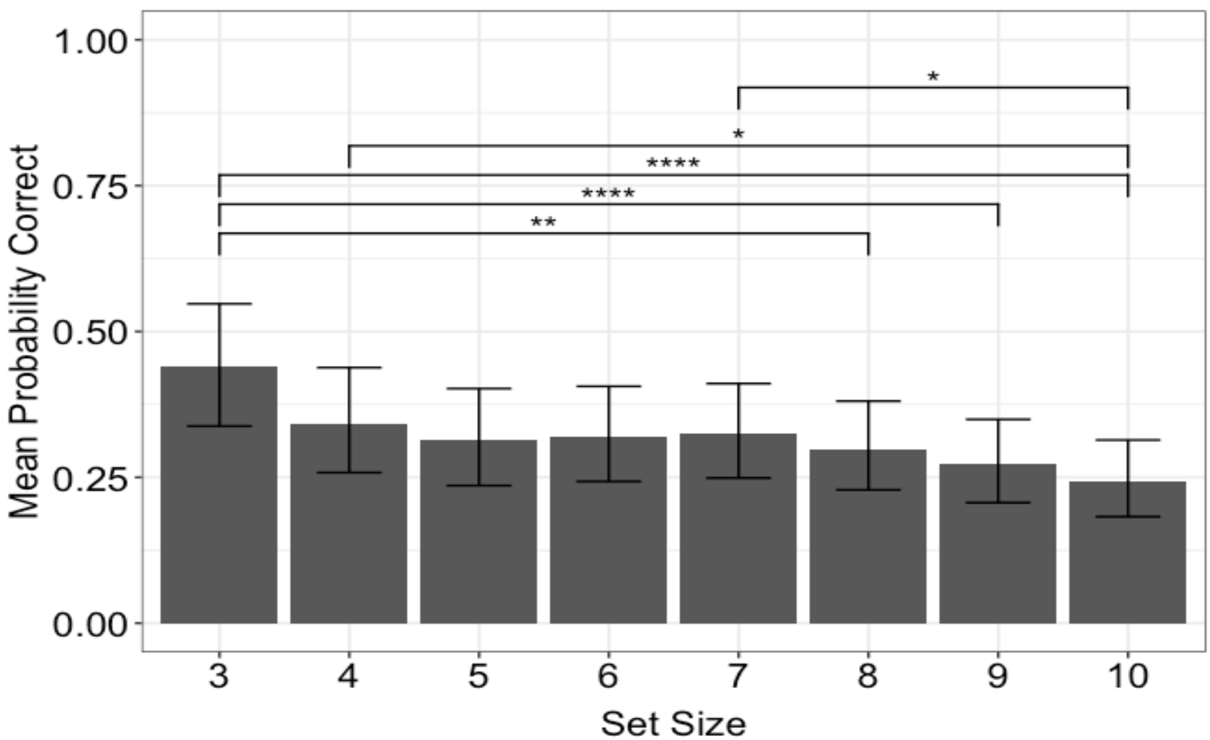
Note: Parentheses = standard error of the mean

Results of the multilevel model analysis indicated that there was a significant effect of set size on recall performance[1] (see Figure 38 for the average probability of correct recall by set size). A Tukey HSD post-hoc analyses indicated that recall of set size 3 was significantly different from that of set size 8 ($z = 3.68$, $SE = 0.17$, $p = 0.01$), set size 9 ($z = 4.55$, $SE = 0.16$, $p < 0.001$), and set size 10 ($z = 5.57$, $SE = 0.16$, $p < 0.001$). Recall of set size 4 was significantly different from that of set size 10 ($z = 3.51$, $SE = 0.14$, $p = 0.01$). Lastly, recall of set size 7 was also significantly different from that of set size 10 ($z = 3.51$, $SE = 0.12$, $p = 0.01$).

No other paired comparison differed significantly. As expected, participants had a higher probability correctly recalling words from smaller set sizes (e.g., sets of 3 and 4) compared to larger set sizes (e.g., sets of 9 and 10).

Figure 38

Set size was predictive of recall performance



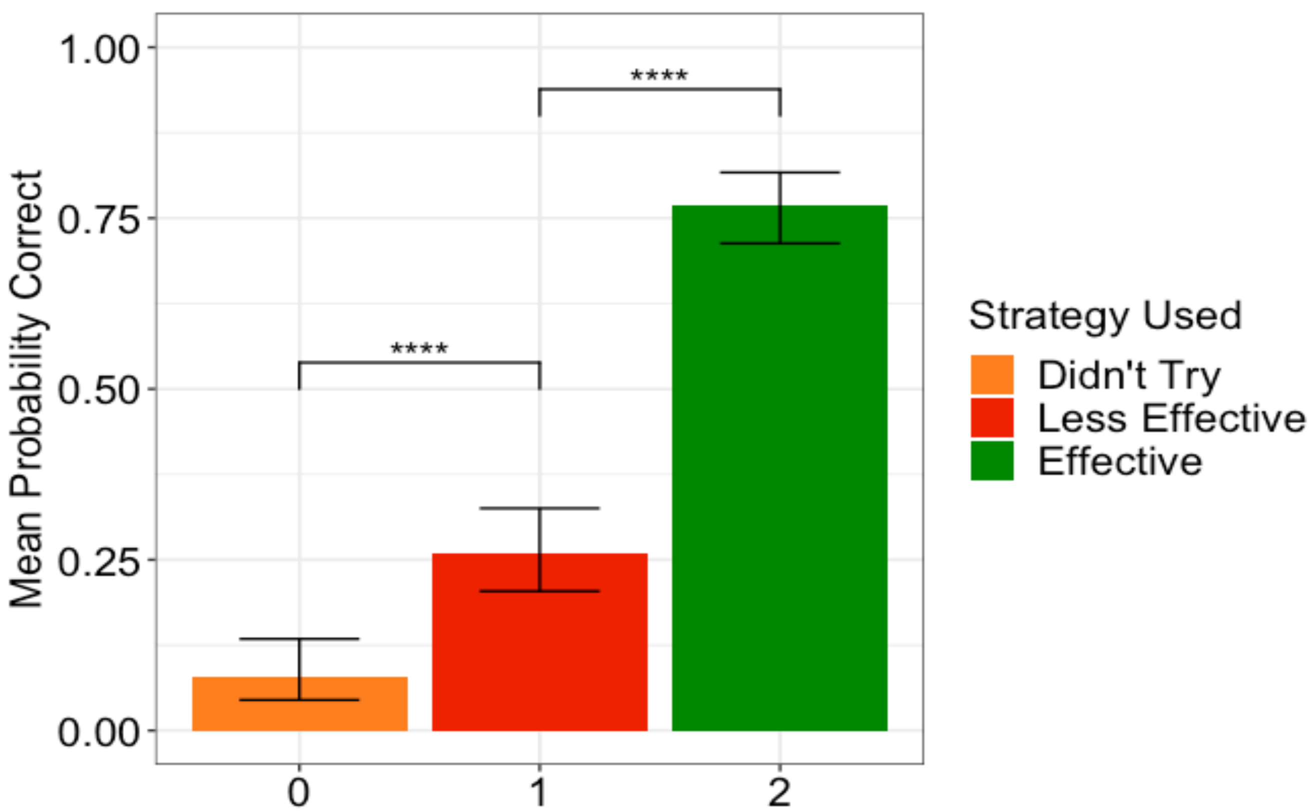
Note. Lower set sizes (e.g., 3 and 4) produced better recall than higher set sizes (e.g., 9 and 10). The errors bars represent standard error.

Further, as expected, the omnibus test showed a significant effect of strategy on recall performance[2] (see Figure 39 for the average probability correct by strategy). A Tukey HSD post-hoc analyses indicated that effective strategies yielded significantly higher probability of correctly recalled words than less effective strategies (z = -25.49, SE = 0.09, $p < 0.001$) and when participants didn't try (z = -13.42, SE = 0.27, $p < 0.001$).

Although less effective strategies yielded lower probability of recalling words than effective strategies, less effective strategies had higher probability of correctly remembering words than when participants did not try (z = -5.11, SE = 0.28, $p < 0.001$).

Figure 39

Type of strategy to remember words was predictive of recall performance



Note. When participants used effective strategies (e.g., "Imagery" or "Sentence Generation"), they recalled words better than when they used less effective strategies ("Repeat" or "Read)

The observed patterns in the behavioral data replicated previous studies (Dunlosky & Kane, 2001; Dunlosky & Hertzog, 2001; Bailey et al., 2011) such that memory performance was better on smaller set sizes and also when participants reported using effective strategies. Further, our participants reported using both normatively effective and less effective strategies. Next, we will analyze the EEG data to investigate theta, gamma, and alpha band power differences between performance (e.g., correct vs. incorrect), strategy type (e.g., effective vs. less effective), the interaction between (e.g., correct vs. incorrect effective & correct vs. incorrect effective), and set size (e.g., sets 3-10).

Appendix C - ERSPs