The roles of the hippocampus and prefrontal cortex during visual long-term memory:

Author: Brittany M. Jeye

THE ROLES OF THE HIPPOCAMPUS AND PREFRONTAL CORTEX DURING VISUAL LONG-TERM MEMORY

Brittany M. Jeye, M.A.

A dissertation

submitted to the Faculty of

the department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

May 2019

THE ROLES OF THE HIPPOCAMPUS AND PREFRONTAL CORTEX DURING VISUAL LONG-TERM MEMORY

Brittany M. Jeye

Advisor: Scott D. Slotnick, Ph.D.

We are able to consciously remember an incredible amount of information for long periods of time (Brady et al., 2008, 2013). Furthermore, we often think about our memories in terms of how successful we are in retrieving them, such as vividly recalling the smell of your grandmother's cooking. However, we can also identify the times when we have forgotten information, such as misremembering the name of an acquaintance or misplacing your car keys. Such instances of forgetting have been suggested to be caused by inhibitory processes acting on associated information, such as the inhibitory processing shown in retrieval-induced forgetting where the retrieval of specific items leads to forgetting related information (Anderson et al., 2004; Wimber et al., 2015). Thus, long-term memory is said to rely on both accurately retrieving specific details and inhibiting potentially distracting information. In Chapter 1, I demonstrate that specificity of long-term memory depends on inhibiting related information through a series of behavioral experiments investigating item memory for faces and abstract shapes. In Chapter 2 and Chapter 3, I examine the neural regions associated with long-term memory specificity and inhibitory processing by focusing on the functional roles of the hippocampus and the prefrontal cortex, two key regions associated with long-term memory. In Chapter 2, I provide evidence that the hippocampus is associated with memory specificity by demonstrating that distinct regions of the hippocampus are associated with memory for different visual field locations. Furthermore, I provide evidence that the hippocampus operates in continuous manner during recollection (i.e., conscious retrieval of details). In Chapter 3, I demonstrate that the prefrontal cortex can inhibit both the hippocampus and language processing regions during retrieval of distracting information during episodic and semantic memory, respectively.

TABLE OF CONTENTS

GENERAL INTRODUCTION

Human visual long-term memory commonly refers to our ability to consciously remember previously seen information over long periods of time (Squire, 1992). However, this broad description does not fully capture the nuances of the human memory system. For instance, long-term memory can be broken down into various distinctions. One important distinction is between item memory and context memory, where item memory refers to the ability to discriminate between things that have been seen before and those that are novel. Context memory, on the other hand, refers to the ability to determine relational information surrounding items (also called source memory, associative memory or relational memory). This contextual information can be spatial in nature (such as determining whether an item was previously presented on the left or right side of a screen) or non-spatial in nature (such as linking an item to a particular color or retrieving temporal information about when an item was shown; Stark, Reagh, Yassa, & Stark, 2017). Of importance, retrieving items and their contexts from memory is intricately tied to our ability to 'mentally time travel' (Tulving, 1972). That is, we are able to re-experience past events, such as the who, what, where and when of a previous situation, to essentially transport our mind through time (i.e., episodic memory).

In addition, the cognitive processes that support item and context memory can further help us understand the specificity of visual long-term memories (i.e., how detailed our memories can get). For example, long-term memory can be based on either detailed recollection or non-detailed familiarity. It is thought that recollection is required during context memory, as one needs to retrieve details surrounding the previous experience (e.g., the correct side of the screen).

Familiarity, on the other hand, occurs during item memory, as one may have a lack of specific details of a previously seen item (e.g., when you know you have seen someone before but cannot place when or where you met them). Thus, visual long-term memory can rely both on recalling the gist (i.e., without specific details) of past events and accurately retrieving specific details about past events.

Of relevance to the studies in this dissertation, visual long-term memory is thought to be a constructive process made up of control and sensory regions in the brain. These control and sensory regions are further thought to mediate the specificity of our long-term memories. Sensory cortical regions (such as the visual cortex) within the brain reflect the contents of memories. For example, when retrieving a memory there is reactivation of sensory-specific regions, such that a visually encoded stimulus will reactivate visual cortex and an auditory stimulus will reactivate auditory cortex (Wheeler et al., 2000). The control regions associated with long-term memory include the medial temporal lobe (i.e., the hippocampus, parahippocampal cortex, and perirhinal cortex), the dorsolateral prefrontal cortex, and the parietal cortex (Rugg & Vilberg, 2013; Wagner, Shannon, Kahn, & Buckner, 2005). The roles of the medial temporal lobe structures are relatively well understood. The parahippocampal cortex processes contextual information (such as the spatial components of a previous scene), while the perirhinal cortex processes item information (Diana, Yonelinas, & Ranganath, 2007). However, of particular importance, the hippocampus sits at the center of this memory control network. Decades of neuroimaging and patient research has demonstrated how critical the hippocampus is to human memory (for a review, see Simons & Spiers, 2003). Its key role in long-term memory is best highlighted with the research on patient HM, who became amnesic after his hippocampus was removed from both hemispheres in a surgical procedure to treat epilepsy (Scoville & Milner, 1957; Corkin, 2002). The hippocampus has also been shown to be necessary for recollection. For example, the hippocampus is thought to bind item information and contextual information together to support detailed long-term memory (i.e., the binding-in-context model; Ranganath, 2010; Schiller et al.,

2015). That is, the hippocampus binds the contextual information processed by the parahippocampal cortex and the item information processed by the perirhinal cortex into a unified memory. This role of the hippocampus is consistent with animal research, which has demonstrated that specific hippocampal neurons (i.e., place cells) are active when an animal is in a particular location in its environment (O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978). That is, the hippocampus supports the spatiotemporal contexts of memory.

On the other hand, the roles of the dorsolateral prefrontal cortex and parietal cortex in long-term memory are less well understood. The parietal cortex has been hypothesized to support conscious episodic retrieval (e.g., possibly through directing internal attention; Wagner et al., 2005; Cabeza et al., 2008), whereas the dorsolateral prefrontal cortex is involved with relational memory and has been hypothesized to support control processes in memory. These control processes include source monitoring (i.e., post-retrieval monitoring) and inhibitory processes (Diamond & Levine, 2018; Gilboa, 2004; Mitchell & Johnson, 2009). Inhibitory processing is thought to aid in long-term memory by suppressing irrelevant information and is of particular relevance to the studies in this dissertation. For example, it has been argued that inhibitory processing is a critical mechanism that leads to forgetting, as inhibition may be flexibly directed at related, interfering, or competing memories during different stages of mnemonic processing. Weakening or deactivating these related memory traces through inhibition can then make them more susceptible to being later forgotten (Anderson & Hanslmayr, 2014; Depue, 2012; Levy & Anderson, 2002). This link between inhibitory processing, the dorsolateral prefrontal cortex, and long-term memory has previously been demonstrated through research into retrieval-induced forgetting (i.e., through unintentional forgetting tasks like the retrieval-practice paradigm) and suppression-induced forgetting (i.e., through intentional forgetting tasks like the think/no-think paradigm; Anderson & Hanslymayr, 2014; Bauml, Pastotter & Hanslmayr, 2010; Depue, 2012). For example, the dorsolateral prefrontal cortex has been shown modulate hippocampal activity during intentional memory suppression, where participants are given an explicit cue to forget

previously learned information (Benoit, Hulbert, Huddleston, & Anderson, 2015). Furthermore, the dorsolateral prefrontal cortex has been shown to be associated with inhibitory processing in retrieval-induced forgetting, where the retrieval of specific items in memory leads to the forgetting of related items (Wimber et al., 2008; Wimber et al., 2009; Wimber et al., 2015). Taken together, these studies indicate that one role of the dorsolateral prefrontal cortex is top-down modulation of the medial temporal lobe in long-term memory.

While previous research has shown that long-term memory relies on both accurately retrieving specific details and inhibiting potentially distracting information, much remains unclear about how these cognitive processes and the neural regions that support them are linked. In Chapter 1, I demonstrate that specificity of long-term memory depends on inhibiting related information through a series of behavioral experiments investigating item memory for faces and abstract shapes. In Chapters 2 and 3, I examine the neural regions associated with long-term memory specificity and inhibitory processing by focusing on the functional roles of the hippocampus and the prefrontal cortex, two key regions associated with long-term memory. In Chapter 2, I provide evidence that the hippocampus is associated with memory specificity by demonstrating that distinct regions of the hippocampus are associated with memory for different visual field locations. Furthermore, I provide evidence that the hippocampus operates in continuous manner during recollection (i.e., conscious retrieval of details). In Chapter 3, I demonstrate that the prefrontal cortex can inhibit both the hippocampus and language processing regions during retrieval of distracting information during episodic and semantic memory.

PART I

BEHAVIORAL EVIDENCE FOR VISUAL LONG-TERM MEMORY SPECIFICITY

CHAPTER 1.0: MEMORY SPECIFICITY FOR VISUAL ITEMS

Long-term memory specificity depends on inhibition of related items
Brittany M. Jeye, Cassidy R. McCarthy and Scott D. Slotnick

Long-term memory relies on both accurately retrieving specific details and inhibiting competing information. In the current investigation, we evaluated long-term memory specificity. During each study phase, participants were presented with either abstract shapes or faces. During the corresponding test phase, old items, related items, and new items were presented and participants made "old"–"new" recognition judgments. For abstract shapes, related items were created by distorting shapes along a five-point continuum (i.e., 0%, 50%, 100%, 150%, and 200% distortions), where independent raters signified 100% distortions were perceptually "different" than old items. For faces, related items were created by morphing two faces together in steps of 20% (i.e., 0%, 20%, 40%, 60% and 80% morphs). Memory representations were detailed as the "old" response rate differed between old items and closely related items for both shapes and faces (i.e., 50% distortions and 20% morphs, respectively). Furthermore, there was evidence of memory inhibition for related items. For shapes, there was a lower "old" response rate for distantly related items (200% distortions) than new items, while for faces, there was a lower "old" response rate for closely related items (20% morphs) than less related items (40% morphs). This may reflect an evolutionary mechanism for recognizing specific faces, which may require inhibition of closely related faces, as compared to shapes that have less specific category boundaries.

We are capable of consciously remembering a massive amount of information (Tulving, 1985; Brady, Konkle, Alvarez, & Oliva, 2008). However, memories can differ objectively in their precision or specificity (i.e., the amount of details remembered) or can differ subjectively in their vividness or confidence (Richter, Cooper, Bays, & Simons, 2016). Such nuanced measures of memory quality beyond simple memory success are more frequently being evaluated (Harlow & Donaldson, 2013; Harlow & Yonelinas, 2016; Qin, van Marle, Hermans, & Fernandez, 2011). Recent findings have suggested that visual long-term memories can be recalled with an incredible amount of precision, on par with visual working memory (Brady, Konkle, Gill, Oliva, & Alvarez, 2013). However, the cognitive processes mediating such detailed visual long-term memories are largely unknown.

Inhibition may be a critical component of detailed long-term memory, as inhibitory mechanisms have long been known to play a role in tasks requiring executive control, such as in selective attention and decision-making (Anderson & Spellman, 1995; Knight, Staines, Swick, & Chao, 1999; Miller & Cohen, 2001). Thus, inhibition may be flexibly directed at interfering and/or competing memories during different stages of mnemonic processing in order to select the appropriate memory from related memories. For example, research utilizing the retrieval-practice paradigm has investigated the role of inhibition during memory selection (Anderson, Bjork, & Bjork, 1994; Wimber, Bäuml, Bergström, Markopoulos, Heinze, & Richardson-Klavehn, 2008; Wimber, Alink, Charest, Kriegeskorte, & Anderson, 2015). In this paradigm, participants study category-exemplar pairs from several different semantic categories, such as ‘fruit-apple’, ‘fruit-orange’, etc. They are then directed to practice only a subset of the previously studied category-exemplar pairs (e.g. ‘fruit-apple’ but not ‘fruit-orange’; baseline categories are not practiced). Finally, participants are given a test where they are asked to retrieve all the exemplars from the first phase. As expected, performance for the practiced items is typically greater than that of non-practiced/baseline items. Of particular relevance, memory for the unpracticed items from

practiced categories is worse than that of baseline items. This appears to reflect inhibition of non-practiced items that are related to practiced items.

In the current study, we employed a novel paradigm to investigate long-term memory specificity. During each study phase, participants viewed either abstract shapes or faces. During the corresponding test phase, participants viewed old items, related items (constructed by distorting/morphing old items), and new items and made an "old"–"new" recognition response. It is notable that this was a standard memory paradigm with two phases (study and test), as compared to the retrieval-induced forgetting paradigm that also requires an intermediate retrieval-practice phase. By varying the amount of distortion added to the related items (along a five-step continuum), we were able to systematically assess memory specificity and inhibitory processes. To anticipate the results, we found a high degree of memory specificity, which replicated previous findings, and also observed evidence for inhibition of distantly related shapes and closely related faces.

Experiment 1a

Methods

Participants

One hundred and four adults aged 18-69 participated in the study. Eight participants were excluded from the analysis due to either below baseline performance on the task (scoring below 50% in the "old" responses for old shapes) or computer error, leaving 96 participants that were included in the analysis (53 females, M = 32 years). This sample size was selected assuming a power of .8 and a small to medium effect size (this effect size was based on the results of a pilot study; Tierney & Slotnick, 2008). The Boston College Institutional Review Board approved the protocol, and informed written consent was obtained prior to each session. This research was conducted in Living Laboratory® at the Museum of Science, Boston, MA.

Materials and Procedure

Each participant completed a one-third length practice run and three full-length runs. Each run consisted of a study and a test phase (see Fig. 1 top). During the study phase, 15 abstract shapes spanning 5.37 degrees of visual angle were presented in the center of the computer screen one at a time (for information on shape construction, see the Supplementary Methods; Slotnick & Schacter, 2004). Participants were instructed to remember each shape. Shape sets were randomized and presented sequentially three times. Each shape was presented for 2.5 seconds followed by a blank screen for 0.5 seconds. Following the study phase, there was a brief delay of 8 seconds in which the screen displayed a prompt reminding the participant of the task instructions. During the test phase, 18 shapes were sequentially presented in the center of the screen that included old shapes, related shapes, and new shapes (see Fig. 1 bottom for an example of related shapes). Related shapes were created by systematically distorting old shapes between 50% and 200%, leading to four steps of varying relatedness (i.e., 50%, 100%, 150%, and 200% distortions). Independent ratings indicated that 100% distortions were perceptually "different" from corresponding old shapes. The shapes presented during the test phase were made up of three of each type of shape (old shapes/0% distortions, 50% distortions, 100% distortions, 150% distortions, 200% distortions, and new shapes). Shapes were pseudo-randomized such that no more than two shapes of a given type were presented sequentially. During the test phase, each shape was presented for 3.0 seconds followed by a confidence rating reminder screen for 2.5 seconds and a 0.5 second blank screen. Participants were instructed to make an "old"–"new" response to each shape, where an "old" response indicated the shape was exactly the same as a shape presented during the study phase and a "new" response indicated the shape was never seen before or it was similar to a shape seen during the study phase. Participants were told that some shapes would be very similar to those shown previously but they should only respond "old" if the shape was exactly the same as a shape shown in the study phase. Each "old"–"new" response was followed by a confidence judgment in which participants indicated whether they were "unsure,"

"sure," or "very sure" that the previous "old"–"new" response was accurate. To make the responses, participants pressed buttons on an external numerical keypad with the fingers of their preferred hand using only the keys 1, 2, and 3. Subjects were allowed to select their response hand. Sets of shapes (old items, each related item set, new items) were counterbalanced across participants using a Latin Square design. Stimuli were presented electronically using E-prime software (Psychology Software Tools, Pittsburgh, PA) on a Dell laptop.

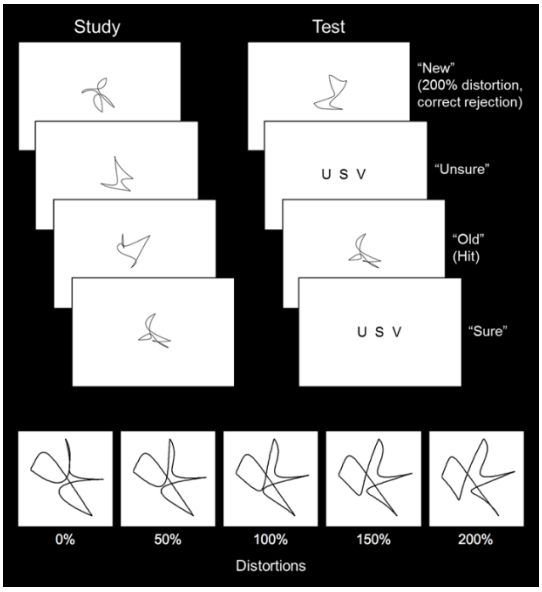

Fig. 1. Top. In the study phase, abstract shapes were presented in the center of the screen. In the test phase, old shapes, related shapes, and new shapes were presented and participants classified each shape as "old" or "new" followed by an "unsure"–"sure"–"very sure" confidence rating. Bottom. An example of the five step continuum of shape distortions used as related shapes in the test phase. 0% indicates the shape was not changed from the study phase.

Analyses

Data analysis was conducted by calculating the percentage of "old" responses out of the total number of responses for each shape type (old items, each related item set, and new items). The "old" response rate for old items (i.e., 0% distortions) was compared to the "old" response rate for different levels of related items (e.g., 50% distortions) to indicate the level of relatedness at which shapes could be distinguished (i.e., the level of memory specificity). In a previous pilot study, we found that the "old" response rate for 200% distortions was significantly lower than the

"old" response rate for both 150% distortions and new items (Tierney & Slotnick, 2008), which was suggestive of inhibitory processing. Thus, in an effort to replicate these findings in the present study, we specifically focused on the rate of "old" responses to 200% distortions as compared to 150% distortions and new items. A *d'* analysis was also conducted to ensure the memory effects were not due to response biases. Cohen's *d* was used to measure effect size and reported for significant results.

Results

Memory specificity was examined by comparing the percent "old" response for old shapes (0% distortions) and 50% distortions (shapes that were closely related to the old shapes). Memory representations were very specific as the "old" response rate differed between old shapes and 50% distortions ($t(95) = 9.71$, $p < .001$, $d = 1.12$; Fig. 2), indicating that participants were able to correctly reject closely related items. The "old" response rate was lower for 200% distortions than for 150% distortions ($t(95) = 3.10$, $p < .01$, $d = 0.33$) and new shapes ($t(95) = -3.11$, $p < .01$, $d = 0.34$), which likely reflects memory inhibition of distantly related items. The *d'* analyses yielded a similar pattern of results (see Supplementary Fig. S1). There was a significant difference between *d'* for old and 50% distortions ($t(95) = 8.10$, $p < .001$, $d = 0.75$), and the 200% distortion d' was significantly negative ($d' = -0.41$; $t(95) = -3.22$, $p < .01$, $d = 0.34$). The same pattern of results was observed when participants were not excluded based on performance (see Supplementary Figs. S2 and S3). These results suggest that long-term memory specificity depends on detailed memory for specific items and inhibition of distantly related items.

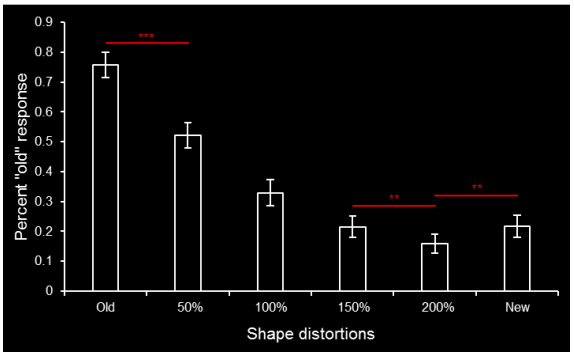

Fig. 2. The percent "old" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions) and new shapes. Error bars represent 95% confidence intervals. *** = $p < .001$, ** = $p < .01$.

Experiment 1b

It is known that long-term memory can also be mediated by consciously remembering a related studied item to reject a new item (i.e., recall-to-reject; Rotello & Heit, 2000). In the previous experiment, participants may have utilized recall-to-reject when viewing related items. In particular, it is possible that participants were cued by distantly related items (i.e., 200% distortions) to remember the corresponding old shapes such that they responded "new" more often to distantly related shapes than new shapes. Experiment 1b assessed whether recall-to-reject was driving this decrease in "old" response rate for distantly related items as compared to new items.

Methods

Unless otherwise specified, the materials and methods of Experiment 1b were identical to those of Experiment 1a.

Participants

A total of 54 Boston College undergraduate students between the ages of 18 and 24 participated in this study. Six participants were excluded from the analysis due to below baseline performance on the task (scoring below 50% in the "old" responses for old shapes), leaving 48 participants in the analyses (33 females, $M = 19$ years). This sample size was selected assuming a

power of .8 and a medium effect size (this effect size was based on the results of Experiment 1a). Participants received research credit for enrolling in the study. The Boston College Institutional Review Board approved the protocol. Informed written consent was obtained prior to each session.

Materials, Procedure, and Analyses

Each participant completed a one-third practice run and six full-length runs (see Fig. 3). During the study phase, 15 shapes spanning 7.15 degrees of the visual angle were presented sequentially. During the test phase, participants were presented with a second set of 18 shapes, which included three old shapes, three of each type of related shape (50%, 100%, 150%, and 200% distortions), and three new shapes. Participants were instructed to make an "old"–"new" response to each shape, where an "old" response indicated the shape was exactly the same as a shape presented during the study phase and a "new" response indicated the shape was never seen before or it was similar to a shape seen during the study phase. Participants then made a second response. If the participant responded "old" for the first response, they then made a "remember" or "know" response. A "remember" response indicated the participant could consciously recollect specific details of the experience of previously seeing the exact shape, while a "know" response indicated the participants were confident the exact shape had been seen before, but were not able to consciously recollect specific details about what was experienced (Slotnick, 2010). If the participant responded "new" for the first response, they then decided whether the shape was "related" or "unrelated" to a shape they saw previously. A shape that was "related" to a shape seen in the study phase session would bring to mind the details of the previously presented shape in order for the participant to determine that the shape was similar to a shape seen in the study session (reflecting recall-to-reject). A shape that was "unrelated" to a shape seen in the study phase would be completely novel (i.e., not at all similar to a shape seen in the study phase) and did not bring to mind any details of a previously presented shape. Thus, there were four possible

response types: “old-remember”, “old-know”, “new-related” and “new-unrelated”. To make the responses, participants pressed buttons on a keyboard with the fingers of their left hand using keys 1 and 2. Sets of shapes (old items, each related item type, new items) were counterbalanced across participants using a Latin Square design.

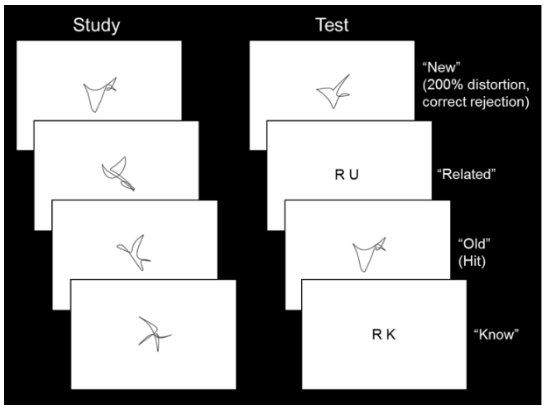

Fig. 3. Left. In the study phase, abstract shapes were presented in the center of the screen. Right. In the test phase, old shapes, related shapes, and new shapes were presented and participants classified each shape as “old”–“remember”, “old”–“know”, “new”–“related” or “new”–“unrelated”.

In addition to assessing memory specificity (as described in Experiment 1a), the primary aim of Experiment 1b was to evaluate whether participants were utilizing either recall-to-reject or inhibition to correctly reject the distantly related shape distortions (i.e., 200% distortions). Therefore, the percentage of “new-related” responses out of total number of responses was calculated and compared between 200% distortions and new shapes. If participants were using a recall-to-reject strategy, they would have a higher “new-related” response rate for 200% morphs than for new shapes. Two participants did not make any “new-related” responses for old shapes and their data were not included in the corresponding calculation.

Results

The “old” response rate for 50% distortions and old distortions (0% morphs) were significantly different ($t(47) = 10.47$, $p < .001$, $d = 1.59$; Fig. 4), which demonstrates that the

memory for shapes were detailed, replicating the results from Experiment 1a. The "old" response rate for 200% distortions was also significantly lower than the "old" response rate for 150% distortions ($t(47) = 4.04$, $p < .001$, $d = 0.53$). Although the difference between the "old" response rate for 200% distortions and new shapes were in the same direction as in Experiment 1a, this difference did not reach significance ($t(47) = -1.60$, $p = .12$; see Supplementary Fig. S4 for the *d'* results). To assess whether participants utilized a recall-to-reject strategy for each level of relatedness, we calculated the percentage of responses where participants responded "new" followed by a "related" response. Critically, there was no difference between the "new"–"related" response rate for 200% distortions and new shapes ($t(47) = 1.70$, $p = .10$; Fig. 5). The latter result suggests that the lower "old" response rate for related shapes that were distorted 200%, as compared to new items, was due to memory inhibition rather than recall-to-reject. The same pattern of results was observed when participants were not excluded based on performance (see Supplementary Figs. S5, S6 and S7).

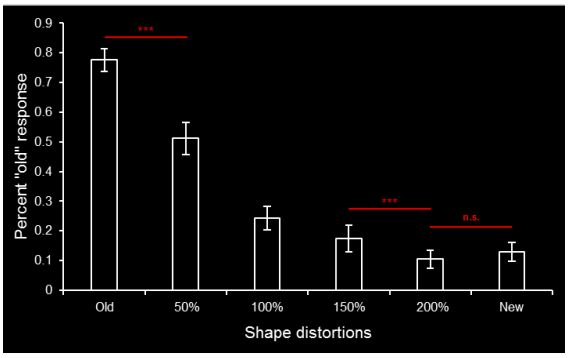

Fig. 4. The percent "old" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions) and new shapes. Error bars represent 95% confidence intervals. *** = $p < .001$, n.s. = not significant.

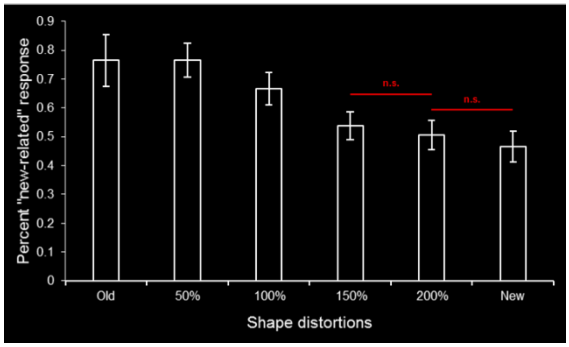

Fig. 5. The percent "new"–"related" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions) and new shapes. Error bars represent 95% confidence intervals. * = $p < .05$, n.s. = not significant.

A split-half analysis utilizing the participants who had lower "old" response rates for 200% distortions compared to new shapes was also conducted. As expected, the "old" response rates for 50% distortions and old shapes (0% distortions) were significantly different ($t(22) = 7.59$, $p < .001$, $d = 1.56$; Fig. 6). Furthermore, the "old" response rate for 200% distortions was significantly lower than the "old" response rate for 150% distortions ($t(22) = 5.14$, $p < .001$, $d = 1.22$) and the "old" response rate for new distortions ($t(22) = 7.50$, $p < .001$, $d = 1.49$). Of importance, for these items that showed a robust decrease in "old" responses for 200% distortions, the "new-related" response rate for 200% distortions was not significantly different than the "new-related" response rate for new shapes ($t(20) = 1.14$, $p = .267$; Fig. 7). These results further indicate that long-term memory specificity depends on inhibition of distantly related items rather than recall-to-reject.

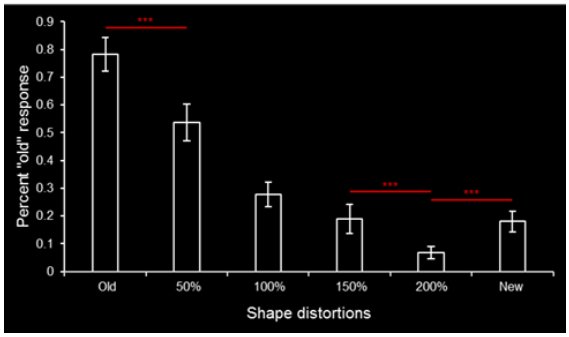

Fig. 6. The percent "old" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape morphs), and new shapes in Experiment 1b for participants who had lower "old" response rates for 200% morphs compared to new shapes. Error bars represent 95% confidence interval. *** = $p < .001$.

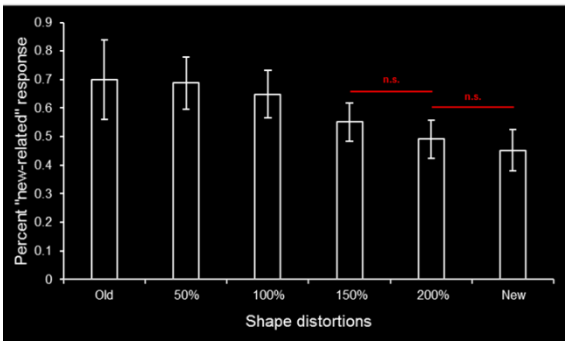

Fig. 7. The percent "new"–"related" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape morphs) and new shapes of Experiment 1b in participants who had lower "old" response rates for 200% morphs compared to new shapes. Error bars represent 95% confidence interval. n.s. = not significant.

Experiment 1c

The previous findings suggested that long-term memory depended on both detailed memories for specific items and inhibition of distantly related items. However, it is possible that the lower "old" response rate for distantly related shapes (i.e., 200% distortions) was due to perceptual differences between the 200% distortions and new shapes, rather than due to inhibition. For example, 200% distortions may have been perceived as completely novel (and not related to the original shapes), while the new shapes may have been perceived as more related to the original shape (even though they were not related). To investigate this possibility, a follow-up perceptual experiment was conducted utilizing old items, 200% distortions, and new items to investigate whether the response pattern was due to perceptual differences.

Methods

Participants

Twenty-seven Boston College undergraduate students participated in this study. Participants received research credit for enrolling in the study. This sample size was selected assuming a power of .9 and a large effect size (the relatively high power/effect size was expected given that this was a perception experiment). The Boston College Institutional Review Board approved the protocol. Informed written consent was obtained prior to each session.

Materials, Procedure, and Analyses

Participants were shown three shapes at a time on the screen, arranged in a triangle (see Fig. 8). The top shape was always an old shape (0% distortion), while the bottom left and right shapes were the 200% distortion of the old shape or a new shape. Shapes were presented for 2.5 seconds followed by a blank screen for 0.5 seconds. Participants were instructed to always look at the center of the screen and indicate which of the bottom shapes ("left" or "right") was most similar to the top shape. To make the responses, participants pressed buttons on a keyboard with the fingers of their left hand using keys 1 and 2. Shapes were taken from Experiment 1b such that the old shapes, the 200% distortions and new shapes were the same ones utilized during each counterbalanced run. To prevent fatigue, shapes were split into two runs such that half the shapes were presented followed by a small break and then the second half of the shapes were presented. New shapes and 200% distortions were counterbalanced based on spatial location (left or right). For the analyses, the percentage of trials in which participants classified 200% distortions as more similar to the old shape was computed.

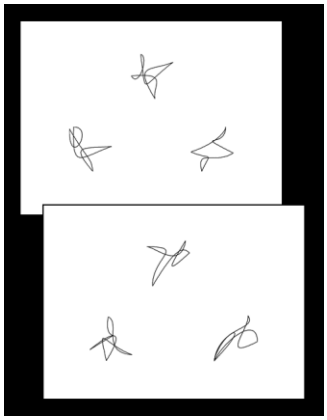

Fig. 8. Example stimuli in the abstract shape perception task. The top shape was always old (i.e., 0% distortion), while the bottom shapes were either the 200% distortion of the old shape or a new shape. Participants decided which of the bottom shapes was more similar to the top old shape.

Results

Participants were significantly above chance at identifying 200% distortions as being more similar to the corresponding old shapes (74.11%, chance = 50%, $t(26) = 27.70$, $p < .001$, $d =$

5.33). This suggests that our previous results demonstrating a lower "old" response rate for distantly related shapes (i.e., 200% distortions) was due to inhibition rather than perceptual differences between the 200% distortions and new shapes.

Experiment 2

The previous findings indicated that memory specificity depends on detailed retrieval of old items and inhibition of distantly related items. However, these paradigms utilized abstract shapes, which while easily manipulated to create precise levels of relatedness, have limited ecological relevance. Faces have obvious ecological relevance and have long been used to investigate specificity in perception and visual working memory. For instance, it has been shown that we are very sensitive at detecting small differences in facial features (Webster, Kaping, Mizokami, & Duhamel, 2004). In the current experiment, we employed faces as stimuli to evaluate specificity and inhibitory effects during long-term memory.

Methods

Unless otherwise specified, the methods of Experiment 2 are identical to those of Experiment 1a.

Participants

One hundred and twelve adults ages 18–76 participated in the study. Sixteen participants were excluded from the analysis due to either below baseline performance on the task (scoring below 50% in the "old" responses for old faces) or computer error, leaving 96 participants in the analysis (42 females, $M = 35$ years). This sample size was selected assuming a power of .8 and a small to medium effect size and to equate the number of participants to Experiment 1a. The Boston College Institutional Review Board approved the protocol, and informed written consent was obtained prior to each session. This research was conducted in Living Laboratory® at the Museum of Science, Boston, MA.

Materials, Procedure and Analyses

Participants completed a one-third practice run and two full-length runs (see Fig. 9 top). During the study phase, 20 faces (10 male, 10 female), spanning approximately 6.55 degrees of the visual angle in height and 4.18 degrees of the visual angle in width, were presented sequentially. Each face set was randomized and presented twice and no more than two male or female faces were presented in sequence. During the test phase, participants were presented with a second set of 24 faces (split equally between male and female faces). These included four old faces, four of each type of related face (20%, 40%, 60%, and 80% morphs; see Fig. 9 bottom for an example of the face morphs), and four new faces (for details on face stimuli construction, see the Supplementary Methods). Participants were instructed to make an "old"–"new" response followed by a confidence judgment in which participants indicated whether they were "unsure," "sure," or "very sure" that the preceding response was accurate. To make the responses, participants pressed buttons on an external numerical keypad with the fingers of their preferred hand using keys 1, 2, and 3. Sets of faces (old items, each related item type, new items) were counterbalanced across participants using a Latin Square design.

Given that faces are a distinct stimulus type as compared to abstract shapes, it was uncertain whether and at what level of face morph inhibitory processing might occur. To determine this empirically, we fit a series of increasing polynomial functions to the group-average plot of "old" response rate as a function of relatedness (i.e., 0%, 20%, 40%, 60%, 80% face morphs, and new items). The correlation for a linear function ($R^2(4) = .895$) was similar to quadratic ($R^2(3) = .912$) and cubic ($R^2(2) = .925$) functions. However, the correlation for a quartic function was very high ($R^2(1) = .993$), which indicates multiple reversals in the function that may reflect inhibition (rather than a monotonic decrease that would be well fit by a linear function alone). A subtraction of the linear function from the quartic function revealed a marked decrease, well below 0% "old" response, corresponding to the 20% morphs (see Supplemental Figure S8).

As such, for the present experiment, we focused on comparing the 20% morphs to old items and 40% morphs to assess whether there were inhibitory effects.

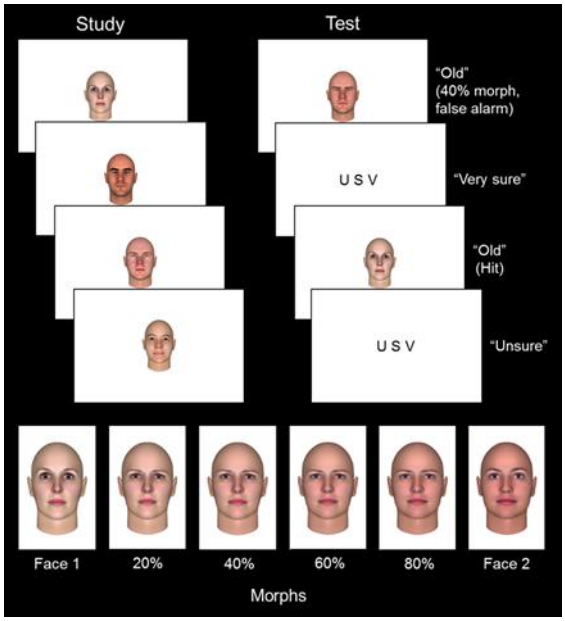

Fig. 9. Top. During the study phase, faces were presented in the center of the screen. During the test phase, old faces, related faces and new faces were presented and participants classified each face as "old" or "new" followed by an "unsure"–"sure"–"very sure" confidence rating. Bottom. An example of the five step continuum of face morphs used as related shapes in the test phase (20% indicates the face was 20% similar to Face 2 and 80% similar to Face 1).

Results

Similar to the previous studies that utilized abstract shapes, memory representations for faces were very specific as the "old" response rate was significantly greater for old faces than 20% morphs ($t(95) = 5.58$, $p < .001$, $d = .76$; Fig. 10). However, unlike the abstract shapes, which had a lower "old" response rate for distantly related shapes than for new shapes, the "old" response rate was significantly lower for 20% face morphs than for 40% face morphs ($t(95) = -2.01$, $p < .05$, $d = .24$). This can be assumed to reflect memory inhibition of closely related faces. An additional *d'* analysis demonstrated that there was a significant difference between old faces and 20% face morphs ($t(95) = 4.52$, $p < .001$, $d = 0.44$; Supplementary Fig. S9); however, although the d' for 20% morphs was numerically lower than the d' for 40% morphs, this difference was not significant ($t(95) < 1$). The same pattern of results was observed when

participants were not excluded based on performance (see Supplementary Figs. S10 and S11). These results suggest that long-term memory specificity depends on detailed memory for specific faces and inhibition of closely related faces.

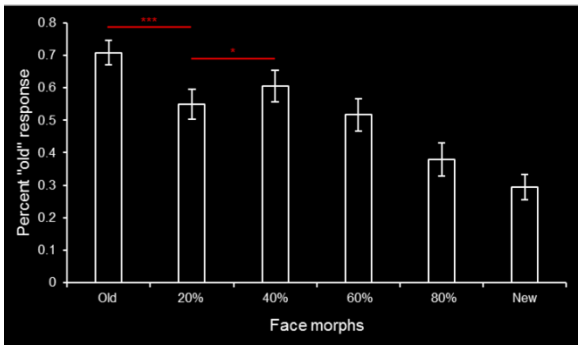

Fig. 10. The percent "old" response for old faces, related faces (20%, 40%, 60%, and 80% face morphs) and new faces. Error bars represent 95% confidence intervals. *** = $p < .001$, * = $p < .05$.

Discussion

The results of the current experiments demonstrate that memory representations can be very detailed. This was shown in Experiment 1a and 1b where the "old" response rate for old shapes was significantly greater than for closely related shapes (i.e., 50% morphs) and in Experiment 2 where the "old" response rate for old faces was significantly greater than for the closely related faces (i.e., 20% morphs). This high memory specificity is consistent with previous research evaluating the precision of visual long-term memory (Brady et al., 2008, 2013). In Experiment 1a, the "old" response rate for distantly related shapes (i.e., 200% morphs) was significantly less than new shapes, and in Experiment 2, the "old" response rate for closely related faces (i.e., 20% morphs) was significantly less than more distantly related faces (i.e., 40% faces). As demonstrated in Experiment 1b, these findings are indicative of an inhibitory mechanism that acts on these related items rather than due to a recall-to-reject strategy.

The pattern of inhibitory activity observed in the present study can be described as a center-surround organization in long-term memory, which is similar to the center-surround organization that has been observed during visual attention (Dagenbach & Carr, 1994; Slotnick,

Hopfinger, Klein, & Sutter, 2002; Slotnick, Schwarzbach, & Yantis, 2003). One difference, however, is that the center-surround organization in selective attention refers to the external focusing of attention on stimuli in the environment, while a center-surround organization in long-term memory refers to the internal focusing of attention onto memory representations. In Experiment 1, this inhibitory center-surround mechanism was observed for distantly related shapes, as the "old" response rate was significantly less than that of new shapes. Future research will need to be conducted utilizing more distantly related shapes (e.g., 250%, 300%, etc.), however, based on our findings, we would expect the "old" response rate for these items to increase again, in line with a center-surround pattern. This pattern was evident in Experiment 2, where the "old" response rate for closely related faces was significantly less than more distantly related faces. The concept of a center-surround organization in memory is not new, as previous research utilizing semantic memory paradigms (Dagenbach & Carr, 1994; Barnhardt, et al., 1996) and visual working memory (Fan & Turk-Browne, 2013; Kiyonaga & Egner, 2016) have also shown a center-surround organization. However, the present study provides the first evidence, to our knowledge, that a center-surround organization exists utilizing a standard long-term memory task.

Several paradigms, including the directed forgetting and think/no-think paradigms, in addition to the retrieval-practice paradigm, have also indicated that inhibitory processes are involved in long-term memory (Anderson & Green, 2001; Bjork, 1970; for a review, see Anderson & Hanslmayr, 2014). The directed forgetting and think/no-think paradigms have focused on situations in which specific memories needed to be stopped (unlike in the retrieval-practice paradigm in which specific memories needed to be selected). In these situations, memories are suppressed, or excluded from conscious awareness, via inhibitory control given an explicit forget cue. The current paradigm is most similar to the retrieval-practice paradigm, as participants had to select a studied item from amongst related items (without an explicit forget cue). Furthermore, the present results are consistent with the retrieval-induced forgetting effect,

where the act of retrieving an item from memory can inhibit related memories. While the retrieval-practice paradigm has typically used words as stimuli, retrieval-induced forgetting has also been shown for visual stimuli (Wimber et al., 2015) and for episodic memories (Cirranni & Shimamura, 1999). This indicates that long-term visual memory can be modulated by inhibitory processes. Additionally, research utilizing the retrieval-practice paradigm has shown that the amount of relatedness between exemplars within categories can differentially influence retrieval-induced forgetting (Anderson, Green, & McCulloch, 2000). Highly similar exemplars (e.g., for the category 'red', exemplars 'radish' and 'tomato') increased the amount of retrieval-induced forgetting. In our paradigm we used highly related stimuli, which may have increased the amount of inhibitory processing we observed.

Our inhibitory findings appeared to be dependent on the type of stimuli, as there was inhibition of closely related faces and distantly related shapes. We speculate that this could reflect an evolutionary advantage for recognizing specific faces, which may require inhibition of closely related faces, as compared to abstract shapes that have broad category boundaries. Furthermore, this inhibition of closely related faces (20% morphs) rather than for all related faces (i.e., 40%, 60% and 80% morphs) suggests that only high similarity competitors may be the target of inhibitory processes for faces. Future work will be needed in order to assess the degree to which inhibition is needed to suppress similar memories, as it is possible that a similar pattern of inhibition may be flexibly directed at other stimuli types with which we have expertise. Furthermore, inhibition may serve an important role in object categorization and recollection of semantic information (i.e., knowledge of facts). For example, inhibitory mechanisms could give rise to object-specific characteristics that help distinguish similar categories from one another (Johnson & Anderson, 2004; Martin & Chao, 2001).

Taken together, these behavioral findings demonstrate that visual long-term memory specificity is mediated by inhibitory mechanisms. Future research will be need to be conducted to elucidate the particular neural regions that support the center-surround organization in long-term

memory and whether they are similar or distinct from those that support the center-surround organization in selective attention, working memory and semantic memory. One likely possibility is that the lateral prefrontal cortex, which has been shown to be involved with inhibition in other executive control processing, is the source of top-down inhibitory control over medial temporal lobe regions, such as the hippocampus, which support long-term memory (Anderson, Bunce & Barbas, 2016; Depue; 2012; Eichenbaum, 2017; Knight et al., 1999; Miller & Cohen, 2001).

Overall, our findings suggest that human visual long-term memory relies on both detailed retrieval of specific items and inhibition of related items. While we used two different stimuli types (i.e., abstract shapes and faces), future work will be necessary to determine whether the same pattern of inhibitory activity is found for other stimuli types, such as objects, or other stimulus classes that have similar geometries to faces (such as houses or inverted faces). Furthermore, additional research will be needed to assess other factors that may modulate this effect, such as the number of stimuli repetitions during the study phase, or the strength of the memory (as determined by participants' confidence responses or their "remember" or "familiarity" responses). Lastly, future research utilizing functional magnetic resonance imaging and event-related potentials will be needed to assess the underlying brain mechanisms that mediate inhibition during long-term memory.

References

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1063.

Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal–hippocampal pathways underlying inhibitory control over memory. *Neurobiology of Learning and Memory, 134*, 145-161.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*, 279-292.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, *410*(6826), 366.

Anderson, M. C., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: evidence for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1141.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review, 102*, 68

Barnhardt, T. M., Glisky, E. L., Polster, M. R., & Elam, L. (1996). Inhibition of associates and activation of synonyms in the rare-word paradigm: Further evidence for a center-surround mechanism. *Memory & Cognition, 24(1)*, 60-69.

Bjork, R.A. (1970). Positive forgetting: The non-interference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior, 9,* 255-268.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325-14329.

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*(6), 981-990.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25(6*), 1403.

Dagenbach, D. E., & Carr, T. H. (1994*). Inhibitory processes in attention, memory, and language*. Academic Press.

Depue, B. E. (2012). A neuroanatomical model of prefrontal inhibitory modulation of memory retrieval. *Neuroscience & Biobehavioral Reviews, 36(5), 1382-1399.*

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience, 18(9),* 547.

Fan, J. E., & Turk-Browne, N. B. (2013). Internal attention to features in visual short-term

memory guides object learning. *Cognition, 129(2),* 292-308.

Harlow, I. M., & Donaldson, D. I. (2013). Source accuracy data reveal the thresholded nature of human episodic memory. *Psychonomic bulletin & review*, *20*(2), 318-325.

Harlow, I. M., & Yonelinas, A. P. (2016). Distinguishing between the success and precision of recollection. *Memory*, *24*(1), 114-127.

Johnson, S. K., & Anderson, M. C. (2004). The role of inhibitory control in forgetting semantic knowledge. *Psychological Science, 15*, 448-453.

Kiyonaga, A., & Egner, T. (2016). Center-surround inhibition in working memory. *Current Biology, 26(1),* 64-68.

Knight, R. T., Staines, W. R., Swick, D., & Chao, L. L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychologica, 101(2-3*), 159-178.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. Current Opinion in Neurobiology, 11, 194-201.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24(1),* 167-202.

Qin, S., van Marle, H. J., Hermans, E. J., & Fernández, G. (2011). Subjective sense of memory strength and the objective amount of information accurately remembered are related to distinct neural correlates at encoding. *Journal of Neuroscience*, *31*(24), 8920-8927.

Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*, *5*.

Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition, 28(6)*, 907-922.

Slotnick, S. D. (2010). "Remember" source memory ROCs indicate recollection is a continuous process. Memory, 18, 27–39.

Slotnick, S. D., Hopfinger, J. B., Klein, S. A., & Sutter, E. E. (2002). Darkness beyond the light: attentional inhibition surrounding the classic spotlight. *NeuroReport, 13*, 773–778.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7(6),* 664.

Slotnick, S. D., Schwarzbach, J., & Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *NeuroImage, 19*, 1602–1611

Tierney, K. P., & Slotnick, S. D. (2008). The specificity of memory for abstract shapes. Poster at the Charles River Association for Memory Conference.

Tulving, E., & Murray, D. (1985). Elements of episodic memory. *Canadian Psychology*, *26*(3), 235-238.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428(6982*), 557.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience, 18(4),* 582.

Wimber, M., Bäuml, K. H., Bergström, Z., Markopoulos, G., Heinze, H. J., & Richardson-Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *Journal of Neuroscience, 28(50),* 13419-13427.

Supplementary Material

Supplementary Method for Experiment 1

Shape Stimuli Generation

Shapes were created using custom scripts in MATLAB (The MathWorks, Inc., Natick, MA, USA). First, old shapes (i.e., 0% distortions) were generated by randomly placing: 1) four end points on each side of bounding square (i.e., one point on each side of the square) and 2) for each adjacent pair of end points, two additional control points were placed within the bounding square. Bezier curves were then constructed using adjacent pairs of end points and the corresponding two control points. This resulted in the creation of one complete old shape (i.e., 0% distortion).

To create the related shapes, first an increasing magnitude of Gaussian distributed noise (with 0.05, 0.10…0.70 standard deviations) was added to the end and control points of each old shape (with the constraint that the points were on or within the bounding square). Adding this noise either systematically increased or decreased how stretched or compressed different components of the old shape became, which perceptually altered the old shape. Second, five independent observers then rated the level of noise at which the distortions began to appear qualitatively 'different' from the corresponding old shape, which reflected 100% distortion. Shapes that varied too much or too little in the mean 'different' rating were rejected (i.e., shapes that were two standard deviations above or below the mean 'different' rating), as were shapes that at least two ratings indicated looked like an animal or object. Lastly, for each old shape, this magnitude of noise was parametrically varied to create the different level of distortions (i.e., 0%, 50%, 100%, 150% and 200% distortions). For more details on shape construction, see Slotnick and Schacter (2004).

Supplementary Method for Experiment 2

Face Stimuli Generation

Faces were created using the FaceGen Modeller software (Version 3.11, Singular Inversions, Vancouver, BC, Canada). Sets of 30 male and female face stimuli (60 in total) were created using the FaceGen Modeller random face generator. The random generator allows either the software complete control over the face creation or allows the user to set initial sliding scale parameters (e.g., sex, age, caricature, race, and asymmetry) through which the software randomly generates faces. For the current face stimuli, we initially set FaceGen parameters before using the random face generator. For the sex of the faces, FaceGen was allowed to randomly generate faces between Female and HyperFemale and Male and Hypermale. No faces were allowed to be randomly generated that fell between Male and Female on the sliding scale in order to create distinct sets of male and female faces. The age was strictly set at 25 in order to create a coherent set of faces and to further control for age related memory effects. The caricature parameter was set between Normal and Typical, in order to create realistic faces, rather than cartoon faces. The race of the faces was set to European, but was allowed to vary along the European sliding scale. The asymmetry of the faces was set between Normal and Typical to create realistic faces. Beyond the initial parameters, FaceGen has many additional settings which were taken into consideration. Of importance, the current face stimuli all had a neutral facial expression, did not have any hair or identifying facial features (e.g., beards, freckles, moles, etc.), and were all presented straight on. All other facial features, such as face shape, eye color, skin tone, etc., were allowed to vary. Setting these initial parameters permitted sufficient variation in faces that were randomly generated but minimized the amount of variables which could influence memory performance to an extreme. Using these parameters, six sets of 30 male and female faces stimuli were created (180 male,180 female) in order to have enough face stimuli to create unique old, related, and new faces.

To create the related face stimuli, within each set of 30 faces (male and female), the faces were randomly paired with one another. Five independent raters were asked to rate each of the paired faces on a 1-6 scale (with 1 indicating the faces were very similar and 6 indicating the

faces were very dissimilar). Ratings were averaged for each pair in each set, and ranked from highest to lowest (higher ratings indicated the paired faces were more dissimilar). From these pairs, the ten highest unique face pairs in each set were used (all of these face pairs had a score equal to or above a 4, which meant the independent raters thought the faces were all at least slightly dissimilar). Face pairs that had an average score less than 3 were discarded, as this indicated the face pairs were at a minimum slightly similar. As these face pairs were used to create the related faces (by morphing between the two faces), we did not want to use face pairs that were already visually similar to one another. The scores of these ten pairs in each set were then averaged and the three highest female sets and three highest male sets were matched and further used. To create the related faces, the first face in the pair was morphed in steps of 20% to the second face. This led to the creation of six face types: original (Face 1), 20% face morph, 40% face morph, 60% face morph, 80% face morph, and new face (Face 2). A 20% face morph meant that the face was more similar (80% similar) to the original face. That is, a 20% face morph only differed from the original face by 20%.

Supplementary Figures

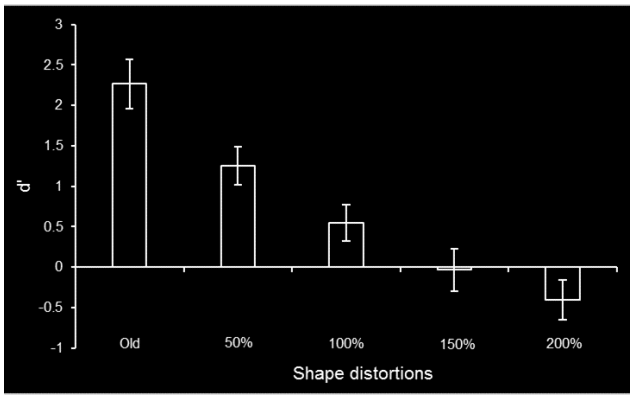

Fig. S1. *d'* results from Experiment 1a for old shapes and related shapes (50%, 100%, 150% and 200% distortions). Error bars represent 95% confidence interval.

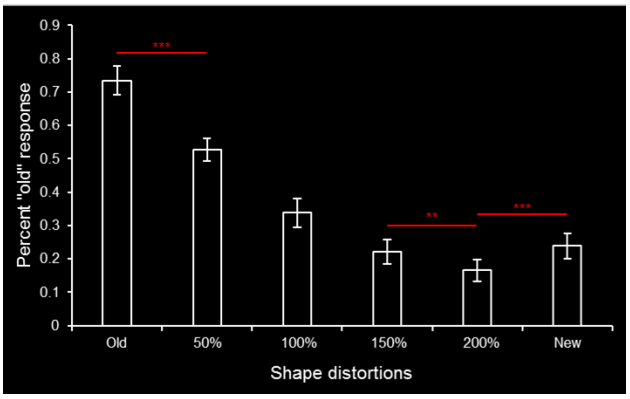

Fig. S2. The percent "old" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions), and new shapes of Experiment 1a for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval. *** = $p <$ .001, ** = $p <$ .01.

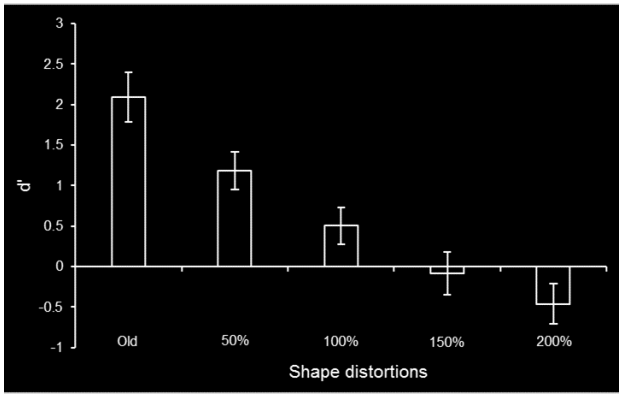

Fig. S3. *d'* results for old shapes and related shapes (50%, 100%, 150% and 200% distortions) from Experiment 1a for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval.

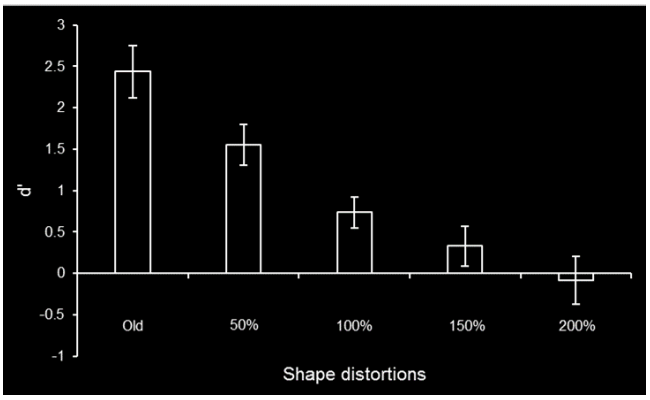

Fig. S4. *d'* results of Experiment 1b for old shapes and related shapes (50%, 100%, 150% and 200% distortions). Error bars represent 95% confidence interval.

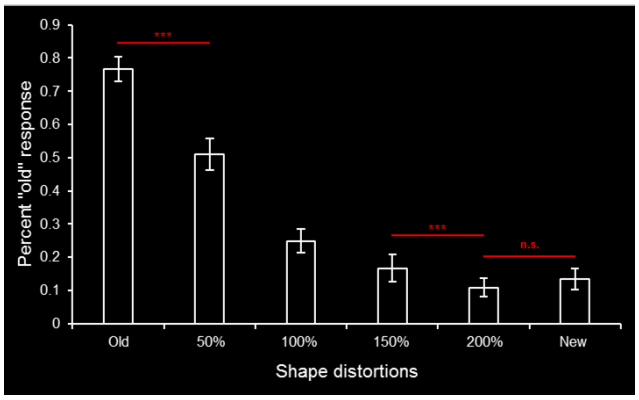

Fig. S5. The percent "old" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions), and new shapes of Experiment 1b for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval. *** = $p < .001$, n.s. = not significant.

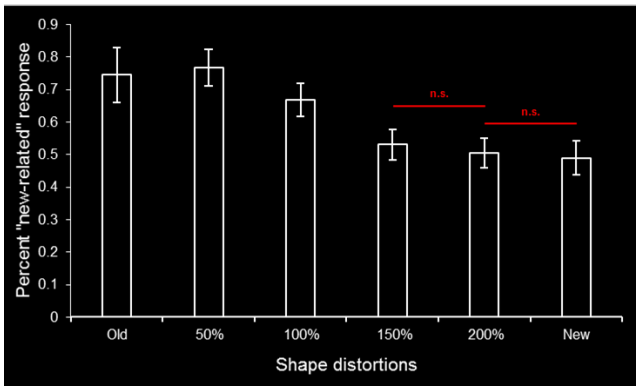

Fig. S6. The percent "new-related" response for old shapes, related shapes (50%, 100%, 150%, and 200% shape distortions) and new shapes of Experiment 1b for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval. n.s. = not significant.

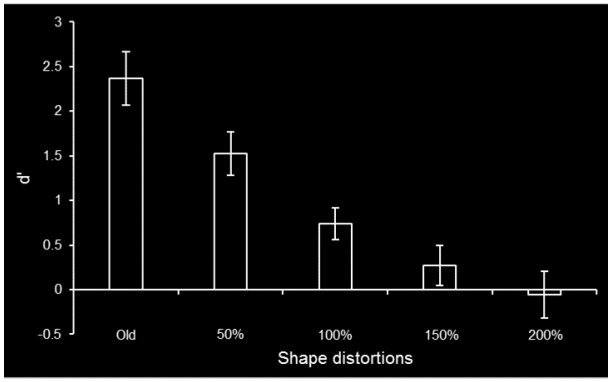

Fig. S7. *d'* results for old shapes and related shapes (50%, 100%, 150% and 200% distortions) in Experiment 1b for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confident.

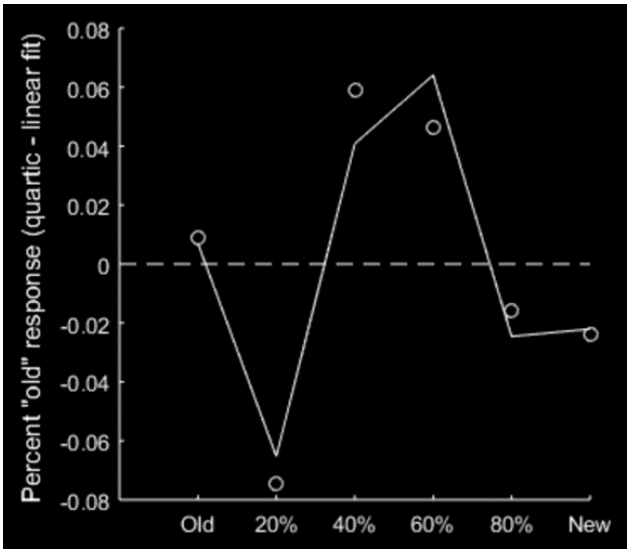

Fig. S8. The polynomial function (subtraction of the linear function from the quartic function) for the "old" responses rate for faces in Experiment 2.

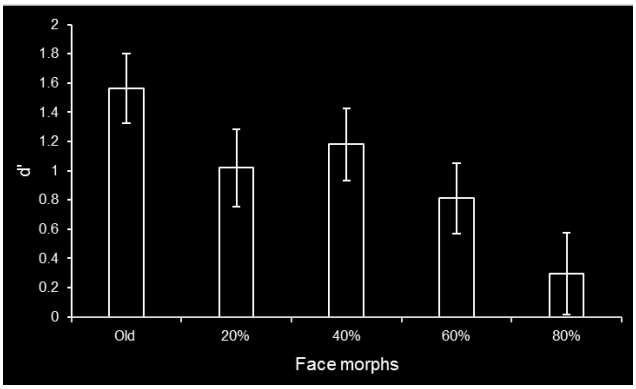

Fig. S9. *d'* results of Experiment 2 for old faces and related faces (20%, 40%, 60% and 80% morphs). Error bars represent 95% confidence interval.

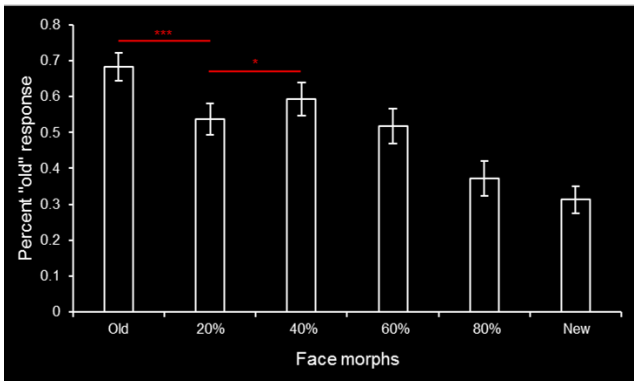

Fig. S10. The percent "old" response in Experiment 2 for old faces, related faces (20%, 40%, 60% and 80% morphs), and new faces for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval. *** = $p < .001$, * = $p < .05$.

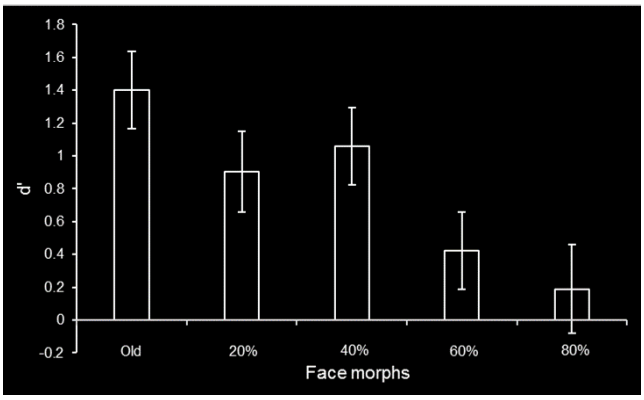

Fig. S11. *d'* results of Experiment 2 for old faces and related faces (20%, 40%, 60% and 80% morphs) for all participants (i.e., not excluding those who had below baseline performance). Error bars represent 95% confidence interval.

PART II

NEURAL EVIDENCE FOR VISUAL LONG-TERM MEMORY SPECIFICITY

CHAPTER 2.0: THE ROLE OF THE HIPPOCAMPUS IN SPATIAL LONG-TERM MEMORY

Chapter 2.1:

Distinct regions of the hippocampus are associated with memory for different spatial locations.
Brittany M. Jeye, Jessica M. Karanian, Sean P. MacEvoy, and Scott D. Slotnick

Published in *Brain Research.*
https://doi.org/10.1016/j.brainres.2018.02.029

In the present functional magnetic resonance imaging (fMRI) study, we aimed to evaluate whether distinct regions of the hippocampus were associated with spatial memory for items presented in different locations of the visual field. In Experiment 1, during the study phase, participants viewed abstract shapes in the left or right visual field while maintaining central fixation. At test, old shapes were presented at fixation and participants classified each shape as previously in the ‘‘left” or ‘‘right” visual field followed by an ‘‘unsure”-‘‘sure”-‘‘very sure” confidence rating. Accurate spatial memory for shapes in the left visual field was isolated by contrasting accurate versus inaccurate spatial location responses. This contrast produced one hippocampal activation in which the interaction between item type and accuracy was significant. The analogous contrast for right visual field shapes did not produce activity in the hippocampus; however, the contrast of high confidence versus low confidence right-hits produced one hippocampal activation in which the interaction between item type and confidence was significant. In Experiment 2, the same paradigm was used but shapes were presented in each quadrant of the visual field during the study phase. Accurate memory for shapes in each quadrant, exclusively masked by accurate memory for shapes in the other quadrants, produced a distinct activation in the hippocampus. A multi-voxel pattern analysis (MVPA) of hippocampal activity revealed a significant correlation between behavioral spatial location accuracy and hippocampal MVPA accuracy across participants. The findings of both experiments indicate that distinct hippocampal regions are associated with memory for different visual field locations.

It has long been known that hippocampal place cells in rodents fire when an animal is in a particular location in space (O'Keefe and Dostrovsky, 1971). Hippocampal place cells have also been identified in other animals, such as bats and primates (for a review, see Hartley et al., 2014). For example, single-cell recording from the monkey hippocampus has demonstrated that there are place cells that are activated by both specific locations in an experimental room and specific locations on a computer screen (Matsumura et al., 1999). Such findings in animals have given rise to the cognitive map theory of hippocampal function, where distinct hippocampal regions are associated with different spatial locations (O'Keefe and Nadel, 1978). Functional magnetic resonance imaging (fMRI) studies with humans have also suggested that the hippocampus is involved during spatial navigation (Maguire et al., 1998; Burgess et al., 2002; Ekstrom et al., 2003; Maguire et al., 2006; Zhang and Ekstrom, 2013; Howard et al., 2014). Additionally, single-cell recording from the human hippocampus has revealed that place cells are active during virtual navigation tasks (Ekstrom et al., 2003; Miller et al., 2013).

Although the evidence from fMRI studies and single-cell recording studies in humans has consistently shown that distinct hippocampal regions are associated with different spatial locations (in line with the cognitive map theory), these studies employed spatial navigation tasks which utilized maps, mazes, or movies of real-world environments. Such spatial navigation tasks involve many cognitive processes in addition to spatial memory, such as the perceptual processing of sensory cues (e.g., environmental cues and self-motion cues) and several executive mechanisms (e.g., setting navigational goals, route planning, and maintaining spatial representations; McNamara et al., 2008; Wolbers and Hegarty, 2010; Rodriguez, 2011; Chersi and Burgess, 2015; Spiers and Barry, 2015; Wolbers, 2015). These cognitive processes are further involved in regulating navigational spatial computations such as path integration, spatial updating, and wayfinding, and, critically, they are associated with brain regions that extend beyond the hippocampus, such as the striatum, the precuneus, and the entorhinal cortex (for a review, see Wolbers, 2015). Thus, the navigation tasks used in previous studies with humans have

confounded spatial memory with other spatial computations. Of importance, there has been no evidence that different regions of the human hippocampus code for different spatial locations during a task that has only involved spatial memory.

In the current fMRI study, we aimed to evaluate whether distinct regions of the human hippocampus were involved with memory for different spatial locations by utilizing paradigms that isolated visual spatial memory (to eliminate the confounds associated with spatial navigation). In Experiment 1, we evaluated whether the hippocampus was differentially associated with memory for items presented along the horizontal meridian in the left visual field or the right visual field. During the study phase, abstract shapes were presented to the left or right of fixation (Fig. 1, left). During the test phase, old shapes were presented at fixation and participants classified each shape as previously in the ‘‘left” or ‘‘right” visual field (Fig. 1, right). Experiment 1 was a re-analysis of a dataset from a previous study where we found that the magnitudes of activity in the anterior prefrontal cortex and the hippocampus were negatively correlated during false memories (Jeye et al., 2017). The current study extended our prior findings as only the present analysis assessed whether there were differential spatial location effects in the hippocampus. In Experiment 2, we evaluated whether the hippocampus was differentially associated with memory for items presented in each quadrant of the visual field.

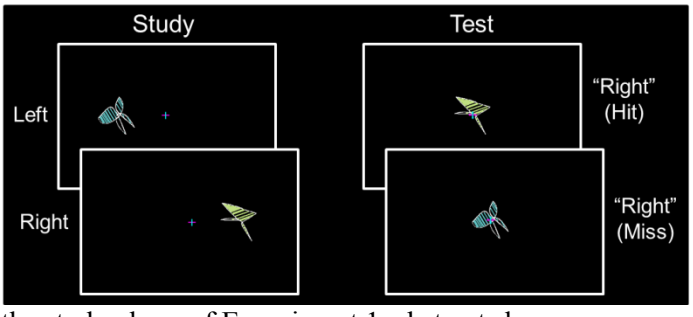

Fig. 1. Left, during the study phase of Experiment 1, abstract shapes were presented in the left or right visual field. Right, during the test phase, shapes were presented at fixation and participants classified each shape as previously on the ‘‘left” or ‘‘right” followed by an ‘‘unsure”-‘‘sure”-‘‘very sure” confidence response. Example spatial location responses are shown to the right (with response types in parenthesis).

Our analytic plan consisted of multiple tests to uncover distinct regions of the hippocampus associated with memory for different spatial locations. First, we attempted to isolate hippocampal activity associated with accurate memory for each spatial location by comparing correct spatial location responses (hits) with incorrect spatial location responses (misses). If either of the spatial locations were associated with null hippocampal activity, to increase power, we compared all correct spatial location responses and all incorrect spatial location responses. Finally, if both of the previous analyses were associated with null hippocampal activity (for either spatial location), we compared high confidence hits and low confidence hits for each spatial location.

Experiment 1

Method

Participants

Sixteen Boston College students who were right-handed, had normal or corrected-to-normal vision, and were native English speakers participated in the study (12 females, age range 22–28 years). Participants were compensated $10 for the behavioral training session and $25 per hour for the fMRI session. The Boston College Institutional Review Board approved the protocol, and informed consent was obtained prior to each session.

Stimulus protocol

During the behavioral training session, each participant completed a one-quarter length run and a full-length run. During the fMRI session, participants completed seven to eight full-length study-test runs in the scanner (Fig. 1). Instructions prior to each study phase reminded participants to remember each shape and its spatial location (i.e., whether it was on the left or right side of the screen). During the study phase, 32 shapes spanning 6.7 of visual angle were presented with their nearest edge 3.6 of visual angle in the left or right visual field (for

information on shape construction, see Slotnick and Schacter, 2004). An equal number of shapes were presented in each visual field. Each shape set was randomized and presented in sequential order and then randomized and presented a second time. Each shape was displayed for 2.5 s followed by 0.5 s of fixation. During the test phase, the shapes from the study phase (i.e., old items) were randomized and presented at fixation for 3.0 s followed by a confidence rating reminder screen for 2.5 s and a 0.5 to 4.5 s fixation period. Participants responded with their left hand to classify each shape as previously presented in the ‘‘left” or ‘‘right” visual field and then made an ‘‘unsure”-‘‘sure”-‘‘very sure” confidence rating. For both the study phase and the test phase, no more than three shapes of a given type were sequentially presented and participants were instructed to maintain central fixation. All of the participants reported that they were able to maintain central fixation. Furthermore, in a previous study that employed a very similar paradigm, with central fixation and lateralized abstract shapes at encoding, eye movements were monitored and participants maintained fixation to within 1 of visual angle from the central fixation cross (Slotnick and Thakral, 2011). In addition, in another study that used a nearly identical paradigm (Slotnick, 2009), activity at encoding was completely lateralized to contralateral early visual regions, which would only have occurred if participants maintained central fixation. Shape location was counterbalanced across participants using a Latin square design and shapes were never repeated across runs.

Data acquisition and analysis

Imaging data were acquired using a Siemens 3 Tesla Trio Scanner with a 32-channel head coil. A magnetization prepared rapid gradient echo sequence was used to acquire anatomic images (T R = 30 ms, TE = 3.3 ms, flip angle = 40°, field-of-view = 256 x 256 mm^2, acquisition matrix = 256 x 256, slices = 128, slice thickness = 1 mm; 1.33 x 1 x 1 mm resolution). An echo planar imaging sequence was used to acquire functional images (TR = 2000, TE = 30 ms, flip angle = 90°, field-of-view = 256 x 256 mm^2, acquisition matrix = 64 x 64, slices = 33, slice

acquisition order = interleaved bottom-to-top, slice thickness = 4 mm, no gap; 4 mm isotropic resolution). BrainVoyager QX (Brain Innovation B.V., Maastricht, the Netherlands) was used to conduct the analysis. Voxels were resampled at 3 mm^3. Functional pre-processing included slice-time correction, motion correction, removal of temporal components below 2 cycles per run length. To maximize spatial resolution, spatial smoothing was not conducted. Anatomic and functional images were transformed into Talairach space.

A random-effect general linear model analysis was conducted. The following event types were included in the general linear model: encoding of items in the left visual field, encoding of items in the right visual field, accurate retrieval of items in the left visual field (left-hits), accurate retrieval of items in the right visual field (right-hits), inaccurate retrieval of items in the left visual field (left-misses), inaccurate retrieval of items in the right visual field (right-misses), no response, and a constant. Unless otherwise stated, we collapsed over confidence responses to maximize power. For all contrasts, an individual voxel threshold of $p < .001$ was enforced, which yielded false discovery rate correction for multiple comparisons of $p < .05$. False discovery rate correction for multiple comparisons does not require a minimal cluster extent but rather, for a given individual voxel threshold, ensures an acceptable rate of false positives across the entire brain (Logan and Rowe, 2004). Hippocampal activations were localized on the group average anatomic volume, based on the known anatomical distinctions within the medial temporal lobe (Insausti et al., 1998; Pruessner et al., 2000; Bernasconi et al., 2003; Malykhin et al., 2007). These anatomical distinctions included the crus fornix to identify the anterior border of the hippocampal tail and the uncal apex to identify the posterior border of the hippocampus head, along with the white matter of the parahippocampal gryus to delineate the inferior borders of the hippocampus body and head.

For regions-of-interest, event-related activations timecourses were extracted from voxels within a 5 mm cube at the center of the activation from –1 to 6 s after stimulus onset (baseline corrected from –1 to 0 s). For each event type, the mean magnitude of activity from 4 to 5 s after

stimulus onset was used for statistical analysis (i.e., the expected maximum amplitude of the hemodynamic response). As the direction of comparisons were known a priori (i.e., hits > misses), one tailed t-tests were employed. To ensure statistical independence, comparisons were not performed on event-related magnitudes that defined regions-of-interest.

Results

Spatial location accuracy did not differ between shapes previously presented in the left visual field (75.5%, chance = 50%) and shapes previously presented in the right visual field (78.6%; $t(15) < 1$). The present analysis focused on hippocampal activity associated with spatial memory during the retrieval phase.

Hippocampal activity associated with accurate spatial memory for shapes in the left visual field was isolated by contrasting correct spatial location responses (left-hits) with incorrect spatial location responses (left-misses; i.e., ‘‘left”/left > ‘‘right”/left). This contrast produced two activations in the body of the left hippocampus (Fig. 2, top left, coordinates, x = –27, y = –14, z = –15, size = 54 mm^3; bottom left, x = –24, y = –19, z = –11, size = 27 mm^3; see Supplementary Material Figs. S1 and S2 for activations projected on individual participant anatomic images). Event-related activation magnitudes were extracted from both hippocampal activations. In the more anterior hippocampal activation (Fig. 2, top right), the interaction between item location (left, right) and accuracy (hits, misses) was not significant ($t(15) < 1$). Of importance, in the more posterior hippocampal activation (Fig. 2, bottom right), there was a significant interaction between item location and accuracy ($t(15) = 1.90$, $p < .05$). Hippocampal activity associated with accurate spatial memory for shapes in the right visual field was isolated by contrasting right-hits and right-misses (i.e., ‘‘right”/right > ‘‘left”/right). Unexpectedly, this contrast did not produce any activations in the hippocampus. This null finding is consistent with the event-related activation profiles corresponding to the previous contrast (Fig. 2, right), as the activations associated with right hits were more negative in magnitude than those associated with right-

misses. The results thus far suggest that the hippocampus is preferentially associated with spatial memory for items in the left visual field.

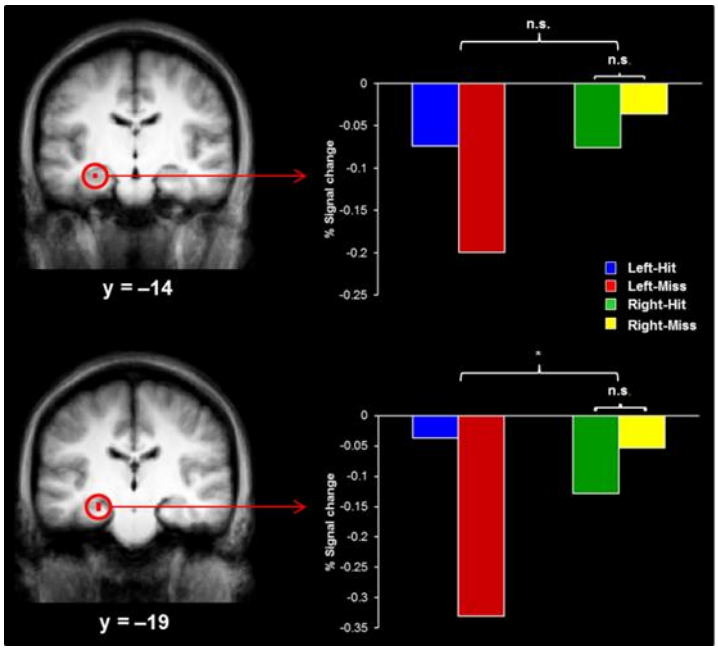

Fig. 2. Left, hippocampal activations associated with accurate spatial memory for items in the left visual field (left-hits > left-misses). Right, event-related activity extracted from the corresponding activations to the left (key to the right; * = $p < .05$, n.s. = not significant).

We conducted additional analyses in an effort to uncover hippocampal activity associated with memory for shapes in the right visual field. To increase power, all correct spatial location responses were contrasted with all incorrect spatial location responses (i.e., all-hits > all-misses). At a threshold of $p < .001$, uncorrected, this contrast produced two activations in the hippocampus (Supplementary Material Fig. S3, top left, coordinates x = –24, y = –16, z = –14, size = 54 mm^3; bottom left, coordinates x = 18, y = –34, z = 4, size = 27 mm^3). In the more anterior hippocampal activation (Supplementary Material Fig. S3, top right), there was a marginally significant difference between left-hits and left-misses ($t(15) = 1.47$, p = .082), no significant difference between right-hits and right-misses ($t(15) < 1$), and no significant interaction between item location and accuracy ($t(15) < 1$). In the more posterior hippocampal activation (Supplementary Materia lFig. S3, bottom right), there was a marginally significant

difference between left-hits and left-misses ($t(15) = 1.67$, $p = .058$), no significant difference between right-hits and right-misses ($t(15) < 1$), and no significant interaction between item location and accuracy ($t(15) < 1$). For both activations, it is worth noting that the differences in activity between hits and misses for shapes in the left visual field were marginally significant, while the differences for shapes in the right visual field were not significant, which is consistent with the findings above that the hippocampus is preferentiall associated with memory for items in the left visual field.

In a further effort to identify hippocampal activity associated with accurate spatial memory for items in the right visual field, we contrasted high confidence (‘‘very sure”) right-hits and low confidence (‘‘unsure”) right-hits. This subjective memory contrast is similar to the ‘‘remember” versus ‘‘know” contrast that has been shown to activate the hippocampus (Eldridge et al., 2000; Yonelinas et al., 2005; Montaldi et al., 2006). This contrast produced four activations in the hippocampus (Fig. 3A, coordinates x = 27, y = –16, z = –14, size = 27 mm^3, see Supplementary Material Fig. S4 for activations projected on individual participant anatomic images; B, coordinates x = –26, y = –20, z = –11, size = 89 mm^3; C, coordinates x = 26, y = –29, z = –3, size = 45 mm^3; D, coordinates x = –25, y = –31, z = –5, size = 51 mm^3). Of importance, in the activation within the body of the right hippocampus (Fig. 3A, top right), there was no significant difference between left-hits and left-misses ($t(15) < 1$) and there was a significant interaction between item location and confidence ($t(15) = 2.63$, $p < .05$). For the other three hippocampal activations, there were significant differences between high confidence left-hits and low confidence left-hits (Fig. 3B, $t(15) = 2.66$, $p < .05$; Fig. 3C, $t(15) = 2.79$, $p < .05$; Fig. 3D, $t(15) = 2.84$, $p < .05$), and the interactions between location and confidence were not significant (all $t(15)$-values < 1). The contrast of high confidence (‘‘very sure”) left-hits and low confidence (‘‘unsure”) left-hits produced a single activation in the body of the hippocampus (Supplementary Material Fig. S5, left, coordinates x = –24, y = –21, z = –11, size = 65 mm^3). For this activation, there was a significant difference between high confidence right-hits and low confidence right-

misses ($t(15) = 4.14$, $p < .05$), and there was no significant interaction between location and confidence ($t(15) < 1$). These results show that the contrast of high confidence versus low confidence right-hits can produce activity in the hippocampus.

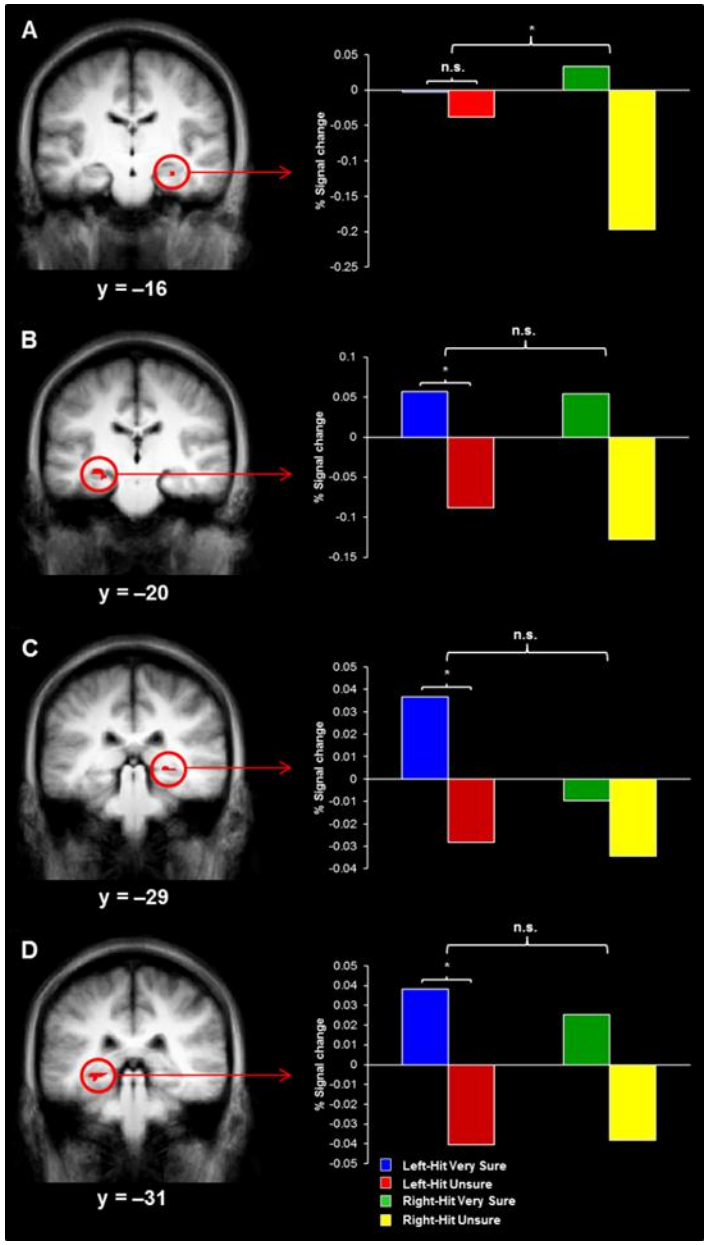

Fig. 3. Left, hippocampal activity associated with spatial memory for high confidence right-hits > low confidence right-hits. Right, event-related activity extracted from the activation to the left (key at the bottom; * = $p < .05$, n.s. = not significant).

Although we were not able to uncover hippocampal activity associated with accurate versus inaccurate spatial memory for items in the right visual field collapsed over confidence, activity was observed for the comparison of high confidence and low confidence right-hits. That the contrast of right-hits and right-misses did not produce activity in the hippocampus suggests that this region may be preferentially associated with spatial memory for items in the left visual field. This differential hippocampal activity is seemingly at odds with the behavioral results, which showed similar levels of spatial memory performance for items in both visual fields. This similar behavioral performance suggests that there are brain regions beyond the hippocampus that underlie accurate spatial memory for items in the right visual field. To isolate this activity, a whole-brain analysis was conducted using the conjunction (right-hits > right-misses) ∩ (right-hits > left-hits). This conjunction produced one activation that was located in language processing cortex/Wernicke's area (Fig. 4, BA 40, x = –57, y = –34, z = 28, size = 27 mm^3). The analogous conjunction (left-hits > left-misses) ∩ (left-hits > right-hits) did not produce any activations. These whole-brain results suggest that accurate spatial memory for items in the right visual field is mediated by language processing cortex to a greater degree than accurate spatial memory for items in the left visual field.

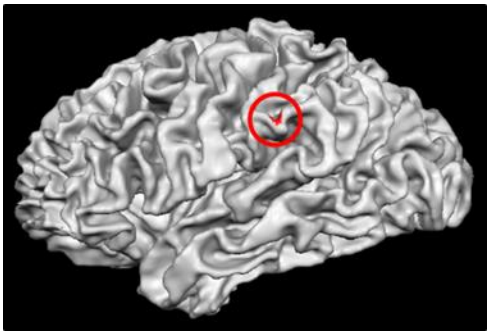

Fig. 4. Whole-brain activity associated with accurate spatial memory for shapes in the right visual field identified using the conjunction (right-hits > right-misses) ∩ (right-hits > left-hits). The activation is circled.

Experiment 2

The right-hit versus right-miss null findings in Experiment 1 might have been due, in part, to more efficient processing of items presented along the horizontal meridian coupled with

language dominance in the contralateral/left hemisphere. Providing some evidence for more efficient processing along the horizontal meridian, orientations near the horizontal and vertical orientations have been shown to be more efficiently processed (i.e., associated with higher accuracy and lower reaction times) than oblique orientations (i.e., the ‘‘oblique effect”; Appelle, 1972). Thus, we predicted that stimuli presented in the center of each quadrant (far from the meridians) might be less susceptible to verbal encoding strategies and thus the hit versus miss contrast could produce hippocampal activations for these stimulus locations. In addition, using four stimulus locations more closely mirrors the rodent literature, which typically utilizes numerous different locations.

During the study phase of Experiment 2, participants viewed abstract shapes presented in the upper-left quadrant, the lower-left quadrant, the upper-right quadrant, or the lower-right quadrant (Fig. 5, left). During the test phase, old shapes were presented and participants classified each shape as previously in the ‘‘upper-left”, ‘‘lower-left”, ‘‘upper-right”, or ‘‘lower-right” (Fig. 5, right).

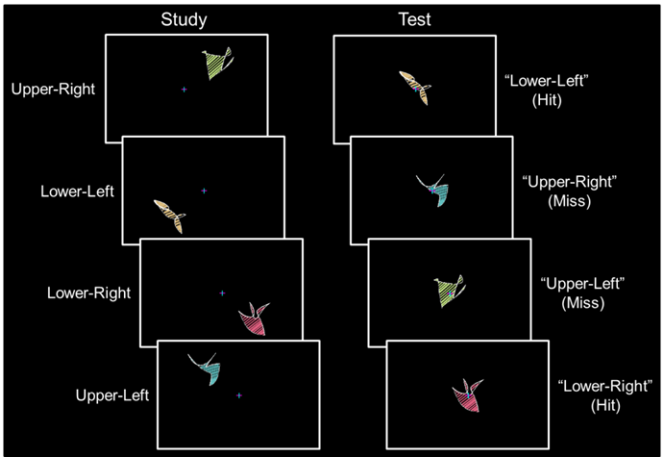

Fig. 5. Left, during the study phase of Experiment 2, abstract shapes were presented in the upper left, lower left, upper right or lower right quadrant of the visual field. Right, during the test phase, shapes were presented at fixation and participants classified each shape as previously in the ‘‘upper-left”, ‘‘lower-left”, ‘‘upper-right” or ‘‘lower-right” followed by an ‘‘unsure”-‘‘sure”-‘‘very sure” confidence response. Example spatial location responses are shown to the right (with response types in parenthesis).

Methods

Unless otherwise specified, the materials and methods of Experiment 2 were identical to those of Experiment 1.

Participants

Sixteen right-handed participants recruited from the Boston College community completed the study (13 females, age range 20–29 years).

Stimulus protocol

During the fMRI session, participants completed seven to eight runs. During the study phase of each run 32 shapes spanning 3.8° of visual angle were presented with their nearest edge 2.1° of visual angle up or down and to the left or right of fixation in the upper-left, lower-left, upper-right or lower-right visual field quadrant. Participants pressed response buttons with their left hand to classify each shape as previously presented in the ‘‘upper-left”, ‘‘lower-left”, ‘‘upper-right” or ‘‘lower-right” quadrant of the visual field.

General linear model analysis

To isolate unique hippocampal activity associated with spatial memory for each quadrant in the visual field, we compared hits and misses at $p < .001$, false discovery rate corrected for multiple comparisons to $p < .05$, and exclusively masked that activity with the contrast of hits and misses for the other three quadrants at a liberal threshold of $p < .05$, uncorrected.

Multi-voxel pattern analysis

Data were analyzed using SPM 12 (Wellcome Trust Centre for Neuroimaging, London, United Kingdom). Pre-processing included slice-time correction, motion correction (registered to first image of each run), and spatial normalization to the Montreal Neurological Institute (MNI)

template (voxels were resampled at 2 mm^3). Spatial smoothing was not conducted. Anatomic images were also transformed into MNI space. A general linear model analysis was conducted for each run of each participant (using a high-pass filter cutoff of 128 s). For each of the runs, we isolated activity associated with hits at retrieval corresponding to each of the four quadrants. The patterns of activity in the hippocampus corresponding to accurate memory in each quadrant for each run were used as vectors in the MVPA analysis.

The MVPA analysis was conducted using custom scripts written in MATLAB (MathWorks, Natick, Massachusetts). The hippocampal voxels were identified based on the group anatomic image. Specifically, all the hippocampal voxels were selected that spanned the most posterior activation (y = –33) and the most anterior activation (y = –9) identified in the general linear model analysis.

For each participant, response patterns were subsequently limited to those voxels that had non-zero values for all stimulus positions and runs. Individual voxel values within the pattern for each position in each run were then normalized such that each pattern had a length of 1 when considered as a vector in high-dimensional space. This was done in order to eliminate potential contamination by any differences in overall signal magnitude across patterns. The set of response patterns was then split into halves by run (e.g., even runs versus odd runs) for classification analysis. Each of the following steps was repeated for each possible run-wise data split. Patterns for each stimulus position in each data half were averaged and then de-trended by subtracting the average magnitude. To assess whether patterns in one data half could be classified based on patterns in the opposite half, we simply asked whether the Euclidean distance (i.e., square root of the sum of the squared voxel-by-voxel pattern differences) between the patterns evoked by one position in the two data halves was shorter than the distance between the patterns for that position and some other position in the two data halves (Haxby et al., 2001; MacEvoy and Epstein, 2009, 2011). If it was, a correct classification decision was recorded. The number of correct decisions was accumulated across all pairwise position matchups and divided by the total number of

matchups to generate an accuracy rate for a singlerun-wise data split. Each participant's accuracy rate was their average across all splits. Because this analysis was conducted as a series of pairwise comparisons between stimulus positions, chance accuracy was 50%.

Results

A repeated measures analysis of variance revealed that spatial location accuracy differed between the upper-left quadrant (‘‘upper-left”/upper-left = 58.2%, chance = 25%), the lower-left quadrant (‘‘lower-left”/lower-left = 65.1%), the upper-right quadrant (‘‘upper-right”/upper-right = 66.7%), and the lower-right quadrant (‘‘lower-right”/lower-right = 58.1%; $F(1, 15) = 7.58$, $p < .05$). Collapsing across stimuli within the left visual field (upper-left, lower-left) and right visual field (upper-right, lower-right), there was no significant difference in spatial location accuracy between the left visual field (61.6%, chance = 50%) and the right visual field (62.4%; $t(15) < 1$), which is consistent with the behavioral results from the Experiment 1.

The contrasts of hits and misses for each quadrant, exclusively masked with the contrast of hits and misses for the other three quadrants, revealed that accurate spatial memory for shapes in each quadrant of the visual field was associated with a distinct hippocampal region. For the upper-left quadrant, there was one activation in the body of the hippocampus (Fig. 6, top left, coordinates x = –24, y = –16, z = –14, size = 54 mm^3; Supplementary Material Fig. S6 for activations projected on individual participant anatomic images). For the lower-left quadrant, there was one activation in the head of the hippocampus (Fig. 6, bottom left, coordinates x = 12, y = –9, z = –14, size = 27 mm^3; Supplementary Material Fig. S7 for activations projected on individual participant anatomic images), for the upper-right quadrant, there was one activation in the body of the hippocampus (Fig. 6, top right, coordinates x = 27, y = –19, z = –10, size = 27 mm^3; Supplementary Material Fig. S8 for activations projected on individual participant

anatomic images), and for the lower-right quadrant, there was one activation in the tail of the hippocampus (Fig. 6, bottom right, coordinates x = 18, y = –33, z = 1, size = 54 mm^3; Supplementary Material Fig. S9 for activations projected on individual participant anatomic images).

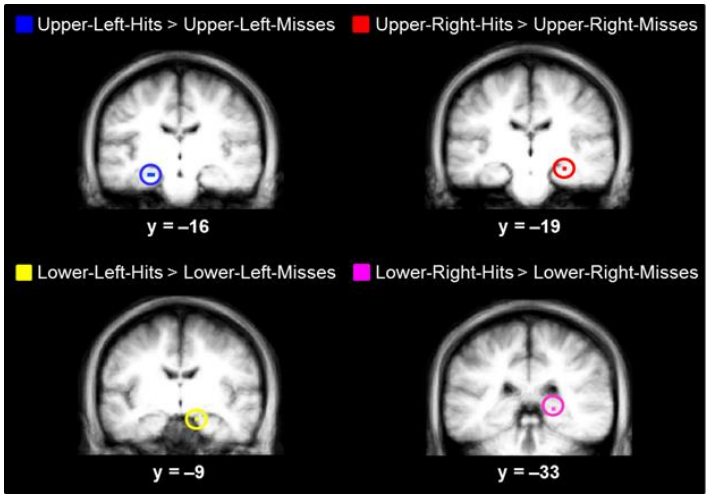

Fig. 6. Hippocampal activity associated with accurate spatial memory for different quadrants of the visual field (hits > misses for each quadrant exclusively masked by hits > misses for the other quadrants).

The previous general linear model analysis results provide evidence that one region of the hippocampus is associated with one spatial location. However, individual spatial locations may be represented by a pattern of activity across the hippocampus. To investigate this possibility, a follow-up multi-voxel pattern analysis (MVPA) assessed whether there were unique patterns of hippocampal activity associated with hits in each quadrant. A pattern classifier was unable to distinguish patterns evoked by items in each quadrant at a rate above chance accuracy (48.6%, chance = 50%, $t(15) < 1$). However, there was a significant correlation between behavioral spatial location accuracy and hippocampal MVPA accuracy (Fig. 7, $r(14) = .50$, $p < .05$). This suggests that the patterns of activity in the hippocampus contain information about spatial location, particularly for participants with higher spatial memory accuracy, which is a topic of future research.

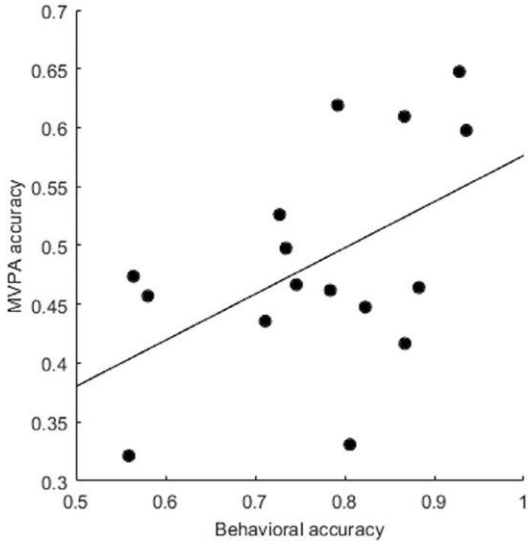

Fig. 7. For each participant, hippocampal MVPA accuracy is plotted as a function of behavioral spatial location accuracy (the best-fit line is shown).

Discussion

The present results indicate that unique hippocampal regions are associated with different visual field locations during memory. In Experiment 1, the contrast between left-hits and left-missesproduced one hippocampal activation in which there was a significant interaction between item location and accuracy and the contrast between high confidence right-hits and low confidence right-hits produced one activation in which there was a significant interaction between item location and confidence. In Experiment 2, the contrast of hits and misses for each quadrant, exclusively masked by the contrast of hits and misses for the other quadrants, produced distinct hippocampal activations. Of direct relevance to our aim, these findings demonstrate that distinct regions of the human hippocampus are involved with memory for different spatial locations. Our paradigms also isolated the process of spatial memory and thus eliminated confounds, such as the executive mechanisms involved in setting goals, route planning, and maintaining spatial representations that have been associated with previous spatial navigation studies.

The results from Experiment 1, where left-hits versus left-misses but not right-hits versus right-misses produced hippocampal activity, suggest that the hippocampus may be preferentially

associated with spatial memory for items presented on the horizontal meridian in the left visual field. These findings replicate the results of a previous fMRI study in which the hippocampus was only associated with spatial memory for items presented on the horizontal meridian in the left visual field (Slotnick and Thakral, 2013). These results support Kosslyn's (1987) hemispheric processing distinction in which categorical processing has been associated with the left hemisphere and coordinate processing has been associated with the right hemisphere (for reviews, see Slotnick et al., 2001; Baumann et al., 2012). This hemispheric processing distinction has also been extended to memory, as the left dorsolateral prefrontal cortex has been associated with categorical visual-spatial memory and the right dorsolateral prefrontal cortex has been associated with coordinate visual-spatial memory (Slotnick and Moo, 2006). In this framework, categorical processing refers to general spatial location processing between an item and a spatial reference (e.g., the lemonade is on the ''left" side of the table), while coordinate processing refers to specific spatial location processing between an item and a spatial reference (e.g., the lemonade is 5 in. from the left side of the table). Based on the known retinotopic organization of the visual processing stream, the left visual field initially maps onto the right hemisphere. As the right hemisphere is preferentially associated with coordinate processing, this could explain why the hippocampus was preferentially associated with items in the left visual field in Experiment 1. In contrast, since the right visual field initially maps onto the left hemisphere, which is associated with categorical/language processing, this could explain why the language processing cortex was preferentially associated with items in the right visual field in Experiment 1. It is also notable that, in Experiment 1, accurate memory for items in the left visual field produced two activations in the left hippocampus (Fig. 2), which replicated Slotnick and Thakral (2013) who similarly found that accurate memory for items in the left visual field produced one activation in the left hippocampus. However, in Experiment 2, accurate memory for items in the left visual field produced one activation in the left hippocampus and one activation in the right hippocampus.

Across all these experiments, the relative number of activations in the left hippocampus (4/5) was not significant (Binomial test, $p = .19$).

In Experiment 2, we found that memory for items in each quadrant of the visual field produced distinct activations in the hippocampus. This is consistent with current neuroimaging literature on the oblique effect, where oblique orientations, as compared to orientations near the horizontal or vertical meridians, have been associated with relatively greater activity in primary visual cortex (V1) (Mannion et al., 2010; Koelewijn et al., 2011; Maloney and Clifford, 2015). The increase in V1 activity for oblique orientations may translate into more visual spatial processing at oblique locations (i.e., locations that are far from the meridians) and less visual spatial processing along the meridians. The lower degree of visual spatial processing along the horizontal meridian along with the language dominance of the contralateral/left hemisphere could explain the null right-hit versus right-miss hippocampal findings in Experiment 1.

Previous research has suggested that there is a long-axis gradient in terms of hippocampal anatomy, connectivity, and function (for a review, see Strange et al., 2014). For example, the anterior hippocampus has been associated with more global representations and the posterior hippocampus has been associated with more local representations (Poppenk et al., 2013). However, we found that accurate spatial memory for shapes in different visual field quadrants produced activity distributed throughout the hippocampus in the anterior-posterior direction. Although our results do not support functional heterogeneity of the hippocampus along the long axis, this may have been due to our particular analysis strategy (i.e., our results are not mutually exclusive with the functional heterogeneity view).

The current findings provide the first fMRI evidence in humans that distinct regions of the hippocampus are associated with memory for different spatial locations. As fMRI averages activity over millimeters, this suggests that there are patches of hippocampal cortex that respond similarly to memory for a specific location in the visual field. If place cells had been more randomly distributed across the hippocampus, the activity could not have been detected with

fMRI. One line of future research will be to measure the connectivity/interactions between the distinct regions of the hippocampus associated with memory for items in different visual field locations and other cortical regions.

While it is possible that participants utilized different strategies for remembering an item's visual field location, such as assigning verbal labels to each item, our results show that they, in large part, visualized each item during retrieval. In Experiment 1, spatial memory for items in the left visual field produced activity in one region of the hippocampus (identified by contrasting left-hits and left-misses) and spatial memory for items in the right visual field produced activity in another region of the hippocampus (identified by contrasting high versus low confident right-hits). In Experiment 2, accurate spatial memory for items in each quadrant was associated with distinct regions of the hippocampus. If participants had used verbal encoding strategies for items in different visual field locations, no hippocampal differences would have been expected. Moreover, in a previous study that used the same type of stimuli and paradigm as Experiment 1, accurate spatial memory for shapes that were presented both hemifields activated contralateral/retinotopic early visual regions (Slotnick, 2009). These findings suggest that participants predominantly used visual-spatial strategies in the present study. Another possibility is that response hand congruency (e.g., making a response with the left hand for items previously presented in the left visual field) might have produced differential hippocampal activation results (i.e., in Experiment 1, left-hits versus left-misses but not right-hits versus right-misses produced hippocampal activity). However, in Experiment 1, there was no significant interaction between item spatial location (left, right) and the distribution of confidence ratings (‘‘unsure”, ‘‘sure”, ‘‘very sure”; $F(5, 30) < 1$). In Experiment 2, hand congruency was not a factor as memory for items in each of the four spatial locations produced activity in the hippocampus (i.e., activity was observed for each quadrant, regardless of response hand congruency). Furthermore, in both experiments, there was no significant difference in spatial location accuracy between the left

visual field and the right visual field. These results indicate that response hand congruency did not affect the hippocampal activations observed in the present study.

In the field of human memory, the hippocampus has been hypothesized to be associated with general relational memory, where this region is thought to bind item information and contextual information (i.e., the binding-in-context model; Eichenbaum et al., 2007; Ranganath, 2010; Schiller et al., 2015). The relation between an item and its context can be spatial in nature, such as linking an item to its location on the screen (Cansino et al., 2002; Ross and Slotnick, 2008; Slotnick, 2010), or non-spatial in nature, such as associating an item to its color (Ranganath et al., 2004; Weis et al., 2004; Staresina and Davachi, 2008; Tendolkar et al., 2008). While the current study investigated spatial memories, it is possible that distinct regions of the hippocampus may also code for non-spatial memories. For example, different regions of the hippocampus may be associated with memory for the previous color of an item (such as when items were previously presented in either red or green). In rodents, the hippocampus has also been shown to be associated with memory for odors (e.g., Fortin et al., 2004), memory for temporal information (for a review, see Eichenbaum, 2014), and social processing (for a review, see Montagrin et al., 2017). As such, there may be distinct hippocampal regions in humans associated with different color, odor, temporal, or social contexts. Another possibility is that distinct hippocampal regions are only associated with spatial memory, while the same regions of the hippocampus are associated with non-spatial memory. Although distinct regions of the hippocampus have not been previously associated with memory for different types of contextual information, it is uncertain whether this is because the corresponding contrasts were not conducted or whether they were conducted and produced null findings. Future fMRI studies that employ context memory paradigms should investigate this issue.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the" oblique effect" in man and animals. *Psychological Bulletin, 78*, 266.

Baumann, O., Chan, E., & Mattingley, J. B. (2012). Distinct neural networks underlie encoding of categorical versus coordinate spatial relations during active navigation. *NeuroImage, 60*, 1630–1637.

Bernasconi, N., Bernasconi, A., Caramanos, Z., Antel, S. B., Andermann, F., & Arnold, D. L. (2003). Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain, 126*, 462–469.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron, 35*, 625–641.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex, 12*, 1048–1056.

Chersi, F., & Burgess, N. (2015). The Cognitive Architecture of Spatial Navigation: Hippocampal and Striatal Contributions. *Neuron, 88*, 64–77.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience, 15*, 732–744.

Eichenbaum, H., Yonelinas, A. R., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123.

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature, 425*, 184–188.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience, 3*, 1149-1152.

Hartley, T., Lever, C., Burgess, N., & O'Keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*, 20120510.

Howard, L.R., Javadi, A.H., Yu, Y., Mill, R.D., Morrison, L.C., Knight, R., Loftus, M.M., Staskute, L., & Spiers, H.J. (2014). The hippocampus and entorhinal cortex encode the path and Euclidean distances to goals during navigation. *Current Biology, 24*, 1331–1340.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., Laakso, M. P., & Pitkänen, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology, 19*, 659–671.

Jeye, B. M., Karanian, J. M., & Slotnick, S. D. (2017). The anterior prefrontal cortex and the hippocampus are negatively correlated during false memories. *Brain Sciences, 7*, 13.

Koelewijn, L., Dumont, J. R., Muthukumaraswamy, S. D., Rich, A. N., & Singh, K. D. (2011). Induced and evoked neural correlates of orientation selectivity in human visual cortex. *NeuroImage, 54*, 2983-2993.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review, 94*, 148.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: a human navigation network. *Science, 280*, 921–924.

Maguire, E.A., Woollett, K., & Spiers, H.J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus, 16*, 1091–1101.

Maloney, R. T., & Clifford, C. W. (2015). Orientation anisotropies in human primary visual cortex depend on contrast. *NeuroImage, 119*, 129-145.

Malykhin, N. V., Bouchard, T. P., Ogilvie, C. J., Coupland, N. J., Seres, P., & Camicioli, R. (2007). Three-dimensional volumetric analysis and reconstruction of amygdala and hippocampal head, body and tail. *Psychiatry Research: Neuroimaging, 155*, 155–165.

Mannion, D. J., McDonald, J. S., & Clifford, C. W. (2010). Orientation anisotropies in human visual cortex. *Journal of Neurophysiology, 103*, 3465-3471.

Matsumura, N., Nishijo, H., Tamura, R., Eifuku, S., Endo, S., & Ono, T. (1999). Spatial-and task-dependent neuronal responses during real and virtual translocation in the monkey hippocampal formation. *The Journal of Neuroscience, 19*, 2381–2393.

McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human spatial memory and navigation. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory*. Volume 2 in J. Byrne (Editor-in-chief), Learning and memory: A comprehensive reference (pp. 157-178). Oxford: Elsevier.

Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S. M., Jacobs, J., Kahana, M. J., & Schulze-Bonhage, A. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science, 342,* 1111–1114.

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus, 16*, 504-520.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research, 34*, 171–175.

Pruessner, J. C., Li, L. M., Serles, W., Pruessner, M., Collins, D. L., Kabani, N., Lupiem, S., & Evans, A. C. (2000). Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cerebral Cortex, 10*, 433–442.

Ranganath, C. (2010). A unified framework for the functional organization of the medial

temporal lobes and the phenomenology of episodic memory. *Hippocampus, 20*, 1263–1290.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42*, 2–13.

Rodriguez P.F. (2011) Neural decoding of goal locations in spatial navigation in humans with fMRI. *Human Brain Mapping, 31*, 391–397

Ross, R. S., & Slotnick, S. D. (2008). The hippocampus is preferentially associated with memory for spatial context. *Journal of Cognitive Neuroscience, 20*, 432–446.

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and Space: Towards an Understanding of the Cognitive Map. *The Journal of Neuroscience, 35*, 13904–13911.

Slotnick, S. D. (2010). Does the hippocampus mediate objective binding or subjective remembering? *NeuroImage, 49*, 1769–1776.

Slotnick, S. D. (2009). Rapid retinotopic reactivation during spatial memory. *Brain Research, 1268*, 97–111.

Slotnick, S. D., & Moo, L. R. (2006). Prefrontal cortex hemispheric specialization for categorical and coordinate visual spatial memory. *Neuropsychologia, 44*, 1560–1568.

Slotnick, S. D., Moo, L. R., Tesoro, M. A., & Hart, J. Jr. (2001). Hemispheric asymmetry in categorical versus coordinate visuospatial processing revealed by temporary cortical deactivation. *Journal of Cognitive Neuroscience, 13*, 1088–1096.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7,* 664–672.

Slotnick, S. D., & Thakral, P. P. (2013). The hippocampus operates in a threshold manner during spatial source memory. *NeuroReport, 24*, 265–269.

Spiers, H. J., & Barry, C. (2015). Neural systems supporting navigation. *Current Opinion in Behavioral Sciences, 1*, 47–55.

Staresina, B. P., & Davachi, L. (2008). Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. *Journal of Cognitive Neuroscience, 20*, 1478–1489.

Tendolkar, I., Arnold, J., Petersson, K. M., Weis, S., Brockhaus-Dumke, A., Van Eijndhoven, P., Buitelaar, J., & Fernández, G. (2008). Contributions of the medial temporal lobe to declarative memory retrieval: Manipulating the amount of contextual retrieval. *Learning & Memory, 15*, 611–617.

Weis, S., Specht, K., Klaver, P., Tendolkar, I., Willmes, K., Ruhlmann, J., Elger, C.E., & Fernández, G. (2004). Process dissociation between contextual retrieval and item recognition. *NeuroReport, 15*, 2729-2733.

Wolbers, T. (2015). Spatial Navigation. In: James D. Wright (editor-in-chief), *International Encyclopedia of the Social & Behavioral Science*s, 2nd edition, Vol 23. (pp. 161–171). Oxford: Elsevier.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences, 14*, 138–146.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience, 25*, 3002-3008.

Zhang, H., & Ekstrom, A.D. (2013). Human neural systems underlying rigid and flexible forms of allocentric spatial representation. *Human Brain Mapping, 34*, 1070–1087.

Supplementary Material

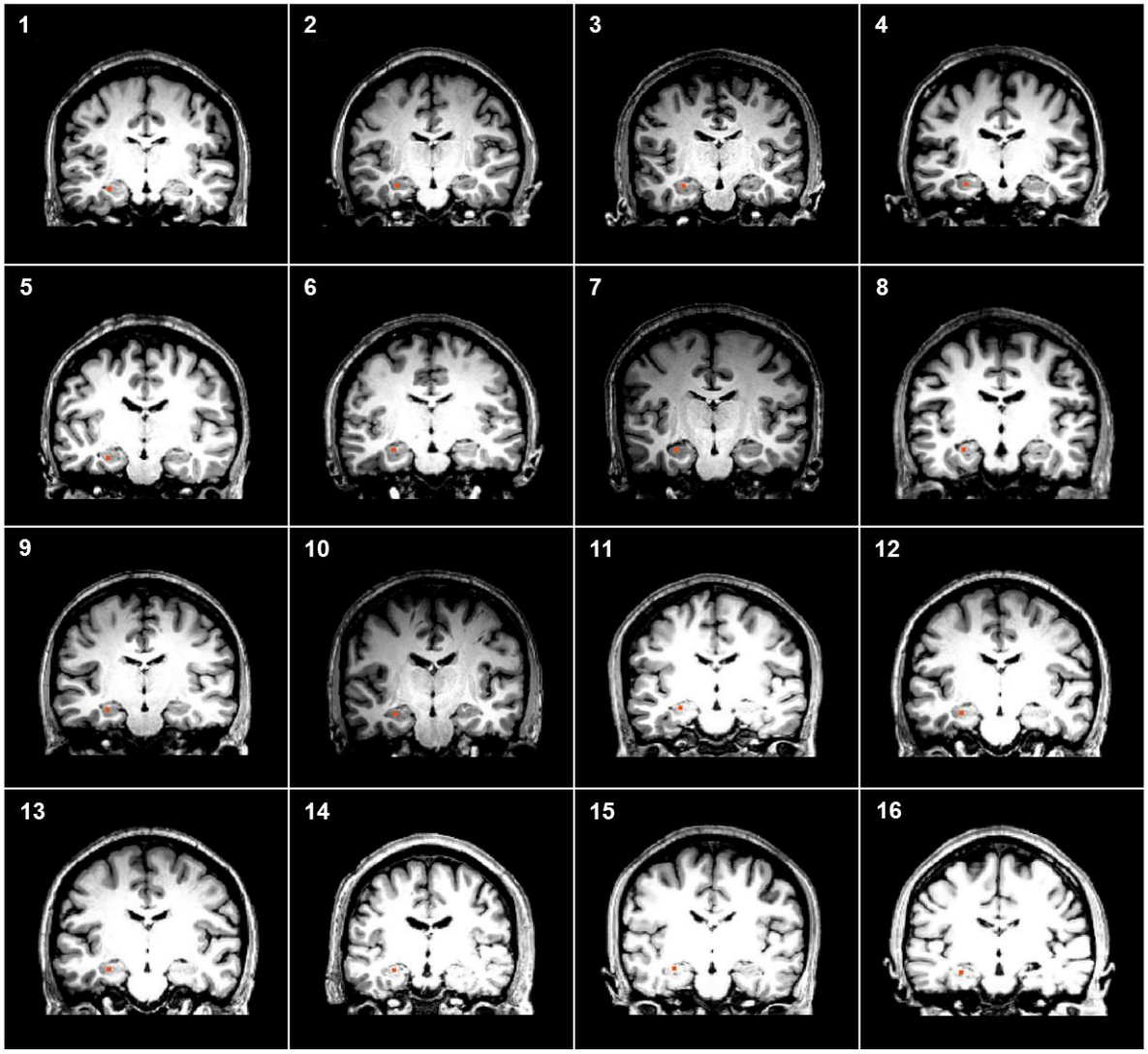

Fig. S1. More anterior group hippocampal activation associated with left-hits > left-misses projected onto each participant's anatomic image (coordinates, x = –27, y = –14, z = –15).

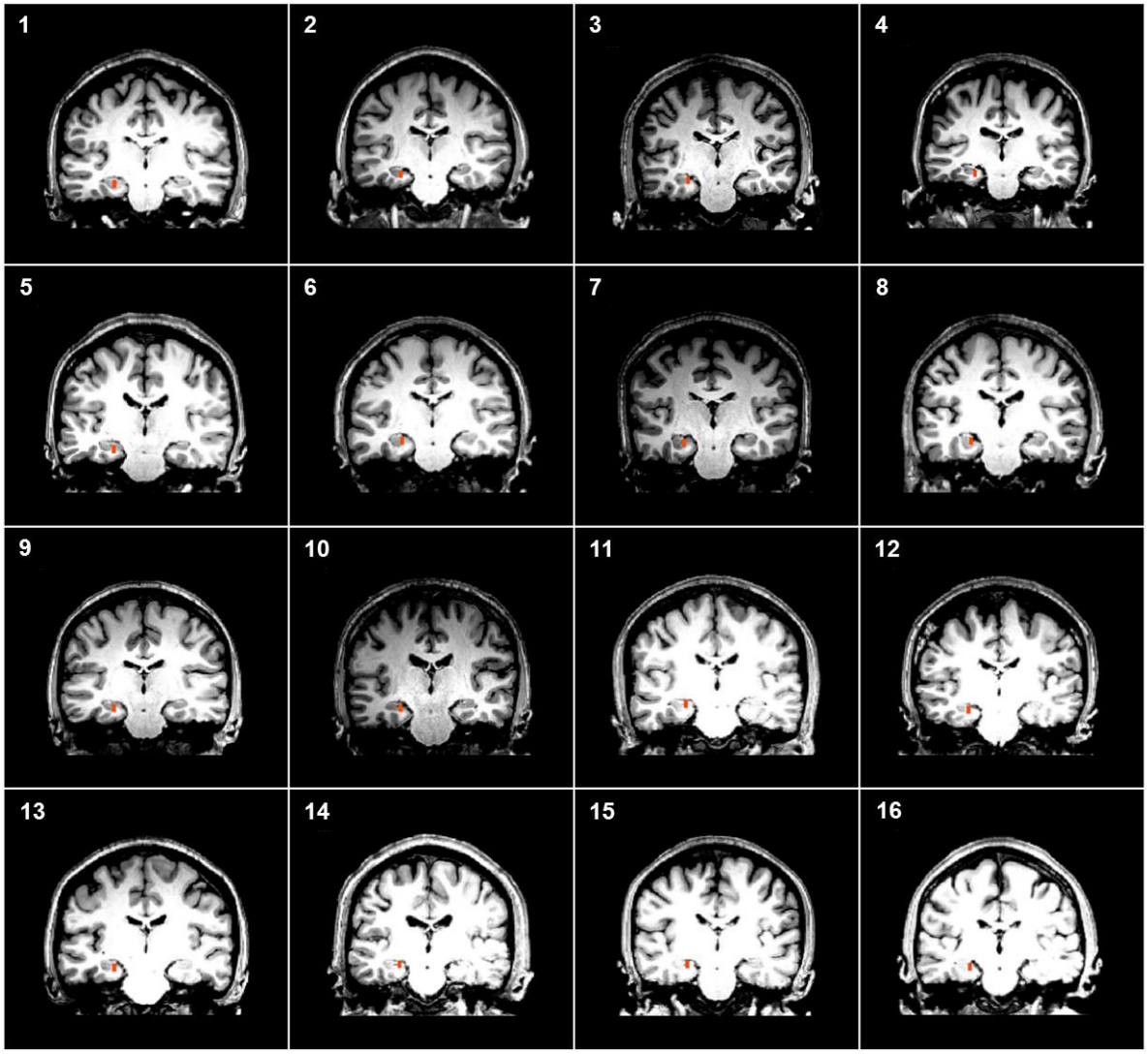

Fig. S2. More posterior group hippocampal activation associated with left-hits > left-misses projected on each participant's anatomical image (coordinates, x = –24, y = –19, z = –11).

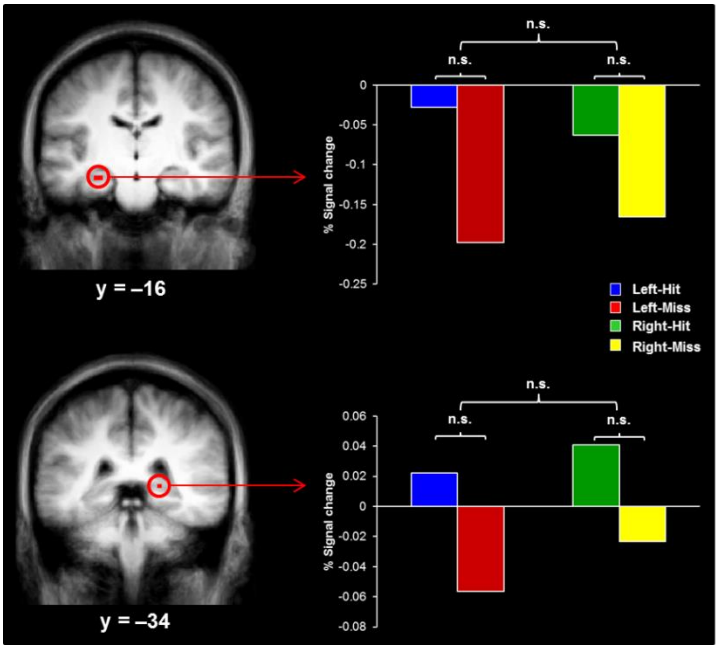

Fig. S3. Left, hippocampal activations associated with all-hits > all-misses (p <.001, uncorrected). Right, event-related activity extracted from the corresponding activations to the left (key to the right; * = p < .05, n.s. = not significant).

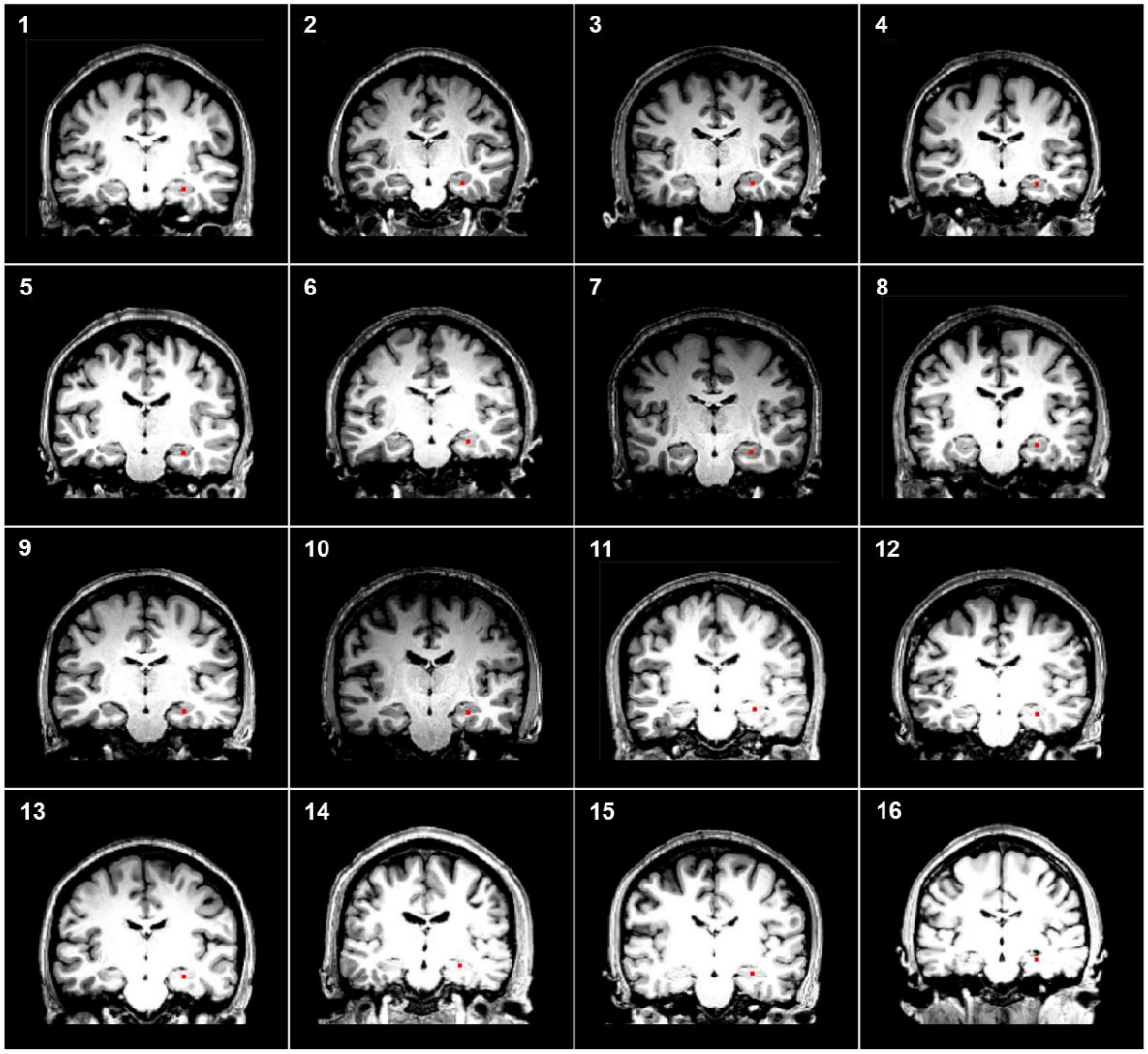

Fig. S4. Most anterior group hippocampal activation associated with high confidence right-hits > low confidence right-hits projected on each participant's anatomical image (coordinates, x = 27, y = –16, z = –14).

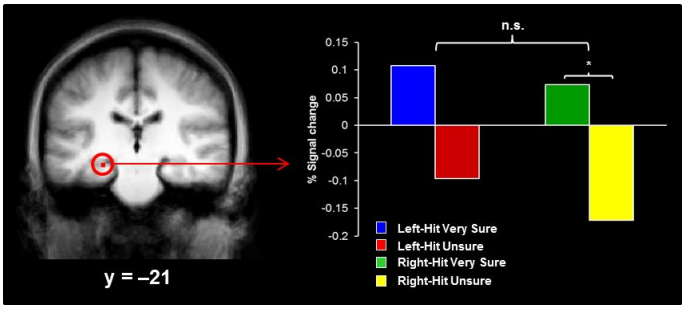

Fig. S5. Left, hippocampal activity associated with high confidence left-hits > low confidence left-hits. Right, event-related activity extracted from the corresponding activations to the left (key at the bottom; * = $p < .05$, n.s. = not significant).

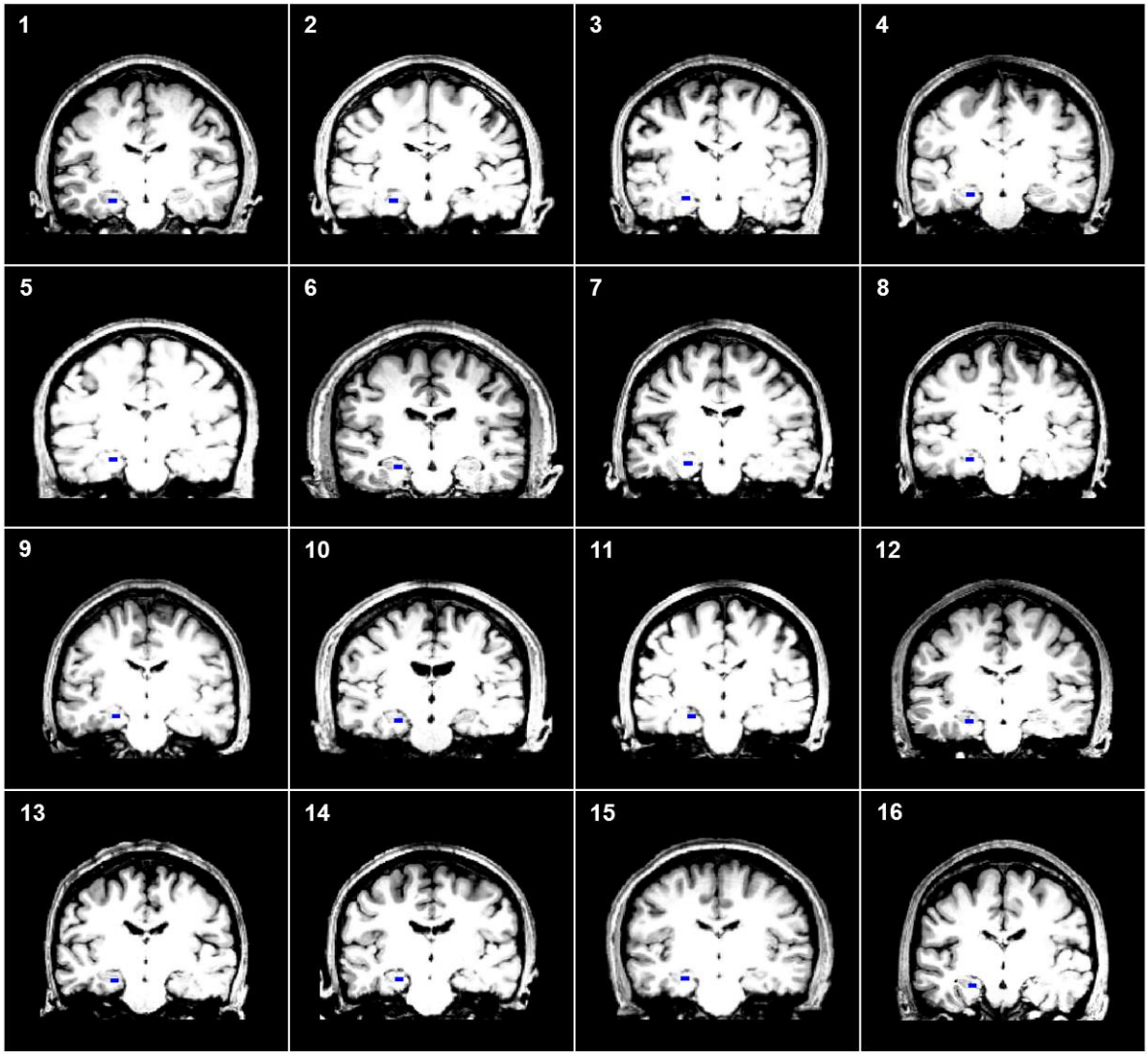

Fig. S6. Group hippocampal activation associated the contrast of hits versus misses for shapes in the upper-left quadrant, exclusively masked with the contrast of hits and misses for the other three quadrants projected onto each participant's anatomic image (coordinates, $x = -24$, $y = -16$, $z = -14$).

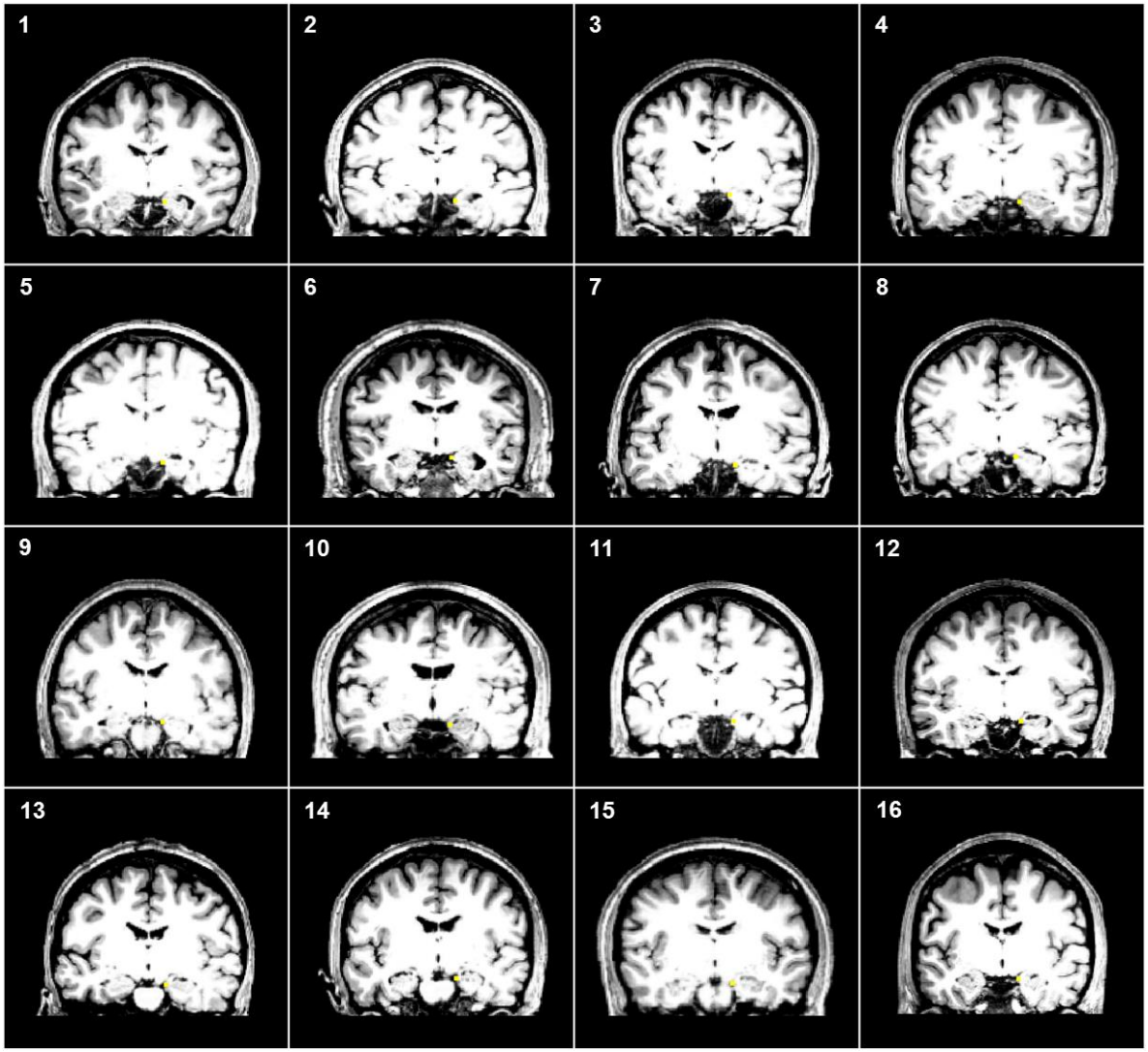

Fig. S7. Group hippocampal activation associated the contrast of hits versus misses for shapes in the lower-left quadrant, exclusively masked with the contrast of hits and misses for the other three quadrants projected onto each participant's anatomic image (coordinates, $x = 12$, $y = -9$, $z = -14$).

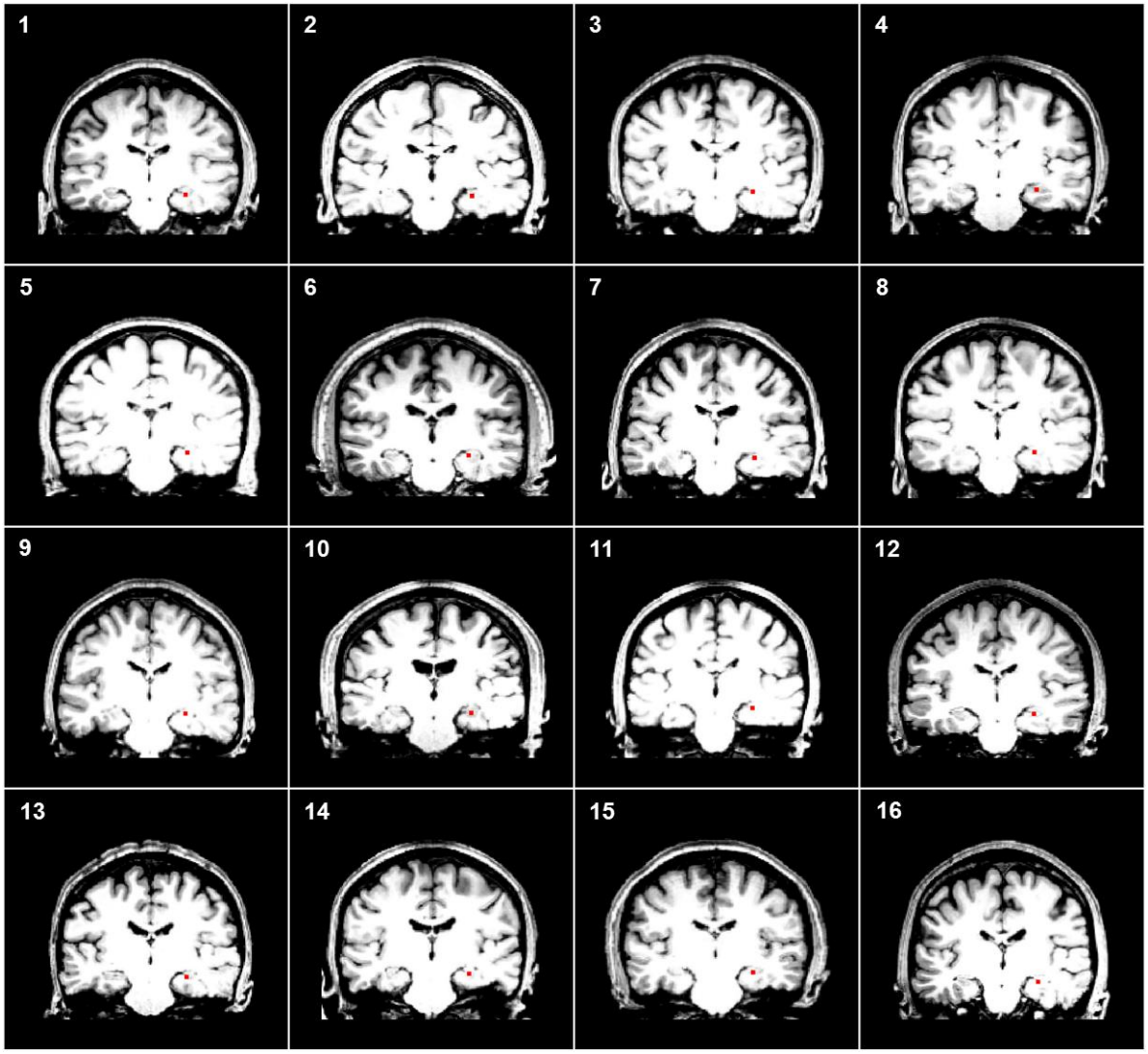

Fig. S8. Group hippocampal activation associated the contrast of hits versus misses for shapes in the upper-right quadrant, exclusively masked with the contrast of hits and misses for the other three quadrants projected onto each participant's anatomic image (coordinates, $x = 27$, $y = -19$, $z = -10$).

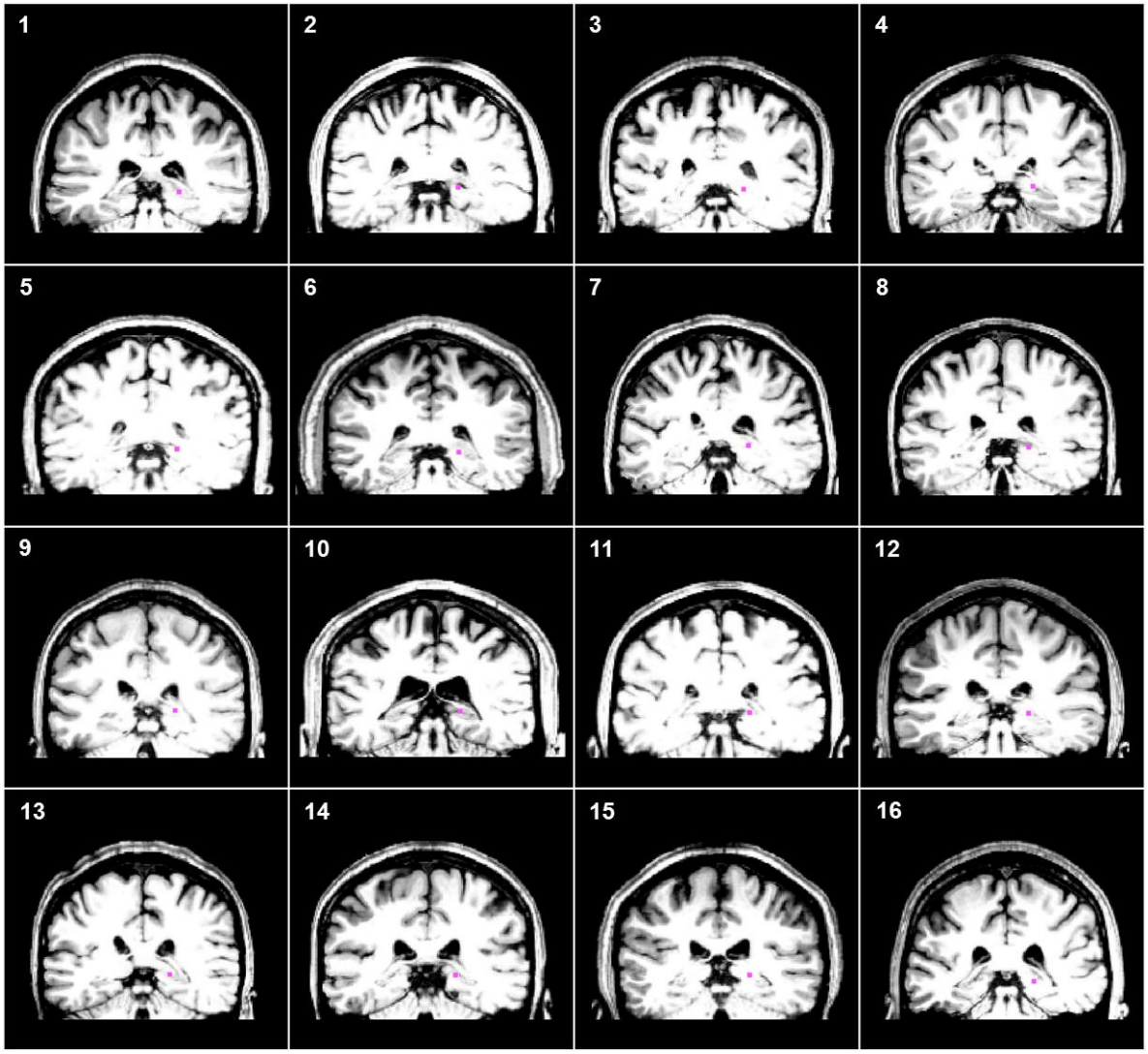

Fig. S9. Group hippocampal activation associated the contrast of hits versus misses for shapes in the lower-right quadrant, exclusively masked with the contrast of hits and misses for the other three quadrants projected onto each participant's anatomic image (coordinates, $x = 18$, $y = -33$, $z = 1$).

Chapter 2.2:

Spatial memory activity distributions indicate the hippocampus operates in a continuous manner.
Brittany M. Jeye, Jessica M. Karanian, and Scott D. Slotnick

Published in *Brain Sciences*.
https://doi.org/10.3390/brainsci6030037

There is a long-standing debate as to whether recollection is a continuous/graded process or a threshold/all-or-none process. In the current spatial memory functional magnetic resonance imaging (fMRI) study, we examined the hippocampal activity distributions—the magnitude of activity as a function of memory strength—to determine the nature of processing in this region. During encoding, participants viewed abstract shapes in the left or right visual field. During retrieval, old shapes were presented at fixation and participants classified each shape as previously in the "left" or "right" visual field followed by an "unsure"–"sure"–"very sure" confidence rating. The contrast of left-hits and left-misses produced two activations in the hippocampus. The hippocampal activity distributions for left shapes and right shapes were completely overlapping. Critically, the magnitude of activity associated with right-miss-very sure responses was significantly greater than zero. These results support the continuous model of recollection, which predicts overlapping activity distributions, and contradict the threshold model of recollection, which predicts a threshold above which only one distribution exists. Receiver operating characteristic analysis did not distinguish between models. The present results demonstrate that the hippocampus operates in a continuous manner during recollection and highlight the utility of analyzing activity distributions to determine the nature of neural processing.

Long-term memory can be based on non-detailed familiarity or detailed recollection. Familiarity is widely believed to be a continuous process, ranging in strength from weak to intermediate to strong. However, the nature of recollection has been a topic of debate. Until about a decade ago, recollection was widely thought to be an all-or-none threshold process, where memories are either completely remembered or forgotten [1–3]. However, a growing body of recent behavioral evidence indicates that recollection is a continuous process [4–6].

The two models of recollection are formally referred to as the continuous unequal variance model and the two-high threshold model [7]. Figure 1, left, illustrates both of these models during memory for one of two sources/contexts. For example, during encoding, items could be presented in green (source 1) or red (source 2). During retrieval, the same items could be presented at fixation in gray and participants would make a confidence rating ranging from "very sure green" to "very sure red". Each confidence rating depends on an item's source memory strength and criteria placement (in this illustration, C_1, C_2, C_3, C_4, and C_5). Memory strength greater than C_5 would yield a "very sure source 2" response, memory strength between C_4 and C_5 would yield a "sure source 2" response, memory strength between C_3 and C_4 would yield an "unsure source 2" response, memory strength between variance model dictates that the sources have Gaussian distributions of memory strength that can have unequal variance (Figure 1, top left). The two-high threshold model dictates that there are two thresholds ($threshold_1$ and $threshold_2$) beyond which only one source distribution exists (Figure 1, bottom left). Figure 1, right, shows the percentage associated with each event type generated from each model to the left (e.g., each rightmost bar is the area under the corresponding distribution to the right of C_5). Correct and incorrect source memory responses are referred to as hits and misses, respectively. The continuous unequal variance model predicts that the event distributions (i.e., source 1 and source 2 hits and misses; Figure 1, right) will be completely overlapping (i.e., all magnitudes will be greater than zero), whereas the two-high threshold model predicts that there is a threshold

above which high confidence hits but not high confidence misses will have magnitudes greater than zero (key differential predictions are illustrated within the dashed boxes).

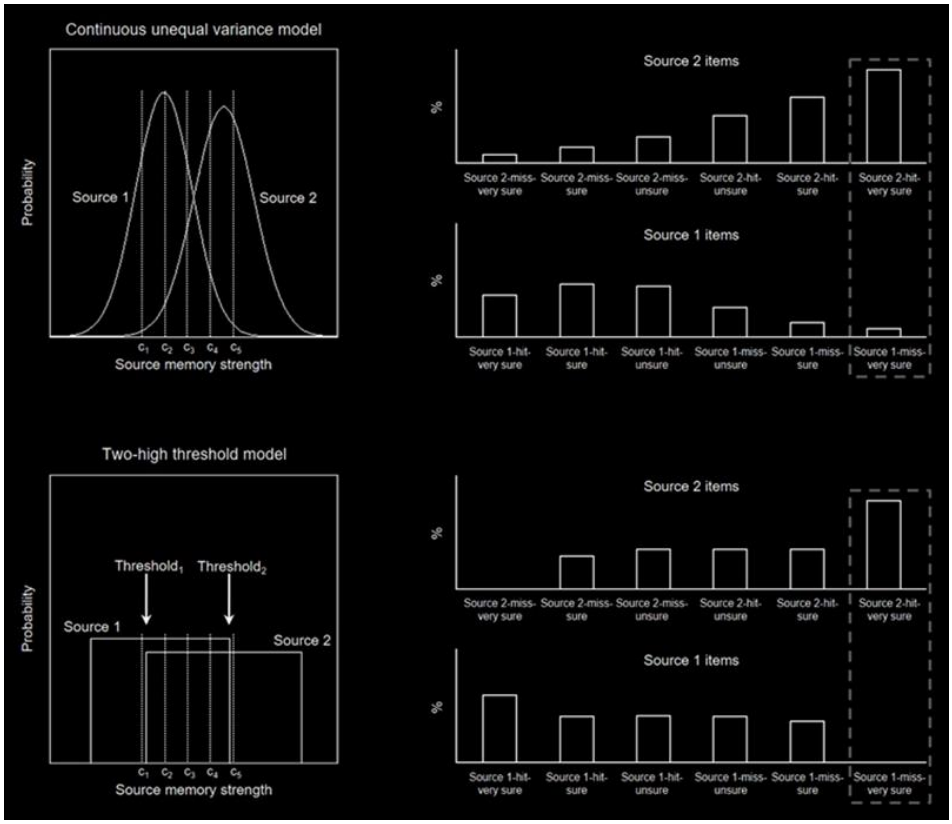

Figure 1. Models of recollection and event distributions. Top, continuous unequal variance model and corresponding percentage for each event type. Bottom, two-high threshold model and corresponding percentage for each event type.

The nature of processing in the hippocampus is of importance as this region is known to be associated with recollection [6,8]. One study assessed whether the hippocampus operated in a continuous manner or a threshold manner by evaluating the activity distributions from this region during spatial memory [9]. Abstract shapes were shown to the left or right of fixation during encoding. During retrieval, old and new shapes were presented at fixation and participants classified each shape as "old-left", "old-right", or "new", followed by an "unsure"–"sure" confidence rating.

The contrast of old-left-hits and old-left-misses produced an increase in activity within one region of the hippocampus. The magnitude of activity associated with high confidence misses

(old-right-miss-sure responses) was not significantly greater than zero, which was taken to support the threshold model. Thus, although previous behavioral results have supported the continuous model of recollection, previous hippocampal results have suggested that this region operates in a threshold manner during recollection [6]. However, in the previous study that evaluated hippocampal activity distributions [9], the activity associated with old-right-miss-sure responses was positive in magnitude and the standard error was large (see Section 4). As compared to that study, the current study was designed to have relatively smaller standard errors. First, we increased the number of participants from 12 to 16. Second, we increased the number of confidence ratings from two to three ("unsure"–"sure"–"very sure"). The "sure" response in the previous study can be assumed to reflect a mixture of cognitive processes, which would increase the standard error associated with this event type. In the current study, the three confidence ratings can be assumed to reflect more isolated cognitive processes and result in smaller standard errors. It is imperative to understand the operating model of the hippocampus, as this provides insight into the type of processing conducted by this region.

In the current spatial memory functional magnetic resonance imaging (fMRI) study, we analyzed the source memory distributions generated from behavioral responses and hippocampal activity to assess whether recollection operated in a continuous manner or a threshold manner. To anticipate the results, the behavioral response distributions and hippocampal activity distributions supported the continuous model of recollection and contradicted the threshold model of recollection.

Methods

Participants

Sixteen participants from the Boston College community completed the study (13 females, age range 22–28 years). Participants were right-handed, had normal or corrected-to-normal vision, were between the ages of 18–35, were native English speakers, were not pregnant,

and had no metal in their bodies. Each participant was compensated $10 for the behavioral training session and $25 per hour (approximately $100) for the fMRI session. The Boston College Institutional Review Board approved the protocol (identification code: 10.008, initial approval date: 9 December 2009). Informed and written consent was obtained prior to the behavioral training session.

Stimulus Protocol

Participants completed a one-quarter length run and a full-length run during the behavioral training session and seven to eight full-length runs during the fMRI session. During fMRI, one participant completed seven runs due to time limitations. The remaining participants completed eight runs; however, for one participant, a stimulus protocol was accidentally repeated and the repeated run was discarded.

During the encoding phase of each full-length run, 32 abstract shapes (half in the left visual field and half in the right visual field) spanning 6.7° of visual angle were presented with their nearest edge 3.6° of visual angle from a central fixation cross (Figure 2, left). The shapes were designed to minimize visual encoding strategies (for information on shape construction, see Slotnick and Schacter [10]). Each shape was displayed for 2.5 s followed by a 0.5 s fixation period. Shape sets were presented three times, with each shape set randomized and presented sequentially. Participants were instructed to remember each shape and its spatial location.

Before each retrieval phase, an instruction screen was displayed for 8 s followed by a 2 s fixation period. During the retrieval phase of each full-length run, the 32 (old) shapes from encoding were randomized and each shape was presented at fixation for 3.0 s followed by a confidence rating reminder screen for 2.5 s and a fixation period of 0.5 to 4.5 s (Figure 2, right). This resulted in an inter-trial-interval of 6.0 to 10.0 s, which is sufficient to allow for the deconvolution of the hemodynamic which is sufficient to allow for the deconvolution of the hemodynamic response. Although the previous study that evaluated hippocampal activity

distributions presented new shapes during retrieval [9], only old shapes were employed in the present study to increase the number of these critical event types. Although the lack of new items in the present study could affect criteria placement, this would not affect the distribution shapes and the predictions of each model (see Figure 1). Participants pressed response buttons with their left hand to classify each shape as previously presented in the "left" or "right" visual field followed by an "unsure"–"sure"–"very sure" confidence rating. In the previous study that evaluated hippocampal activity distributions "unsure" confidence ratings could correspond to forgotten old items or new items [9], while in the present study "unsure" confidence ratings could only correspond to forgotten old items. Of importance, this difference was not of importance as the key analyses were only conducted with confident responses.

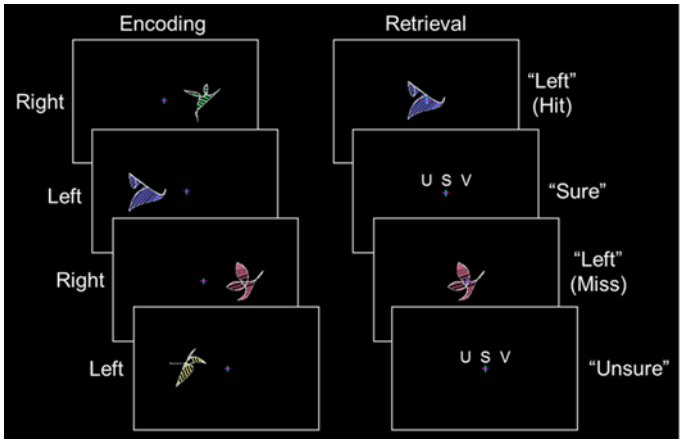

Figure 2. Experimental paradigm. Left, during encoding, participants viewed abstract shapes to the left or right of fixation (item types are shown to the left). Right, during retrieval, old items from encoding were presented at fixation and participants classified each shape as previously on the "left" or "right" followed by an "unsure"-"sure"-"very sure" confidence rating (possible responses and corresponding event types are shown to the right).

For both the encoding phase and the retrieval phase, no more than three shapes of a given type were sequentially presented and participants were instructed to maintain fixation. Shapes were never repeated across runs and shape location (i.e. left and right) was counterbalanced across participants using a Latin square design.

Data Acquisition and Pre-processing

A Siemens 3 Tesla Trio Scanner with a 32-channel head coil was used to acquire imaging data. Functional images were acquired with an echo planar imaging sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, field-of-view = 256 × 256 mm2, acquisition matrix = 64 × 64, slices = 33, slice acquisition order = interleaved bottom-to-top, slice thickness = 4 mm, no gap; 4 mm isotropic resolution). Anatomic images were acquired with a magnetization prepared rapid gradient echo sequence (TR = 30 ms, TE = 3.3 ms, flip angle = 40°, field-of-view = 256 × 256 mm2, acquisition matrix = 256 × 256, slices = 128, slice thickness = 1 mm; 1.33 × 1 × 1 mm resolution). Analyses were conducted with BrainVoyager 20.0 (Brain Innovation B.V., Maastricht, the Netherlands). Functional pre-processing included slice-time correction, motion correction, and removal of temporal components below 2 cycles per run length (using a general linear model to remove low frequency Fourier basis sets). Voxels were resampled at 3 mm3. To maximize spatial resolution, spatial smoothing was not conducted. Anatomic and functional images were transformed into Talairach space.

General Linear Model Analysis

A random-effect general linear model analysis was conducted. Each event type was modeled based on its onset and the subsequent behavioral response (if a response was made). Encoding trials and no-response trials were assumed to have durations of 2.5 s and the mean level of activity for each run was modeled with a constant. The contrast of correct spatial location memory (left-hits) and incorrect spatial location memory (left-misses; i.e., "left"/left > "right"/left) was used to isolate activity associated with spatial memory for shapes in the left visual field and the analogous contrast (i.e., "right"/right > "left"/right) was used to isolate activity associated with spatial memory for shapes in the right visual field. The previous study that evaluated hippocampal activity distributions isolated hippocampal activity using the conjunction of left-hits > left-misses and left-hits > right-hits in addition to the analogous

conjunction for right visual field stimuli [9]. However, these relatively conservative conjunctions did not produce any significant hippocampal activity using the present data; therefore, we used the standard hit > miss contrast in each visual field in an effort to identify multiple hippocampal activations. For all contrasts, an individual voxel threshold of $p < 0.001$ was enforced, false discovery rate corrected for multiple comparisons to $p < 0.05$. False discovery rate correction for multiple comparisons does not require a minimal cluster extent but rather, for a given individual voxel threshold, ensures an acceptable rate of false positives across the entire brain [11]. Known anatomical distinctions within the medial temporal lobe were used to localize hippocampal activations [12-15]. Activations were localized on the group average anatomic volume. Each Talairach coordinate refers to the voxel with peak activity.

Hippocampal Activity Distribution Analysis

For each hippocampal activation, event-related magnitudes were extracted from active voxels within a 5 mm cube (centered on the activation) from -2 to 12 s after stimulus onset (baseline corrected from -2 to 0 s). To ensure activation magnitudes were greater than, or equal to, baseline, which corresponds to a lower boundary on neural firing at zero spikes per second, the minimum activation magnitude across all event types was subtracted from each activation timecourse [9]. As fMRI activity can be assumed to reflect the underlying neural activity [16], we subtracted the minimum activation magnitude in an effort to make the zero point in the magnitude of fMRI activity correspond to the zero point of neural activity. This zero point in the magnitude of fMRI activity is analogous to no responses in behavior. To ensure the hippocampal activity distribution results did not depend on baseline correction, we also compared the magnitude of activity associated with miss-sure responses and miss-very sure responses as both of these trial types have the same baseline (which were subtracted out in the comparison). The threshold model predicts that the magnitude of activity associated with miss-sure responses will be significantly greater than the magnitude of activity associated with miss-very sure responses (see Figure 1).

For each event type, the mean magnitude of activity from 4 to 6 s after stimulus onset was used for analysis (i.e., the expected maximum amplitude of the hemodynamic response). Hippocampal activity distributions were generated by plotting the magnitude of activity as a function of memory strength where, from left to right, responses were "right-very sure", "right-sure", "right-unsure", "left-unsure", "left-sure", and "left-very sure". This corresponded to the following event types (mean trial numbers are shown in parenthesis) for left shapes: left-miss-very sure (3.1), left-miss-sure (9.5), left-miss-unsure (18.1), left-hit-unsure (21.3), left-hit-sure (31.1), left-hit-very sure (42.6), and the following event types for right shapes: right-hit-very sure (42.9), right-hit-sure (33.6), right-hit-unsure (22.2), right-miss-unsure (14.8), right-miss-sure (10.6), right-miss-very sure (1.6). The analysis was conducted after excluding three participants who made no right-miss-very sure responses. Although the numbers of trials associated with high confidence misses were relatively low, this would be expected to increase the corresponding standard errors and produce null results. Since significant results were observed (see Section 3.2), this was not of concern. As the direction of the statistical tests were known a priori, given that only an increase in the magnitude of activity relative to baseline reflects a memory-related activation, one tailed t-tests were employed. The behavioral response distribution analysis mirrored the hippocampal activity distribution analysis, except for the comparison between miss-sure responses and miss-very sure responses conducted as a test that was independent of baseline.

ROC Analysis

To compute activation percentages as a function of memory strength, each left shape activation magnitude was divided by the sum of all left shape activation magnitudes, and each right shape activation magnitude was divided by the sum of all right shape activation magnitudes. Hit rates were then computed by cumulating the probabilities from the highest to lowest memory strength for that stimulus type (e.g., left), and false alarm rates were computed by cumulating the

probabilities from lowest to highest memory strength for the other stimulus type (e.g., right). Each hippocampal ROC was generated by plotting these hit rates versus false alarm rates. The behavioral ROC was generated by plotting the hit rates versus false alarm rates based on the percentage of responses as a function of memory strength for each event type (left items and right items were arbitrary defined as source 1 and source 2, respectively; see Figure 1). For hippocampal ROC analysis we have assumed that the parametric estimates of neuronal activity are analogous to the parametric estimates of behavioral responses. That is, behavioral ROC analysis is based on the number of responses associated with each event type, as is commonly done, while hippocampal ROC analysis is based on the magnitude of activity associated with each event type. Hippocampal ROC results are only valid if this assumption is correct. The two-high threshold recollection model (with parameters R1 and R2) and the continuous unequal variance model (with parameters *d'* and σ_{s2}/σ_{s1}, the ratio of source distribution standard deviations) were fit to each ROC by adjusting model parameters using maximum likelihood estimation. The log-likelihood chi-square value was used to assess the adequacy of each model, where a lower chi-square value reflects a better fit ($p > 0.05$ indicates an adequate fit).

Results

Behavioral Results

Figure 3 shows the behavioral response distributions (in percentage of responses) for each left shape and right shape event type. Of importance, the percentage of highest confidence misses were both significantly greater than zero (left-miss-very sure, $t(15) = 6.61$, $p < 0.001$; right-miss-very sure, $t(15) = 4.95$, $p < 0.001$). These behavioral findings support the continuous unequal variance model of recollection and contradict the two-high threshold model of recollection. A chi-square analysis revealed that the behavioral ROC was not adequately fit by either the continuous unequal variance model (χ^2 (3) = 30.14, $p < 0.001$) or the two-high

threshold model (χ^2 (3) = 525.40, $p < 0.001$; Figure A1). It is notable that when forgotten items are included in the analysis, this can artifactually flatten the ROC and result in an inadequate fit for the continuous model [6,7]. Recollection-based ROCs that do not include forgotten items in the analysis are adequately fit by the continuous model but not the threshold model [7,17,18]. As a significant proportion of "unsure" responses can be assumed to reflect forgotten items, the present behavioral ROC results are not inconsistent with the continuous model. Still, as neither model adequately fit the behavioral ROC, the chi-square analysis results did not distinguish between the continuous model of recollection or the threshold model of recollection.

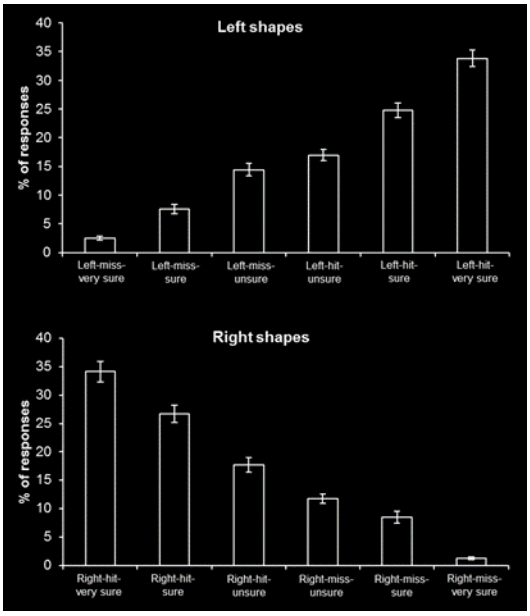

Figure 3. Behavioral response distributions. Percentage of responses for each left shape and right shape event type (mean ± S.E.).

Hippocampal Results

The contrast of left-hits and left-misses produced two activations in the hippocampus (Figure 4, top; x = –27, y = –14, z = –15, size = 54 mm^3; x = –24, y = –19, z = –11, size = 27 mm^3). The contrast of right-hits and right-misses did not produce any activity in the hippocampus, even at a reduced threshold of $p < 0.01$, uncorrected. This preferential hippocampal activity during memory for items in the left visual field has been observed in previous fMRI

studies [9,19]. Figure 4, bottom, shows the hippocampal activity distributions (in percent signal change) for left shapes and right shapes. Of importance, for both activations, the magnitude of activity associated with the highest confidence misses (right-miss-very sure responses) was significantly greater than zero (bottom left, $t(12) = 1.98$, $p < 0.05$; bottom right, $t(12) = 2.54$, $p < 0.05$). It should be highlighted that left-miss-very sure responses were not expected to have magnitudes that were significantly greater than zero in these regions because they were identified by contrasting left-hits and left-misses (i.e., left-miss was the baseline event and was expected to have a relatively low magnitude of activity). In addition, for both activations, the magnitude of activity associated with right-miss-sure responses was significantly greater than the magnitude of activity associated with right-miss-very sure responses (both ts$(12) < 1$). These hippocampal findings support the continuous model of recollection and contradict the threshold model of recollection.

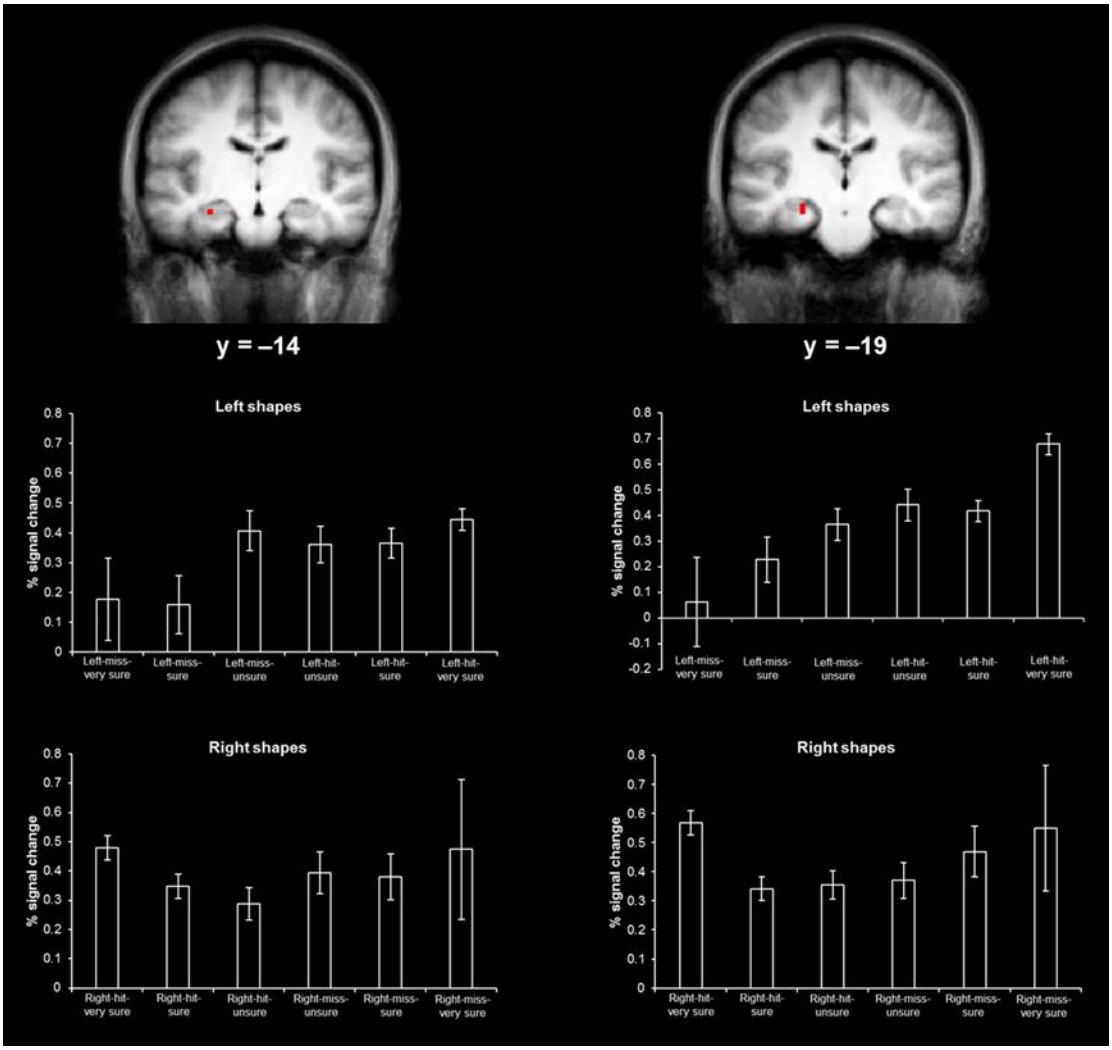

Figure 4. Hippocampal activity associated with spatial memory and hippocampal activity distributions. Top, hippocampal activity associated with left-hits versus left-misses (in red; coronal views). Bottom, hippocampal activity distributions (percent signal change for each event type) for left-shapes and right shapes corresponding to each hippocampal activation above.

A chi-square analysis revealed that both hippocampal ROCs were adequately fit by the continuous unequal variance model (y = –14 region, $\chi2(3) = 4.67$, $p = 0.20$; y = –19 region, $\chi2(3) = 5.73$, $p = 0.13$) and the two-high threshold model (y = –14 region, $\chi2(3) = 4.00$, $p > 0.20$; y = –19 region, $\chi2(3) = 3.91$, $p > 0.20$; Figure S2). The adequate fit for both models is likely due to the relatively low signal strength in both regions (y = –14 region, d' = 0.23; y = –19 region, d' = 0.38), which corresponds to an ROC that lies close to the diagonal (i.e., the chance line) and should be well fit by both models. As both models adequately fit the hippocampal ROCs, the chi-

square analysis results did not distinguish between the continuous model of recollection and the threshold model of recollection.

Discussion

In the present behavioral response distribution and the hippocampal activity distributions, there was no threshold above which only one source distribution existed. Specifically, in the behavioral response distribution, the percentage of left-miss-very sure and right-miss-very sure responses were significantly greater than zero. In the hippocampal activity distributions, the magnitude of activity associated with left-miss-very sure responses was significantly greater than zero. These findings support the continuous model of recollection and contradict the threshold model of recollection.

Although analysis of the behavioral response distributions and hippocampal activity distributions distinguished between the continuous model and the threshold model of recollection, the ROC analysis did not distinguish between these models. The behavioral ROC was not adequately fit by either model and the hippocampal ROCs were adequately fit by both models. This indicates that behavioral and hippocampal distribution analysis is a more sensitive measure than ROC analysis in distinguishing between the continuous model of recollection and the threshold model of recollection. This is likely because distribution analysis focuses on the differential prediction of the single critical event type (i.e., whether or not the magnitude of high confidence misses is significantly greater than zero). By contrast, the ROC is generated from all the event types, which could mask differential effects that exist. As mentioned previously, the hippocampal ROC analysis is based on the assumption that the parametric estimates of neural activity are analogous to the parametric estimates of behavioral responses and the hippocampal ROC results are only valid if this assumption is correct. This assumption is not critical to the present results as the chi-square analysis did not distinguish between the models of recollection.

As in the present study, Slotnick and Thakral analyzed the spatial memory hippocampal activity distribution to distinguish between the continuous model of recollection and the threshold model of recollection [9]. As mentioned in the Introduction, in that study, the conjunction of left-hits > left-misses and left-hits > right-hits produced one activation in the hippocampus and the magnitude of activity associated with old-right-miss-sure responses was not significantly greater than zero. However, the old-right-miss-sure activity was positive in magnitude (0.16 % signal change) and the standard error was large (0.16). Therefore, this null finding can be attributed to the large standard error rather than old-right-miss-sure activity being zero in magnitude. The current study was designed to reduce this large standard error by increasing the number of participants and requiring three confidence responses. These modifications can explain why the present results were significant, which should be favored over the null finding of Slotnick and Thakral [9]. Slotnick and Thakral also reported that the threshold model but not the continuous model adequately fit the hippocampal activity ROC [9]. However, an equal variance continuous model was fit to the ROC, which assumes the variance for old-left items and old-right items are identical. Recent evidence indicates that the hippocampus is associated with memory for items previously presented in the left visual field to a greater degree than memory for items previously presented in the right visual field [18], which would be expected to produce unequal variances for old-left items and old-right items. As such, the unequal variance continuous model should have been fit to the hippocampal ROC. We fit the continuous unequal variance model and the two-high threshold model to the hippocampal ROC from Slotnick and Thakral [9] and a chi-square analysis revealed that both models provided an adequate fit (continuous unequal variance model, $\chi 2(1) = 2.89$, $p = 0.089$; two-high threshold model, $\chi 2(1) < 1$; Supplementary Figure 3). Thus, as in the current study, ROC analysis did not distinguish between the continuous model of recollection and the threshold model of recollection.

The present behavioral response distribution and hippocampal activity distributions suggest that recollection is a continuous process. These findings support recent behavioral

evidence that recollection is a continuous process [4-6]. Our findings further suggest that continuous processing in the hippocampus contributes to continuous behavioral processing. Future work will be needed to evaluate the nature of processing in other neural regions to determine how processing across the brain gives rise to continuous behavioral processing during recollection.

The current findings indicate that the hippocampus operates in a continuous manner during recollection. This has implications for other lines of memory research. For instance, computational models of hippocampal function should not assume that this region operates in a threshold/all-or-none manner [20]. In addition, prospective memory (i.e., imagining the future) relies on episodic memory in which specific details from past events are recalled and recombined to form imagined scenarios, and prospective memory has been associated with activity in the hippocampus [21,22]. The present results indicate that during prospective memory, hippocampal activity does not reflect retrieval of details in a threshold/all-or-none manner, but rather reflects retrieval of graded details, ranging in strength from weak to intermediate to strong. While the current hippocampal activity distributions support the continuous model of recollection, future research could manipulate experimental factors such as stimulus type, number of repetitions, and encoding-retrieval delay to evaluate hippocampal activity distributions under different conditions. It is predicted that these findings will also support the present findings that the hippocampus operates in a continuous manner during recollection.

References

1. Mandler, G. Recognizing: The judgment of previous occurrence. *Psychol. Rev*. 1980, 87, 252–271. DOI: 10.1037/0033-295X.87.3.252

2. Yonelinas, A. P. The nature of recollection and familiarity: A review of 30 years of research. *J. Mem. Lang*. 2002, 46, 441–517. DOI: 10.1006/jmla.2002.2864

3. Yonelinas, A. P., & Parks, C. M. Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychol. Bull*. 2007, 133, 800–832. DOI: 10.1037/0033-2909.133.5.800

4. Wixted, J. T. Dual-process theory and signal-detection theory of recognition memory. *Psychol. Rev*. 2007, 114, 152–176. DOI: 10.1037/0033-295X.114.1.152

5. Wixted, J. T., & Mickes, L. A continuous dual-process model of remember/know judgments. *Psychol. Rev*. 2010, 117, 1025–1054. DOI: 10.1037/a0020874

6. Slotnick, S. D. The nature of recollection in behavior and the brain. *NeuroReport* 2013, 24, 663–670. DOI: 10.1097/WNR.0b013e328362e47e

7. Slotnick, S. D., & Dodson, C. S. Support for a continuous (single–process) model of recognition memory and source memory. *Mem. Cognition* 2005, 33, 151–170. DOI: 10.3758/BF03195305

8. Diana, R. A., Yonelinas, A. P., & Ranganath, C. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn. Sci*. 2007, 11, 379–386. DOI: 10.1016/j.tics.2007.08.001

9. Slotnick, S. D., & Thakral, P. P. The hippocampus operates in a threshold manner during spatial source memory. *NeuroReport* 2013, 24, 265–269. DOI: 10.1097/WNR.0b013e32835f282d

10. Slotnick, S. D., & Schacter, D. L. A sensory signature that distinguishes true from false memories. *Nat. Neurosci*. 2004, 7, 664–672. DOI: 10.1038/nn1252

11. Logan, B. R., & Rowe, D. B. An evaluation of thresholding techniques in fMRI analysis. *NeuroImage* 2004, 22, 95-108. DOI: 10.1016/j.neuroimage.2003.12.047

12. Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., Laakso, M. P., & Pitkänen, A. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am. J. Neuroradiol. 1998*, 19, 659–671. ISSN: 0195-6108

13. Pruessner, J. C., Li, L. M., Serles, W., Pruessner, M., Collins, D. L., Kabani, N., Lupiem, S., & Evans, A. C. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cereb. Cortex* 2000, 10, 433–442. ISSN: 1047-3211

14. Bernasconi, N., Bernasconi, A., Caramanos, Z., Antel, S. B., Andermann, F., & Arnold, D. L. Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain* 2003, 126, 462–469. DOI: 10.1093/brain/awg034

15. Malykhin, N. V., Bouchard, T. P., Ogilvie, C. J., Coupland, N. J., Seres, P., & Camicioli, R. Three-dimensional volumetric analysis and reconstruction of amygdala and hippocampal head, body and tail. *Psychiat. Res.-Neuroim*. 2007, 155, 155–165. DOI: 10.1016/j.pscychresns.2006.11.011

16. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 2001, 412, 150-157.

17. Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A.P. An analysis of signal detection and threshold models of source memory. *J. Exp. Psychol. Learn*. 2000, 26, 1499–1517. DOI: 10.1037/0278-7393.26.6.1499

18. Slotnick, S. D. "Remember" source memory ROCs indicate recollection is a continuous process. *Memory* 2010, 18, 27–39. DOI: 10.1080/09658210903390061

19. Jeye, B. M., Karanian, J. M., & Slotnick, S. D. (2018) Distinct regions of the hippocampus are associated with memory for different spatial locations. *Brain Research*.

20. Elfman, K. W., Aly, M., & Yonelinas, A. P. Neurocomputational account of memory and perception: Thresholded and graded signals in the hippocampus. *Hippocampus* 2014, 24, 1672–1686. DOI: 10.1002/hipo.22345

21. Schacter, D. L., Addis, D. R., & Buckner, R. L. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci*. 2007, 8, 657-661. DOI: 10.1038/nrn2213

22. Addis, D. R., & Schacter, D. The hippocampus and imagining the future: where do we stand?. *Front. Hum. Neurosci*. 2012, 5, 173. DOI: 10.3389/fnhum.2011.00173

Supplementary Materials

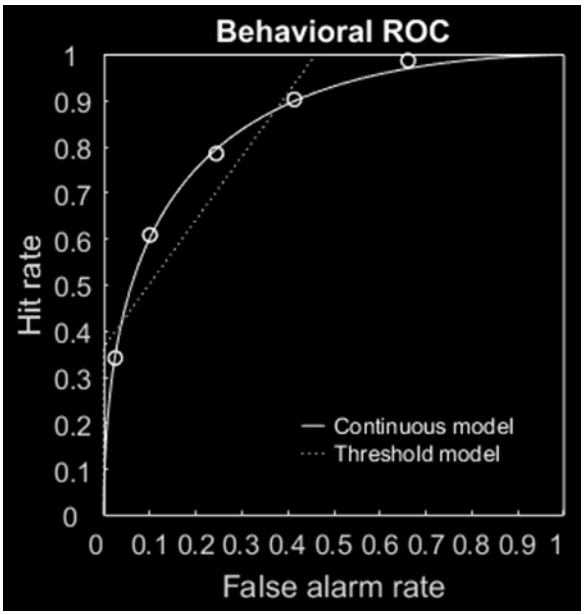

Figure S1. Behavioral ROC (circles) with the best-fit continuous unequal variance model and two high threshold model ROCs (key at the bottom right).

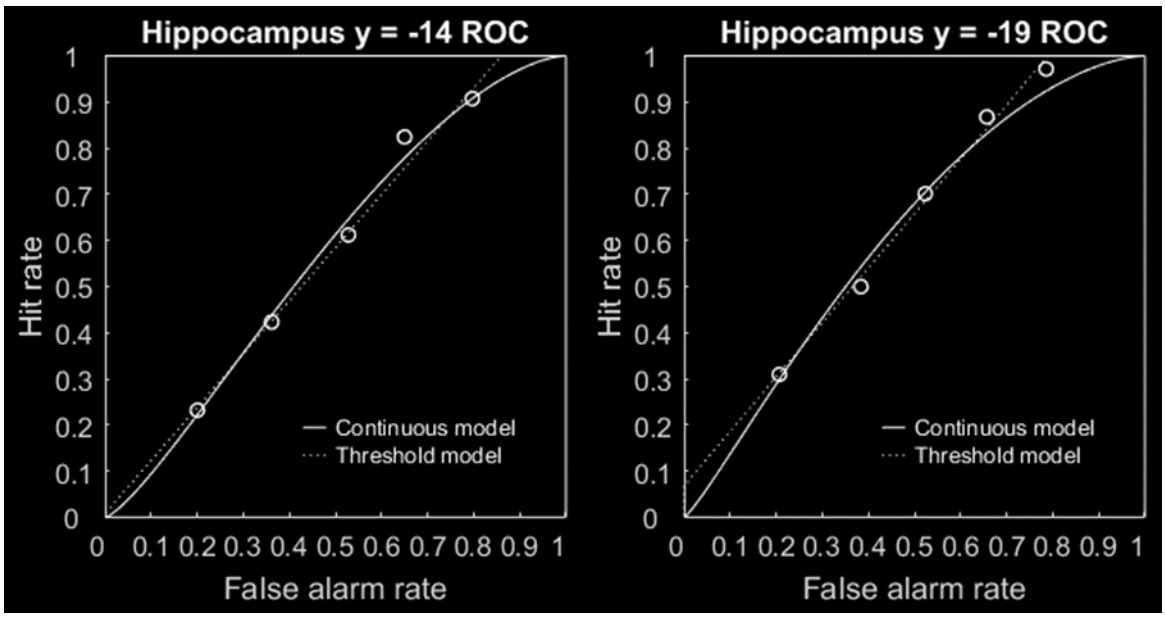

Figure S2. Hippocampal ROCs (circles) with the best-fit continuous unequal variance model and two-high threshold model ROCs (key at the bottom right of each ROC).

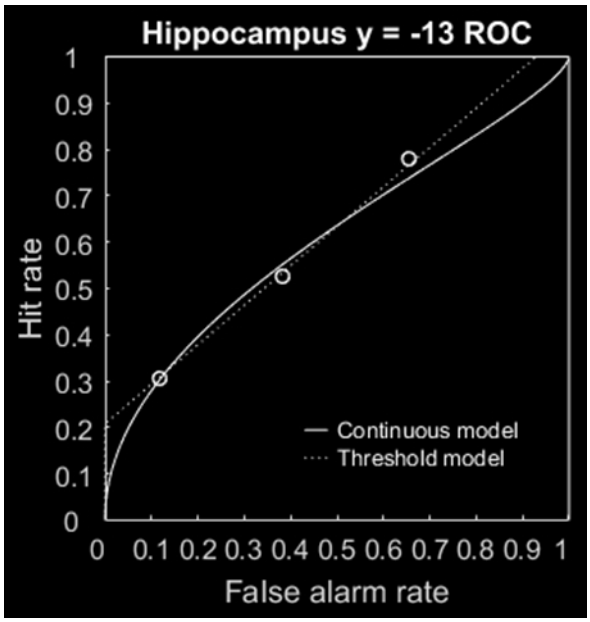

Figure S3. Hippocampal ROC from Slotnick and Thakral (2013) with the best-fit continuous unequal variance model and two-high threshold model ROCs (key at the bottom right).

CHAPTER 3.0: THE ROLE OF THE PREFRONTAL CORTEX IN VISUAL LONG-TERM MEMORY

Chapter 3.1:

The anterior prefrontal cortex and the hippocampus are negatively correlated during false memories.

Brittany M. Jeye, Jessica M. Karanian and Scott D. Slotnick

Published in *Brain Science.*
https://doi.org/10.3390/brainsci7010013

False memories commonly activate the anterior/dorsolateral prefrontal cortex (A/DLPFC) and the hippocampus. These regions are assumed to work in concert during false memories, which would predict a positive correlation between the magnitudes of activity in these regions across participants. However, the A/DLPFC may also inhibit the hippocampus, which would predict a negative correlation between the magnitudes of activity in these regions. In the present fMRI study, during encoding, participants viewed abstract shapes in the left or right visual field. During retrieval, participants classified each old shape as previously in the "left" or "right" visual field followed by an "unsure"-"sure"-"very sure" confidence rating. The contrast of left-hits and left-misses produced two activations in the hippocampus and three activations in the left A/DLPFC. For each participant, activity associated with false memories (right-"left"-"very sure" responses) from the two hippocampal regions were plotted as a function of activity in each A/DLPFC region. Across participants, for one region in the left APFC, there was a negative correlation between the magnitudes of activity in this region and the hippocampus. This suggests that the APFC might inhibit the hippocampus during false memories and that participants engage either the APFC or the hippocampus during false memories.

The hippocampus has long been known to play a critical role in accurate long-term memories (i.e., true memories). There is also evidence that the hippocampus can be involved in the construction of false memories – memories for events that never occurred [1-3]. Previous functional magnetic resonance imaging (fMRI) studies have demonstrated that there can be overlapping neural activity in the hippocampus during true memory and false memory [4-12]. For example, in our recent spatial memory fMRI study [13], participants viewed abstract shapes in either the left or right visual field during the study phase. During the test phase, old shapes from the study phase were presented at fixation and participants identified whether each shape was previously presented in the "left" or "right" visual field and made an "unsure"-"sure"-"very sure" confidence rating. The same regions of the hippocampus were found to be associated with true memory for spatial location and false memory for spatial location. Such hippocampal activations are thought to reflect the binding of item information and context information during memory [14,15]. That is, during true memory, the hippocampus appears to bind item information with the correct context (e.g., the correct spatial location), and during false memory, the hippocampus appears to bind item information with the incorrect context (e.g., the incorrect spatial location).

Like the hippocampus, the anterior/dorsolateral prefrontal cortex has also been associated with both true memory and false memory [3,16]. True memory and false memory activity in the left anterior/dorsolateral prefrontal cortex may reflect context memory [17,18]. Specifically, true memories can involve retrieval of the correct context and false memories can involve retrieval of the incorrect context [2]. The anterior/dorsolateral prefrontal cortex has also been associated with the subjective confidence during memory [19,20]. As true memories and false memories are often associated with high confidence, activity in the anterior/dorsolateral prefrontal cortex may also reflect this cognitive function.

The previous evidence indicates that the left anterior/dorsolateral prefrontal cortex and the hippocampus are associated with false memory. These regions are generally thought to work in concert during false memories, which would predict a positive correlation between the

magnitudes of activity in these regions across participants. However, there is evidence that the anterior/dorsolateral prefrontal cortex may inhibit the hippocampus during retrieval, such as during motivated forgetting [21,22] and retrieval-induced forgetting [23]. If the anterior/dorsolateral prefrontal cortex inhibits the hippocampus during false memories, this would predict a negative correlation between the magnitudes of activity in these regions across participants.

In the current spatial memory fMRI study, to distinguish between the previous hypotheses, we evaluated the correlation between the magnitudes of activity in the left anterior/dorsolateral prefrontal cortex and the hippocampus during false memory. To anticipate the results, we found that the magnitude of false memory activity across participants was negatively correlated between these regions.

Methods

Participants

Sixteen right-handed Boston College students who had normal or corrected-to-normal vision and were native English speakers participated in the study (12 females, age range 22–28 years). The Boston College Institutional Review Board approved the protocol (identification code: 10.008, initial approval date: 9 December 2009) and informed consent was obtained prior to the behavioral training session. Each participant was compensated $10 for the behavioral training session and $25 per hour for the fMRI session. The results of the current study are an extension of our recent fMRI study [13] (the same participants, paradigm, and analysis protocol was employed).

Stimulus Protocol

Participants completed a behavioral training session, which included a one-quarter length run and a full-length run, and 7 to 8 full-length runs during the fMRI session. During the study

phase of each full-length run, 32 abstract shapes spanning 6.7° of visual angle were presented with their nearest edge 3.6° of visual angle from a central fixation cross (Figure 1, left; for information on shape construction, see [5]). Each shape was displayed for 2.5 s followed by a 0.5 s fixation period. An equal number of shapes were presented to the left and right visual field. Participants were instructed to remember each shape and its spatial location while maintaining fixation. Shape sets were presented three times, with each shape set randomized and presented sequentially.

Before each test phase, an instruction screen was displayed for 8 s followed by a 2 s fixation period. During the test phase of each full-length run, the 32 shapes from encoding were presented in a random order at fixation for 3.0 s followed by a confidence rating reminder screen for 2.5 s and a fixation period of 0.5 to 4.5 s (Figure 1, right). Participants responded by pressing buttons with the fingers of their left hand to classify each shape as previously presented in the "left" or "right" visual field followed by a subsequent "unsure"-"sure"-"very sure" confidence rating. No more than three shapes of a given type were sequentially presented in the study phase or the test phase, shapes were never repeated across runs, and shape location (i.e., left and right) was counterbalanced across participants using a Latin square design.

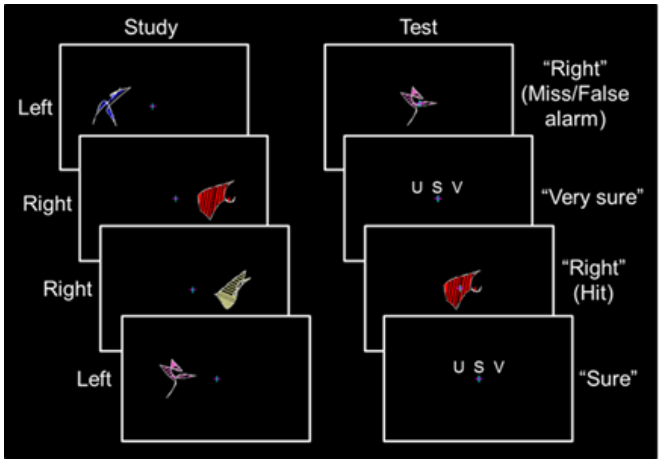

Figure 1. Stimulus protocol. Left, during the study phase, participants viewed abstract shapes to the left or right of fixation (labeled to the left). Right, during the test phase, old shapes were presented at fixation and participants classified each shape as previously on the "left" or "right" and made an "unsure"-"sure"-"very sure" confidence rating (possible responses and corresponding event types are shown to the right).

Data Acquisition and Analysis

A Siemens 3 Tesla Trio Scanner (Siemens, Erlangen, Germany) with a 32-channel head coil was used to acquire imaging data. Anatomic images were acquired with a magnetization prepared rapid gradient echo sequence (TR = 30 ms, TE = 3.3 ms, flip angle = 40°, field-of-view = 256 mm × 256 mm, acquisition matrix = 256 × 256, slices = 128, slice thickness = 1 mm; 1.33 × 1 × 1 mm resolution). Functional images were acquired with an echo planar imaging sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, field-of-view = 256 mm × 256 mm, acquisition matrix = 64 × 64, slices = 33, slice acquisition order = interleaved bottom-to-top, slice thickness = 4 mm, no gap; 4 mm isotropic resolution). BrainVoyager 20.0 (Brain Innovation B.V., Maastricht, the Netherlands) was used to conduct the analyses. Pre-processing of the functional images included motion correction, slice-time correction, and removal of temporal components below two cycles per run length (using a general linear model to remove low frequency Fourier basis sets). Voxels were resampled at 3 mm3. To maximize spatial resolution, spatial smoothing was not conducted. Anatomic and functional images were transformed into Talairach space.

A random-effect general linear model analysis was conducted. Each event type was modeled based on its onset and the subsequent behavioral response (if a response was made). It was assumed that encoding trials and no-response trials had durations of 2.5 s. This produced the following event types: encoding location, accurate memory for spatial location, inaccurate memory for spatial location, no response, and a constant. The contrast of left-"right"-"very sure" and left-"left"-"very sure" was used to isolate activity associated with false memory for shapes in the "right" visual field and the contrast of right-"left"-"very sure" and right-"right"-"very sure" was used to isolate activity associated with false memory for shapes in the "left" visual field. These contrasts did not produce any activity in the hippocampus. Therefore, as neural activity for both true memory and false memory overlap in the hippocampus and the anterior/dorsolateral prefrontal cortex (see the Introduction), the contrast of left-"left" and left-"right" (i.e., true memory spatial location hits versus misses, collapsed over confidence) was used to isolate

activity associated with spatial memory for shapes in the left visual field and the contrast of right-"right" and right-"left" was used to isolate activity associated with spatial memory for shapes in the right visual field. The true memory activations in the hippocampus and the anterior/dorsolateral prefrontal cortex served as regions of interest to extract and analyze false memory activity. For all contrasts, an individual voxel threshold of $p < 0.001$ was enforced, false discovery rate corrected for multiple comparisons to $p < 0.05$. Hippocampal activations within the medial temporal lobe were localized based on established anatomical distinctions [24-27]. All activations were localized on the mean group anatomic volume and each Talairach coordinate refers to the voxel with peak activity.

For each region of interest identified using the preceding analysis, event-related magnitudes were extracted from active voxels within a 5 mm cube (centered on the activation) from –2 to 12 s after stimulus onset (baseline corrected from –2 to 0 s). To ensure activation magnitudes were greater than or equal to baseline (corresponding to a lower boundary on neural firing of zero spikes per second), the minimum activation magnitude across all event types was subtracted from each activation timecourse [13,28]. As fMRI activity can be assumed to reflect the underlying neural activity [29], we subtracted the minimum activation magnitude in an effort to make the zero point in the magnitude of fMRI activity correspond to the zero point of neural activity. Of importance, baseline correction resulted in a constant shift for all magnitudes in a given region and thus did not influence the correlation results. For each participant, the mean event-related magnitude of activity associated with false memories (i.e., high confidence false alarms) from 4 to 6 s after stimulus onset was used for analysis (i.e., the expected maximum amplitude of the hemodynamic response). Three participants (two males) who made no right-"left"-"very sure" responses were excluded from the fMRI analysis.

Results

Behavioral accuracy was at an intermediate level and did not differ for shapes previously presented in the left visual field (75.5% correct) and shapes previously presented in the right visual field (78.6% correct; $t(15) < 1$). The contrast of left-hits and left-misses, which was used to isolate activity associated with true memory, produced two activations in the hippocampus (Figure 2, left; x = −27, y = −14, z = −15, size = 54 mm^3; x = −24, y = −19, z = −11, size = 27 mm^3), while the analogous contrast of right-hits and right-misses did not produce any activity in the hippocampus, even at a reduced threshold of $p < 0.01$, uncorrected.

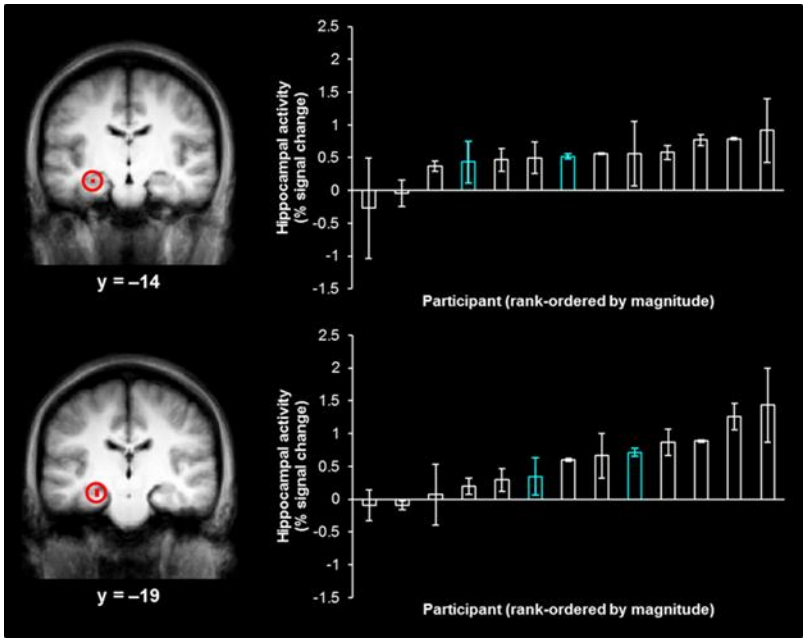

Figure 2. Hippocampal activity associated with true memory for items in the left visual field and the corresponding individual-participant magnitudes of hippocampal activity associated with false memory. Left, hippocampal activations associated with left-hits versus left-misses (circled in red; coronal views). Right, individual-participant magnitudes of activity (percent signal change) associated with false memories (right-"left"-"very sure" responses), rank ordered for the lowest to the highest magnitude of activity, corresponding to each hippocampal activation to the left (results from male participants are shown in blue).

As the neural activity associated with true memory and false memory overlap in the hippocampus (see the Introduction), we extracted individual participant magnitudes of activity associated with false memory (e.g., right-"left"-"very sure" responses) from each hippocampal activation (Figure 2, right). As the activations were identified by contrasting left-hits and left-misses, only false memories for items in the "left" visual field (i.e., right-"left"-"very sure"

responses) were expected to produce activity in these regions and were employed in the correlation analysis. Such false memories were no more likely to stem from shapes that were presented in the first or last 5 trials of each study phase than the middle 22 trials of each study phase (i.e., there was no evidence of primacy/recency effects; $\chi 2 < 1$). The range of values shown by the distribution of the magnitudes of activity associated with right-"left"-"very sure" responses demonstrates the variability in the magnitude of hippocampal activity during false memories across participants. This distribution suggests there are some participants with hippocampal-dependent processing during false memories (associated with higher magnitudes of activity in this region) and some participants with hippocampal-independent processing during false memories (associated with lower magnitudes of activity in this region).

The contrast of left-hits and left-misses also produced three activations in the left anterior/dorsolateral prefrontal cortex (x = –21, y = 35, z = 37, BA 9 within the superior frontal sulcus, size = 27 mm^3; x = –9, y = 47, z = 28, BA9 within the anterior prefrontal cortex, size = 27 mm^3; x = –36, y = 41, z = 10, BA46 within the inferior frontal sulcus, size = 27 mm^3). For each participant, we plotted the magnitude of activity associated with right-"left"-"very sure" responses in each of the two hippocampal regions as a function of the magnitude of activity in each anterior/dorsolateral prefrontal cortex region. For the anterior prefrontal cortex region (Figure 3, left), there was a significant negative correlation between the magnitude of activity in the hippocampus and the magnitude of activity in this region (Figure 3, right; r = –0.48, $p < 0.05$, Bonferroni corrected for the three hippocampal-prefrontal cortex correlations). Although the present results were not powered to assess gender effects, the correlation was nearly identical (r = –0.50, $p < 0.05$) after removing males from the analysis. It should be highlighted that all of the activations evaluated were associated with "very sure" responses (i.e., confidence was held constant); thus, the activations were not correlated with confidence. One limitation of the current study is that our sample size was relatively small; however, this would be expected to produce null results. As significant results were observed, the sample size was not of major concern.

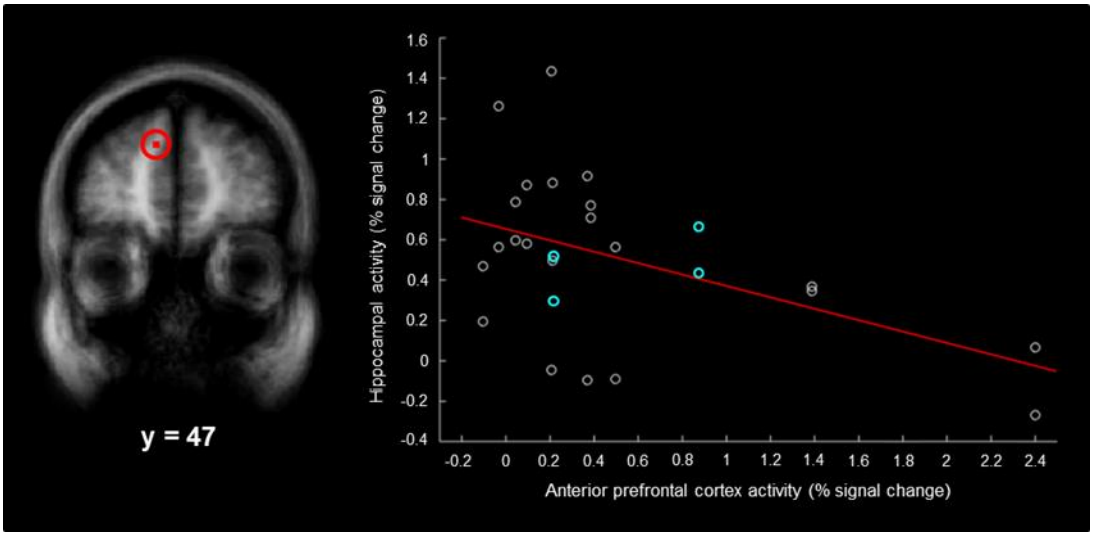

Figure 3. Relationship between the magnitude of spatial memory activity in the left anterior prefrontal cortex and the hippocampus. Left, left anterior prefrontal cortex activity associated with left-hits and left-misses (circled in red; coronal view). Right, for each participant, the magnitude of hippocampal activity associated with false memories as a function of the magnitude of left anterior prefrontal cortex activity associated with false memories (the best-fit line is shown in red; results from male participants are shown in blue).

We conducted additional analyses to assess whether the negative correlation between activity in the anterior dorsolateral prefrontal cortex and the hippocampus was specific to false memories (i.e., right-"left"-"very sure" responses). The correlation between these regions was not significant for either right-"left"-"unsure" responses ($r = –0.20$, $p > 0.20$) or right-"right"-"very sure" responses ($r = –0.13$, $p > 0.20$). These findings indicate that the anterior dorsolateral prefrontal cortex and the hippocampus were not correlated during the analogous low confidence responses or during confident true memories.

To determine whether there were behavioral differences between participants with hippocampal based false memories and participants with anterior dorsolateral prefrontal cortex based false memories, we conducted a post hoc split-half analysis. The behavioral performance of the participants with higher magnitudes of anterior prefrontal cortex activity was compared with the behavioral performance of the participants with lower magnitudes (the participant with an intermediate magnitude of activity was left out such that there were equal numbers in each group). There was no difference between these groups of participants in either overall behavioral

accuracy ($t(11) < 1$) or the rate of false memories for items in the "left" visual field (i.e., right-"left"-"very sure" responses/all right-"very sure" responses; $t(11) < 1$).

Discussion

In the present study, we found that the magnitude of activity in the hippocampus was negatively correlated with the magnitude of activity in the left anterior prefrontal cortex during false memories. These findings suggest that false memories may be mediated by the hippocampus (and not the anterior prefrontal cortex) in some participants and the anterior prefrontal cortex (and not the hippocampus) in other participants. This is the first time, to our knowledge, that participants have been shown to engage either the hippocampus or the anterior prefrontal cortex during false memories.

It is notable that the contrast of left-hits and left-misses only produced activations in the left hippocampus, which replicates a previous study [28]. As the left hippocampus has been associated with verbal memory [30,31], these activations might have reflected language processing associated with accurate memory for the spatial location of each shape (i.e., the verbal label "left"). Alternatively, the hemispheric laterality in the hippocampus may have been a consequence of limited power.

The current findings may shed light on the variable nature of false memory activity previously reported in the hippocampus [2,3,32]. In the present study, the magnitudes of hippocampal false memory activity ranged from −0.27 to 1.43 percent across participants, the magnitudes of left anterior false memory activity ranged from −0.10 to 2.40 percent, and there was a negative correlation between these regions (Figure 3, right). This demonstrates that false memories were only based on hippocampal activity in some participants and were only based on left anterior prefrontal cortex activity in other participants. If some groups of participants engage the anterior prefrontal cortex and not the hippocampus during false memories, this would predict relatively low magnitudes of hippocampal activity during false memory, which has previously

been observed [2,8,9]. On the other hand, if some groups of participants engage the hippocampus and not the anterior prefrontal cortex, this would predict relatively high magnitudes of hippocampal activity during false memory, which has also been observed [4-6,10]. Future research will be needed to determine the specific stimulus or task conditions under which the hippocampus is more or less strongly associated with false memory.

The present negative correlation between the magnitude of activity in the hippocampus and the magnitude of activity in the anterior prefrontal cortex suggests that these regions interact during false memories. One possibility is that, during false memories, the left anterior/dorsolateral prefrontal cortex may be activated due to (incorrect) context memory or high confidence and this region may inhibit the hippocampus to reduce the amount of potentially conflicting information. (Note that it is also possible that another region of the dorsolateral prefrontal cortex, such as the left inferior dorsolateral prefrontal cortex (BA46) activation in the present study, may reflect language processing, which can also give rise to false memories [33,34]). Conversely, as correlation does not confer directionality, the hippocampus may be activated due to (incorrect) binding and this region may inhibit the anterior/dorsolateral prefrontal cortex to reduce the amount of potentially conflicting information. Although the direction of the interaction between the anterior/dorsolateral prefrontal cortex and the hippocampus is uncertain, the anterior/dorsolateral prefrontal cortex has been generally associated with inhibition [22,35], and thus it is likely that this region inhibited the hippocampus during false memories.

Single-cell recording evidence in non-human animals also indicates that the prefrontal cortex and the hippocampus interact during memory for item and context information [36]. For instance, a recent study in rats demonstrated that information flowed from the hippocampus to the prefrontal cortex during item-in-context memory encoding and information flowed from the prefrontal cortex to the hippocampus during item-in-context memory retrieval [37]. As the present findings were observed during retrieval, these behavioral neuroscience findings provide additional evidence that, for some participants, the anterior prefrontal cortex inhibited the

hippocampus during false memories. Future studies could employ simultaneous depth electrode recording in the hippocampus, such as in patients with intractable epilepsy, and scalp electrophysiological recording to investigate the nature of the interactions between these regions during false memory construction.

References

1. Schacter, D. L., & Slotnick, S. D. The cognitive neuroscience of memory distortion. *Neuron* 2004, 44, 149–160.

2. Karanian, J. M.; Slotnick, S. D. False memory for context activates the parahippocampal cortex. *Cogn. Neurosci*. 2014, 5, 186-192. DOI: 10.1080/17588928.2014.938035

3. Dennis, N. A.; Bowman, C. R.; Turney, I. C. Functional neuroimaging of false memories. In *The Wiley Handbook on the Cognitive Neuroscience of Memory*; Duarte, A., Addis, D.R, Barense, M., Eds.; Wiley-Blackwell: West Sussex, United Kingdom, 2015; pp. 150-171.

4. Cabeza, R.; Rao, S. M.; Wagner, A. D.; Mayer, A. R.; Schacter, D. L. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proc. Natl. Acad. Sci*. 2001, 98, 4805-4810. ISSN: 0027-8424

5. Slotnick, S. D.; Schacter, D. L. A sensory signature that distinguishes true from false memories. *Nat. Neurosci*. 2004, 7, 664-672. DOI: 10.1038/nn1252

6. Kensinger, E. A.; Schacter, D. L. Neural processes underlying memory attribution on a reality-monitoring task. *Cereb. Cortex* 2006, 16, 1126-1133. DOI: 10.1093/cercor/bhj054

7. Kim, H.; Cabeza, R. Trusting our memories: dissociating the neural correlates of confidence in veridical versus illusory memories. *J. Neurosci*. 2007, 27, 12190-12197. E-ISSN: 1529-2401

8. Paz-Alonso, P. M.; Ghetti, S.; Donohue, S. E.; Goodman, G. S.; Bunge, S. A. Neurodevelopmental correlates of true and false recognition. *Cereb. Cortex* 2008, 18, 2208-2216. DOI: 10.1093/cercor/bhm246

9. Giovanello, K. S.; Kensinger, E. A.; Wong, A. T.; Schacter, D. L. Age-related neural changes during memory conjunction errors. *J. Cogn. Neurosci*. 2009, 22, 1348-1361. DOI: 10.1162/jocn.2009.21274

10. Stark, C. E. L.; Okado, Y.; Loftus, E. F. Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learn. Mem*. 2010, 17, 485-488. DOI: 10.1101/lm.1845710

11. Dennis, N. A.; Bowman, C. R.; Vandekar, S. N. True and phantom recollection: An fMRI investigation of similar and distinct neural correlates and connectivity. *Neuroimage* 2012, 59, 2982-2993. DOI: 10.1016/j.neuroimage.2011.09.079

12. Gutchess, A. H.; Schacter, D. L. The neural correlates of gist-based true and false recognition. *Neuroimage* 2012, 59, 3418-3426. DOI: 10.1016/j.neuroimage.2011.11.078

13. Jeye, B. M.; Karanian, J. M.; Slotnick, S. D. Spatial memory activity distributions indicate the hippocampus operates in a continuous manner. *Brain Sci*. 2016, 6, 37. DOI: 10.3390/brainsci6030037

14. Eichenbaum, H.; Yonelinas, A. R.; Ranganath, C. The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci*. 2007, 30, 123-152. DOI: 10.1146/annurev.neuro.30.051606.094328

15. Ranganath, C. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus* 2010, 20, 1263-1290. DOI: 10.1002/hipo.20852

16. Slotnick, S. D.; Schacter, D. L. The cognitive neuroscience of memory and consciousness. In *The Cambridge Handbook of Consciousness*; Zelazo, P.D., Moscovitch, M., Thompson, E., Eds.; Cambridge University Press: New York, New York, 2007; pp. 809-827.

17. Slotnick, S. D.; Moo, L. R.; Segal, J. B.; Hart, J. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cogn. Brain Res*. 2003, 17, 75-82. DOI: 10.1016/S0926-6410(03)00082-X

18. Mitchell, K. J.; Johnson, M. K. Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psych. Bull*. 2009, 135, 638. DOI: 10.1037/a0015849

19. Chua, E. F.; Rand-Giovannetti, E.; Schacter, D. L.; Albert, M. S.; Sperling, R. A. Dissociating confidence and accuracy: Functional magnetic resonance imaging shows origins of the subjective memory experience. *J. Cogn. Neurosci*. 2004, 16, 1131-1142. ISSN: 0898-929X

20. Kao, Y.; Davis, E. S.; Gabrieli, J. D. Neural correlates of actual and predicted memory formation. *Nat. Neurosci*. 2005, 8, 1776-1783. DOI: 10.1038/nn1595

21. Anderson, M. C.; Ochsner, K. N.; Kuhl, B.; Cooper, J.; Robertson, E.; Gabrieli, S. W.; Glover, G. H.; Gabrieli, J. D. Neural systems underlying the suppression of unwanted memories. *Science* 2004, 303, 232-235. E-ISSN: 1095-9203

22. Anderson, M. C.; Hanslmayr, S. Neural mechanisms of motivated forgetting. *Trends Cogn. Sci*. 2014, 18, 279-292. DOI: 10.1016/j.tics.2014.03.002

23. Wimber, M.; Alink, A.; Charest, I.; Kriegeskorte, N.; Anderson, M. C. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nat. Neurosci*. 2015, 18, 582-589. DOI: 10.1038/nn.3973

24. Insausti, R.; Juottonen, K.; Soininen, H.; Insausti, A. M.; Partanen, K.; Vainio, P.; Laakso, M. P.; Pitknen, A. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am. J. Neuroradiol*. 1998, 19, 659-671. ISSN: 0195-6108

25. Pruessner, J. C.; Li, L. M.; Serles, W.; Pruessner, M.; Collins, D. L.; Kabani, N.; Lupien, S.; Evans, A. C. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: Minimizing the discrepancies between laboratories. *Cereb. Cortex* 2000, 10, 433-442. ISSN: 1460-2199

26. Bernasconi, N.; Bernasconi, A.; Caramanos, Z.; Antel, S. B.; Andermann, F.; Arnold, D. L. Mesial temporal damage in temporal lobe epilepsy: A volumetric MRI study of the

hippocampus, amygdala and parahippocampal region. *Brain* 2003, 126, 462-469. ISSN: 0006-8950

27. Malykhin, N. V.; Bouchard, T. P.; Ogilvie, C. J.; Coupland, N. J.; Seres, P.; Camicioli, R. Three-dimensional volumetric analysis and reconstruction of amygdala and hippocampal head, body and tail. *Psychiatry Res*. 2007, 155, 155-165. DOI: 10.1016/j.pscychresns.2006.11.011

28. Slotnick, S. D.; Thakral, P. P. The hippocampus operates in a threshold manner during spatial source memory. *Neuroreport* 2013, 24, 265-269. DOI: 10.1097/WNR.0b013e32835f282d

29. Logothetis, N. K.; Pauls, J.; Augath, M.; Trinath, T.; Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 2001, 412, 150-157. DOI: 10.1038/35084005

30. Milner, B. Disorders of memory after brain lesions in man. *Neuropsychologia* 1968, 6, 175-179.

31. Glosser, G.; Deutsch, G. K.; Cole, L. C.; Corwin, J.; Saykin, A. J. Differential lateralization of memory discrimination and response bias in temporal lobe epilepsy patients. *J. Int. Neuropsychol. Soc.* 1998, 4, 502-511.

32. Kurkela, K. A.; Dennis, N. A. Event-related fMRI studies of false memory: An activation likelihood estimation meta-analysis. *Neuropsychologia* 2016, 81, 149-167. DOI: 10.1016/j.neuropsychologia.2015.12.006

33. Garoff-Eaton, R. J..; Slotnick. S. D.; Schacter, D. L. Not all false memories are created equal: The neural basis of false recognition. *Cereb. Cortex* 2006, 16, 1645-1652. DOI: 10.1093/cercor/bhj101

34. Karanian, J. M.; Slotnick, S. D. The cortical basis of true memory and false memory for motion. *Neuropsychologia* 2014, 54, 53-58. DOI: 10.1016/j.neuropsychologia.2013.12.019

35. Anderson, M. C.; Ochsner, K. N.; Kuhl, B.; Cooper, J.; Robertson, E.; Gabrieli, S. W.; Glover, G. H.; Gabrieli, J. D. Neural systems underlying the suppression of unwanted memories. *Science* 2004, 303, 232-235. E-ISSN: 1095-9203

36. Eichenbaum, H. Memory: Organization and control. *Annu. Rev. Psychol*. in press.

37. Place, R.; Farovik, A.; Brockmann, M.; Eichenbaum, H. Bidirectional prefrontal-hippocampal interactions support context-guided memory. *Nat. Neurosci.* 2016, 19, 992-994. DOI: 10.1038/nn.4327

Chapter 3.1:

Support for an inhibitory model of semantic memory retrieval.
Brittany M. Jeye, Sarah K. Kark, Elizabeth A. Kensinger, Lauren R. Moo, and Scott D. Slotnick

Semantic memory retrieval may involve an inhibitory process in which a target word is activated and related words are suppressed. In the current functional magnetic resonance imaging (fMRI) study, we examined the inhibition of language processing cortex by the left dorsolateral prefrontal cortex (dlPFC) during memory retrieval using an anagram solving paradigm. Participants were presented with a distractor that was read aloud followed by a to-be-solved anagram. Distractor types were defined relative to orthographic overlap with the subsequent anagram solution and included related words with one letter different (e.g., "gripe" for the anagram of "price"), related non-words, and unrelated words (i.e., all five letters were different). The anagram solution reaction time was slower in both the related word and related non-word distractor conditions as compared to the unrelated word distractor condition, which can be attributed to greater semantic memory inhibition following related distractors. The contrast of related words and unrelated words produced one activation in the left dlPFC, a region that has been associated with memory inhibition. To identify the regions that were negatively correlated with activity in the left dlPFC for related distractors, we conducted a functional connectivity analysis between this left dlPFC region and the rest of the brain. We found negatively correlated activity between the dlPFC and language processing cortex for the related word distractor condition (and the related non-word distractor condition at a relaxed threshold). These findings suggest that that the left dlPFC may inhibit related word representations in language processing cortex during semantic memory retrieval.

Inhibitory processing in semantic memory has been demonstrated with the retrieval-practice paradigm, where the retrieval of studied items leads to forgetting of related items (i.e., retrieval-induced forgetting; Anderson, Bjork & Bjork, 1994; Anderson, Bjork & Bjork, 2000; Anderson & Spellman, 1995; Levy & Anderson, 2002). The retrieval-practice paradigm typically consists of three phases. In the first phase, participants study lists composed of category-exemplar word pairs (e.g., fruit-banana, fruit-apple, furniture-table, furniture-chair, etc.). This is followed by a retrieval-practice phase in which participants are directed to practice a subset of the category-exemplar pairs through a cued-stem recall test (e.g., fruit-ap___). In the third phase, participants are asked to retrieve all exemplars from the first phase. Memory performance is then examined for practiced exemplars (which were seen in the first two phases), non-practiced exemplars from practiced categories (which were only seen in the first phase, but are related to items practiced in the second phase), and non-practiced exemplars from non-practiced categories (which were only seen in the first phase, i.e., control/baseline items). The key behavioral finding is that memory for the non-practiced exemplars from practiced categories is worse than that of baseline exemplars. This finding has been termed retrieval-induced forgetting and it is thought to reflect inhibition of non-practiced items during the second phase, as inhibitory processing may be flexibly directed at interfering and/or competing memories during retrieval-practice (e.g., 'banana' is inhibited during retrieval of 'apple'), which leads to an increased rate of forgetting (Levy & Anderson, 2002; Anderson & Hanslmayr, 2014). Inhibitory mechanisms during the retrieval-practice paradigm have been associated with activity that include the left dorsolateral prefrontal cortex (dlPFC; Wimber, Bäuml, Bergström, Markopoulos, Heinze & Richardson-Klavehn, 2008; Wimber, Rutschmann, Greenlee & Bäuml, 2009; Wimber, Alink, Charest, Kriegeskorte & Anderson, 2015), which is consistent with this region mediating inhibition during other cognitive processes such as selective attention, decision making, and motor planning (Knight, Staines, Swick & Chao, 1999; Miller & Cohen, 2001).

Retrieval-induced forgetting has been extensively studied utilizing a wide variety of stimuli (for a review, see Levy & Anderson, 2002). Of importance, the large majority of studies have used stimuli that are semantic in nature. This includes words in different categories (as in the example above), propositions (where retrieving facts about a topic impairs the recall for related facts; Anderson & Bell, 2001; Radvansky, 1999), and semantic generation (where generating/retrieving general related knowledge about a study item can cause forgetting of studied items; Bauml, 2002; Johnson & Anderson, 2004). Critically, all of these studies have used variations of the retrieval-practice paradigm, where unconscious inhibitory processing (i.e., without an explicit cue) is thought to operate during the retrieval practice phase. However, if inhibition of related information occurs during semantic memory retrieval, it need not to be limited to the retrieval-practice phase and may occur at any stage requiring retrieval.

In the current study, to investigate inhibitory processes during semantic memory retrieval, we employed an anagram-solving paradigm consisting of a word/non-word distractor reading task followed by an anagram task (Fig. 1). On each trial, participants vocalized a distractor and then solved a subsequent anagram, pressing a button when a solution was obtained. Distractors varied in orthographic similarity to the solution of the anagram, from four letters in common (related distractors, either words or non-words) to no letters in common (unrelated distractors).

We hypothesized that the related distractor conditions (i.e., related word and related non-word distractor conditions) would produce inhibition of subsequent anagram-solution words, which would be evidenced by increased left dlPFC activity (Wimber et al., 2008; Wimber et al., 2009; Wimber et al., 2015) in these distractor conditions as compared to the unrelated word distractor condition. Furthermore, we hypothesized that language processing regions would be inhibited during the related word distractor condition as compared to the unrelated word distractor condition, as orthographically similar words to the distractor (including the anagram solution) should be inhibited. As non-words (e.g., pseudo-words, random letter strings) have been

shown to activate language processing cortex (Cohen, Lehericy, Chochon, Lemerm Rivaud & Dahaene, 2002; Vigneau, Jobard, Mazoyer & Tzourio-Mazoyer, 2005; Glezer, Jiang & Riesenhuber, 2009), we also hypothesized that language processing cortex may be inhibited during the non-word condition as well.

Methods

Participants

Ten right-handed adults with normal or corrected-to-normal vision participated in the study (6 females, M = 32.8 years, range 24-41 years). The Johns Hopkins Institutional Review Board approved of the experimental protocol and informed written consent was obtained prior to each session.

Stimulus Protocol and Behavioral Analysis

The fMRI study consisted of a practice session followed by an experimental session. Both sessions were completed in the scanner (trials shown during the practice session were not used during the experimental session). During each session, participants were shown trials consisting of a pronounceable five-letter distractor word or non-word followed by a five-letter, to-be-solved anagram (Fig. 1). Three distractor types were shown that varied in orthographic similarity relative to the to-be-solved anagram: related words that had four letters in common (e.g., distractor "gripe" for the anagram or "price"), non-words with four letters in common (e.g., distractor "letry" for the anagram "style"), and unrelated words with no letters in common (e.g., distractor "claim" for the anagram "press"). Distractor words and non-words were displayed in green capital letters on a white background for 1.5 s, followed by the to-be-solved anagram that was displayed in black capital letters on a white background for 10 s and a .5 s blank fixation period. The second and fourth letters of the to-be-solved anagrams were anchored in the correct position as indicted by an underline (e.g., the anagram "novel" was presented as "VOLEN"). To

reduce potential head movement, participants were instructed to vocalize the distractor words or non-words with minimal head or jaw movement while they were presented on the screen. When the to-be-solved anagrams were on the screen, participants were instructed to press a response button with their left hand when they had obtained a solution. The anagram remained on screen for the entire 10 s duration regardless of when the button was pressed. The to-be-solved anagrams ranged from 50-125 on Kucera and Francis Frequency scale. Distractors were randomized with the constraint that no more than three of the same type were presented sequentially.

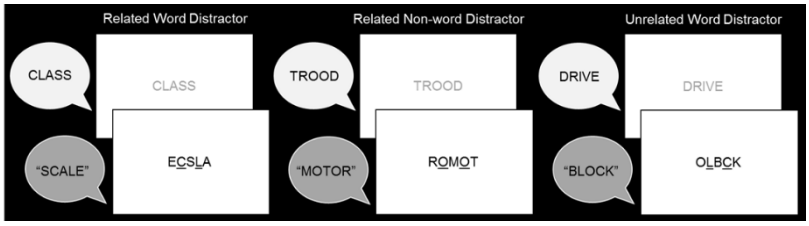

Fig. 1. Illustration of conditions and stimuli in the fMRI anagram-solving paradigm. Distractors types varied in orthographic similarity to the to-be-solved anagrams from four letters in common (related words and related non-words) to no letters in common (unrelated words). Participants were asked to vocalize the distractors, which was followed by a to-be-solved anagram.

Given our directional hypothesis, a one-tailed paired t-test was utilized to compare anagram solution reaction times between the related word distractor condition and unrelated word distractor condition as well as the related non-word distractor condition and unrelated word distractor condition. As there was no prior hypothesis concerning the magnitude of inhibitory processing between the related word and related non-word distractor conditions, a two-tailed paired t-test was used for this comparison.

Data Acquisition and Analysis

Imaging data were acquired on a 1.5T Philips ACS-NT scanner with a standard birdcage head coil. Functional images were acquired using an echo planar imaging sequence (TR = 2 s, TE = 40 ms, flip angle = 90°, 26 slices, no gap, 4.5 mm isotropic resolution). Whole-brain anatomic images were acquired using a magnetization-prepared rapid gradient-echo sequence (12.4 min

acquisition time, TR = 8.1 ms, TE = 3.7 ms, flip angle = 8°, slices = 256, no gap, 1 mm isotropic resolution). Data were analyzed using SPM8 (Wellcome Department of Imaging Neuroscience). Pre-processing included slice-time correction, motion correction (registered to first image of each run), and spatial normalization to the Montreal Neurological Institute (MNI) template. To maximize spatial resolution, spatial smoothing was not conducted. Anatomic images were also transformed into MNI space. A random-effect general linear model analysis was conducted for each participant (using a high-pass filter cutoff of 128 s). The following trial types were included in the general linear model: related words, related non-words, and unrelated words, those three trial types without responses, and a constant. Trials with responses were modeled from anagram onset to the button press, while no response trials were assumed to last 10.5 s. Activity associated with inhibitory processing in the left dlPFC was assessed through the following a priori contrasts at a threshold of $p < .0005$, with a cluster extent threshold of two voxels: related word distractor condition versus unrelated word distractor condition (one-tailed), related non-word distractor condition versus unrelated word distractor condition (one-tailed), and related word distractor condition versus related non-word distractor condition (two-tailed). Signal intensity magnitudes (beta values) for each trial type were extracted from the left dlPFC actviations of interest using the REX toolbox (web.mit.edu/swg/software.htm) and the MarsBar toolbox (marbar.sourceforge.net/) to build the ROIs.

To assess which regions were negatively correlated with activity in the left dlPFC, a functional connectivity analyses was conducted using the generalized psychological interactions (gPPI) toolbox (brainmap.wisc.edu/PPI; McLaren, Ries, Xu & Johnson, 2012). The left dlPFC seed region was defined using the activation from the group contrast between the related word distractor condition and the unrelated word distractor condition. For each participant at the first level, task regressors were created to estimate the magnitude of activity within the left dlPFC activation, and then psychophysiological interaction was calculated using the gPPI toolbox. Regions that showed a negative relationship with left dlPFC seed region across the entire brain

were identified for the contrast of unrelated word distractor and related word distractor conditions, which should reflect inhibition during the related word condition. Language processing cortex was broadly defined as including the left superior temporal cortex (including Wernicke's area in the posterior superior temporal gyrus), left inferior frontal gyrus (i.e., Broca's area), extrasylvian temporo-parietal cortex, and left occipito-temporal sulcus (i.e., the visual word form area; Cohen, Dehaene, Naccache, Lehéricy, Dehaene-Lambertz, Hénaff & Michel, 2000; Price, 2000). The right hemisphere homologues of these areas were also considered part of language processing cortex (Glezer, Jiang & Riesenhuber, 2009; Price, 2000; Vigneau, Jobard, Mazoyer & Tzourio-Mazoyer, 2005; Vigneau et al., 2011). Contrast files from each participant were entered into single-sample t-tests at the group level and thresholded at $p < .005$ (Kark, Slotnick & Kensinger, 2016) with a cluster extent threshold of 10 voxels. This contrast was also inclusively masked by the contrast of the unrelated word distractor (baseline) condition and related word distractor condition (using a threshold of $p < .05$ with a cluster extent threshold of 10 voxels) to ensure the magnitude of activity during the related word condition was below the magnitude associated with baseline word processing. The same procedure was used to conduct the gPPI analysis for the related non-word distractor condition. Activation coordinates are reported in Talairach space.

Results

Behavioral Results

In support of our hypothesis that related distractors would be associated with semantic memory inhibition, anagram solution reaction times were greater in the related word distractor condition (5.51 s) than the unrelated word distractor condition (4.98 s; $t(9) = 1.87$, $p < .05$). Anagram solution reaction times were also greater in the related non-word distractor condition (5.75 s) than the unrelated distractor condition ($t(9) = 4.02$, $p < .01$), and the related word and non-word distractor conditions did not significantly differ ($t(9) < 1$).

Neuroimaging Results

The contrast of related word and unrelated word distractor conditions produced one activation in the left dlPFC/BA8 (Fig. 2a, coordinates, x = -29, y =4, z = 54, BA 8). There were no significant activations in the left dlPFC for the contrasts of related non-word and unrelated word distractor conditions or the related word and related non-word distractor conditions.

To identify the regions that were negatively correlated with activity in the left dlPFC (which corresponds to a lower magnitude of activity for the related-word than unrelated and non-word distractor conditions), we conducted a functional connectivity analysis between the left dlPFC region of interest identified above and the rest of the brain (see Supplementary Table S1 for a list of all activations). In support of our hypothesis that left dlPFC would inhibit language processing cortex during the related word distractor condition, we found negatively correlated activity between left dlPFC and the left anterior superior temporal gyrus/BA22 (Fig 2b) and right superior temporal gyrus/BA22/41/42 (Figs. 2b, middle and bottom activations).

In an effort to determine if there was similar inhibitory processing for the related non-word distractor condition, we extracted the magnitude of activity associated with this condition from the left dlPFC activation of interest above. The magnitude of activity associated with related non-word distractors was significantly greater than zero in this region ($t(9) = 1.83$, $p < .05$), which suggests that this dlPFC region was also associated with inhibitory processing during the related non-word distractor condition. A functional connectivity analysis between the left dlPFC region and the rest of the brain did not identify any significantly negatively correlated activations within the language processing cortex at our original threshold. However, at a reduced threshold to $p < .01$ (with the same cluster extent), there was significantly negatively correlated activity between dlPFC and the supramarginal gyrus/BA40 (Fig 3c, top and middle) and superior temporal gyrus/BA41/42 (Fig. 2c, bottom; see Supplementary Table S1).

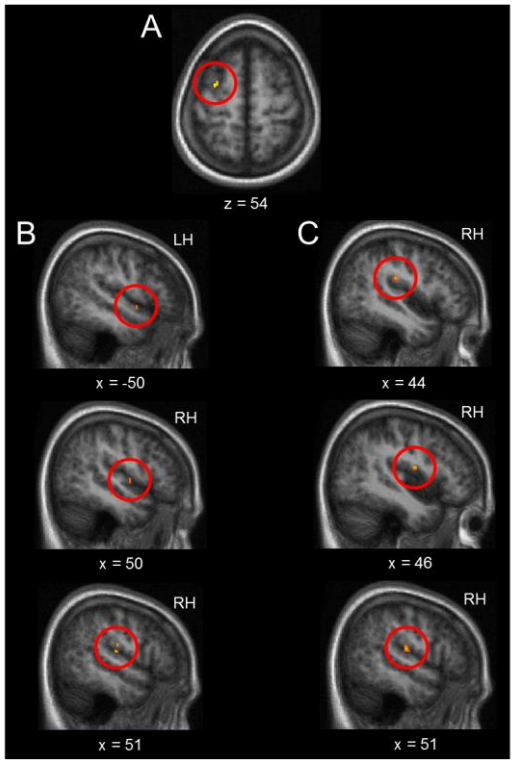

Fig. 2. fMRI results. A) Left dlPFC activity associated with inhibitory processing identified by contrasting the related word and unrelated word distractor conditions. B) Activations in language processing cortex that were negatively correlated with the left dlPFC region for the related word distractor condition (activations circled, LH = left hemisphere, RH = right hemisphere, Talairach coordinates). C) Activations in the language processing cortex that were negatively correlated with the left dlPFC region for the related non-word distractor condition.

Discussion

In support of our hypotheses, the contrast of related and unrelated word distractor conditions produced an activation in the left dlPFC, a region known to be involved with inhibitory processing during long-term memory (Wimber et al, 2008, Wimber et al., 2009; Wimber et al., 2015). Furthermore, we also found negative functional connectivity between this left dlPFC activation and language processing cortex for the related word distractor condition. Similar to the related word distractor condition, the magnitude of activity for the related non-word distractor condition in this left dlPFC region was also significantly greater than zero and this region was negatively connected to language processing cortex (although only with a relaxed threshold). These findings suggest that the left dlPFC inhibited related word representations (in

both the related word and non-word distractor conditions) in language processing cortex during semantic memory retrieval during the current anagram-solving task.

The present results are consistent with previous research indicating that the prefrontal cortex is involved with inhibition of episodic memory (Anderson, Bunce & Barbas, 2015; Jeye, Karanian & Slotnick, 2017; Wimber et al., 2015). Furthermore, anatomical pathways have been identified that suggest the prefrontal cortex can inhibit sensory and memory processing regions, such as the hippocampus (Anderson, Bunce & Barbas, 2015; Eichenbaum, 2017). The current findings extend these previous results and demonstrate that the left dlPFC can inhibit language processing cortex during semantic memory retrieval.

The present anagram-solving paradigm provides a novel demonstration of unconscious inhibition during long-term memory, which has previously been shown only during the retrieval-practice paradigm. Unlike the retrieval practice paradigm, where unconscious inhibition is thought to act on related information during the retrieval-practice phase (Anderson, Bjork & Bjork, 1994), inhibition in the current study operated during semantic memory retrieval. Furthermore, participants were not explicitly trained on the anagram task (as they are in the retrieval-practice paradigm); therefore, this paradigm may more directly measures unconscious inhibition of related information that occurs in everyday life.

Unconscious inhibitory processing was also demonstrated through the related non-word distractor condition, which indicates that non-words may inhibit orthographically similar words. However, future research will need to be conducted in order to determine the pathways involved with inhibitory processing at the orthographic level of word processing, as our functional connectivity analyses revealed distinct regions were negatively connected with the left dlPFC region for related non-word distractors as compared to related word distractors.

The current results indicate that semantic memory retrieval may depend on unconscious inhibitory processing that suppresses related information. These findings also set the stage for related lines of research investigating the role of inhibitory processing more broadly in the

retrieval of related memories, both episodic and semantic, and the regions involved in mediating this processing.

References

Anderson, M. C., & Bell, T. (2001). Forgetting our facts: the role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General, 130(3), 544.*

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic bulletin & review, 7(3*), 522-530.

Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal–hippocampal pathways underlying inhibitory control over memory. *Neurobiology of learning and memory, 134*, 145-161.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological review, 102(1)*, 68.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in cognitive sciences, 18(6*), 279-292.

Bäuml, K. H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science,* 13(4), 356-360.

Cohen L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*, 291-307.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience,* 18(9), 547.

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron, 62(2),* 199-204.

Jeye, B., Karanian, J., & Slotnick, S. (2017). The anterior prefrontal cortex and the hippocampus are negatively correlated during false memories. *Brain sciences, 7(1),* 13.

Johnson, S. K., & Anderson, M. C. (2004). The role of inhibitory control in forgetting semantic knowledge. *Psychological Science, 15(7),* 448-453.

Kark, S. M., Slotnick, S. D., & Kensinger, E. A. (2016). Repetition enhancement of amygdala and visual cortex functional connectivity reflects nonconscious memory for negative visual stimuli. *Journal of cognitive neuroscience, 28(12),* 1933-1946.

Knight, R. T., Staines, W. R., Swick, D., & Chao, L. L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta psychologica, 101(2-3),* 159-178.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in cognitive sciences, 6(7)*, 299-305.

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage, 61(4)*, 1277-1286.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience, 24(1)*, 167-202.

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, 197, 335-59.

Radvansky, G. A. (1999). Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General*, 128(4), 563.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage, 54(1),* 577-593.

Vigneau, M., Jobard, G., Mazoyer, B., & Tzourio-Mazoyer, N. (2005). Word and non-word reading: what role for the visual word form area?. *Neuroimage, 27(3)*, 694-705.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature neuroscience, 18(4), 582.*

Wimber, M., Bäuml, K. H., Bergström, Z., Markopoulos, G., Heinze, H. J., & Richardson-Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *Journal of Neuroscience, 28(50),* 13419-13427.

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K. H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience*, 21(3), 538-549.

Supplementary Material

Table S1. Functional connectivity whole brain activations

Region	BA	x	y	z
(Unrelated Word > Related Word) ∩ (0 > Related Word)				
*Left Anterior Superior Temporal Gyrus	22	-50	2	-5
*Right Superior Temporal Gyrus	22	50	-4	0
*Right Heschel's Gyrus	41/42	51	-17	12
Right Supramarginal Gyrus	40	51	-25	42
Right Precuneus	7	8	-41	67
Right Precentral Sulcus	6	40	6	37
Left Inferior Frontal Sulcus	46	-36	25	26
Left Superior Occipital Gyrus	19	-18	-72	29
(Unrelated Word > Related Non-word) ∩ (0 > Related Non-word)				
*#Right Supramarginal Gyrus	40	44	-3	15
*#Right Heschel's Gyrus	41/42	51	-17	14
*#Right Supramarginal Gyrus	40	46	-22	23
Right Precentral Sulcus	6	36	11	33
Left Anterior Prefrontal Cortex	10	−34	47	7

BA refers to Brodmann area and Talairach coordinates (x, y, z) are reported. * = language processing cortex. # = appear at a reduced threshold of $p < .001$.

GENERAL DISCUSSION

The goal of the studies within this dissertation was to investigate the cognitive and neural processes that support visual long-term memory specificity. That is, how are we able to experience many events that may be similar or related, yet are able to form discrete memories (e.g., remembering where you parked your car today versus remembering where you parked your car yesterday). In particular, it is thought that this ability relies on both accurately retrieving specific details and inhibiting potentially distracting information. To this end, in Chapter 1, I demonstrated that specificity for item memory is dependent on retrieving accurate details about unique items and inhibiting related information, and that this inhibitory pattern differs depending on stimuli type (i.e., faces or abstract shapes). In Chapters 2 and 3, I investigated the neural processes associated with long-term memory specificity by examining the functional roles of the hippocampus and prefrontal cortex, respectively.

The hippocampus and visual long-term memory specificity

A number of previous studies have indicated that the hippocampus is necessary for binding item and contextual information to create a uniform memory (Eichenbaum, Yonelinas & Ranganath, 2007; Ranganath & Ritchey, 2012). In particular, the hippocampus has been shown to be critically involved in spatial contextual processing. These spatial properties of the hippocampus have been well documented in both animals and humans. For example, place cells in the rodent hippocampus fire when an animal is in a particular location within its environment (O'Keefe & Dostrovsky, 1971). Similar spatial properties have been shown in the human

hippocampus during spatial navigation (Burgess et al., 2002; Ekstrom et al., 2003; Maguire et al., 2006; Miller et al., 2013). However, navigational tasks use cognitive processes other than memory (such as wayfinding, route planning, and setting navigation goals; for a review, see Wolbers, 2015), which may have confounded spatial memory with other spatial computations. Thus, in Chapter 2, I utilized tasks that isolated spatial memory in order to assess the functional properties of the hippocampus during long-term memory. In Chapter 2.1, I demonstrated that distinct regions of the hippocampus were associated with accurate spatial memory for different visual field locations. This research adds to the growing body of literature suggesting that the hippocampus is responsible for binding item and contextual information (in this case, abstract shapes and their locations within the field field) to create detailed memories. This is in line with my behavioral research from Chapter 1, which demonstrated memory specificity for unique items, as participants were able to distinguish previously seen items from highly similar ones.

The process by which the human hippocampus stores, consolidates and retrieves these unique independent memories is thought to rely on the computational mechanisms of pattern separation and pattern completion (Leal & Yassa, 2018; Yassa & Stark, 2011). In this computational framework, pattern completion is the process by which the hippocampus fills-in incomplete mental representations using previously stored experiences and representations, and pattern separation is the ability of the hippocampus to create distinct non-overlapping mental representations of stimuli to reduce the interference of similar stimuli. Electrophysiological and neuroimaging findings suggest that different hippocampal subfields mediate pattern separation and pattern completion, with the dentate gyrus responsible for pattern separation and the CA3 subfield responsible for pattern completion (Leal & Yassa, 2018; Yassa & Stark, 2011). Furthermore, previous research has shown that there is a long-axis gradient in terms of hippocampal anatomy, connectivity and function, with the anterior hippocampus associated with more global spatial representations and the posterior hippocampus associated with more local spatial representations (for a review, see Strange et al., 2014; Poppenk et al., 2013). In Chapter

2.1, I found that accurate spatial memory for shapes in different visual field quadrants produced activity that was distributed throughout the hippocampus in the anterior-posterior direction. While it may seem that our results don't support this heterogeneity of hippocampal function, it may have been due to our specific analyses. This is a topic of future research.

In addition, this binding-in-context model also highlights the important role of the hippocampus in recollection. That is, the hippocampus is important for the retrieval of detailed memories, including item and contextual details, which are unique in the face of competing, highly related information. Thus, in Chapter 2.2, I investigated the underlying mathematical processes through which the hippocampus supports recollection of these specific details. Our findings suggests that the hippocampus operates in a continuous/graded manner during recollection, as hippocampal activity distributions for accurate spatial memory for shapes in left and right visual field were completely overlapping. This is in opposition to previous research which has suggested that recollection is a threshold/all-none-process (Yonelinas, 2002; Yonelinas & Parks, 2007). These results support the role of the hippocampus in maintaining detailed, unique representations in long-term memory specificity, although this region operates in a graded manner during spatial memory.

The prefrontal cortex in visual long-term memory specificity

While the Chapter 2 focused on the accurate retrieval of detailed memories and the role of the hippocampus in long-term memory specificity, there is a growing body of literature that suggests that the prefrontal cortex (the dorsolateral prefrontal cortex, in particular) is necessary for the top-down control of memory. For instance, it has been hypothesized that one role of the dorsolateral prefrontal cortex is to inhibit distracting, interfering, or competing memories (for a review, see Diamond & Levine, 2018). The inhibitory role of the dorsolateral prefrontal cortex was explored in Chapter 3 of this dissertation. In Chapter 3.1, I found that the anterior/dorsolateral prefrontal cortex was negatively correlated with activity in the hippocampus

during false memories (i.e., memories for events that never actually occurred). False memories can be assumed to reflect potentially distracting or irrelevant information; therefore, these results suggesting that the anterior/dorsolateral prefrontal cortex may inhibit the hippocampus. This finding extends previous fMRI results that illustrated dorsolateral prefrontal cortex mediated inhibition of medial temporal lobe regions during long-term memory (Benoit et al., 2015). Anatomical pathways have been identified that support models implicating the prefrontal cortex in top-down inhibitory control of other memory regions, including the hippocampus (Anderson, Bunce, & Barbas, 2016; Depue; 2012; Eichenbaum, 2017). The entorhinal gating hypothesis suggests that the dlPFC suppresses hippocampal activity through the anterior cingulate cortex, which has projections to the entorhinal cortex and that this pathway supports the proactive stopping of retrieval (see Annderson, Bunce, & Barbas, 2016). The thalamo-hippocampal modulation via the nucleus reuniens also suggests that the dlPFC suppresses hippocampal activity through the anterior cingulate cortex, which in turn may modulate the hippocampus directly through bidirectional connections with the thalamic reuniens nucleus.

In Chapter 3.1, I further explored the role of the dorsolateral prefrontal cortex in inhibitory processing by demonstrating that the dorsolateral prefrontal cortex inhibits language processing cortex during semantic memory through functional connectivity analyses. These results indicate that the top-down control of neural regions involved in mnemonic processing by the dorsolateral prefrontal cortex is not limited to other memory control regions, but may extend to sensory cortical regions. Of importance, our studies utilized paradigms in which there was no explicit cue to suppress memory (i.e., unintentional inhibitory processing, such as in the retrieval practice paradigm; Wimber et al., 2015). Therefore, it is possible that other prefrontal cortex regions may be involved with explicit memory suppression (i.e., intentional inhibitory processing, such as in the think/no-think paradigm; Anderson et al., 2004).

General Conclusions

The findings from this dissertation shed light on the cognitive and neural processes that support visual long-term memory specificity. In particular, I demonstrated that long-term memory relies on both accurately retrieving specific details, which is mediated by the hippocampus, and inhibiting potentially distracting information, which is mediated by the dorsolateral prefrontal cortex.

GENERAL REFERENCES

Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal–hippocampal pathways underlying inhibitory control over memory. *Neurobiology of learning and memory*, *134*, 145-161.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in cognitive sciences*, *18*(6), 279-292.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*(5655), 232-235.

Bäuml, K. H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1047-1054.

Benoit, R. G., Hulbert, J. C., Huddleston, E., & Anderson, M. C. (2015). Adaptive top–down suppression of hippocampal activity and the purging of intrusive memories from consciousness. *Journal of Cognitive Neuroscience*, *27*(1), 96-111.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience*, *9*(8), 613.

Corkin, S. (2002). What's new with the amnesic patient HM?. *Nature reviews neuroscience*, *3*(2), 153.

Depue, B. E. (2012). A neuroanatomical model of prefrontal inhibitory modulation of memory retrieval. *Neuroscience & Biobehavioral Reviews*, *36*(5), 1382-1399.

Diamond, N. B., & Levine, B. (2018). The Prefrontal Cortex and Human Memory. In J.L. Cummings (Ed.), *The human frontal lobes.* The Guilford Press. New York, NY.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in cognitive sciences*, *11*(9), 379-386.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, *18*(9), 547.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.*, *30*, 123-152.

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature, 425*, 184.

Gilboa, A. (2004). Autobiographical and episodic memory—one and the same? Evidence from

prefrontal activation in neuroimaging studies. *Neuropsychologia, 42*, 1336-1349.

Leal, S. L., & Yassa, M. A. (2018). Integrating new findings and examining clinical applications of pattern separation. *Nature neuroscience*, 1.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in cognitive sciences*, *6*(7), 299-305.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus*, *16*(12), 1091-1101.

Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., & Kahana, M. J. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, *342*(6162), 1111-1114.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory?. *Psychological bulletin*, *135*(4), 638.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*.

O'Keefe, J., & Nadel, L. (1979). Précis of O'Keefe & Nadel's The hippocampus as a cognitive map. *Behavioral and Brain Sciences*, *2*(4), 487-494.

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in cognitive sciences*, *17*(5), 230-240.

Ranganath, C. (2010). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus*, *20*(11), 1263-1290.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713.

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current opinion in neurobiology*, *23*(2), 255-260.

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

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and space: towards an understanding of the cognitive map. *Journal of Neuroscience*, *35*(41), 13904-13911.

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

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, *99*(2), 195.

Stark, S. M., Reagh, Z. M., Yassa, M. A., & Stark, C. E. (2018). What's in a context? Cautions, limitations, and potential paths forward. *Neuroscience letters*, *680*, 77-87.

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, *15*(10), 655.

Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, *1*, 381-403.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in cognitive sciences*, *9*(9), 445-453.

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivate sensory-specific cortex. *Proceedings of the National Academy of Sciences, 97*, 11125-11129.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature neuroscience*, *18*(4), 582.

Wimber, M., Bäuml, K. H., Bergström, Z., Markopoulos, G., Heinze, H. J., & Richardson-Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *Journal of Neuroscience, 28(50),* 13419-13427.

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K. H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience*, *21*(3), 538-549.

Wolbers, T. (2015). Spatial Navigation. In: James D. Wright (editor-in-chief), *International Encyclopedia of the Social & Behavioral Science*s, 2nd edition, Vol 23. (pp. 161–171). Oxford: Elsevier.

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

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language*, *46*(3), 441-517.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological bulletin*, *133*(5), 800.