2018-12-01

Contributions of Target-Lure Similarity and Sensory Modality to Lure False Alarms

Daniel Kent Bjornn
Brigham Young University

Contributions of Target-Lure Similarity and Sensory Modality to Lure False Alarms

Daniel Kent Bjornn

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

C. Brock Kirwan, Chair
Shawn D. Gale
Ramona O. Hopkins
Steven G. Luke
Anthony O. Vance

Department of Psychology

Brigham Young University

ABSTRACT

Contributions of Target-Lure Similarity and Sensory Modality to Lure False Alarms

Daniel Kent Bjornn
Department of Psychology, BYU
Doctor of Philosophy

The processes of pattern separation and pattern completion are very important in the correct discrimination of similar memories. Much research has been conducted on these processes, but there are some gaps that need to be addressed. First, there is some debate as to whether false alarms to lure items come about because of a failure to accurately encode a memory or a failure to retrieve a memory. Second, much of the research on pattern separation and pattern completion in humans is done with visual stimuli and contributions of stimulus modality to these processes are not well understood as a result.

Study 1 consisted of three experiments conducted using a combination of eye tracking and functional magnetic resonance imaging (fMRI) methods. Analyses of eye tracking data in the experiments examined the contribution of fixation counts at encoding and retrieval, as well as target-lure similarity level, to accuracy on lure trials. Task designs were altered across studies to attempt to replicate specific research previously conducted with a specific answer period, as well as generalize the findings to a broader body of research that allows participants to answer while the stimulus is presented. The three experiments showed mixed support for the contribution of fixation counts at encoding and retrieval to the accurate discrimination of similar lures. Target-lure similarity, however, was a robust predictor of accuracy for all three experiments.

Prior research examining activity in the hippocampus demonstrates a reduction of fMRI activity to repetitions of a stimulus. Greater activity is also observed in the dentate gyrus/CA3 (DG/CA3) subregions for correct rejections of lure items compared to lure false alarms. There should be a greater reduction in the DG/CA3 as a function of encoding for lure false alarms than for lure correct rejections if memory encoding drives the activity differences between these outcomes. The fMRI data showed a marked reduction of activity in the left hippocampus to repetition trials as a function of encoding trial fixation count. There was no significant difference between activity as a function of encoding fixation count in the DG/CA3 for lure correct rejections and lure false alarms. There was also no difference in activity for the CA1 either. Overall, the results of the eye tracking and fMRI data give support for the contribution of pattern completion to false alarms to lure stimuli rather than poor encoding.

Study 2 examined the contribution of sensory modality to accurate discriminations of lure stimuli. A behavioral task was developed to directly compare discrimination of similar lures on visual and auditory stimuli. Participants were significantly more accurate and more confident of their responses when discriminating visual stimuli as compared to discriminating auditory stimuli.

Keywords: pattern separation, pattern completion, sensory modality, memory encoding, memory specificity

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my dear wife, Shantel Bjornn, for all her support through the many years of schooling. Her patience and encouragement helped me get through the hardest times. I also wish to thank my parents, Kent and Valerie Bjornn, for teaching me how to work hard and for supporting me throughout my life, as well as through school. Next, my parents-in-law, Wayne and Jacquie Pierson, provided great support through school in many forms and I could not have completed without their help.

My mentor, Dr. Brock Kirwan, fostered a great learning environment that allowed me to explore on my own, but also get the help I needed to get out of the hole I dug for myself. I greatly appreciate his patience with my many questions as well as his willingness to entertain my sometimes broad range of interests.

Like any major undertaking, I was not the only one who contributed to this body of work. Dr. Steven Luke provided much needed guidance on the design and analysis of the eye tracking experiments. Nathan Muncy provided a great deal of help with the hippocampal segmentation protocol. Ariana Hedges-Muncy answered many, many statistics questions. Cooper Hodges and Ty Bodily also helped by giving feedback on the study design and analysis while allowing me to bounce ideas off them. Dr. Andrew Berrett provided the initial idea for study 2 and assisted with iterations on the design and pilot testing of the behavioral task.

I could also not complete this project without the help of many undergraduate research assistants. Torin Kelley, Elijah Möller, Emily Maxwell, Leila LeSueur, Clara Rodriguez, Caitlyn Towne, Eric Brighton, Caroline Straw, Mia Felin, Jordan Clark, Athena Howell, Annelise Toolson, Todd Winn, and Laurine Giraud-Carrier all helped with the development of experiment tasks, subject recruitment, and data collection.

TABLE OF CONTENTS

TITLE i

ABSTRACT ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

Contributions of Target-Lure Similarity and Sensory Modality to Lure False Alarms 1

Common Research Methods 2

The Involvement of the Hippocampus 4

Hippocampal Subregions 5

Dentate gyrus 5

CA3 8

CA1 9

Research Gaps 10

Study 1 11

Experiment 1 16

Experiment 2 28

Experiment 3 36

General Discussion 50

Study 2 57

Methods 60

Results 65

Discussion 70

Conclusion 73

References 74

LIST OF TABLES

Table 1 *Proportion of Responses by Trial Type and Response for Experiment 1* 23

Table 2 *β-values for Mixed-effects Regression Model in Experiment 1* 26

Table 3 *Proportion of Responses by Trial Type and Response for Experiment 2* 30

Table 4 *β-values for Mixed-effects Regression Model in Experiment 2* 34

Table 5 *Proportion of Responses by Trial Type and Response for Experiment 3* 42

Table 6 *β-values for Mixed-effects Regression Model in Experiment 3* 45

Table 7 *Results of Hypotheses for Study 1* 51

Table 8 *Proportion of Responses by Stimulus Type, Trial Type, and Response* 66

LIST OF FIGURES

Figure 1 *Labels of Trial Types.* 19

Figure 2 *Sample Sequence of Stimuli* 20

Figure 3 *Sample High- and Low-similarity Lure Pairs* 21

Figure 4 *Bias-corrected Scores for Experiment 1 by Outcome* 24

Figure 5 *Memory-based Fixation Counts for Experiment 1 by Presentation and Outcome* 25

Figure 6 *Bias-corrected Scores for Experiment 2 by Outcome* 31

Figure 7 *Memory-based Fixations per Second for Experiment 2 by Presentation and Outcome* .33

Figure 8 *Regions of Interest for fMRI Analyses* 41

Figure 9 *Bias-corrected Scores for Experiment 3 by Outcome* 43

Figure 10 *Memory-based Fixations for Experiment 3 by Presentation and Outcome* 46

Figure 11 *Difference in Hippocampal Activity on Repeat Items* 47

Figure 12 *Difference in Hippocampal Subfield Activity on Lure Items* 49

Figure 13 *Sample Trial from Similarity Rating Questionnaire* 61

Figure 14 *Protocol for Task Used in Study 2* 63

Figure 15 *Correlation of Visual d' Values and Auditory d' Values* 67

Figure 16 *d' Scores Divided by Similarity Level and Stimulus Type* 68

Figure 17 *Proportions of High-confidence Divided by Stimulus Modality and Similarity Level* ..70

Contributions of Target-Lure Similarity and Sensory Modality to Lure False Alarms

Memory is a necessity in every aspect of daily life. It is important for remembering where you left your keys, what your address deficits in memory illustrate this importance. Perhaps the most famous of cases is patient H.M., is, or even who your spouse and children are. Severe who lost the capacity to form new episodic memories after a removal of most of the medial temporal lobe in both hemispheres (Squire, 2009). Patient E.P. also developed severe anterograde amnesia after a case of encephalitis (Insausti, Annese, Amaral, & Squire, 2013). Both of these individuals were greatly impaired in everyday life due to their lack of ability to form new episodic memories. While severe, these cases are rare. More common are small memory mistakes made by healthy individuals. Of particular interest are the mistakes when an individual generalizes similar memories, thinking that something has occurred, when in reality that is not the case.

Memory specificity is the ability to discriminate very similar memories. Pattern separation and pattern completion are computational processes that are involved in memory specificity. Pattern separation allows for similar memory traces to be encoded as distinct and dissimilar as possible (Hunsaker & Kesner, 2013; Yassa & Stark, 2011). Such a process is important in ensuring that distinct memories remain distinct. Pattern completion, on the other hand, makes overlapping traces even more overlapping and allows for quick retrieval of memories when there is only partial information available (Hunsaker & Kesner, 2013; Yassa & Stark, 2011).

Memory specificity is involved in many examples of everyday life. For instance, someone trying to find their car in a parking lot would require them to recognize that the very similar-looking car in nearly the same location is not their car. Successful pattern separation at the time they parked the car would allow for a memory trace that is distinct enough to recognize

that the location they are at is not where they parked their car. However, a failure of pattern separation may cause that individual to pattern complete when they searched for the car and think that the similar-looking car in the similar location was their own. This situation would cause that individual to have the embarrassing experience of trying to unlock someone else's car.

Common Research Methods

The processes of pattern separation and pattern completion are studied using a variety of methods. In rodents, behavioral tasks that involve a rat or mouse learning to find a food cue by discriminating several different stimuli are common. Stimuli in these experiments can be spatial markers (Clelland et al., 2009; Creer, Romberg, Saksida, van Praag, & Bussey, 2010; Kesner, Taylor, Hoge, & Andy, 2015; Morris, Churchwell, Kesner, & Gilbert, 2012), temporal cues (Gilbert, Kesner, & Lee, 2001; Hunsaker, Fieldsted, Rosenberg, & Kesner, 2008; Lee, Jerman, & Kesner, 2005), or odors (Hunsaker, Fieldsted, et al., 2008; Weeden, Hu, Ho, & Kesner, 2014). In lesion studies, neurotoxic lesions are used to destroy various parts of the brain once they have learned to discriminate the stimuli (Dees & Kesner, 2013; Kesner et al., 2015; Weeden et al., 2014). Resulting deficits in behavior give insight into the necessity of that region of the brain to the pattern separation and pattern completion processes. Electrode studies measure the firing patterns of areas of the brain as rodents perform various tasks (Leutgeb et al., 2005; McHugh et al., 2007; Wilson, Ikonen, Gallagher, Eichenbaum, & Tanila, 2005). This method gives insight into how the regions of the brain perform these functions, whether it be through different neurons firing (Leutgeb et al., 2005) or the same neurons changing the firing rate (Leutgeb et al., 2005; Wilson et al., 2005).

In humans, the pattern separation and pattern completion is generally assessed using behavioral testing with healthy subjects (Doxey, Hodges, Bodily, Muncy, & Kirwan, 2018;

Stark, Yassa, & Stark, 2010; Vieweg, Riemer, Berron, & Wolbers, 2018) or members of various clinical populations (Kirwan et al., 2012; Shelton & Kirwan, 2013; Stark et al., 2010). Participants perform tasks that require them to differentiate stimuli that are exact repetitions (i.e., targets) of items they have encountered before or only similar items (i.e., lures). These behavioral tasks can be accompanied by functional magnetic resonance imaging (fMRI; Kirwan & Stark, 2007; Lacy, Yassa, Stark, Muftuler, & Stark, 2011; Leal, Tighe, & Yassa, 2014; Motley & Kirwan, 2012), electroencephalography (Anderson, James, & Kirwan, 2017), or eye tracking (Molitor, Ko, Hussey, & Ally, 2014; Vieweg et al., 2018).

Experiments using fMRI rely on the blood oxygen level dependent (BOLD) response to determine blood flow to various parts of the brain while an individual completes a cognitive task (Kannurpatti, Motes, Rypma, & Biswal, 2010; Yassa et al., 2010). As areas of the brain are more active, they require greater amounts of glucose and oxygen (Huettel, Song, & McCarthy, 2014). Deoxygenated hemoglobin in the blood are receptive to a magnetic field and are therefore able to be manipulated and measured using fMRI (Huettel et al., 2014). Changes in deoxygenated hemoglobin through the brain over time allows for the measurement of the contribution of different regions to a variety of experimental conditions.

In the context of experiments looking at pattern separation and pattern completion, fMRI relies on the repetition suppression effect to examine how the brain discriminates or generalizes similar memories. The repetition suppression effect is a reduction in activity in response to a stimulus when that stimulus is repeated (Kovács, Kaiser, Kaliukhovich, Vidnyánszky, & Vogels, 2013; Liu, Shen, Olsen, & Ryan, 2017; Reagh, Watabe, Ly, Murray, & Yassa, 2014). Relative activity between a target and a lure would remain fairly constant in the presence of pattern separation, because pattern separation allows for memories to be encoded as distinctly as

possible. The lure is therefore treated as an encounter with a new stimulus and activity reflects such by remaining at a similar level (Bakker, Kirwan, Miller, & Stark, 2008; Lacy et al., 2011). In the context of pattern completion, however, relative activity between a target and a lure stimulus would be different. This is because pattern completion allows for memory traces to be as overlapping as possible, so exposures to similar lures are therefore treated as a repetition and a reduction in activity is displayed (Bakker et al., 2008; Lacy et al., 2011).

The use of eye tracking in the study of pattern separation and pattern completion is relatively new, with only a handful of experiments using this method (Molitor et al., 2014; Vieweg et al., 2018). In general, eye tracking uses infrared light to reflect off of the cornea and retina of the eye. The direction and distance between these points are calculated as the eye moves and gaze location on a visual stimulus can be detected with incredible accuracy. Eye tracking experiments can test whether particular areas of interest were fixated (Bonhage, Mueller, Friederici, & Fiebach, 2015; Eisenbarth & Alpers, 2011; Olejarczyk, Luke, & Henderson, 2014) or measure the number of fixations on an object (Kafkas & Montaldi, 2011; Sekiguchi, 2011; Vance, Jenkins, Anderson, Bjornn, & Kirwan, 2018) among a host of other eye movement related metrics.

The Involvement of the Hippocampus

The hippocampus is a common area of interest when examining pattern separation and pattern completion due to the high level of involvement the region has in these processes (Hunsaker & Kesner, 2013; Kirwan & Stark, 2007; Yassa & Stark, 2011). While a relatively small structure, the hippocampus is extremely important in memory (Insausti et al., 2013; Squire, 2009) and provides the mechanisms for rapid learning (McClelland, McNaughton, & O'Reilly, 1995). Rodents with lesions in the hippocampus show impairment when completing spatial tasks

(Kesner et al., 2015; Kirwan, Gilbert, & Kesner, 2005; Lee, Hunsaker, & Kesner, 2005) as well as discriminating the sequence of odors (Fortin, Agster, & Eichenbaum, 2002; Kesner, Hunsaker, & Ziegler, 2010). Human amnesic patients with damage to the hippocampus are impaired in tasks that rely on both recollection and familiarity (Jeneson, Kirwan, Hopkins, Wixted, & Squire, 2010). When looking specifically at pattern separation, individuals with damage to the hippocampus are more likely to respond to similar lures as old than are healthy controls (Kirwan et al., 2012). Activity in the hippocampus also differs when healthy individuals correctly identify lure items than when they incorrectly classify them as repeated items (Kirwan & Stark, 2007; Motley & Kirwan, 2012).

The hippocampus is divided into several different subregions, of which the present project focuses specifically on the dentate gyrus (DG), CA1, and CA3 regions. Input into the hippocampus is sent by the entorhinal cortex into both the DG and CA3 (Kesner & Rolls, 2015). The DG sends its signals on to the CA3 (Kesner, 2013; Kesner & Rolls, 2015), which then transfers those signals, along with the signals received from the entorhinal cortex to the CA1 (Knierim & Neunuebel, 2016) The CA1 is the main output of the hippocampus via the subiculum, with signals projecting to a variety of other regions of the brain (Kesner & Rolls, 2015; Knierim & Neunuebel, 2016).

Hippocampal Subregions

Dentate gyrus

Research focusing on the DG has shown its role in the execution of pattern separation, which is thought to be due to the sparse firing rate of cells that result in unique representations (Rolls, 2013). The DG is also one of the few places in the brain where neurogenesis occurs in the granule cells (Clelland et al., 2009; Creer et al., 2010). Young granule cells have been implicated

in pattern separation while older granule cells seem to change their function more towards pattern completion as they age (Nakashiba et al., 2012).

Lesion methods in rodents have contributed much to the understanding of this region of the hippocampus. Lesions to the DG impair the discrimination of changes in spatial information (Goodrich-Hunsaker, Hunsaker, & Kesner, 2008; Hunsaker, Rosenberg, & Kesner, 2008; Lee, Hunsaker, et al., 2005). Morris et al. (2012) lesioned the DG of rats and trained them to find a food reward in a radial arm maze. The rodents were randomly assigned to an adjacent arm condition, where the reward and non-reward arms were directly next to each other, or the separate arm condition, where the reward and non-reward arms were separated by two unused arms. Rats in the separate arm condition took fewer trials to reach the predetermined criterion than did the rats in the adjacent arm condition. The authors attributed this to the smaller degree of overlap in the separate arm condition which would have less of a need for pattern separation. This pattern of behavior is not unique to spatial stimuli. Performance on an odor matching-to-sample task diminishes for rats with ventral DG lesions as well (Weeden et al., 2014). This decrease in performance is most pronounced when the difference between odors is a single carbon chain.

With its involvement in pattern separation the DG is critical for the formation of new, non-overlapping memory representations. Rodents with DG lesions show impairment in encoding new memories, but not the retrieval of formed memories (Lee & Kesner, 2004). Further, Weeden et al. (2014) found that rats with lesions in the DG were impaired at delays of 60 seconds, but not at delays of 15 seconds, which the authors interpreted as evidence of the behavior resulting from a memory impairment and not a sensory impairment.

The DG has also been studied across different types of pattern separation. For example, Costa, Bueno, and Xavier (2005) lesioned the DG and tested rats on a task that required them to wait a certain amount of time to respond in order to receive a reward. The DG lesioned rats significantly underestimated the amount of time required to receive the reward, while the sham controls were not impaired. Further, the authors tested spatial pattern separation using a delayed non-matching-to-place task. The DG lesioned rats were impaired in choosing the correct arm of the maze, but controls were unimpaired. Increasing the inter-trial-interval impaired both DG lesioned and control rats, but it only took an inter-trial-interval of 4 minutes for the DG lesioned rats to perform at chance levels while it took delays of 16 minutes for the control rats to perform at chance.

Human experiments using fMRI require the DG to be combined with nearby CA3 because the resolution of fMRI is currently not fine enough to differentiate the BOLD response in the DG from the CA3 (Reagh et al., 2014). Bakker et al. (2008) argue that because the representations that are orthogonalized in the DG are projected to the CA3, it is possible to observe pattern separation related activity in both regions. The research done in humans confirms what has been found in rodents, to the extent possible with the decreased spatial resolution. Multiple studies demonstrate a reduction of activity in the DG/CA3 upon the presentation of a repeated stimulus, but no reduction when a lure stimulus is presented (Bakker et al., 2008; Lacy et al., 2011). The lack of reduction in response is indicative of the DG treating the stimulus as new and is present even in response to a highly similar lure (Lacy et al., 2011). Experiments that monitor subregion activity while participants complete a mnemonic discrimination task report greater activation in DG/CA3 when individuals correctly discriminate similar lures from their targets (Leal, Tighe, Jones, & Yassa, 2014; Reagh & Yassa, 2014). Doxey and Kirwan (2015)

also found that left DG/CA3 volume was highly correlated with behavioral performance on a mnemonic discrimination task. Berron et al. (2016) examined DG and CA3 activity with a 7-tesla MRI machine and were able to separate the two regions due the increased spatial resolution. They reported that multivoxel patterns for similar stimuli did not overlap in the DG while they did overlap in the CA3. The lack of overlap in the DG voxel patterns was taken as evidence that the DG provides memory representations that are less overlapping than those in the CA3.

CA3

The function of the CA3 has been debated over the years with some researchers reporting that the CA3 performs pattern separation (Leutgeb, Leutgeb, Moser, & Moser, 2007) and others reporting that the CA3 performs pattern completion (Lee, Jerman, et al., 2005; Neunuebel & Knierim, 2014). This discrepancy seems to relate to the amount of overlap between target-lure pairs as further research demonstrated that the CA3 performs pattern separation under conditions of high overlap while it performs pattern completion under conditions of low overlap (Vazdarjanova & Guzowski, 2004).

The CA3 receives input from the entorhinal cortex through the perforant path as well as the DG (Kesner, 2013; Kesner & Rolls, 2015). Signals from the mossy fibers of the DG prove to be important in encoding and storing spatial memory representations in CA3 (Lassalle, Bataille, & Halley, 2000; Lee & Kesner, 2004) but are not important for consolidation and recall (Lassalle et al., 2000). Rather, lesions to the perforant path input to CA3 cause impairments regarding retrieval in rodents (Lee & Kesner, 2004). The CA3 is thought to be involved in pattern completion because of the recurrent collaterals that act as an autoassociative network (Kesner & Rolls, 2015), which allow different aspects of a memory to be strengthened (Gilbert & Brushfield, 2009). If part of the recurrent collateral network is activated, then the remaining

areas can be activated through the recurrent loop, allowing a memory to be retrieved from a part (Kesner & Rolls, 2015).

CA1

The CA1 subregion of the hippocampus receives input from the CA3 and projects output to other areas of the brain (Kesner & Rolls, 2015; Knierim & Neunuebel, 2016). There are some differing ideas of the contribution of CA1 to the processes of pattern separation and pattern completion. In rodents, CA1 shows patterns of activity that are sometimes more similar than the CA3 and sometimes less similar than CA3 (Lee, Yoganarasimha, Rao, & Knierim, 2004; Leutgeb, Leutgeb, Treves, Moser, & Moser, 2004; Vazdarjanova & Guzowski, 2004). CA1 also has the same number of inputs and outputs, unlike the DG that has many more outputs than inputs (Knierim & Neunuebel, 2016). Because of these points, the CA1 is thought to neither perform pattern separation nor pattern completion, but shows a linear pattern of activity dependent on the inputs it receives (Knierim & Neunuebel, 2016; Yassa & Stark, 2011).

Research into temporal pattern separation, however, shows that CA1 is important in this process. Rodents with CA1 lesions show impairment in temporal based tasks when the task requires discriminating the temporal order of spatial stimuli (Gilbert et al., 2001; Lee, Jerman, et al., 2005). It was not clear from these experiments whether the important part of it was the temporal order of spatial stimuli specifically or whether the CA1 responded to temporal pattern in general. Further research testing the temporal order of odor presentation gave support for the role of CA1 in temporal pattern separation across stimuli (Hunsaker, Fieldsted, et al., 2008; Kesner et al., 2010).

In humans, CA1 shows a pattern of activity that is consistent with pattern completion, and is many times implicated in the pattern completion process. For example, Bakker et al.

(2008) were the first to look at activity in hippocampal subregions while participants completed an implicit discrimination task. While the DG/CA3 showed patterns consistent with pattern separation, CA1 showed activity more like pattern completion. Further, Lacy et al. (2011) took the experiment one step further by varying the level of similarity between target and lure in their implicit discrimination task. They discovered that the CA1 responds in a graded fashion, with activity increasing as similarity level decreases.

Research Gaps

Pattern separation and pattern completion in the hippocampus have been well studied, but there are still aspects of these processes that are not well understood. For example, there is some debate about the contribution of pattern completion to false alarms to similar lures. Some argue that false alarms come about because of insufficient encoding of stimuli rather than the contribution of pattern completion at the time of testing (Molitor et al., 2014). Another aspect of false alarms that is not known is the effect that stimulus modality has on pattern separation and pattern completion processes. In rodents, much of the research on pattern separation and pattern completion to date has been conducted using visual, spatial, temporal, and olfactory stimuli (Hunsaker, Fieldsted, et al., 2008; Weeden et al., 2014). In humans, visual stimuli are almost always used, but sometimes in a spatial and temporal context (Azab, Stark, & Stark, 2014; Roberts, Ly, Murray, & Yassa, 2014). Auditory pattern separation has not been extensively studied in either rodents or humans and is not understood.

The current project consists of two studies. Study 1 uses three experiments to evaluate the contribution of pattern completion to false alarms by teasing apart the contributions of encoding, retrieval, and target-lure similarity level to false alarms. Study 2 tests the development of a

behavioral task that directly compares visual and auditory mnemonic discrimination, which is thought to tax pattern separation.

Study 1

Pattern separation and pattern completion are often studied in humans using a mnemonic discrimination task where individuals are shown items of everyday objects and asked to indicate whether or not they have seen the item before. A correct response to a lure item requires the execution of the pattern separation process (Kirwan & Stark, 2007). Pattern completion also plays a role in this task, as it is generally accepted that participants use a "recall to reject" strategy in which the individual uses the current item to recall the previously encountered image (Hunsaker & Kesner, 2013). For example, a person may have seen a black teddy bear as a study trial and then sees a brown teddy bear as a test trial. The individual would then need to use the brown teddy bear to recall any instances of a teddy bear that they have seen in the task.

Of particular interest in the mnemonic discrimination task is the case in which an individual incorrectly responds to a lure item as old (i.e., false alarm). One explanation of such mistakes is a lack of pattern separation paired with a propensity toward pattern completion (Kim & Yassa, 2013). Specifically, the amount of interference—the degree of similarity between the study item and the test item—is too high for the distinct memory trace and the individual identifies the item as one they have seen before.

Computational models explain this pattern of behavior as a lack of pattern separation and a propensity toward pattern completion (Norman & O'Reilly, 2003; O'Reilly & Norman, 2002). The amount of overlap between the target and the lure is too great to recognize a distinct item and therefore the lure item is recalled rather than the target. These computational models suggest that the DG is important in pattern separation and is therefore necessary for correct

discrimination of visual stimuli. Activity in the DG is predicted by these models to have a sharp increase to even small changes in input. The CA1 on the other hand, is predicted to have a more graded activity pattern that follows pattern completion. In this context, small changes between the target and lure are not detected because the memory representations are made to overlap in order to facilitate memory retrieval. As the difference between target and lure increases, however, the memory representations overlap less and support correct discrimination to be made. Lastly, CA3 is different than DG and CA1 in the sense that it contributes to pattern completion when there are large amounts of overlap between target and lure but contributes to pattern separation when target and lure differ greatly.

Rodent research shows support for these computational models and demonstrated that the regions of the hippocampus contribute to pattern separation and pattern completion. For example, rodents with DG lesions show impairment to small changes in spatial stimuli (Gilbert et al., 2001; Hunsaker, Fieldsted, et al., 2008). Performance, however, increases as the amount of similarity between target and lure decreases (Gilbert et al., 2001; Hunsaker, Fieldsted, et al., 2008), consistent with behavior predicted by the models for pattern completion. In effect, the overlap between target and lure at high similarity levels is too great and the similar memory is retrieved as if it were a repeated stimulus. Without pattern separation to correctly encode the small amounts of difference in the stimuli as distinct memory representations, there is nothing to stop the pattern completion process from indicating that the stimuli encountered is, in fact, something encountered previously.

Human behavioral research examining accuracy on various discrimination tasks demonstrates that accuracy is greater for target-lure pairs that are less similar. Accuracy rates are routinely greater for items where target-lure similarity is decreased (Hannula, Baym, Warren, &

Cohen, 2012; Kim & Yassa, 2013; Motley & Kirwan, 2012). Individuals are also more confident in their answers and quicker at responding when similarity levels are low than when they are high (Horry & Brewer, 2016). Research examining familiarity and recollection shows that false alarms are not simply due to an item seeming overly familiar. Participants report being able to recall source information relating to the false alarm which increased as target-lure similarity level increased (Kim & Yassa, 2013).

Neuroimaging research in humans also sheds light on the contribution of hippocampal subfields to pattern separation and pattern completion. Lacy et al. (2011) examined the effects of target-lure similarity on hippocampal subfield activity using high resolution fMRI in order to track activity in DG/CA3 and CA1. While presented images of everyday objects, participants indicated whether the object presented was generally encountered indoors or outdoors. No overt answer was given regarding memory for the item. Some of the target-lure pairs had were very similar while others were more distinct. DG/CA3 showed a sharp increase in fMRI activity even for very minor changes. CA1, however, demonstrated a more graded increase in activity, consistent with predictions of computational models for pattern completion.

Another experiment by Yassa, Mattfield, Stark, and Stark (2011) focused exclusively on activity in DG/CA3 when comparing between young and older adults. Older adults were more likely to respond “old” to lure items. Activity in the DG/CA3 for young adults showed that there was very little change in activity when comparing across similarity levels. These results are consistent with predictions made by computational models (Norman & O’Reilly, 2003) as well as the findings previously mentioned by Lacy et al. (2011). Older adults, however, showed a graded pattern of activity in DG/CA3 when comparing across similarity level. Specifically, activity increased as the difference between the items increased. The researchers argued that the

pattern of activity was evidence of a greater need for dissimilarity before pattern completion would occur in the older adults.

Further, Motley & Kirwan (2012) compared implicit and explicit versions of the continuous recognition task. In an implicit version of the task, participants look at images one at a time with no specific requirement to remember the objects. In the case of Motley and Kirwan (2012), participants indicated whether or not the item they saw was a toy. An explicit version of the task requires participants to respond to each item displayed by indicating whether or not they have seen that item before in the experiment. Comparing the fMRI activity in the hippocampus for each version of the task revealed that the implicit version aligned with activity predicted for pattern separation in computational models. The explicit version of the task, however, demonstrated activity that was more indicative of pattern completion, lending support for the "recall to reject" strategy theorized to be used by individuals when completing the task (Hunsaker & Kesner, 2013).

Molitor et al. (2014) argue that pattern completion does not fully explain why individuals false alarm to lure stimuli in mnemonic discrimination tasks. The authors proposed that fuzzy trace theory is a better explanation of this pattern of behavior. The underlying theme of fuzzy trace theory is that a memory includes both a verbatim trace and a gist (or "fuzzy") trace (Brainerd & Reyna, 2002). The verbatim trace involves specific details while the gist trace involves the meaning surrounding the memory. The theory posits that false memories come about because the gist trace is stronger than the verbatim trace and the meaning of the item seems familiar (Brainerd & Reyna, 2002).

Molitor et al. (2014) used eye tracking as an implicit measure of encoding images while participants completed a mnemonic discrimination task. Based on fuzzy trace theory, they

hypothesized that lure items incorrectly identified as old items, would have fewer fixations on the first presentation of the pair, or study trial, than lure items that were correctly identified as similar. Their findings supported their hypothesis; lure false alarms had fewer fixations on the study trial than lure correct rejections. This was interpreted to show that a failure of encoding could also be an explanation of false alarms to lure items rather than just pattern completion.

The findings of Molitor et al. (2014) offer an important criticism of the mnemonic discrimination paradigm, but the authors did not consider several factors that can affect memory performance. Accordingly, we extend this research program. First, the visual complexity of images has a direct impact on the number of fixations, with more visually complex images having more fixations (Henderson, 2003). We therefore accounted for visual complexity by calculating a score of visual complexity for each image (Rosenholtz, Li, & Nakano, 2007) and added those scores as a covariate into the analysis. Second, as initial papers often have over-inflated effect sizes compared to attempted replications (Button et al., 2013), we increased the sample size to ensure sufficient power to avoid type II error. Third, as mentioned previously, the relative similarity between stimuli in lure pairs has a direct impact on how well an individual correctly responds to test trials (Kim & Yassa, 2013; Motley & Kirwan, 2012). We incorporated high- and low-similarity pairs to account for this. If lure pair similarity is a better predictor of test outcome than fixation count of the encoding trial, that would be more indicative of pattern completion's involvement in the false alarm to a lure stimulus rather than a fuzzy memory trace.

Further, the setup of Molitor et al. (2014) was different than most other research examining mnemonic discrimination to visual stimuli. Their experiment allowed participants to respond only after the image was presented for a set duration. While the fixed amount of time makes the analysis on fixation counts easier, it limits the interpretation of the results because

many experiments that examine mnemonic discrimination use a design that allows participants to respond while the stimulus is still presented. The extra time on test trials may have influenced performance on the task. Experiment 2 addresses this issue by changing the design to require participants to respond to an item while the image was still presented.

While eye movement data provide a better picture of encoding than straight behavioral measures, this method is still too far removed from neural processing to accurately make inferences about the involvement of memory related brain structures such as the hippocampus. Experiment 3 further extends the research by collecting fMRI data along with the eye movement data.

Experiment 1

In experiment 1, participants completed a task with the same parameters as the task in the experiment conducted by Molitor et al. (2014). The only differences in task design were that similar lures were divided evenly between high- and low-similarity target-lure pairs and participants completed an additional two blocks of trials. Similarity level proves to be an excellent predictor of accuracy on the Mnemonic Similarity Task (Kim & Yassa, 2013; Motley & Kirwan, 2012), but measures of encoding strength have not been compared to the contribution of similarity level on accuracy directly. This small change in task design allowed us to measure the extent to which an individual fixated on stimuli during study trials and compare it to accuracy rates between similarity levels.

The first hypothesis we tested was that fixations at study differ between response types to attempt to replicate the findings of Molitor et al. (2014). Next, we tested the hypothesis that visual complexity drives fixation counts in the task. If the number of fixations at study no longer differ between response types when visual complexity is accounted for in the analysis, it would

be evidence that the effect was driven by the complexity of the images used on the study trials. Lastly, we hypothesized that target-lure similarity level would be a better predictor of performance on the task than fixations at study or test, as compared by their contribution to predicting an accurate outcome through a logistic regression.

Methods.

Participants. A power analysis was conducted using the effect sizes that Molitor et al. (2014) reported. Using the lowest effect size in their repeated measures ANOVA ($\eta_p^2 = 0.17$ for the main effect of presentation), the analysis revealed that 60 participants would be needed to have sufficient statistical power at the 0.95 level. To allow for attrition, a total of 67 participants (32 male, 35 female) were recruited from Brigham Young University for the experiment. Participants were required to have normal or corrected-to-normal vision and were also excluded if they reported color blindness or any history of traumatic brain injury, learning disability, or psychological or neurological diagnosis.

Data were excluded from four participants. Three participants' data were excluded due to technical problems with the eye tracker. Another participant did not have their vision corrected at the time they participated. After exclusions, a total of 63 participants (29 male, 34 female) were included for data analysis, with an average age of 19.97 years (*SD* = 1.91). Participants were granted course credit for participating in the experiment. All study protocols were approved by the Brigham Young University Institutional Review Board and all participants gave written, informed consent prior to participating.

Apparatus. Stimuli were presented using a BenQ monitor (BenQ Corporation, Taipei, Taiwan) with a resolution of 1600x900 pixels and a refresh rate of 60 Hz. Eye movement data were collected using an EyeLink 1000 Plus eye tracker (SR Research, Ontario, Canada) sampled

at 1000 Hz from the right eye. The eye tracker was connected to a host computer running EyeLink 5.00 tracking software. The experiment was programmed using Experiment Builder software (SR Research, Ontario, Canada) and participant responses were collected using a ResponsePixx Handheld button box (VPixx Technologies, Quebec, Canada).

Prior to each block, a 9-point eye-tracking calibration was performed to ensure accurate data collection. Successful calibration required an average error of less than 0.50° and a maximum error of less than 1.00°.

Stimuli and procedure. Participants viewed images of everyday objects one at a time on a computer screen while eye movement data were collected. Stimuli were grouped into five different types, namely subsequent repeat, subsequent lure, repeat, lure, and foil (see Figure 1). Study trials consisted of subsequent repeats and subsequent lures while test trials consisted of repeats, lures, and foils. Subsequent repeat and repeat items were identical in each respective pair. Subsequent lure and lure items were similar images of the same type of object, but they were not exact repeats (e.g., a light brown teddy bear and a dark brown teddy bear). Foil items were images that were presented only once in the study.

In following the procedure that Molitor et al. (2014) used, each trial began with a drift check that also served as the inter-trial interval. For the drift check, participants were required to look at a circle presented in the middle of the screen and press a button to ensure that the eye position was still calibrated correctly. A calibration was run before proceeding to the next trial if the drift check indicated that the eye position was too far from acceptable measures.

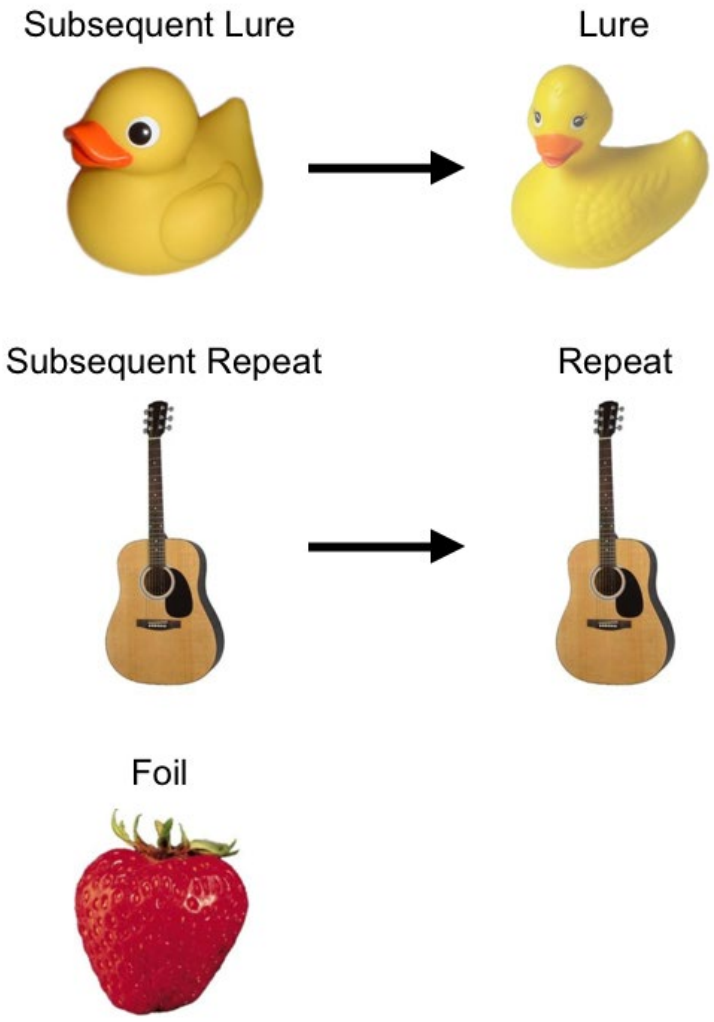

Figure 1. Labels of trial types. Participants saw objects one at a time on the center of a computer screen. Subsequent lures, subsequent repeats, and foils were all new items that were not presented before. Subsequent lures were followed at a variable amount of trials by a similar item of the same type of object. Subsequent repeats were followed at a variable amount of trials by an exact repeat of the image. Foils had not corresponding trial that followed their presentation.

After each drift check, an image was presented for 2.5 seconds, which was followed by an answer period where the response options of "old", "similar", and "new" were presented on the screen. The answer period ended when a participant responded and the task then continued

with the next trial. Each image was sized to 17.5° of visual angle on its longest axis and all images were presented for 2.5 seconds on the center of a screen with a gray background (see Figure 2 for sample sequence of images). Each task block consisted of 116 trials, 44 of which were foil presentations, 18 subsequent repeat presentations, 18 subsequent lure presentations, 18 repeat presentations, and 18 lure presentations (See Figure 2). Lure and repeat presentations were divided into different lag delays of 4, 12, and 40 trials with six trials of each stimulus type falling in each lag delay.

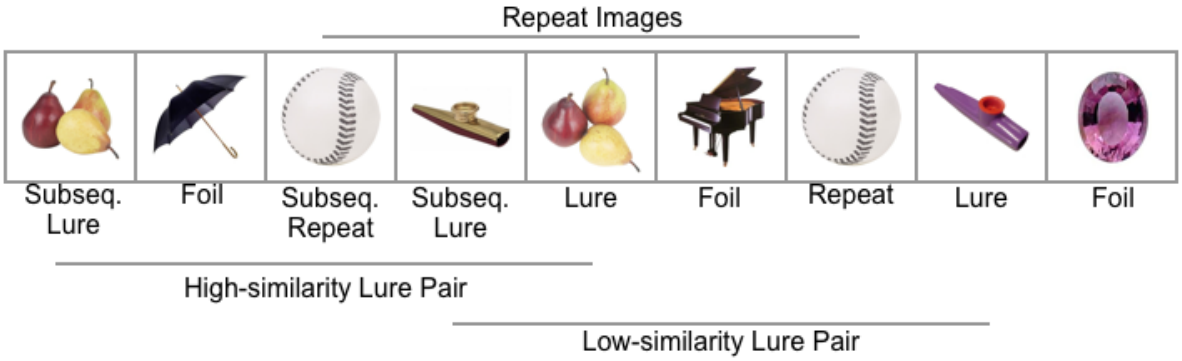

Figure 2. Sample sequence of stimuli. Stimuli were presented one at a time on the center of the screen in a pseudo-random order. Repeat and lure trials were presented with lag times of 4, 12, and 40 trials between first and second presentations. Lag times of 4 trials are presented for simplicity.

Lure pairs were also divided evenly between high similarity pairs and low similarity pairs (see Figure 3 for examples). These pairs were then evenly incorporated into each run so that, of the 18 lure pairs that were presented in each run, 9 were high similarity pairs and nine were low similarity pairs. These were further divided among the lag delays so that half of the trials in each lag were high similarity and half were low similarity.

Participants completed four task blocks. Items only repeated within a block and no items carried over to another block. Participants responded to each trial by indicating whether the item they saw was a new item (i.e., something they had never seen in the experiment), similar item

(i.e., something like an item they had seen before, but not exactly the same), or old item (i.e., an exact repeat of something they had seen before).

Data processing.

Behavioral data. Behavioral data were corrected for response bias prior to analysis. Bias corrected hit scores were calculated by subtracting the proportion of "old" responses to foil trials from the proportion of "old" responses to repeat trials. Lure correct rejection scores were

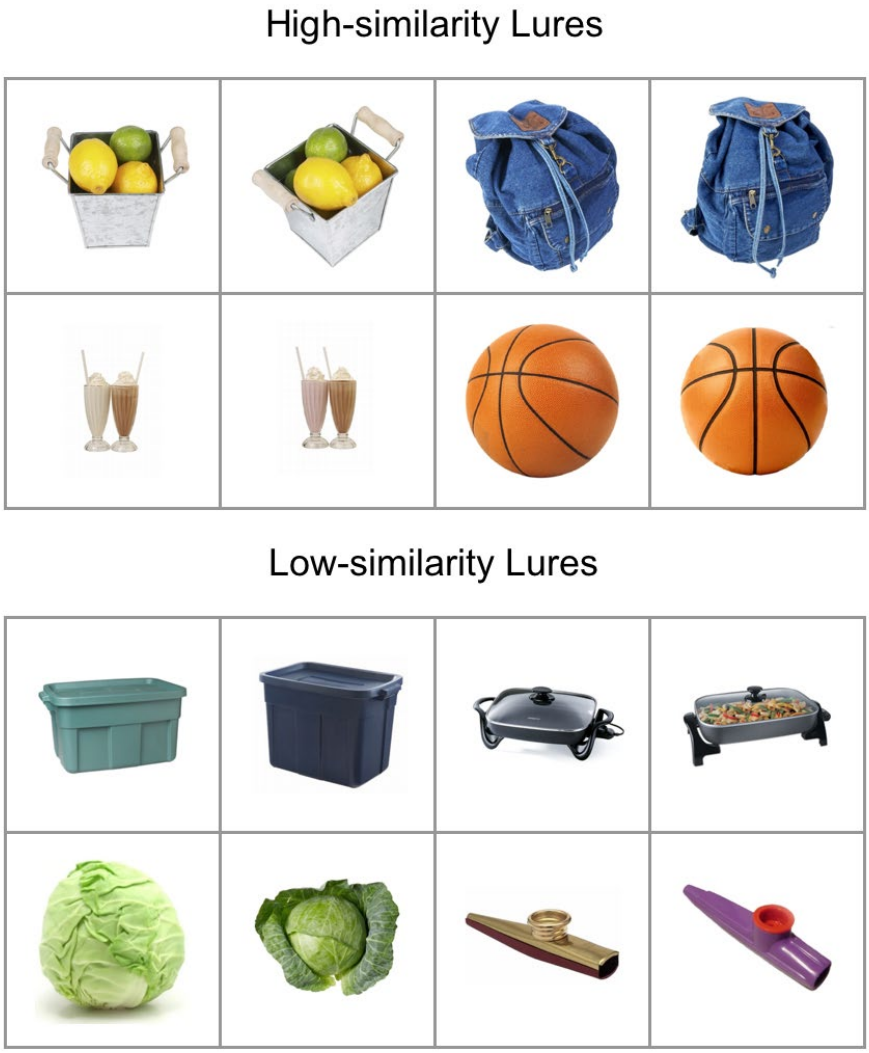

Figure 3. Sample high- and low-similarity lure pairs. Lure pairs were divided evenly between high similarity pairs and low similarity pairs.

calculated by subtracting the proportion of "similar" responses to foil trials from the proportion of "similar" responses to lure trials. Lure false alarm scores were calculated by subtracting the proportion of "old" responses to foil trials from the proportion of "old" responses to lure trials.

Eye movement data. Raw eye tracking data were processed using Data Viewer software (SR Research, Ontario, Canada). Saccades were defined as movement of at least 0.1° with velocity equal to or greater than 30°/sec and acceleration of 8000°/sec^2. Fixations were defined as periods of time between saccades that were not part of blinks. Saccades and fixations were removed from analysis if they were outside the maximum stimulus bounds (17.5° × 17.5° visual angle) or if they were immediately preceding or following a blink. Fixations less than 50 ms or greater than 1500 ms were also removed from analysis. Following the approach used by Molitor et al. (2014), memory-based fixation counts were calculated by subtracting the number of fixations for correct foil trials (i.e., correct rejections) from all other trial types. Thus, positive scores denote more fixations than correct foil trials and negative scores denote fewer fixations than correct foil trials.

Results.

Behavioral data. Overall, participants did well on the task, correctly identifying 94% of targets and 60% of lures. Proportions of responses can be found in Table 1. Bias corrected scores for hits and lure correct rejections were included in a t-test. Outliers were fenced to 1.5 times the interquartile range above or below the 75th and 25th percentiles, respectively. Scores for hits were significantly higher than scores for lure correct rejections, $t(62) = 24.01$, $p < .001$, 95% CI [0.34, 0.40], indicating that participants were better at recognizing repeat items than they were at discriminating similar lures. Figure 4 shows the bias corrected scores for hits, lure correct rejections, and lure false alarms.

Table 1

Proportion of Responses by Trial Type and Response for Experiment 1

Trial Type	Response	Mean Proportion	*SD* Proportion
First	New	0.96	0.03
First	Old	0.00	0.01
First	Similar	0.04	0.03
Foil	New	0.96	0.03
Foil	Old	0.01	0.01
Foil	Similar	0.04	0.03
Lure	New	0.07	0.06
Lure	Old	0.32	0.11
Lure	Similar	0.60	0.12
Repeat	New	0.02	0.04
Repeat	Old	0.94	0.05
Repeat	Similar	0.04	0.03

Note. SD = standard deviation, First = subsequent repeats and lures, Foil = items only presented once, Lure = similar items, Repeat = items presented twice

Eye movement data. To replicate the analysis of Molitor et al. (2014), memory-based fixation counts were entered into a repeated-measures ANOVA with factors coding for presentation (1st, 2nd) and response type (hit, lure correct rejection, lure false alarm). All univariate and multivariate outliers were fenced to 1.5 times the interquartile range above the 75th percentile or below the 25th percentile. The effects for presentation, $F(1, 62) = 9.04$, $p < .01$, $\eta_p^2 = .13$, and response type, $F(2, 124) = 32.30$, $p < .001$, $\eta_p^2 = .34$, were significant. The interaction between presentation and response type was also significant, $F(2, 124) = 11.39$, $p < .001$, $\eta_p^2 = .16$. Fixations on the first presentation of subsequent lure correct rejections were greater than fixations on the first presentation of subsequent lure false alarms, $t(62) = 4.37$, $p <$

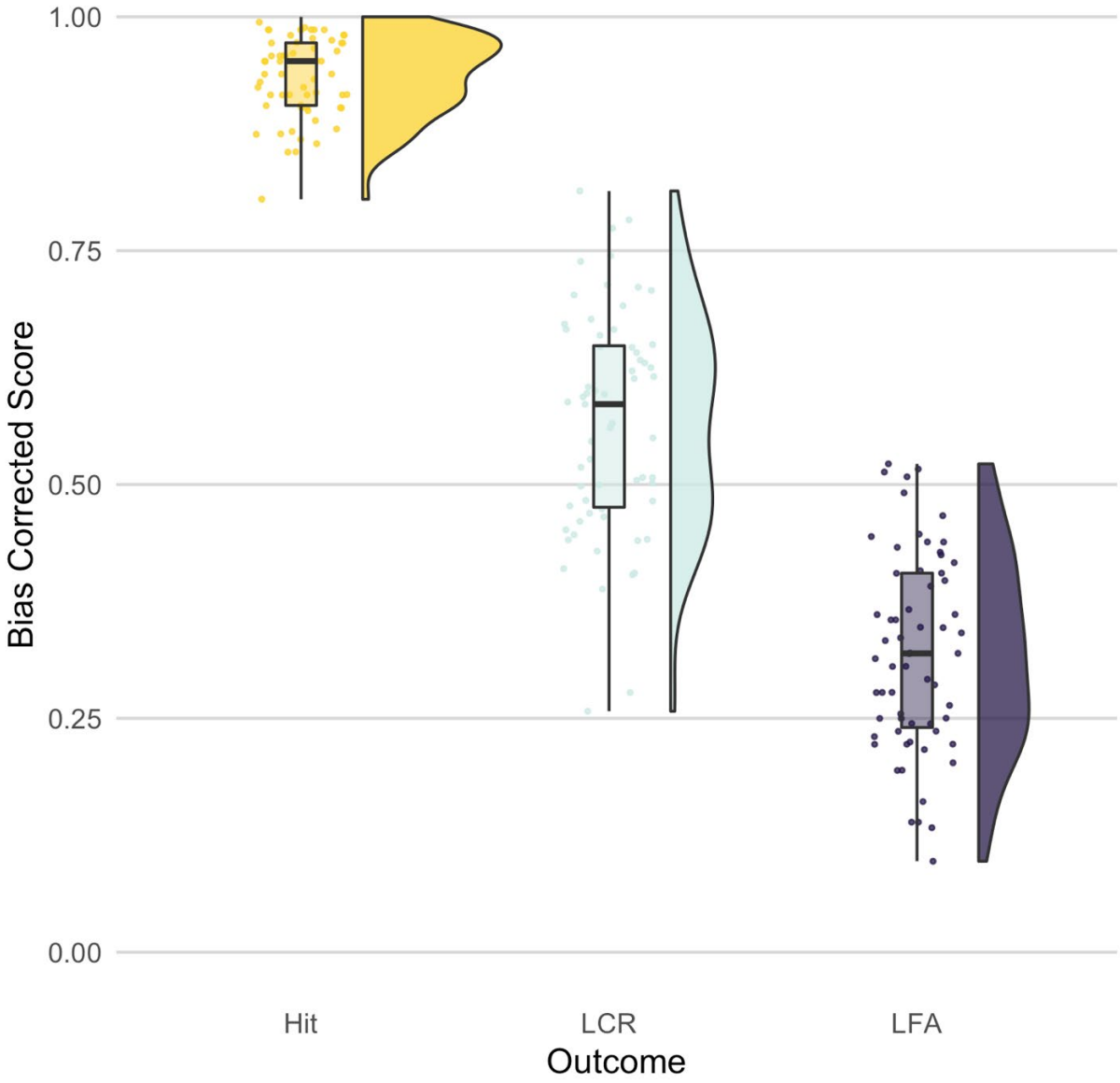

Figure 4. Bias-corrected scores for experiment 1 by outcome. Scores were calculated by subtracting the proportion of foils called old from the proportions of hits and lure false alarms and the proportion of foils called similar from lure correct rejections. LCR = lure correct rejection, LFA = lure false alarm.

.001, 95% CI [0.15, 0.41], thus replicating the findings of Molitor et al. (2014). However, fixations on the second presentation of lure correct rejections were also greater than fixations on the second presentation of lure false alarms, $t(62) = 3.32$, $p < .01$, 95% CI [0.07, 0.30]. There were fewer fixations for first presentations of subsequent lure correct rejections, $t(62) = -3.53$, p

< .01, 95% CI [-0.03, -0.08], and subsequent lure false alarms, $t(62) = -3.77$, $p < .001$, 95% CI [-0.44, -0.13], than the second presentations of the respective response types. There was no such difference for first and second presentation of hits, $t(62) = 1.05$, $p = .30$, 95% CI [-0.06, 0.18]. See Figure 5 for memory-based fixation counts by presentation and response type.

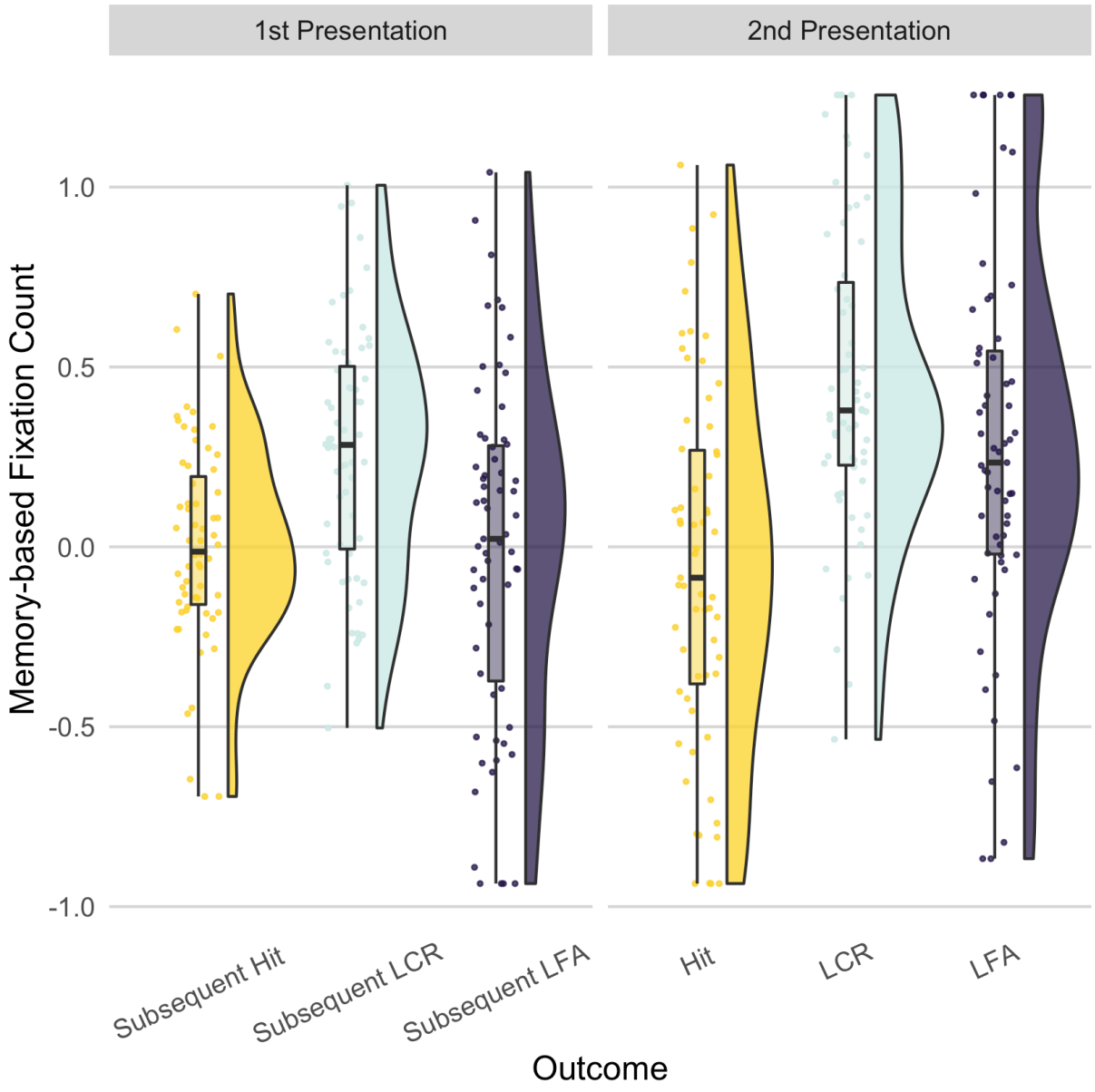

Figure 5. Memory-based fixation counts for experiment 1 by presentation and outcome. Fixation counts for subsequent LCRs were greater than for subsequent LFAs. LCRs also had a greater number of fixations than LFAs. LCR = lure correct rejection, LFA = lure false alarm

The more visually complex an image is, the greater the number of fixations on that image (Henderson, 2003). This could be a confounding factor in the present study, so we wanted to ensure that the effects remained significant when we controlled for visual complexity. If they did not remain significant, this would be evidence that visual complexity was driving the fixation counts rather than the presentation or response type. To test this, memory-based fixation counts were entered into a linear mixed effect model with presentation number and outcome as fixed effects along with visual complexity scores. Presentation number and outcome were also entered into the model as random slopes and subject was entered as a random intercept.

Visual complexity did significantly predict fixation counts, $\beta = 0.06$, $p < .001$, 95% CI [0.04, 0.08] when other variables were held at baseline. However, presentation number and response type still significantly predicted fixation count when visual complexity was accounted for. Table 2 contains information from the model.

To test the hypothesis that target-lure similarity level would be a better predictor of accuracy on the task than fixation counts at study or test, we entered data for lure trials into a logistic mixed effects model with accuracy as the dependent variable (i.e., 1 for correct and 0 for

Table 2

β-values for Mixed-effects Regression Model in Experiment 1

Effect	β-values	95% CI	
Subsequent LCR	0.26	0.14	0.37
Subsequent LFA	-0.08	-0.22	0.06
Hits	-0.08	-0.19	0.03
LCR	0.27	0.15	0.40
LFA	0.41	0.26	0.57
Visual Complexity	0.06	0.04	0.08

Note. Subsequent hits were used as the baseline.

incorrect). Raw fixation count at study, raw fixation count at test, and similarity level were entered as fixed effects. All three were also entered as random slopes and subject was entered as a random intercept. Fixation counts were centered at the mean. The model failed to converge, and further examination revealed that some subjects had no or few false alarms to low-similarity lure trials. The interactions with similarity level were therefore removed and similarity level was removed as a random slope. The successful model showed that both fixations at study, $\beta = 0.11$, $p < .001$, 95% CI [0.06, 0.17], and fixations at test, $\beta = 0.06$, $p = 0.02$, 95% CI [0.01, 0.12], significantly predicted a correct response when similarity level was held to high similarity trials. These effects, however, were quite small, with each fixation increasing the odds of a correct response by 12.09% (95% CI [6.28%, 18.21%]) and 6.96% (95% CI [1.17%, 13.29%]), respectively. Similarity level was by far the strongest predictor of a correct response, $\beta = 3.45$, $p < .001$, 95% CI [3.23, 3.68], with the odds of a correct response to a low similarity lure trial being 3,053.92% (95% CI [2,424.03%, 3,880.34%]) more likely than high similarity lure trials. The interaction between fixations at study and fixations at test was not significant. Model fit was tested by calculating a confusion matrix and 80.70% of the trials were correctly predicted.

Discussion. Experiment 1 sought to replicate the findings of Molitor et al. (2014) by using the same study design, with the exception of a greater number of trials and controlling for lure pair similarity level. We replicated the finding that fixations at study were significantly greater for subsequent lure correct rejections than for subsequent lure false alarms. However, we also found the same pattern between lure correct rejections and lure false alarms on test trials. Interestingly, there was a greater number of fixations at test than there was for study for both lure correct rejections and lure false alarms. The pattern of fixations is consistent with the suggestion that a participant recognizes that the gist of the image is consistent with something that they have

seen before and then spends time looking in greater detail for the aspects that would indicate whether it is an exact repeat and whether it is only similar.

The logistic regression demonstrates that fixations at study predicted the correct response of a trial at a greater rate than fixations at test. This is consistent with the interpretation of Molitor et al. (2014) that it is important for an individual to have an accurate memory trace with which the new stimulus may be compared. Even more compelling, however, and not considered by Molitor et al. (2014) is that target-lure similarity level is a greater predictor of accuracy on the task than both fixations at study and fixations at test. This finding is consistent with the idea that pattern completion is a large contributor to lure false alarms. This will be discussed in greater detail in the general discussion.

Experiment 2

The design of experiment 1 is different from designs that are most commonly used for a mnemonic discrimination task by requiring that participants respond to trials on a separate answer period after the image was no longer displayed. It is more common to have participants respond while the image is still displayed (Kirwan & Stark, 2007), thus reducing the amount of time that participants view the stimulus before making a memory judgment about it. The difference in study design makes it difficult to generalize the findings to other experiments conducted on mnemonic discrimination and pattern separation.

Experiment 2 addresses this issue by requiring participants were required to respond to trials while the image was displayed. All the hypotheses for experiment 2 were the same as experiment 1. Specifically, we hypothesized that fixations would differ between lure correct rejections and lure false alarms. We also hypothesized that visual complexity would significantly

predict fixations on trials. Lastly, we hypothesized that similarity level would predict accuracy on trials at a greater rate than fixations at study or test.

Methods.

Participants. The same power analysis used for experiment 1 was used to determine the number of subjects needed for experiment 2. A total of 68 participants (34 male, 34 female) were recruited from Brigham Young University for the experiment. All participation requirements and exclusion criteria were the same as experiment 1 as well as the requirement that participants had not participated in experiment 1.

Data were excluded from five participants. Two participants were excluded because the eye tracker could not calibrate correctly. Two more participants were excluded because of technical issues with the display computer. Lastly, one participant was excluded for not following task instructions. After exclusions, a total of 63 participants (33 male, 30 female) were included for data analysis. Participants had an average age of 20.57 years (SD = 2.37). All participants were granted course credit for participating in the experiment. All study protocols were approved by the Brigham Young University Institutional Review Board and all participants gave written informed consent prior to participating.

Apparatus. The same apparatus used for experiment 1 was used for experiment 2.

Stimuli and procedure. All stimuli and procedures were the same as experiment 1 with the exception of the removal of the drift check between each trial and the separate answer period. Instead, a drift check was performed after every 10 trials. In between trials, participants viewed a fixation cross at the center of the screen for 0.5 seconds. The image was presented for 2.5 seconds and participants responded while the image was presented. All data processing was the same as in experiment 1.

Results.

Behavioral data. Participants correctly identified 92% of targets and 55% of lures (see Table 3 for proportion of responses). Like experiment 1, bias corrected scores for hits and lure correct rejections were compared with a t-test. Outliers were fenced to 1.5 times the interquartile range above or below the 75th and 25th percentiles, respectively. Scores for hits were significantly higher than scores for lure correct rejections, $t(62) = 24.30$, $p < .001$, 95% CI [0.37, 0.43]. Figure 6 shows the bias corrected scores for hits, lure correct rejections, and lure false alarms.

Table 3

Proportion of Responses by Trial Type and Response for Experiment 2

Trial Type	Response	Mean Proportion	*SD* Proportion
First	New	0.95	0.06
First	Old	0.01	0.03
First	Similar	0.04	0.04
Foil	New	0.95	0.06
Foil	Old	0.01	0.03
Foil	Similar	0.04	0.04
Lure	New	0.09	0.07
Lure	Old	0.35	0.1
Lure	Similar	0.55	0.13
Repeat	New	0.04	0.05
Repeat	Old	0.92	0.06
Repeat	Similar	0.05	0.04

Note. SD = standard deviation, First = subsequent repeats and lures, Foil = items only presented once, Lure = similar items, Repeat = items presented twice

Eye movement data. Like experiment 1, the hypothesis that fixation counts would differ between lure correct rejections and lure false alarms was tested by calculating memory-based fixation counts and then dividing them by the trial time to calculate memory-based fixations per

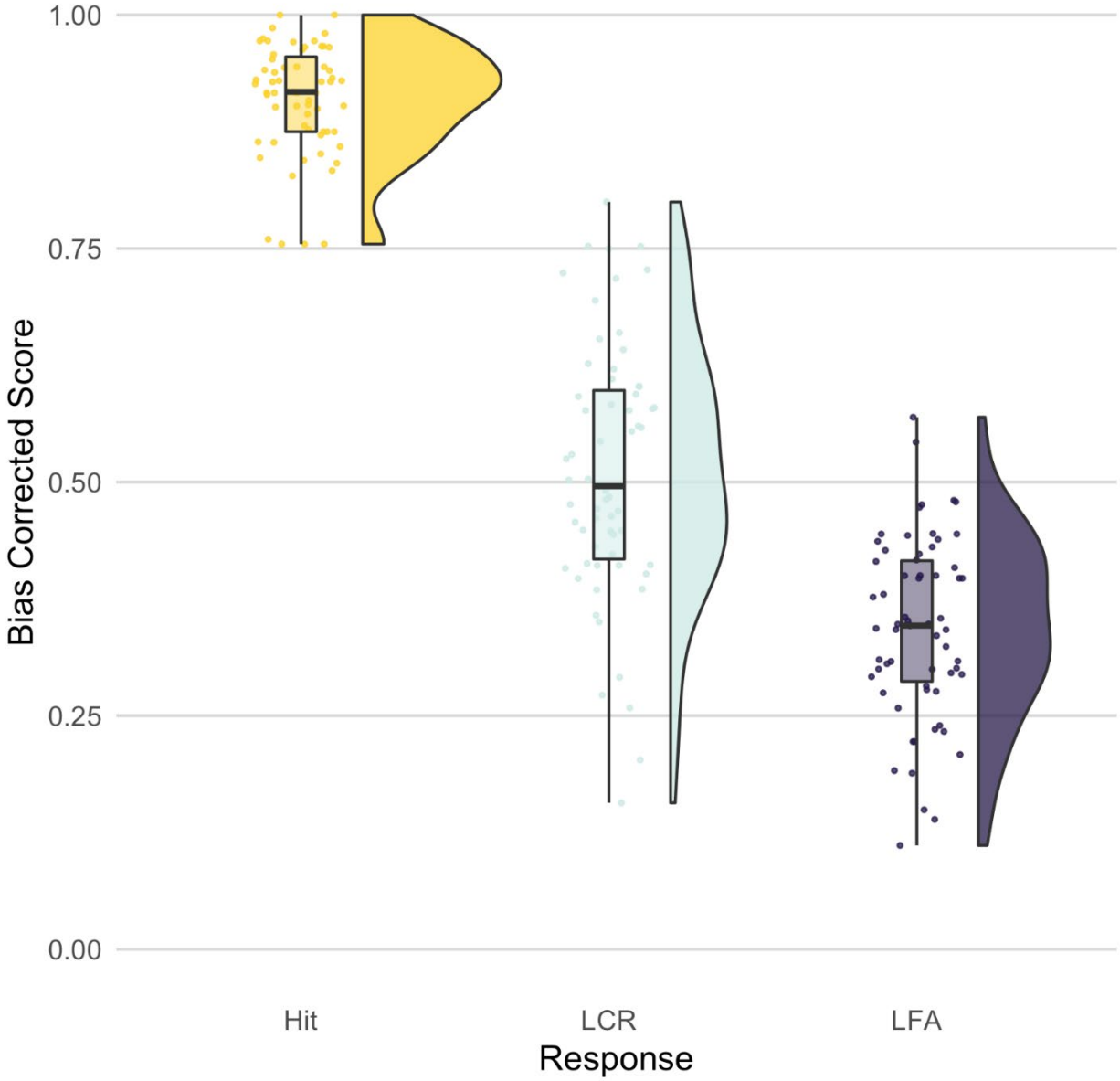

Figure 6. Bias-corrected scores for experiment 2 by outcome. Scores were calculated by subtracting the proportion of foils called old from the proportions of hits and lure false alarms and the proportion of foils called similar from lure correct rejections. LCR = lure correct rejection, LFA = lure false alarm.

second and then entering those data into a repeated-measures ANOVA with factors coding for presentation (1st, 2nd) and response type (hit, lure correct rejection, lure false alarm). All univariate and multivariate outliers were fenced to 1.5 times the interquartile range above the 75th percentile or below the 25th percentile. The main effects of presentation, $F(1, 62) = 163.31$, p

< .001, η_p^2 = .73, and response, $F(2, 124) = 8.49$, $p < .001$, η_p^2 = .12, were significant, but the interaction between presentation and response was not significant, $F(2, 124) = 0.64$, $p = .53$, η_p^2 = .01. Post-hoc analyses on the effect of response revealed fewer fixations per second for hits than lure correct rejections, $t(62) = -4.26$, $p < .001$, 95% CI [-0.12, -0.04]. Fixations per second for hits and lure false alarms did not differ, $t(62) = -1.49$, $p = .14$, 95% CI [-0.07, 0.01]. Lure correct rejections had a greater number of fixations per second that lure false alarms, $t(62)$ = 2.44, $p = .02$, 95% CI [0.01, 0.09]. Figure 7 shows the fixations per second by presentation and response type.

We then entered the data into a linear mixed-effects model to test the hypothesis that visual complexity of the images would predict fixation count. Presentation number, response type, and visual complexity scores were entered as fixed effects. Presentation number and response type were allowed to interact. Presentation number and response type were also entered as random slopes along with subject as a random intercept. Visual complexity did significantly predict fixation counts when other variables were held at baseline, $\beta = 0.02$, $p < .001$, 95% CI [0.01, 0.02]. Controlling for visual complexity, lure correct rejections significantly differed from hits, $\beta = 0.07$, $p < .01$, 95% CI [0.02, 0.12], but lure false alarms did not, $\beta = -0.002$, $p = .94$, 95% CI [-0.05, 0.05]. Interaction effects between presentation number and response type were not statistically significant. Table 4 contains data from the model.

We then entered the data into a logistic mixed-effects model with accuracy as the dependent variable to test the hypothesis that target-lure similarity level would be a better predictor of performance on the task than fixations at study or test. Fixations per second at study, fixations per second at test, and similarity level were entered into the model. Fixations per second at study and fixations per second at test were allowed to interact and were also entered as

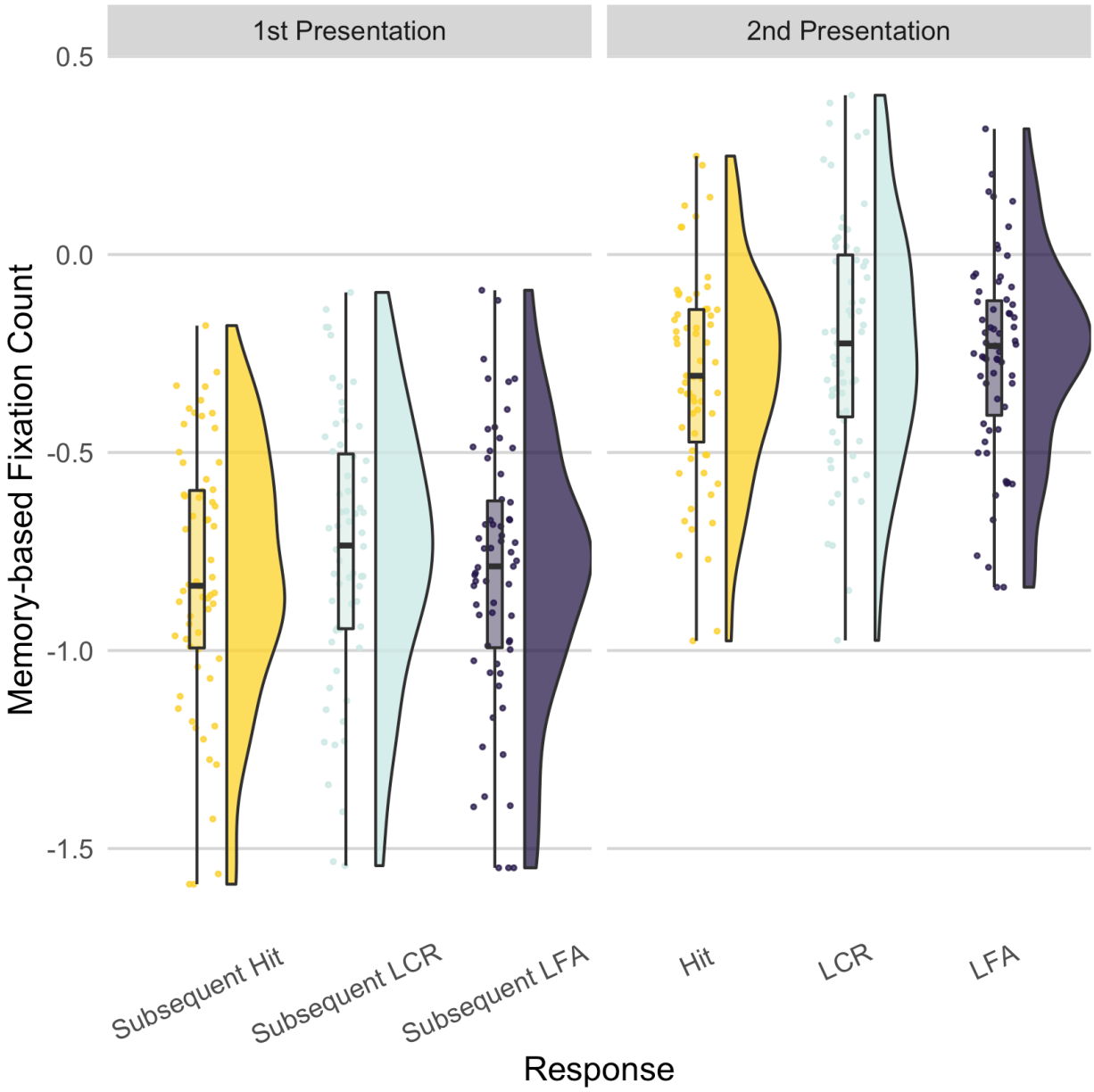

Figure 7. Memory-based fixations per second for experiment 2 by presentation and outcome. Unlike experiment 1, fixation counts for subsequent LCRs and subsequent LFAs did not differ, nor did counts for LCRs and LFAs. LCR = lure correct rejection, LFA = lure false alarm.

random slopes. Subject was entered as a random intercept. The model failed to converge, so the random slopes were removed. On the converged model, fixations at study significantly predicted accuracy on lure trials, $\beta = 0.18$, $p < .01$, 95% CI [0.07, 0.29], but fixations at test did not, $\beta = 0.04$, $p = .39$, 95% CI [-0.06, 0.14].Like experiment 1, similarity level predicted accuracy on lure

Table 4

β-values for Mixed-effects Regression Model in Experiment 2

Effect	β-values	95% CI	
Subsequent LCR	0.07	0.02	0.12
Subsequent LFA	0.00	-0.05	0.05
Hits	0.51	0.43	0.60
LCR	0.01	-0.05	0.07
LFA	0.05	-0.02	0.12
Visual Complexity	0.02	0.01	0.02

Note. Subsequent hits were used as the baseline.

trials at a much higher rate than fixations, β = 3.09, $p < .001$, 95% CI [2.90, 3.29], increasing the odds of a correct response on lure trials by 2,106.89% (95% CI [1,718.05%, 2,595.34]) as compared to 19.49% (95% CI [7.04%, 33.42%]) for fixations at study. The interaction between fixations at study and fixations at test was not statistically significant, β = -0.07, $p = 0.15$, 95% CI [-0.15, 0.02]. Model fit was tested by calculating a confusion matrix and 79.09% of the trials were correctly predicted.

To compare the results of experiment 1 with experiment 2 we redid the analyses conducted in experiment 1 with fixations per second as the dependent variable. The results of all the tests with fixations per second were the same as the results presented in experiment 1 that used the number of fixations as the dependent variable.

Discussion. Experiment 2 sought to replicate the findings of Molitor et al. (2014) using a more traditional mnemonic discrimination task, specifically by requiring that participants answer during the presentation of the stimulus rather than during a set answer period. Experiment 2 did not replicate the finding of an interaction between response type and presentation number on memory-based fixation counts. Interaction effects in the linear mixed-effects model that

controlled for visual complexity were also not significant, while still observing a significant effect of visual complexity. Further, the logistic regression model that used fixations at study, fixations at test, and similarity level showed evidence of fixations at study predicting performance, but still at a much lower rate than target-lure similarity level.

The findings of experiment 2 suggest that some of the finding of Molitor et al. (2014) could be unique to their study design. Small differences in study design can change outcomes for other measures as well. For example, Motley & Kirwan (2012) found that fMRI activity patterns differed when participants completed an implicit task that required them to respond whether or not and item was a toy as opposed to a task that required them to explicitly respond to an object as old, similar, or new. The present experiment varied when participants could respond to the items, allowing them to respond while the stimulus was presented rather than waiting for the answer period like in experiment 1. The different design in experiment 2 decreased the amount of time that participants were able to view the test item.

We did observe the same limitation of low number of correct responses on low-similarity lure trials in experiment 2. It is possible that this contributed to why the logistic regression model failed to converge in experiment 2, which required adjustment to the model before it would work. Further research could address this weakness by either increasing the number of trials in the task, or slightly increasing the difficulty of the low-similarity lure trials to ensure that there are enough false alarms for a complete analysis.

The results of the logistic regression model in the present experiment are also difficult to directly compare to the results of experiment 1 due to the fact that the same model could not be used on both sets of data. While the same effects were used, the random slopes differed and could change the effects of the experiment. It is likely, however, that adding the random slopes

would decrease the significance of the effects in the model since more of the variance is accounted for in the error terms of the model.

Overall, experiment 1 and experiment 2 show different results in terms of the number of fixations at study for lure correct rejections and lure false alarms. Both experiments, however, strongly demonstrated that target-lure similarity was a much more reliable predictor of accuracy on lure trials than eye movements at study. These patterns of results across experiments give stronger support for the contribution of pattern completion in false alarms to lure stimuli.

Experiment 3

The previous two experiments show interesting results with the eye movement data. As stated previously, experiment 3 combines eye movement data with fMRI data. The benefit of gathering both of these simultaneously is that we can examine activity in the hippocampus as a function of the number of fixations an individual makes on an item. This approach allows us to see how fixation counts affect hippocampal activity and whether this effect is different between lure correct rejections and lure false alarms. A recent study (Liu et al., 2017) demonstrated that hippocampal activation for repeated images changed as a function of the number of fixations on the first presentation of the stimulus, with more fixations at study predicting a larger decrease in activation in the hippocampus at retrieval. Experiment 3 extends the findings of experiments 1 and 2 in an effort to determine if a similar effect occurs in the context of a difficult mnemonic discrimination task (i.e., the MST).

Neuroimaging research examining mnemonic discrimination relies on the repetition suppression effect in order to differentiate how items are processed by the hippocampus. Repeat items show a marked decrease in activity as compared to novel items, much like what was found by Liu et al. (2017). False alarms to lure items follow this same pattern, but activity for correct

rejections of lure items looks more like the activity for novel items. If fixations to lure items are driving this reduction in activity, we should see a greater contribution of study trial fixations for lure false alarms than for lure correct rejections.

Methods.

Participants. We recruited 49 individuals (22 male, 27 female) to participate in the study. All participants were required to have normal to corrected to normal vision and were excluded if they reported color blindness or any history of traumatic brain injury, learning disability, or psychological or neurological diagnosis. Because this study used fMRI as well as eye tracking, participants were also required to be right handed, native English speakers to control for any potential confounds for those factors. To ensure the safety of participants in the scanner, they were excluded if they had any contraindications for an MRI scan as indicated by the BYU MRI Research Facility screening form.

Data were excluded from 14 individuals. Four participants were removed due to difficulty calibrating the eye tracker. Six other participants were removed because of technical difficulties with the MRI scanner, having participated in an experiment with same stimulus set previously, safety considerations, not following task instructions, or time constraints causing the experiment to not finish. Four participants were also excluded for excessive motion while in the scanner. The limit for excessive motion was set at 1.5 times the interquartile range above the 75th percentile of the number of TRs lost for subjects. After exclusions, a total of 35 participants (18 male, 17 female) were included for data analysis. Participants had an average age of 22.06 years (SD = 2.41). Participants were given the choice of monetary compensation or a small 3D printed copy of their brain. All study protocols were approved by the Brigham Young University

Institutional Review Board and all participants gave written informed consent prior to participating.

Apparatus. Stimuli were presented using an MRI compatible BOLDscreen monitor (Cambridge Research Systems Ltd., Rochester, United Kingdom) with a resolution of 1920 × 1200 pixels and a refresh rate of 60 Hz. Eye movement data were collected using an MRI compatible EyeLink 1000 Plus long-range eye tracker (SR Research, Ontario, Canada) with a spatial resolution of 0.01° and sampled at 1000 Hz from the right eye. The eye tracker was connected to a host computer running EyeLink 5.00 tracking software. The experiment was programmed with PsychoPy (version 1.85.6; Peirce, 2007, 2008) and used the EyeLink Developers Toolkit (SR Research, Ontario, Canada) to integrate the eye tracker. Participant responses were collected using a 4-button cylinder MRI compatible button box (Cambridge Research Systems Ltd, Rochester, United Kingdom).

Prior to each block, a 9-point calibration was run to ensure accurate data collection. Successful calibration required an average error of less than 0.50° and a maximum error of less than 1.00°.

MRI data acquisition. MRI scans were conducted using a Siemens 3T Tim-Trio scanner at the Brigham Young University MRI Research Facility. A structural scan was acquired for each subject for functional localization using a T1-weighted magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence with the following parameters: TR = 1900 ms; TE = 2.26 ms; number of slices = 176; slice thickness = 1 mm; matrix size = 256 × 215; field of view = 256 × 215 mm; voxel size = 1 mm^3; flip angle = 9°. High resolution structural images of the medial temporal lobe were also obtained for examination of sub-fields. A T2-weighted pulse sequence was acquired with the following parameters: TR = 6260 ms; TE = 64 ms; 19

interleaved slices; slice thickness = 3 mm; matrix size = 512 × 512; field of view = 200 mm^2; voxel size = 0.39 × 0.39 × 3 mm; flip angle = 178°; averages = 2. This scan was aligned perpendicular to the longitudinal axis of the hippocampus prior to acquisition. If a SAR warning came up prior to the scan, TR length was extended to preserve the number of slices while ensuring participant safety. Functional scans were acquired using a gradient-echo, echo-planar, T2*-weighted pulse sequence with the following parameters: TR = 1900 ms; TE = 28 ms; number of slices = 29 interleaved; slice thickness = 1.8 mm; matrix size = 128 × 128; field of view = 230.4 mm^2; voxel size = 1.8 mm^3; flip angle = 90°. Slices were positioned parallel to the longitudinal axis of the hippocampus prior to acquisition.

Stimuli and procedure. The stimuli and procedure were the same as experiment 1 with some exceptions. First, participants completed five blocks of 116 trials rather than four blocks. The increased number of blocks was to ensure an adequate number of trials for fMRI analysis. Second, the task completed in the scanner did not contain a drift check prior to each trial due to the need to keep trial timing consistent for fMRI analysis. Third, again to keep trial timing consistent, the answer period consisted of 1.5 seconds rather than a self-paced answer period.

Data processing.

Behavioral data. Behavioral data were processed the same way as in experiment 1.

Eye movement data. Eye movement data were processed using the same procedures as experiment 1.

MRI data. Data were analyzed using the Analysis of Functional Neuroimaging (AFNI) suite of programs (Cox, 1996). Specifically, structural scans were coregistered to the first functional run. Functional scans were slice time corrected and within scan motion was corrected by registering to the middle volume of each run. Between scan motion was corrected by

registering the second functional run to the first. Volumes with large motion events, defined as more than 0.3° rotation or 0.6 mm translation in any direction, were not included in the data analysis. A single subject regression was conducted for each participant that included regressors for motion and scanner drift, along with regressors coding for the task conditions of subsequent hits, hits, subsequent misses, misses, subsequent lure correct rejections, lure correct rejections, subsequent lure false alarms, lure false alarms, inter-trial interval and answer periods, and all other trials. Correct foil trials were used as the baseline. A parametric modulator coding for the number of fixations for each trial was also included for regressors where fixation counts were recorded. The regression analysis gave two different β values for each regressor. The first was a regression analysis that ignored the parametric modulator and the second included the parametric modulator in the analysis.

Hippocampal subfields were segmented for each subject using Freesurfer 6.0 (Iglesias et al., 2015) with both the T1 and T2 structural images. Regions of interest (ROI) for DG/CA3were created by combining the CA3, CA4, and DG output of the segmentation process for each hemisphere of each subject. The ROI for CA1 was created by using the CA1 output of the segmentation process for each hemisphere of each subject. An ROI for the whole hippocampus was created by combining all subfields in the segmentation output (see Figure 8). Binary masks were then created for each of the ROIs and were then used to extract the mean β -value from the single subject regressions for each ROI. Mean β -values were extracted for all regressors with the exception of inter-trial interval and answer periods as well as the "other" category of trials. Difference scores were then calculated by subtracting the β-values for test trials from the β-values for study trials within each outcome time within each ROI (Liu et al., 2017). The resulting

score shows the difference in activity from study trials to test trials as a function of fixation counts for that outcome and ROI.

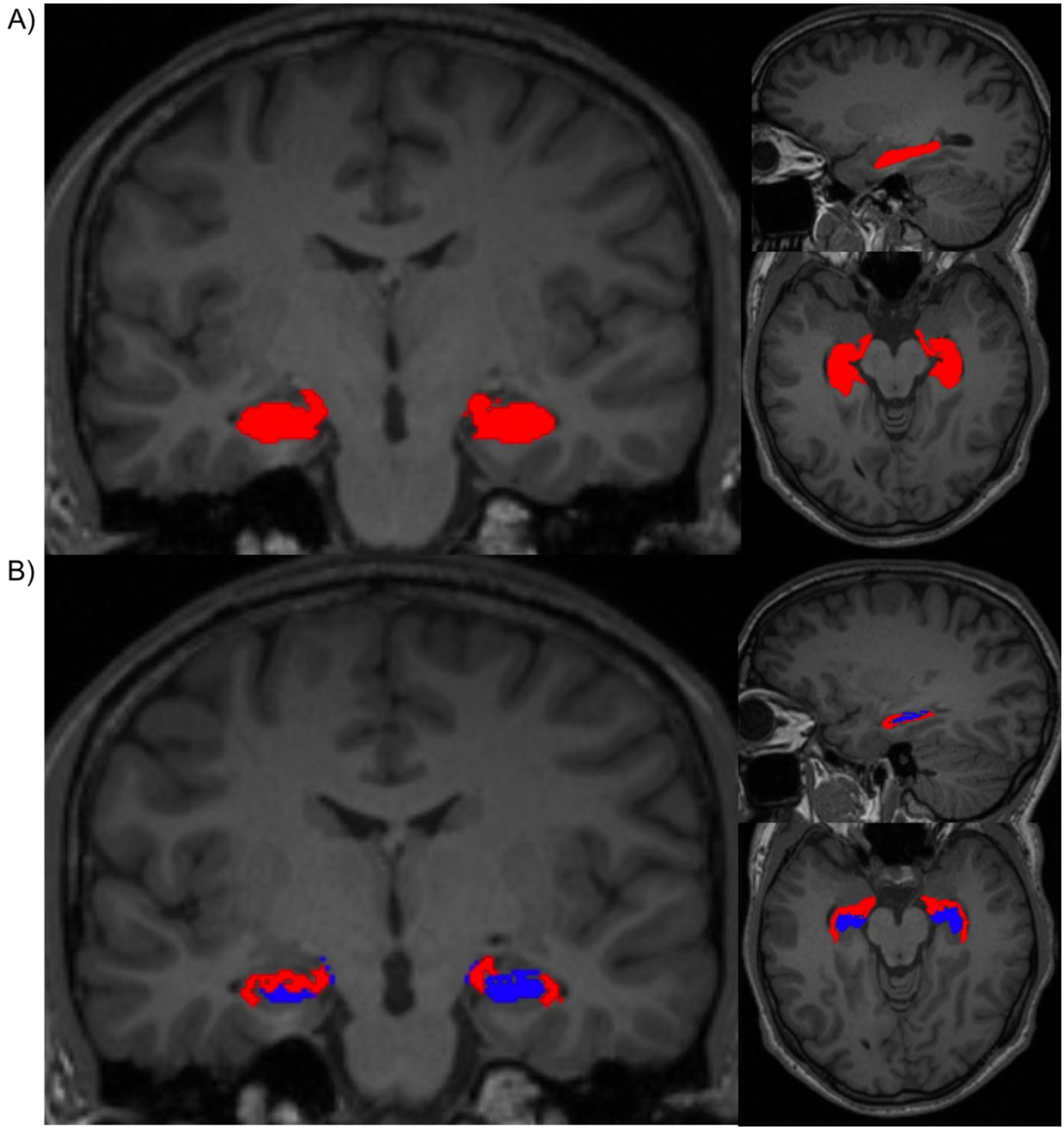

Figure 8. Regions of interest for fMRI analyses. A) Regions of interest for the full hippocampus. B) Subfield regions of interest. Blue = DG/CA3, Red = CA1.

Results.

Behavioral data. Participants correctly responded to 89% of targets and 63% of lures. Table 5 shows the proportion of responses. Bias corrected scores for hits and lure correct rejections were compared using a paired t-test. Outliers were fenced to 1.5 times the interquartile range above or below the 75th and 25th percentiles, respectively. Participants had significantly more hits than they had lure correct rejections, $t(34) = 14.05$, $p < .001$, 95% CI [0.24, 0.32]. Figure 9 shows the bias corrected scores for hits, lure correct rejections, and lure false alarms.

Table 5

Proportion of Responses by Trial Type and Response for Experiment 3

Trial Type	Response	Mean Proportion	*SD* Proportion
First	New	0.98	0.02
First	Old	0.00	0.01
First	Similar	0.02	0.02
Foil	New	0.97	0.03
Foil	Old	0.00	0.01
Foil	Similar	0.02	0.03
Lure	New	0.11	0.11
Lure	Old	0.26	0.09
Lure	Similar	0.63	0.13
Repeat	New	0.04	0.04
Repeat	Old	0.89	0.08
Repeat	Similar	0.08	0.05

Note. SD = standard deviation, First = subsequent repeats and lures, Foil = items only presented once, Lure = similar items, Repeat = items presented twice

Eye movement data. To test the hypothesis that fixations at study would differ between response types we calculated memory-based fixation counts and entered them into a repeated measures ANOVA as done in experiments 1 and 2. Outliers were fenced to 1.5 times the interquartile range above the 75th percentile or below the 25th percentile. The effects of presentation, $F(1, 34) = 26.47$, $p < .001$, $\eta_p^2 = .44$, and response type, $F(2, 68) = 3.65$, $p = .03$, η_p^2

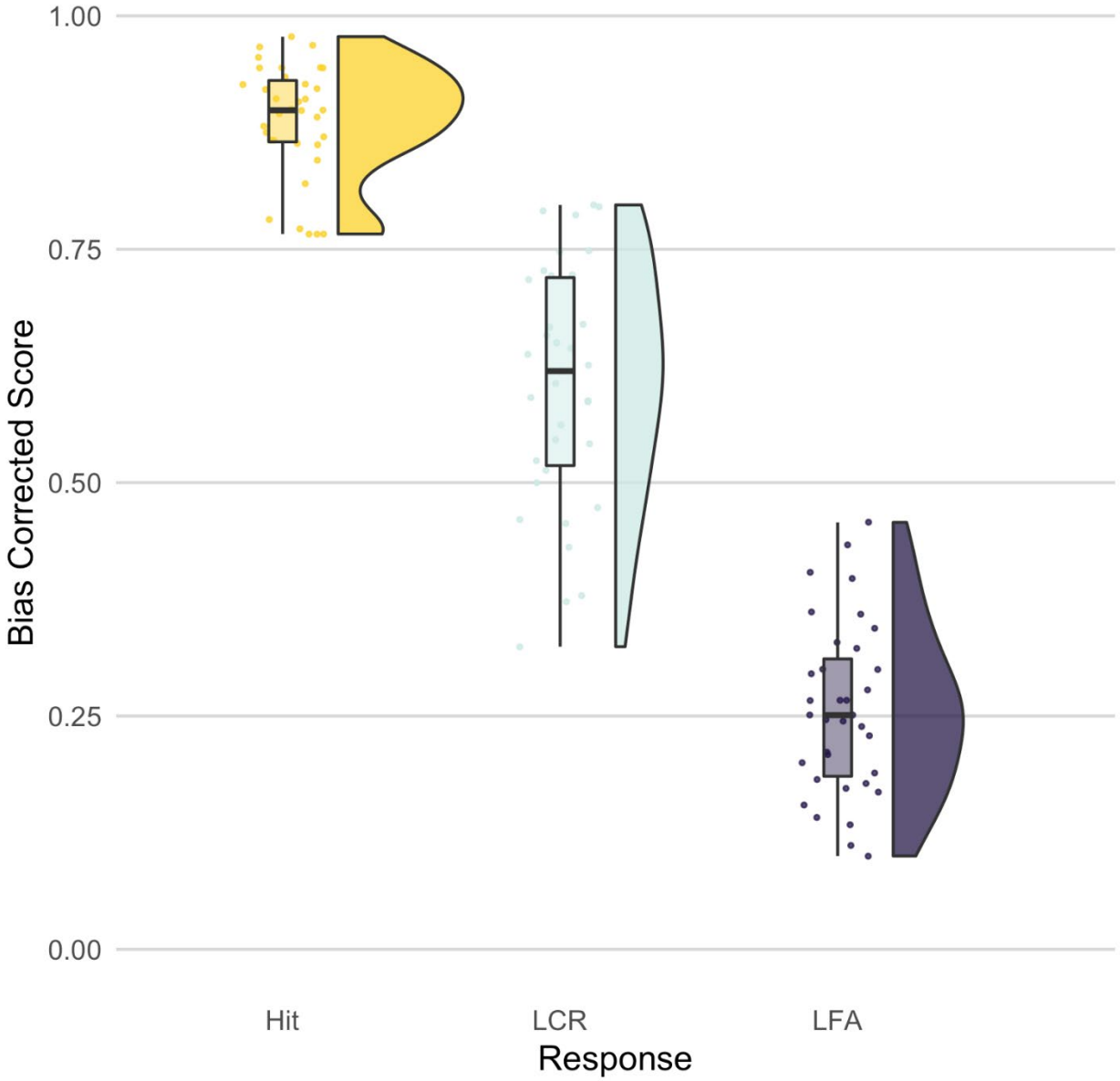

Figure 9. Bias-corrected scores for experiment 3 by outcome. LCR = lure correct rejection, LFA = lure false alarm.

= .10, were significant. The interaction between presentation and response type was not significant, $F(2, 68) = 0.04$, $p = .96$, $\eta_p^2 = .00$. Post-hoc tests revealed significantly fewer fixations for hits than for lure correct rejections, $t(34) = -3.13$, $p < .01$, 95% CI [-0.27, -0.06], and for lure false alarms, $t(34) = -2.04$, $p < .05$, 95% CI [-0.25, 0.00], though the upper limit of the95% confidence interval very closely approaches zero. There was not a significant difference between fixation counts for lure correct rejections and lure false alarms, $t(34) = 0.49$, $p = .63$, 95 % CI [-0.11, 0.19]. Figure 10 shows memory-based fixation counts by presentation and response type.

We then tested the hypothesis that visual complexity would predict the number of fixations on trials. Memory-based fixation counts were entered into a linear mixed-effects model with fixed effects of presentation number, response type, and visual complexity. Presentation number and response type were allowed to interact. Presentation number and response type were also added to the model as random slopes along with subject as a random intercept to replicate the model used in experiments 1 and 2. The model failed to converge, so the random slope for presentation number was removed. The converged model showed that visual complexity significantly predicted fixation counts, $\beta = 0.04$, $p < .001$, 95% CI [0.02, 0.06]. Table 6 displays data from the model.

Lastly, to test the hypothesis that target-lure similarity level would be a better predictor of accuracy on the task than fixation counts at study or test, we entered data for lure trials into the same logistic mixed effects model that was used for experiments 1 and 2. Specifically, accuracy was entered as the dependent variable (i.e., 1 for correct and 0 for incorrect). Raw fixation count at study, raw fixation count at test, and similarity level were entered as fixed effects and fixation counts were allowed to interact. Fixations at study and test were entered as random slopes and

Table 6

β-values for Mixed-effects Regression Model in Experiment 3

Effect	β-values	95% CI	
Subsequent LCR	0.16	0.03	0.28
Subsequent LFA	0.08	-0.06	0.23
Hits	0.26	0.16	0.36
LCR	-0.03	-0.18	0.12
LFA	0.04	-0.16	0.24
Visual Complexity	0.04	0.02	0.06

Note. Subsequent hits were used as the baseline.

subject was entered as a random intercept. In this model, target-lure similarity level was the only significant predictor of a correct response on lure trials, β = 2.84, $p < .001$, 95% CI [2.59, 3.12], increasing the odds of a correct response by 1,617.84% (95% CI [1,226.76%, 2,154.75%]). Fixations at study, β = 0.01, $p = .70$, 95% CI [-0.06, 0.08], and fixations at test, β = 0.004, $p = .91$, 95% CI [-0.07, 0.08], did not significantly predict correct responses, nor did the interaction between fixations at study and fixations at test, β = -0.02, $p = .08$, 95% CI [-0.05, 0.002]. Model fit was tested with a confusion matrix and 78.08% of trials were correctly predicted.

MRI data. We first wanted to ensure that we found the same effect of novel fixations on repetition suppression that Liu et al. (2017) found before proceeding to test trial outcomes and subfield ROIs. We therefore collapsed difference scores across hits and misses to calculate the mean difference score for repeat trials. Outliers were fenced to 1.5 times the interquartile range above the 75th percentile or below the 25th percentile. There was a greater amount of repetition suppression on repeat trials as study trial fixation increased in the left, $t(34) = 2.15$, $p = .04$, 95% CI [0.02, 0.66], but not the right hippocampus, $t(34) = 0.76$, $p = 0.45$, 95% CI [-0.17, 0.37]. Figure 11 shows the difference scores divided by hemisphere.

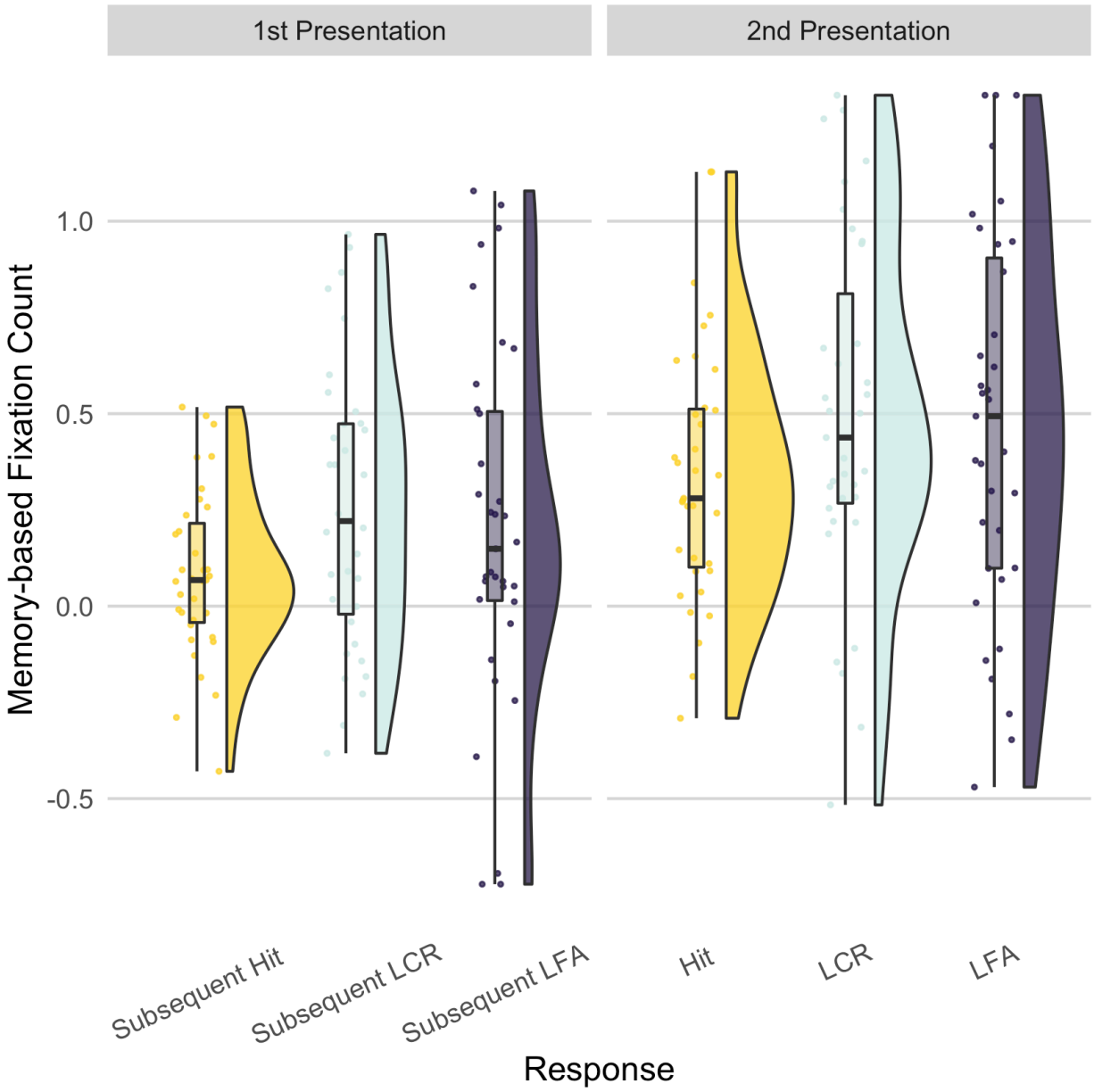

Figure 10. Memory-based fixations for experiment 3 by presentation and outcome. Fixation counts for subsequent LCRs and subsequent LFAs did not differ, nor did counts for LCRs and LFAs. LCR = lure correct rejection, LFA = lure false alarm.

Next, to test the hypothesis that lure false alarms would show a greater reduction in activity as a function of novel fixations, we entered difference scores for the left and right hippocampus into separate repeated measures ANOVA with coded factors for region (CA1,

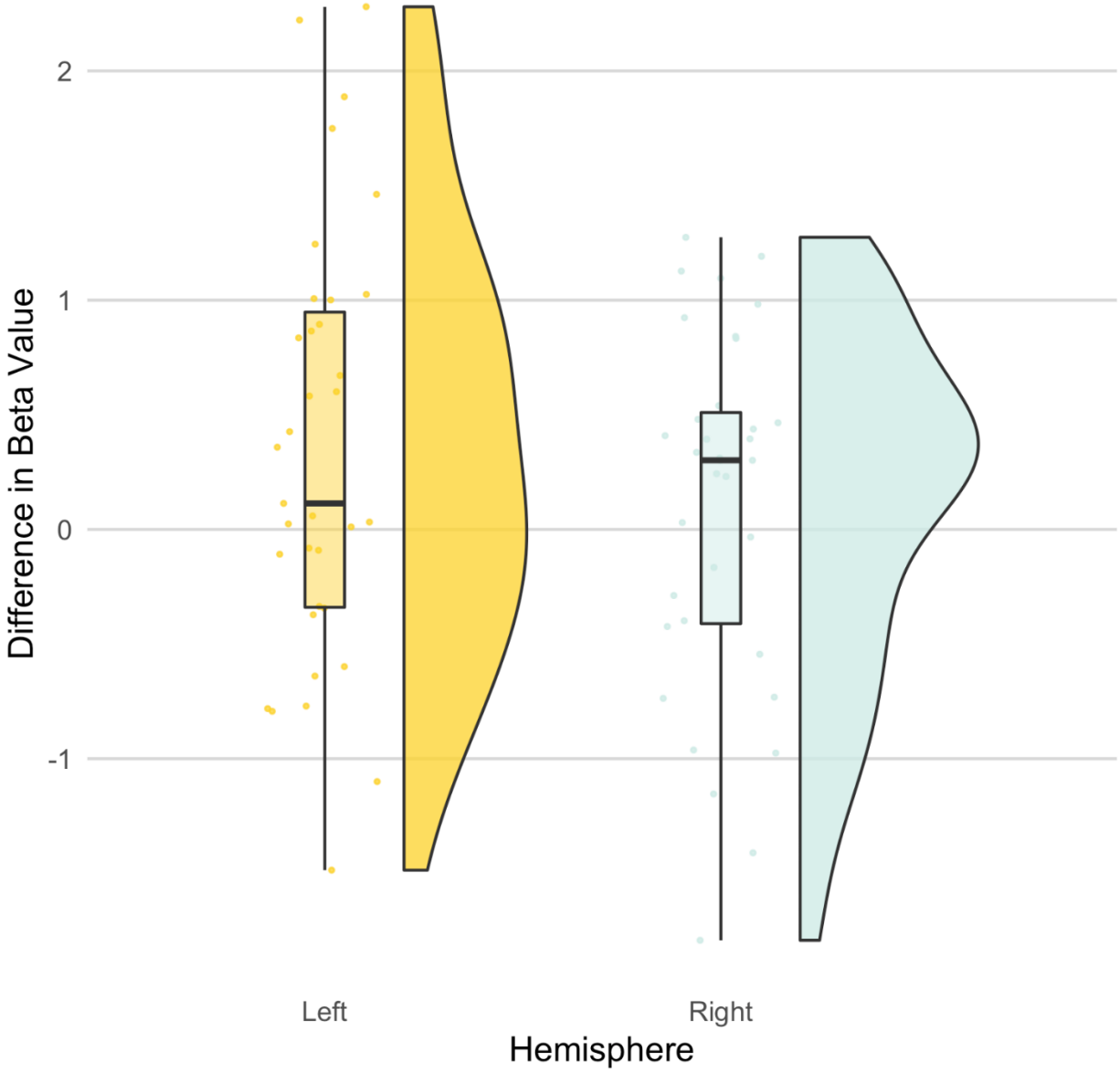

Figure 11. Difference in hippocampal activity on repeat items. Difference scores were calculated by subtracting the test trial β from the study trial β to reveal the amount of decrease in activity from study to test trials as a function of study trial fixations. The left hippocampus showed significant reduction in fMRI activity, but the right hippocampus did not.

DG/CA3) and outcome (LCR, LFA). Scores for each hemisphere were entered into separate ANOVAs due to the difference in significance observed when testing the whole hippocampus data. Outliers were fenced to 1.5 times the interquartile range above the 75th percentile or below

the 25th percentile. In the left hippocampus, the main effects of region, $F(1, 34) = 0.86$, $p = .36$, $\eta_p^2 = .03$, and outcome, $F(1, 34) = 0.02$, $p = .90$, $\eta_p^2 = .00$, were not significant. The interaction between the two was also not significant, $F(1, 34) = 0.10$, $p = 0.75$, $\eta_p^2 = .00$. In the right hippocampus, the main effects of region, $F(1, 34) = 0.01$, $p = .91$, $\eta_p^2 = .00$, and outcome, $F(1, 34) = 0.05$, $p = .82$, $\eta_p^2 = .00$, were not significant. The interaction between the two was also not significant, $F(1, 34) = 0.03$, $p = 0.86$, $\eta_p^2 = .00$. Figure 12 shows the difference scores divided by region, outcome, and hemisphere.

Discussion. Experiment 3 sought to examine the contribution of eye fixations to hippocampal activity as well as replicate the findings of the previous two experiments. While we replicated the finding of Liu et al. (2017) in the left hippocampus, experiment 3 failed to demonstrate a greater reduction in fMRI activity for lure false alarms than for lure correct rejections. More on this finding will be addressed in the general discussion. With regard to the eye tracking data, the hypothesis that the number of fixations at study would differ between lure false alarms and lure correct rejections was not supported. This finding, along with the failure of the number of fixations at study to predict accuracy on lure trials does not show support for the contribution of the number of fixations at study to lure false alarms. The eye movement data showed results closer to experiment 2 rather than experiment 1 even though the study design was closer to the design of experiment 1.

Three limitations should be addressed. First, the same pattern of few numbers of lure false alarms on low similarity trials showed up in experiment 3. We attempted to address this possibility in the study design by adding a fifth block of trials, but it was not enough. We therefore were not able to examine possible differences in activity between high and low similarity lure trials. Second, there was an abnormally large number of participants that needed

to be excluded for a variety of reasons. It is possible that with a greater number of participants, there would be increased statistical power in the analyses, which could have led to significant findings where there were not. However, the mixed support for the contributions of the number

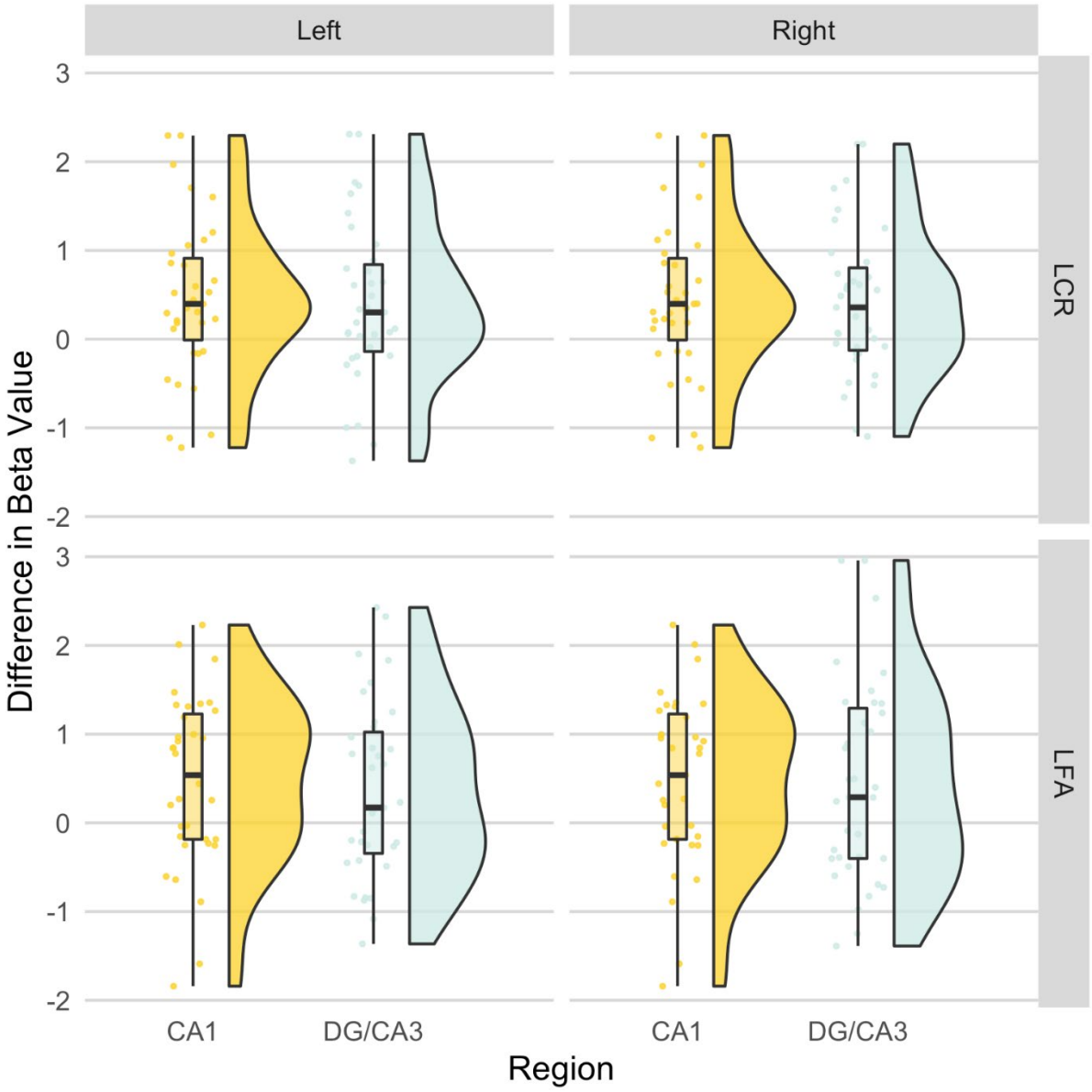

Figure 12. Difference in hippocampal subfield activity on lure items. Difference scores were calculated the same way as done for repeat items and compared across outcome for hippocampal subregions. All comparisons were not significant. LCR = Lure Correct Rejections, LFA = Lure False Alarms

of fixations at study to false alarms in experiments 1 and 2, which were both sufficiently powered, does not give confidence in this explanation. Third, there was also an abnormally large amount of motion in the fMRI scans that needed to be removed. This could have resulted in a reduction of the number of usable trials in each bin for the fMRI analysis, thus reducing reliability of the fMRI signal. Further research can address these limitations.

General Discussion

The current experiments sought to examine the relative contributions of fixations at study and test as well as target-lure similarity level to mnemonic discrimination. Experiment 1 expanded the study design of Molitor et al. (2014) by manipulating target-lure similarity level and adding two extra trial blocks. Experiment 2 sought to replicate the finding that fixations for lure correct rejections are higher than lure false alarms at both study and test in a more common method of administering a mnemonic discrimination task, namely allowing the participant to respond while the stimulus is presented rather than in a separate answer period. Last, experiment 3 included fMRI with the design of experiment 1 to examine the contribution of fixations at study to hippocampal activity. Table 7 shows the hypotheses tested and outcomes of each experiment.

The eye movement data show mixed results for the contribution of fixations at study and test in the accurate discrimination of lure items. First of all, when conducting the same repeated measures ANOVA that Molitor et al. (2014) used, we found differing results across experiments. Experiment 1 replicated the finding that fixations for subsequent lure correct rejections were greater than fixations for subsequent false alarms and also found the same pattern with lure correct rejections and lure false alarms. Experiments 2 and 3 did not find the interaction between presentation and response type that was expected.

It is possible that the lack of an interaction on experiment 2 was because of the change in task design, where participants responded to each item while the image was displayed rather than during a separate answer period. Fixation duration is decreased and saccade amplitude is increased when participants are asked to complete more specific goal oriented tasks than when they are given more general instructions or allowed to freely view an image (Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011). It is possible that the fixation data in experiments 1 and 2 differed simply because participants were responding at a different point in time. The requirement to respond before the image disappeared may have caused participants to change their eye movement patterns. Future research could design experiments that center around examining this possibility.

Table 7

Results of Hypotheses for Study 1

Hypothesis	Experiment 1	Experiment 2	Experiment 3
The number of fixations at study differ in the interaction between response types (LCR & LFA) and presentation	Supported	Not Supported	Not Supported
Visual complexity drives fixation counts in the task	Supported	Supported	Supported
The number of fixations at study differ between response types and presentation when visual complexity is accounted for	Supported	Not Supported	Not Supported
Target-lure similarity level is a better predictor of task performance than the number of fixations at study or test	Supported	Supported	Supported
There is a greater reduction in hippocampal fMRI activity as the number of fixations at study increases	NA	NA	Supported
The reduction in activity as a function of study fixation count will be greater for lure false alarms than for lure correct rejections	NA	NA	Not Supported

It is not clear why experiment 3 found differing results than experiment 1. The task design was the same with the exception of an extra block to allow for a greater number of trials needed for the fMRI analysis. The fact that participants were completing the task inside the MRI scanner may have altered the eye movement patterns due to the changed environment. There is a smaller number of participants in experiment 3 as compared to experiment 1, specifically about half as many. We did a power analysis for the first two experiments which revealed that we would need 60 participants for sufficient power in this analysis and it is possible that the analysis is simply underpowered and would find the same results had we not had the common constraints that fMRI presents on sample size (e.g., cost per scan and further restrictions on participant compatibility in the MRI environment). Alternatively, it is possible that the result is not reliable and does not hold up under replication. This last explanation seems most likely when comparing the mixed support for the contribution of the number of fixations at study across experiments 1 and 2, both of which had enough participants for sufficient statistical power. Future replications with the appropriate sample size are necessary to evaluate these two possible explanations in detail.

Fixation counts can be influenced by many variables, and another possible explanation of a greater number of fixations for subsequent lure correct rejections than for subsequent lure false alarms is differing levels of visual complexity. Fixation counts are greater when individuals are presented with stimuli that is more visually complex (Henderson, 2003). We tested this possible explanation by entering visual complexity scores along with the fixation data into a regression equation. All three experiments showed a significant effect of visual complexity, with greater visual complexity resulting in a greater number of fixations to the image. The addition of visual complexity to the model did not change any of the significant results that the experiments found

for effects of presentation, response type, or an interaction between the two, so it appears that significant differences between subsequent lure false alarms and subsequent lure correct rejections cannot be explained by a difference in visual complexity between the target and the lure.

We then looked at the relative contributions of fixations at study, fixations at test, and target-lure similarity level to accuracy on lure trials. Through all three experiments, target-lure similarity level strongly contributed to accurate responses on lure trials, with low-similarity lures many times more likely to be correctly answer than high-similarity lures. Fixations at study and fixations at test significantly contributed to the outcome in experiment 1, while experiment 2 only showed support for fixations at study and experiment 3 did not support a significant contribution in either.

There is support in other literature that the number of fixations at study can increase accuracy on memory tests. For example, the number of fixations at encoding are greater when individuals report greater memory strength at test (Kafkas & Montaldi, 2011). Hits on an immediate recognition memory test have a greater number of fixations than misses, though there is no difference between the two when testing is done after a 15 minute delay (Parag & Vakil, 2018). Further, memory impairment is demonstrated when fixation duration is constrained to 75 ms, but not at 150 ms, suggesting that there is a minimum amount of time required for fixations to gather information (Rayner, Smith, Malcolm, & Henderson, 2009). These experiments, however, do not use similar lures in their set of stimuli, so these findings may not generalize well to tests of memory specificity.

Other measures of eye movement may be more sensitive to the role that visual sampling during encoding or retrieval predicts mnemonic discrimination. Research into change blindness

shows that the location of fixations plays a particular role in discovering the change (Henderson & Hollingworth, 1999; Hollingworth, Schrock, & Henderson, 2001). Individuals must fixate upon the change at both study and test in order to recognize it. Further, individuals demonstrate similar fixation patterns at retrieval when asked to describe from memory while looking at a blank screen (Johansson & Johansson, 2014; Laeng, Bloem, D'Ascenzo, & Tommasi, 2014). Further, reaction times are longer when fixations are constrained to the center of a screen and not allowed to freely move to the area of the screen that the object was present (Johansson & Johansson, 2014; Laeng et al., 2014). When presented with images of scenes scan paths at retrieval that were more similar to the scan paths at encoding led to better accuracy on a memory test (Foulsham et al., 2012). Older adults that show a propensity to pattern completion have fixation patterns for new items at test that overlap more with the studied items they incorrectly identified the new items as (Vieweg et al., 2018). Future research could examine fixation location and scan paths to examine their contributions to false alarms.

When we look at the influence of fixations at study on hippocampal activity, we replicated the finding of greater reduction on repetitions as a function of study fixations in the left hippocampus but did not replicate it in the right hemisphere. We do not find support that fixations at study differentially affect hippocampal activity for lure correct rejections and lure false alarms in the DG/CA3 and CA1. The DG/CA3 shows greater activity for lure correct rejections than lure false alarms (Leal, Tighe, Jones, et al., 2014). The DG/CA3 also shows a greater amount of activity than CA1 when comparing lure correct rejections (Reagh & Yassa, 2014). The results of the current experiment show that the differences in fMRI activity demonstrated in these regions across trial outcomes is not driven by the number of fixations at study.

Target-lure similarity was the only predictor of accuracy that was significant across all three experiments. Such a finding supports the idea that pattern completion drives false alarms to lure stimuli. When an explicit answer is required on a mnemonic discrimination task, a "recall to reject" strategy is commonly thought to be used (Hunsaker & Kesner, 2013). Such a strategy requires pattern completion because an individual needs to use the image presented to them to recall the previously encountered image and compare them. If the overlap between the two images is too great, pattern completion causes an individual to indicate that the item is a repeat rather than a similar lure.

Behaviorally, this is observed in many uses of the Mnemonic Similarity Task with varying degrees of similarity. Performance is consistently worse for items that have a greater amount of overlap (Hannula et al., 2012; Kim & Yassa, 2013; Motley & Kirwan, 2012). Research into pattern separation deficits in aging show that older adults have a propensity to pattern completion, with greater rates of "old" responses to lure items than do young adults (Huffman & Stark, 2017; Reagh et al., 2016; Stark, Stevenson, Wu, Rutledge, & Stark, 2015). Vieweg et al. (2018) developed a task, the Memory Image Completion Task (MIC), that specifically taxes pattern completion, requiring participants to identify an object from a masked cue. Participants learned a set of images and then viewed images with varying amounts of it masked. They were required to identify the scene of the image or indicate that none of the choices present identified the scene if the image was completely new. Older adults showed an increased propensity to pattern completion than young adults, similar to those findings using a mnemonic discrimination task.

Further, a recent case study by Baker et al. (2016) directly compared the performance on the MST and the MIC in an individual, patient B.L., with damage to the hippocampus that is

largely confined to the DG. Patient B.L.'s performance on the MST was consistent with others who have hippocampal lesions (Kirwan et al., 2012) with decreased accuracy in discriminating similar lures. Interestingly, however, patient B.L. was not impaired on recognizing targets in the MIC but was impaired on foil trials that required participants to indicate that the item was not seen before. When calculating response bias on the MIC, patient B.L. showed a high level of bias toward responding that partial cues were images that had been seen before whereas controls stayed closer to levels that would indicate a lack of bias. When considering patient B.L.'s performance on the MST and MIC as well as the specificity of his damage to the hippocampus, there is great evidence that false alarms to the MST come about because of a lack of pattern separation with an increased propensity toward pattern completion.

Limitations for the individual experiments are noted in the separate discussion sections, but one limitation for the current study should be discussed further. There were few false alarms on low-similarity lure trials through all three experiments. Such a bias toward lure correct rejections on these trials may have biased the findings of our analyses. Such a limitation does point to the strong effect of lure pair similarity level, however, as participants did much better on the low-similarity lure trials than on high-similarity lure trials. Further the contributions of fixations at study and test to task performance were constrained to high-similarity lure trials, of which there were a good amount of false alarms. Still, in the case of being thorough, further research could alter the similarity level or stimulus presentation time to decrease performance slightly so that there is a better representation of false alarms in the low similarity trials.

In conclusion, we found some evidence for the contribution of poor encoding in mnemonic discrimination errors, though the effect did not replicate across all experiments. This difference in findings across experiments could be accounted for by the change in study design

in experiment 2 and the lack of power in experiment 3 but could also simply not be strong enough to replicate across multiple studies. Further research can address these issues. Fixations at study did not predict hippocampal activity differences commonly seen between lure correct rejections and lure false alarms. Overall target-lure similarity level was the greatest predictor of performance on the task, giving support for the contribution of pattern completion in false alarms to lure stimuli.

Study 2

To date, much of the research on pattern separation has focused on visual stimuli. Auditory pattern separation, as a result, is not as well understood in humans. The hippocampus receives input from all sensory modalities (Rolls, 1989), so it is reasonable that a common computational process such as pattern separation underlies memory specificity across sensory modalities. In rodents, pattern separation of olfactory and visual stimuli has been researched and these modalities present similar findings. For instance, research into olfactory and visual pattern separation at the sensory and perceptual level show that rodents are able to detect small differences in stimuli regardless of modality (Barnes, Hofacer, Zaman, Rennaker, & Wilson, 2008; Hunsaker, Rosenberg, et al., 2008). Rodents are able to correctly find a food reward among different odors even when the odor cues differ by a single carbon chain and accuracy increases with the amount of difference between the odors (Weeden et al., 2014).

Performance on visual tasks is similar with rodents performing quicker for stimuli that have greater differences than stimuli that are very similar (Gilbert & Kesner, 2003). While the rodent literature seems to suggest that mnemonic discrimination behavior is similar across modalities there is evidence that at a neural level pattern separation for different modalities depends on different brain structures such as the involvement of the olfactory bulb in olfactory

pattern separation and the perirhinal cortex in object pattern separation (Hunsaker & Kesner, 2013). For this reason, it is important not to generalize findings from one modality to another without direct testing.

Research focused on other forms of memory in humans has found differences for various modalities. For instance, Herz (2004) found that autobiographical memories cued by odors were rated by participants to be much more emotional than those memories cued by sounds or images. Research into serial position effects shows differences in memory for different modalities as well. For example, Johnson and Miles (2009) tested serial position effects for various types of stimuli and found that visual stimuli showed both a primacy and recency effect, auditory stimuli only showed a recency effect, while olfactory stimuli showed no serial position effect. Another study by Johnson, Volp, and Miles (2014) tested serial position effects on gustatory stimuli using wine. In this experiment, participants tasted a series of different wines and then were given a test wine and asked to recall the position of that wine in the sequence that they previously consumed. Researchers found that participants were better at correctly determining the position of the wine in the sequence when it was either the first or last wine presented.

Further, research focused on recognition memory shows that memory for auditory stimuli is impaired compared to visual stimuli. For example, Cohen, Horowitz, and Wolfe (2009) conducted a series of experiments in which they tested recognition memory in both auditory and visual domains. In all but one of the experiments, recognition memory to visual stimuli was superior to memory to auditory stimuli. The experiment where performance was equal required visual stimuli to be blurred almost to the point of being unrecognizable. A follow up to this study also showed that memory to auditory stimuli cannot reach levels of performance seen with visual stimuli even in highly specialized populations such as musicians, who have greater exposure to

auditory stimuli (Cohen, Evans, Horowitz, & Wolfe, 2011). Gloede, Paulauskas, & Gregg (2017) replicated previous findings by Cohen et al. (2009) as well as discovered that memory training for auditory memory did not improve performance to the level of visual memory.

While the research on recognition memory (Cohen et al., 2011, 2009; Gloede et al., 2017) measured discrimination, they did not use similar lures in their set of stimuli, only targets and foils. To our knowledge, only one study has tested mnemonic discrimination to similar auditory stimuli in humans. In this study, Trier, Lacy, and Marsh (2016) tested various aspects of auditory mnemonic discrimination of words. In the first experiment, participants were presented with a series of words and were asked to indicate whether each word was new, old, or similar. Old words consisted of the same word read by the same voice, while similar words consisted of the same word read by a different voice. The researchers found that accuracy was less for items that were read by a different voice rather than the same voice. In the second experiment, participants were asked to indicate whether the voice they heard was new, old, or similar. The voices that were presented read two different phrases, and the corresponding test trials had either the same phrase (i.e., old) or a different phrase (i.e., similar) read by the voice. Accuracy was impaired when repetition trials included a different phrase as compared to repetition trials that included the same phrase. These two experiments show that both the word itself or the auditory characteristics of the voice reading the word can play a part in how well an individual is able to accurately discriminate similar auditory stimuli.

We wished to directly compare mnemonic discrimination performance between auditory and visual stimuli while using similar lures. We first hypothesized that discrimination would be correlated within subjects between visual and auditory stimuli, giving evidence of common neuro-cognitive processes that influence mnemonic discrimination to both modalities. We further

expected that we would see similar patterns of performance to those previously demonstrated by research conducted on recognition memory (Cohen et al., 2011, 2009; Gloede et al., 2017). Specifically, we hypothesized that we would observe greater performance, as shown by higher d′ values, to visual stimuli than to auditory stimuli. In addition, we hypothesized that participants would be more confident of their answers on visual trials than on auditory trials.

Methods

Similarity ratings. We recruited 102 individuals (29 male, 73 female) from the Brigham Young University community to rate the similarity of the sounds within each lure pair. Participants had a mean age of 21.21 years (standard deviation, *SD* = 4.89 years) with a mean education level of 13.89 years (*SD* = 1.36) and were required to have normal or corrected-to-normal hearing. All participants were compensated with course credit for participating in the study. All procedures were approved by the Brigham Young University Institutional Review Board.

Each participant completed an online questionnaire using Qualtrics software (Qualtrics, Provo, Utah) on their own computers in which they were presented with sound pairs. The software presented two audio clips on the screen that participants could click to play (see Figure 13 for an example of a single trial). Participants were asked to rate the degree of similarity on a six-point scale with 1 being least similar and 6 being most similar. The task was self-paced, and participants were allowed to freely play each sound clip as needed before making their response.

The sample of pairs that each participant received was pseudo-randomized so that participants rated half of the total 235 pairs and each pair would have a roughly equal number of ratings across participants. The mean number of observations for each pair was 51.43 with a minimum of 48 and a maximum of 58. For the final stimulus set, we selected 50 pairs with the

lowest similarity ratings and 50 pairs with the highest similarity ratings as lure trials. All other sounds were used as repeat trials in the task.

Participants. A total of 173 participants (79 male, 94 female) were recruited from Brigham Young University for the experiment. Participants were required to have no history of traumatic brain injury or diagnosis of a psychiatric or neurological condition as well as normal hearing. All participants were compensated with course credit for their participation in the experiment. All study protocols were approved by the Brigham Young University Institutional

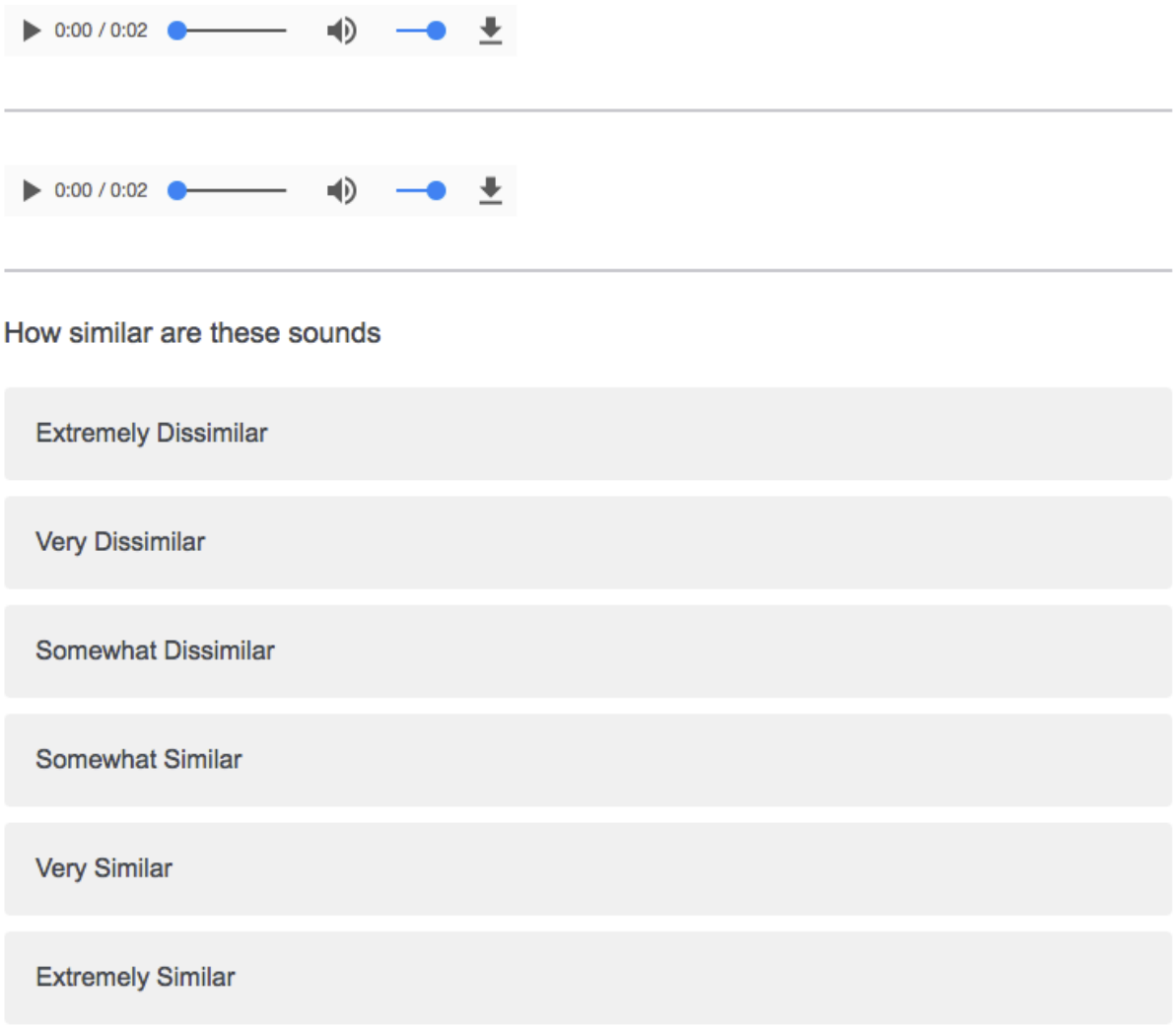

Figure 13. Sample trial from similarity rating questionnaire. Participants were allowed to freely play the sounds of the two items in the pair and then were asked to rate the level of similarity on a 6-point scale.

Review Board and all participants gave written informed consent prior to participating. Following data collection, we excluded data for 27 participants from our analyses. Six participants were excluded due to failure to follow task instructions (e.g., mobile phone use during the task or failure to use all answer options). Another three participants were removed due to technical problems during data collection (e.g., a testing computer restarting in the middle of the task or computer monitors becoming unplugged). Two participants fell asleep during data collection. Two participants chose not to complete the study. Lastly, 14 participants were removed because of a high number of trials with no responses. This threshold for exclusion based on non-responses was set at 1.5 times the interquartile range above the 3rd quartile. After exclusions, a total of 146 participants (63 male, 83 female) were included in our data analysis. Participants had an average age of 20.53 years (*SD* = 2.01).

Apparatus. The task was presented using PsychoPy software (version 1.85.6; Peirce, 2007, 2008) on Dell PCs running Windows 10. Monitors were set to a resolution of 1920 × 1080 pixels. Participants responded by pressing the 'v', 'b', 'n', and 'm' keys on the keyboard.

Task. The task was a variation of the visual study-test Mnemonic Similarity Task (Kirwan & Stark, 2007; Stark, Yassa, Lacy, & Stark, 2013; Yassa et al., 2011). Figure 14 gives a visual representation of the task. Each study-test block consisted of 40 trials in the study portion and 40 trials in the test portion and stimuli were randomly distributed across blocks. The trials in each portion were split equally so that 20 trials were repeat and 20 trials were lure. Repeat stimuli were identical in both the study and test portions of the task. Lure stimuli were two variations of the same theme (e.g. knocking on a door), one presented in the study portion and the other presented in the test portion. The lures were further divided evenly between similarity level so that 10 trials were high similarity lures and 10 trials were low similarity lures. Stimuli

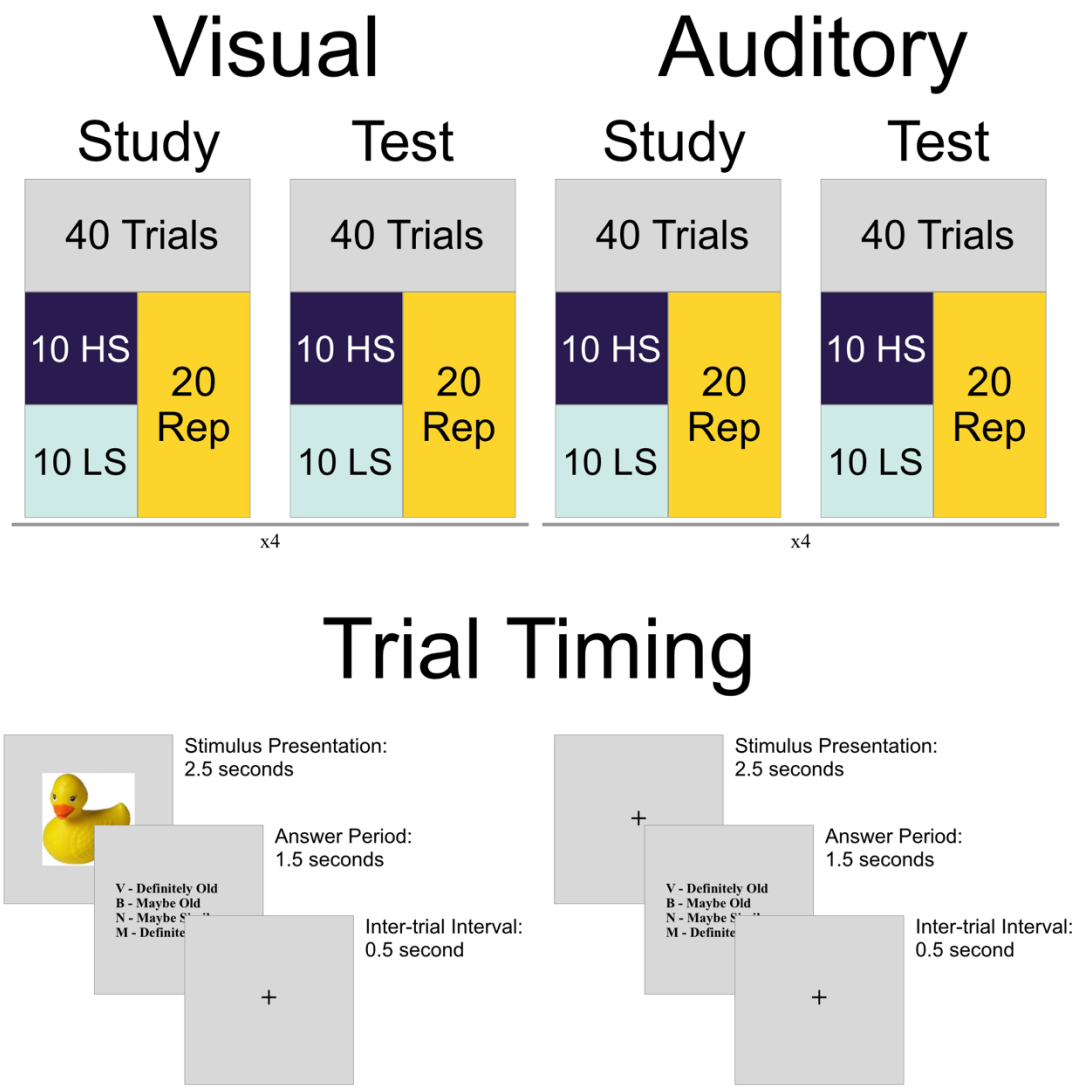

Figure 14. Protocol for task used in study 2. Participants were presented a single stimulus at a time on the computer for 2.5 seconds. On study blocks, participants indicated whether the item is generally found indoors or outdoors. On test blocks, participants responded whether the item was something they encountered previously in the experiment as well as their confidence in their answer.

alternated between sounds and images across each study-test block and the type of stimulus for the first block was counterbalanced so that half of the participants saw images first and half of the participants heard sounds first. Each test portion of a block occurred directly after the study portion of a block.

For both study and test portions, each trial started with a fixation cross presented for 0.5 seconds, after which the stimulus was presented for 2.5 seconds. During stimulus presentation for image blocks, the image was presented on the center of the screen. During stimulus presentation for sound blocks, a fixation cross was centered on the screen and the sound was played to the participants through headphones.

Upon completion of the presentation of the stimulus, answer options appeared on the screen for 1.5 seconds, during which participants could give their answer. For study portions, participants indicated whether the object or sound presented is generally encountered indoors or outdoors. For test portions, participants indicated whether the stimulus was identical to what was presented during the study portion (i.e., "old" response) or if the stimulus was similar, but not identical to what was presented during the study portion (i.e., "similar" response). At the same time, participants also rated their confidence in their answer by indicating that they were definitely sure or maybe sure. Accordingly, the four answer options were definitely old, maybe old, maybe similar, and definitely similar.

Procedure. Data were collected in a computer lab that allowed simultaneous data collection from multiple participants. Upon arriving at the testing location, participants were greeted by a research assistant who obtained their written, informed consent to participate. Participants were then assigned to a computer. Once all the participants for the session had arrived, the researcher instructed participants to check the volume level on the headphones by playing a sample sound and adjusting the volume to a comfortable level.

Participants were instructed to watch an instruction video for the first study portion of first block. The research assistant then asked if anyone had questions upon completion of the video. After questions were answered, participants completed the study portion of the first block.

This process was repeated so that participants received instructions on study and test portions for both sounds and images (i.e., 4 instruction videos total). Once participants completed the second test portion, they were instructed to continue the experiment until they finished.

Data processing and cleaning. We examined and cleaned the data prior to analysis using several steps. First, as previously mentioned, the data were examined for high rates of non-responses. The threshold was set at 1.5 times the interquartile range above the third quartile. Any participants who had a greater number of non-responses than this were excluded from analysis. Second, we checked the number of non-responses for each stimulus. The highest number of non-responses to any single item was 13, which was less than 10% of the total possible (n = 147). Accordingly, all items were left in for analysis.

Test trials were classified as hits, misses, correct rejections, or false alarms. Hits consisted of repeat trials that were correctly identified as “old”. Misses were repeat trials that were incorrectly classified as “similar”. Correct rejections were lure trials that were correctly identified as “similar”. Lastly, false alarms were lure trials that were incorrectly classified as “old”. Study trials were coded as eliciting a response or not and test trials were only included in the analysis if the corresponding study trial elicited a response. We calculated d' scores by subtracting the z-score of the false alarm rate (i.e., number of false alarms divided by the total number of lure trials) from the z-score of the hit rate (i.e., the number of hits divided by the total number of repeat trials).

Results

Participants responded to 97.19% of the trials and response rates were similar between visual and auditory stimuli (i.e., 97.49% and 96.89%, respectively). See Table 8 for a summary of the proportion of responses by stimulus type, trial type, and response. Discrimination on both

the visual, $t(145) = 50.00$, $p < .001$, 95% CI[1.98, 2.15], and auditory blocks, $t(145) = 38.89$, $p < .001$, 95% CI [0.86, 0.96], was significantly better than chance, with d′ values of 2.07 (*SD* = 0.50) and 0.91 (*SD* = 0.32) on visual and auditory blocks, respectively.

Table 8

Proportion of Responses by Stimulus Type, Trial Type, and Response

Stimulus Type	Trial Type	Response	Mean Proportion	*SD* Proportion
Image	Lure	Old	0.33	0.11
Image	Lure	Similar	0.67	0.11
Image	Repeat	Old	0.93	0.06
Image	Repeat	Similar	0.08	0.05
Sound	Lure	Old	0.45	0.10
Sound	Lure	Similar	0.55	0.10
Sound	Repeat	Old	0.77	0.09
Sound	Repeat	Similar	0.23	0.09

Note. SD = standard deviation, Lure = similar items, Repeat = items presented twice

We tested the correlation between d′ for visual stimuli and d′ for auditory stimuli to test the hypothesis that discrimination of visual and auditory stimuli relies on common neuro-cognitive processes. The data were not normally distributed and attempts to apply a transformation did not successfully normalize the data. For this reason, we used a Spearman Rank correlation which showed that performance on visual and auditory blocks were significantly correlated, $r = 0.35$, $p < 0.001$ (see Figure 15 for a visual display of the correlation), thus supporting our hypothesis and indicating that auditory and visual mnemonic discrimination rely on common neuro-cognitive processes.

Next, to test the hypothesis that we would observe greater performance to visual stimuli than to auditory stimuli, we compared performance between auditory and visual blocks by

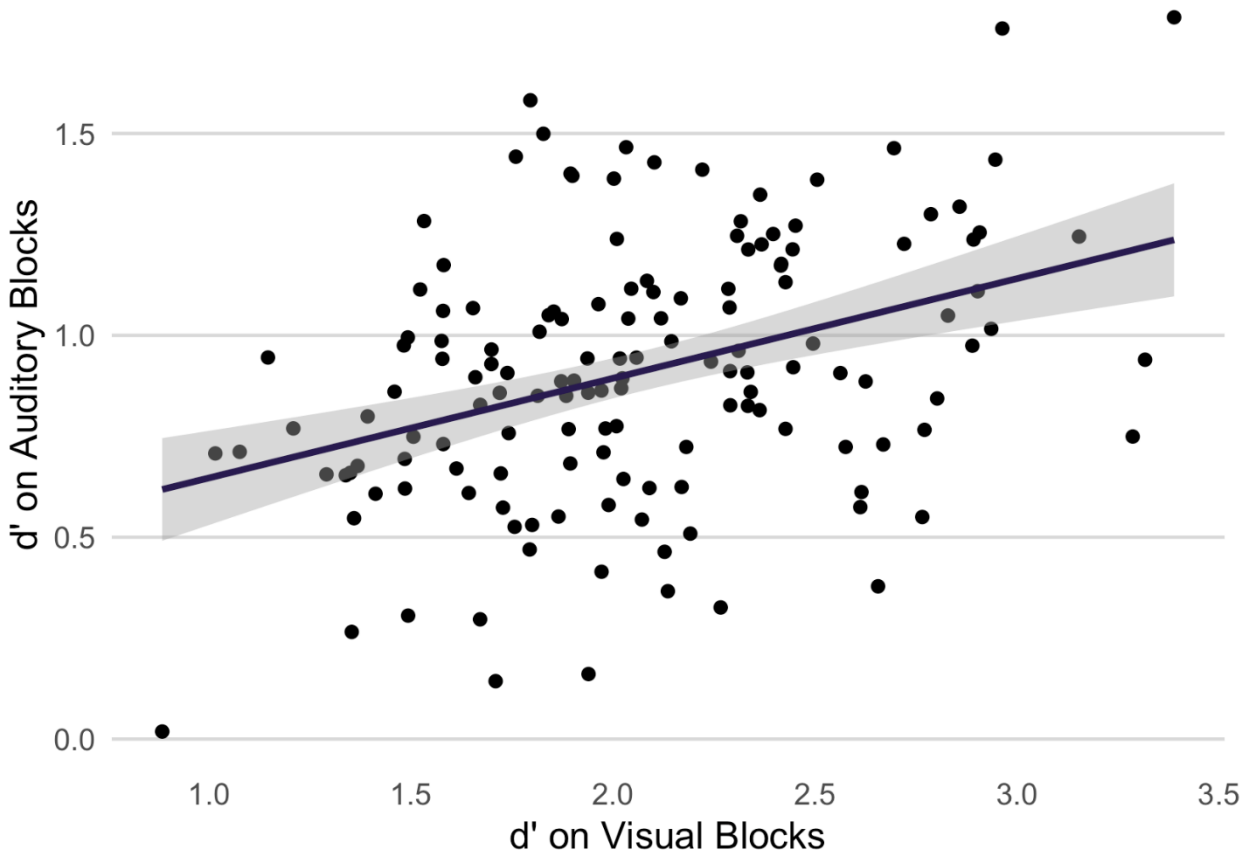

Figure 15. Correlation of visual d′ values and auditory d′ values.

entering discrimination scores into a linear mixed-effects model with fixed effects of stimulus type (visual and auditory) and similarity level (high and low). Subject was entered as a random intercept and stimulus type and similarity level were entered as random slopes.

Overall, we observed that there was greater performance on visual trials than on auditory trials, thus supporting our hypothesis. Sound high-similarity trials were correctly discriminated at a lower rate than image high-similarity trials, β = -0.76, $t(235.9) = -14.67$, $p < .001$, 95% CI [-0.86, -0.66]. On the other hand, image low-similarity trials were discriminated at a much higher rate than image high-similarity trials, β = 1.86, $t(289.33) = 50.99$, $p < .001$, 95% CI [1.79, 1.93]. The interaction term was also significant, with accuracy on sound low-similarity trials β = -1.18, $t(290) = -23.66$, $p < .001$, 95% CI [-1.28, -1.09] at a lower rate than image high-similarity trials.

The assumptions of linearity, absence of collinearity, homoskedasticity, and normality of residuals were all met. Influential data points were checked using Cook's distance and the maximum value returned was 0.03. Figure 16 displays d′ values for each similarity level of each stimulus modality.

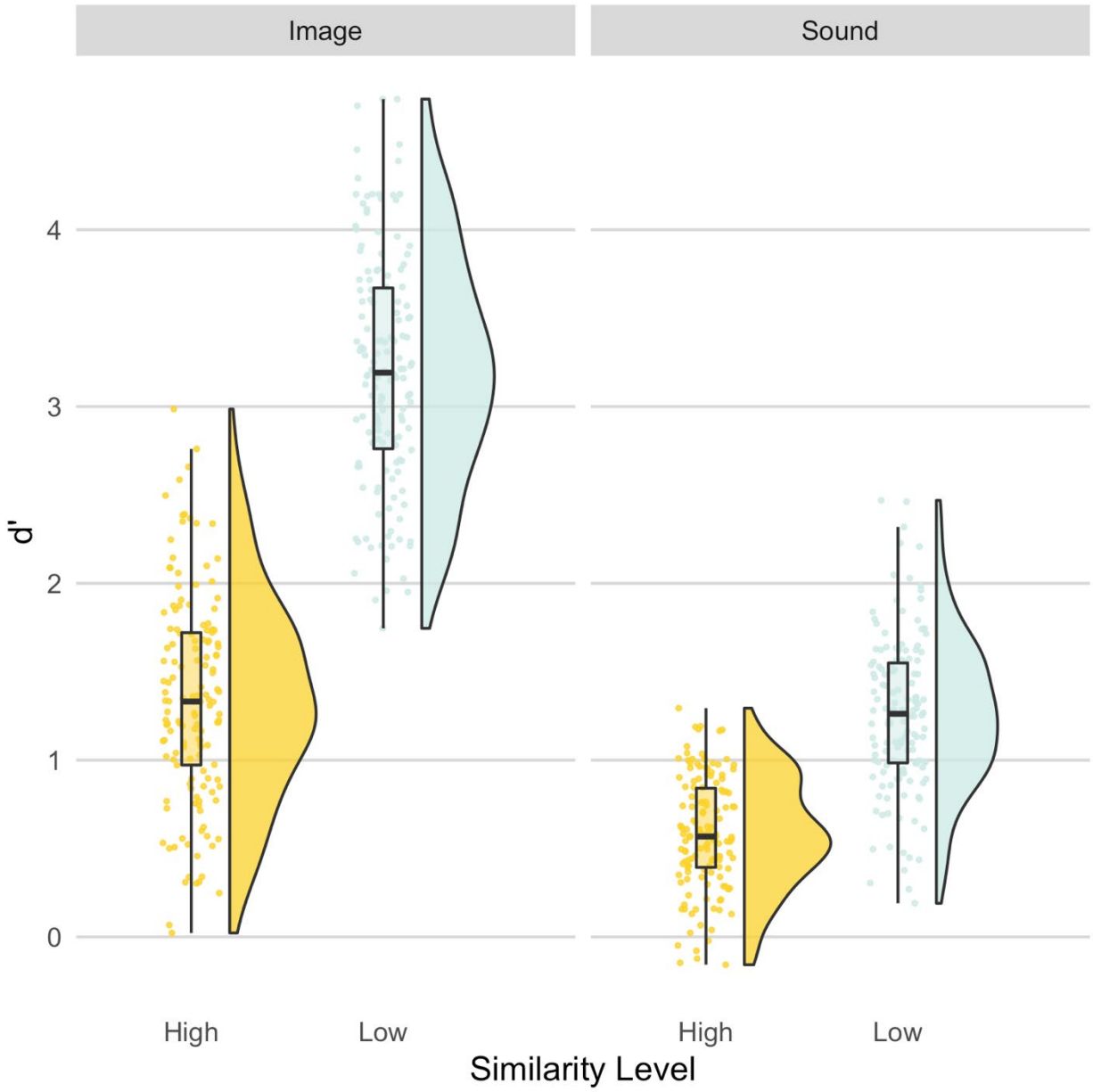

Figure 16. d′ scores divided by similarity level and stimulus type.

To test the hypothesis that participants would be more confident for visual stimuli than auditory stimuli, we entered data into a logistic mixed effects model with confidence level (i.e., 1 = High, 0 = Low) as the dependent variable. Stimulus type (i.e., visual and auditory) and response type (i.e., correct rejection and hit) were entered as fixed effects. Stimulus type and response type were also entered as random slopes and subject was entered as a random intercept.

The model failed to converge, so the random slope for response type was removed because it accounted for less of the variance than stimulus type. On the converged model, proportions of high confidence for sound correct rejections, $\beta = -1.87$, $p < .001$, 95% CI [-0.35, 0.15], OR = 0.15, were significantly lower than proportions of high confidence for image correct rejections. Alternatively, confidence for image hits was significantly greater than image correct rejections, $\beta = 0.61$, $p < .001$, 95% CI [0.43, 0.62], OR = 1.84. The interaction term was also significant, with confidence for sound hits being significantly lower than image correct rejections but higher than sound correct rejections, $\beta = 0.54$, $p < .001$, 95% CI [1.91, 2.23], OR = 0.49. A confusion matrix was calculated to test model fit and 81.79% of the trials were correctly predicted. Figure 17 displays the proportion of high confidence responses as a function of stimulus type and response type. Influential subjects were tested using Cook's distance and the maximum value returned was 0.06.

Discussion

The current study sought to examine potential differences in mnemonic discrimination between visual and auditory sensory modalities. Participants completed the Mnemonic Similarity Task in which blocks alternated between visual and auditory stimuli. Performance on auditory and visual blocks were correlated, but participants were better at discriminating visual stimuli

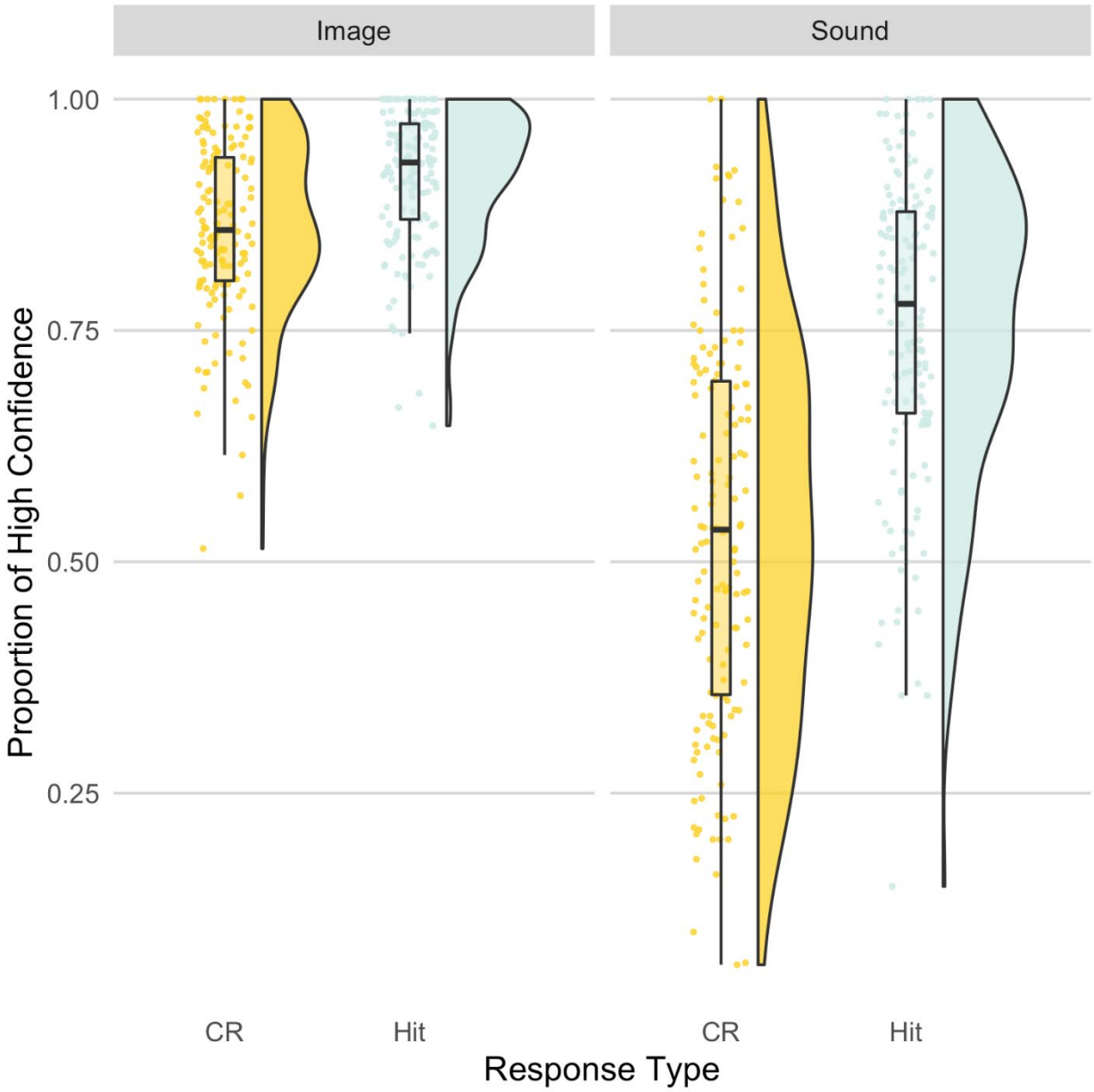

Figure 17. Proportions of high-confidence divided by stimulus modality and similarity level.

than they were at discriminating auditory stimuli. Participants were also more confident of their responses to visual stimuli than to auditory stimuli. The significant correlation between performance on visual and auditory blocks indicate that the discrimination of both types of stimuli share some neuro-cognitive resources. Such a finding is to be expected, since the hippocampus receives input from all sensory modalities (Rolls, 1989).

The finding that participants perform better for visual stimuli than for auditory stimuli is consistent with previous research comparing the sensory modalities in recognition memory (Cohen et al., 2011, 2009; Gloede et al., 2017). A notable addition to these studies that the current study offers is the control of similarity level among lures. When performance is compared between similarity levels, we find poorer performance for high-similarity lures as compared to low-similarity lures, a finding common in the pattern separation literature. The interaction between similarity level and sensory modality is of greater interest, however. There is a greater difference in performance with high- and low-similarity lures for visual stimuli than there is for auditory stimuli. Such an effect appears to be primarily driven by the performance on the low-similarity lures of the different modalities. For the visual stimuli, performance on low-similarity lures was very high with a mean discrimination score of 3.20 (SD = 0.66), while performance on auditory low-similarity lures was much lower at 1.26 (SD = 0.43). The difference in performance between visual high-similarity lures (M = 1.35, SE = 0.59) and auditory high-similarity lures (M = 0.59, SE = 0.31) was not as extreme.

The confidence ratings track very well with accuracy rates, consistent with previous studies of recognition memory (Wixted, 2007). Participants were much more confident for hits than correct rejections, which followed their accuracy rates and is also consistent with other research examining confidence to mnemonic discrimination. The interaction between stimulus modality and response type for confidence ratings also followed the same pattern as the interaction between stimulus modality and similarity level for accuracy. There was not as much of a difference between the confidence ratings of hits across modalities as there was between confidence ratings of correct rejections across modalities.

One limitation in this study should be noted. Participants did exceptionally well on the visual low-similarity lure trials, in which 25 participants correctly answered 100% of the trials presented to them. None of the participants responded with 100% accuracy on visual high-similarity lure trials or either similarity level of the auditory lure trials. With some participants at ceiling performance, it is difficult to accurately judge the full magnitude of the difference between discrimination ability on visual and auditory stimuli. Future research can address this by varying the difficulty of the task by adjusting other aspects, such as presentation time or response time allowed.

With the development of a task that directly compares mnemonic discrimination between auditory and visual stimuli, a natural next step would be to compare the involvement of various regions of the brain for each sensory modality. There are differing opinions about how the hippocampus performs pattern separation across different modalities. Yassa and Stark (2011) suggest that the pattern separation process in the hippocampus is likely to be stimulus independent. Thus, the hippocampus would orthogonalize stimuli the same way, regardless of whether or not the stimuli are visual or auditory.

Another view of how the hippocampus orthogonalizes stimuli is that different forms of pattern separation rely on different subregions of the hippocampus. This idea is illustrated by Hunsaker and Kesner (2013) where they summarize much of the research that has been conducted on different forms of pattern separation, such as spatial and temporal pattern separation. Specifically, research into temporal pattern separation implicates the involvement of CA1 (Kesner et al., 2010). Alternatively, when looking at spatial pattern separation, the dentate gyrus, rather than CA1, is an important structure and lesions decrease pattern separation efficiency (Gilbert et al., 2001). The differences here illustrate that various forms of pattern

separation are reliant on different subregions of the hippocampus and it is therefore likely that differences in subregion activity would be observed when examining pattern separation across different sensory modalities. Further research could use functional magnetic resonance imaging (fMRI) to examine activity in hippocampal subfields, particularly the CA1, CA3, and DG, to test whether activity differs for mnemonic discrimination across sensory modalities.

Conclusion

The two studies of the current project sought to examine the contribution of target-lure similarity and sensory modality to false alarms to similar lures. Study 1 consisted of three experiments that used eye tracking and fMRI to tease apart the relative contributions of encoding, retrieval, and target-lure similarity level to understand how pattern completion contributes to false alarms. The fixation data showed mixed support for the role of fixations at study to the accurate response to lures, but target-lure similarity level proved to be a robust predictor of accuracy. Further, the number of fixations at study did not differentiate between hippocampal activity for lure correct rejections and lure false alarms. The results of these experiments give support to the contribution of pattern completion to false alarms to similar lures rather than poor encoding.

Study 2 used a behavioral task to examine the contribution of sensory modality to false alarms to lure stimuli. Mnemonic discrimination was measured for visual and auditory stimuli and performance was significantly better on visual stimuli, even when target-lure similarity level was taken into account. Participants were also much more confident of their answers to visual stimuli. Results of this study give support for varying levels of accuracy across sensory modalities and further directions are discussed.

References

Anderson, M. L., James, J. R., & Kirwan, C. B. (2017). An event-related potential investigation of pattern separation and pattern completion processes. *Cognitive Neuroscience*, *8*(1), 9–23.

Azab, M., Stark, S. M., & Stark, C. E. L. (2014). Contributions of human hippocampal subfields to spatial and temporal pattern separation. *Hippocampus*, *24*(3), 293–302.

Baker, S., Vieweg, P., Gao, F., Gilboa, A., Wolbers, T., Black, S. E., & Rosenbaum, R. S. (2016). The human dentate gyrus plays a necessary role in discriminating new memories. *Current Biology: CB*, *26*(19), 2629–2634.

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*, *319*(5870), 1640–1642.

Barnes, D. C., Hofacer, R. D., Zaman, A. R., Rennaker, R. L., & Wilson, D. A. (2008). Olfactory perceptual stability and discrimination. *Nature Neuroscience*, *11*(12), 1378–1380.

Berron, D., Schütze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Düzel, E. (2016). Strong evidence for pattern separation in human dentate gyrus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(29), 7569–7579.

Bonhage, C. E., Mueller, J. L., Friederici, A. D., & Fiebach, C. J. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *68*, 33–47.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164–169.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, *14*(5), 365–376.

Clelland, C. D., Choi, M., Romberg, C., Clemenson, G. D., Jr, Fragniere, A., Tyers, P., … Bussey, T. J. (2009). A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science*, *325*(5937), 210–213.

Cohen, M. A., Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychonomic Bulletin & Review*, *18*(3), 586–591.

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(14), 6008–6010.

Costa, V. C. I., Bueno, J. L. O., & Xavier, G. F. (2005). Dentate gyrus-selective colchicine lesion and performance in temporal and spatial tasks. *Behavioural Brain Research*, *160*(2), 286–303.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162-173.

Creer, D. J., Romberg, C., Saksida, L. M., van Praag, H., & Bussey, T. J. (2010). Running enhances spatial pattern separation in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 2367–2372.

Dees, R. L., & Kesner, R. P. (2013). The role of the dorsal dentate gyrus in object and object-context recognition. *Neurobiology of Learning and Memory*, *106*, 112–117.

Doxey, C. R., Hodges, C. B., Bodily, T. A., Muncy, N. M., & Kirwan, C. B. (2018). The effects of sleep on the neural correlates of pattern separation. *Hippocampus*, *28*(2), 108–120.

Doxey, C. R., & Kirwan, C. B. (2015). Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. *Hippocampus*, *25*(4), 524–533.

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion*, *11*(4), 860–865.

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*(5), 458–462.

Foulsham, T., Dewhurst, R., Nyström, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach. *Journal of Eye Movement Research*, *5*(4). Retrieved from https://bop.unibe.ch/JEMR/article/view/2341

Gilbert, P. E., & Brushfield, A. M. (2009). The role of the CA3 hippocampal subregion in spatial memory: A process oriented behavioral assessment. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *33*(5), 774–781.

Gilbert, P. E., & Kesner, R. P. (2003). Recognition memory for complex visual discriminations is influenced by stimulus interference in rodents with perirhinal cortex damage. *Learning & Memory* , *10*(6), 525–530.

Gilbert, P. E., Kesner, R. P., & Lee, I. (2001). Dissociating hippocampal subregions: Double dissociation between dentate gyrus and CA1. *Hippocampus*, *11*(6), 626–636.

Gloede, M. E., Paulauskas, E. E., & Gregg, M. K. (2017). Experience and information loss in auditory and visual memory. *Quarterly Journal of Experimental Psychology*, *70*(7), 1344–1352.

Goodrich-Hunsaker, N. J., Hunsaker, M. R., & Kesner, R. P. (2008). The interactions and dissociations of the dorsal hippocampus subregions: How the dentate gyrus, CA3, and CA1 process spatial information. *Behavioral Neuroscience*, *122*(1), 16–26.

Hannula, D. E., Baym, C. L., Warren, D. E., & Cohen, N. J. (2012). The eyes know: Eye movements as a veridical index of memory. *Psychological Science*, *23*(3), 278–287.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*(5), 438–443.

Herz, R. S. (2004). A naturalistic analysis of autobiographical memories triggered by olfactory visual and auditory stimuli. *Chemical Senses*, *29*(3), 217–224.

Hollingworth, A., Schrock, G., & Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, *29*(2), 296–304.

Horry, R., & Brewer, N. (2016). How target-lure similarity shapes confidence judgments in multiple-alternative decision tasks. *Journal of Experimental Psychology. General*, *145*(12), 1615.

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). *Functional magnetic resonance imaging* (3rd ed.). Oxford, UK: Oxford University Press.

Huffman, D. J., & Stark, C. E. L. (2017). Age-related impairment on a forced-choice version of the Mnemonic Similarity Task. *Behavioral Neuroscience*, *131*(1), 55–67.

Hunsaker, M. R., Fieldsted, P. M., Rosenberg, J. S., & Kesner, R. P. (2008). Dissociating the roles of dorsal and ventral CA1 for the temporal processing of spatial locations, visual objects, and odors. *Behavioral Neuroscience*, *122*(3), 643–650.

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience and Biobehavioral Reviews*, *37*(1), 36–58.

Hunsaker, M. R., Rosenberg, J. S., & Kesner, R. P. (2008). The role of the dentate gyrus, CA3a,b, and CA3c for detecting spatial and environmental novelty. *Hippocampus*, *18*(10), 1064–1073.

Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., … Alzheimer's Disease Neuroimaging Initiative. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *NeuroImage*, *115*, 117–137.

Insausti, R., Annese, J., Amaral, D. G., & Squire, L. R. (2013). Human amnesia and the medial temporal lobe illuminated by neuropsychological and neurohistological findings for patient EP. *Proceedings of the National Academy of Sciences*, 201306244.

Jeneson, A., Kirwan, C. B., Hopkins, R. O., Wixted, J. T., & Squire, L. R. (2010). Recognition memory and the hippocampus: A test of the hippocampal contribution to recollection and familiarity. *Learning & Memory* , *17*(1), 63–70.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25*(1), 236–242.

Johnson, A. J., & Miles, C. (2009). Single-probe serial position recall: Evidence of modularity for olfactory, visual, and auditory short-term memory. *Quarterly Journal of Experimental Psychology* , *62*(2), 267–275.

Johnson, A. J., Volp, A., & Miles, C. (2014). Immediate recognition memory for wine. *Journal of Cognitive Psychology* , *26*(2), 127–134.

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology* , *64*(10), 1971–1989.

Kannurpatti, S. S., Motes, M. A., Rypma, B., & Biswal, B. B. (2010). Neural and vascular variability and the fMRI-BOLD response in normal aging. *Magnetic Resonance Imaging*, *28*(4), 466–476.

Kesner, R. P. (2013). A process analysis of the CA3 subregion of the hippocampus. *Frontiers in Cellular Neuroscience*, *7*, 78.

Kesner, R. P., Hunsaker, M. R., & Ziegler, W. (2010). The role of the dorsal CA1 and ventral CA1 in memory for the temporal order of a sequence of odors. *Neurobiology of Learning and Memory*, *93*(1), 111–116.

Kesner, R. P., & Rolls, E. T. (2015). A computational theory of hippocampal function, and tests of the theory: New developments. *Neuroscience and Biobehavioral Reviews*, *48*, 92–147.

Kesner, R. P., Taylor, J. O., Hoge, J., & Andy, F. (2015). Role of the dentate gyrus in mediating object-spatial configuration recognition. *Neurobiology of Learning and Memory*, *118*, 42–48.

Kim, J., & Yassa, M. A. (2013). Assessing recollection and familiarity of similar lures in a behavioral pattern separation task. *Hippocampus*, *23*(4), 287–294.

Kirwan, C. B., Gilbert, P. E., & Kesner, R. P. (2005). The role of the hippocampus in the retrieval of a spatial location. *Neurobiology of Learning and Memory*, *83*(1), 65–71.

Kirwan, C. B., Hartshorn, A., Stark, S. M., Goodrich-Hunsaker, N. J., Hopkins, R. O., & Stark, C. E. L. (2012). Pattern separation deficits following damage to the hippocampus. *Neuropsychologia*, *50*(10), 2408–2414.

Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*. Retrieved from http://learnmem.cshlp.org/content/14/9/625.short

Knierim, J. J., & Neunuebel, J. P. (2016). Tracking the flow of hippocampal computation: Pattern separation, pattern completion, and attractor dynamics. *Neurobiology of Learning and Memory*, *129*, 38–49.

Kovács, G., Kaiser, D., Kaliukhovich, D. A., Vidnyánszky, Z., & Vogels, R. (2013). Repetition probability does not affect fMRI repetition suppression for objects. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(23), 9805–9812.

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. L. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory* , *18*(1), 15–18.

Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, *131*(2), 263–283.

Lassalle, J. M., Bataille, T., & Halley, H. (2000). Reversible inactivation of the hippocampal mossy fiber synapses in mice impairs spatial learning, but neither consolidation nor memory retrieval, in the Morris navigation task. *Neurobiology of Learning and Memory*, *73*(3), 243–257.

Leal, S. L., Tighe, S. K., Jones, C. K., & Yassa, M. A. (2014). Pattern separation of emotional information in hippocampal dentate and CA3. *Hippocampus*, *24*(9), 1146–1155.

Leal, S. L., Tighe, S. K., & Yassa, M. A. (2014). Asymmetric effects of emotion on mnemonic interference. *Neurobiology of Learning and Memory*, *111*, 41–48.

Lee, I., Hunsaker, M. R., & Kesner, R. P. (2005). The role of hippocampal subregions in detecting spatial novelty. *Behavioral Neuroscience*, *119*(1), 145–153.

Lee, I., Jerman, T. S., & Kesner, R. P. (2005). Disruption of delayed memory for a sequence of spatial locations following CA1- or CA3-lesions of the dorsal hippocampus. *Neurobiology of Learning and Memory*, *84*(2), 138–147.

Lee, I., & Kesner, R. P. (2004). Encoding versus retrieval of spatial memory: Double dissociation between the dentate gyrus and the perforant path inputs into CA3 in the dorsal hippocampus. *Hippocampus*, *14*(1), 66–76.

Lee, I., Yoganarasimha, D., Rao, G., & Knierim, J. J. (2004). Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature*, *430*(6998), 456–459.

Leutgeb, J. K., Leutgeb, S., Moser, M. B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science*, *315*(5814), 961–966.

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., & Moser, M. B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science*, *309*(5734), 619–623.

Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M. B., & Moser, E. I. (2004). Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science*, *305*(5688), 1295–1298.

Liu, Z. X., Shen, K., Olsen, R. K., & Ryan, J. D. (2017). Visual sampling predicts hippocampal activity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(3), 599–609.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

McHugh, T. J., Jones, M. W., Quinn, J. J., Balthasar, N., Coppari, R., Elmquist, J. K., … Tonegawa, S. (2007). Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science*, *317*(5834), 94–99.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, *11*(8), 17.

Molitor, R. J., Ko, P. C., Hussey, E. P., & Ally, B. A. (2014). Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus*, *24*(6), 666–672.

Morris, A. M., Churchwell, J. C., Kesner, R. P., & Gilbert, P. E. (2012). Selective lesions of the dentate gyrus produce disruptions in place learning for adjacent spatial locations. *Neurobiology of Learning and Memory*, *97*(3), 326–331.

Motley, S. E., & Kirwan, C. B. (2012). A parametric investigation of pattern separation processes in the medial temporal lobe. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(38), 13076–13085.

Nakashiba, T., Cushman, J. D., Pelkey, K. A., Renaudineau, S., Buhl, D. L., McHugh, T. J., … Tonegawa, S. (2012). Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. *Cell*, *149*(1), 188–201.

Neunuebel, J. P., & Knierim, J. J. (2014). CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron*, *81*(2), 416–427.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646.

Olejarczyk, J. H., Luke, S. G., & Henderson, J. M. (2014). Incidental memory for parts of scenes from eye movements. *Visual Cognition*, *22*(7), 975–995.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*(12), 505–510.

Parag, O., & Vakil, E. (2018). Distinct eye movements for different cognitive processes as expressed in the face recognition task. *Memory* , *26*(4), 524–534.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10.

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science*, *20*(1), 6–10.

Reagh, Z. M., Ho, H. D., Leal, S. L., Noche, J. A., Chun, A., Murray, E. A., & Yassa, M. A. (2016). Greater loss of object than spatial mnemonic discrimination in aged adults. *Hippocampus*, *26*(4), 417–422.

Reagh, Z. M., Watabe, J., Ly, M., Murray, E., & Yassa, M. A. (2014). Dissociated signals in human dentate gyrus and CA3 predict different facets of recognition memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(40), 13301–13313.

Reagh, Z. M., & Yassa, M. A. (2014). Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), E4264–E4273.

Roberts, J. M., Ly, M., Murray, E., & Yassa, M. A. (2014). Temporal discrimination deficits as a function of lag interference in older adults. *Hippocampus*, *24*(10), 1189–1196.

Rolls, E. T. (1989). 13 - Functions of neuronal networks in the hippocampus and neocortex in memory. In J. H. Byrne & W. O. Berry (Eds.), *Neural Models of Plasticity* (pp. 240–265). Academic Press.

Rolls, E. T. (2013). A quantitative theory of the functions of the hippocampal CA3 network in memory. *Frontiers in Cellular Neuroscience*, *7*, 98.

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7*(2), 1–22.

Sekiguchi, T. (2011). Individual differences in face memory and eye fixation patterns during face learning. *Acta Psychologica*, *137*(1), 1–9.

Shelton, D. J., & Kirwan, C. B. (2013). A possible negative influence of depression on the ability to overcome memory interference. *Behavioural Brain Research*, *256*, 20–26.

Squire, L. R. (2009). The legacy of patient HM for neuroscience. *Neuron*, *61*(1), 6–9.

Stark, S. M., Stevenson, R., Wu, C., Rutledge, S., & Stark, C. E. L. (2015). Stability of age-related deficits in the mnemonic similarity task across task variations. *Behavioral Neuroscience*, *129*(3), 257–268.

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, *51*(12), 2442–2449.

Stark, S. M., Yassa, M. A., & Stark, C. E. L. (2010). Individual differences in spatial pattern separation performance associated with healthy aging in humans. *Learning & Memory*, *17*(6), 284–288.

Trier, H. A., Lacy, J. W., & Marsh, B. U. (2016). Limitations of episodic memory for highly similar auditory stimuli. *Journal of Cognitive Psychology* , *28*(7), 843–855.

Vance, A. O., Jenkins, J. L., Anderson, B. B., Bjornn, D. K., & Kirwan, C. B. (2018). Turning out security warnings: A longitudinal examination of habituation through fMRI, eye tracking, and field experiments. *Management Information Systems Quarterly*, *42*(2), 355–380.

Vazdarjanova, A., & Guzowski, J. F. (2004). Differences in hippocampal neuronal population responses to modifications of an environmental context: Evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(29), 6489–6496.

Vieweg, P., Riemer, M., Berron, D., & Wolbers, T. (2018). Memory Image Completion (MIC): Establishing a task to behaviorally assess pattern completion in humans. *Hippocampus*. https://doi.org/10.1002/hipo.23030

Weeden, C. S. S., Hu, N. J., Ho, L. U. N., & Kesner, R. P. (2014). The role of the ventral dentate gyrus in olfactory pattern separation. *Hippocampus*, *24*(5), 553–559.

Wilson, I. A., Ikonen, S., Gallagher, M., Eichenbaum, H., & Tanila, H. (2005). Age-associated alterations of hippocampal place cells are subregion specific. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(29), 6877–6886.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176.

Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. L. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(21), 8873–8878.

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525.

Yassa, M. A., Stark, S. M., Bakker, A., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2010). High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic Mild Cognitive Impairment. *NeuroImage*, *51*(3), 1242–1252.