Mnemonic Discrimination: Correcting False Memories and Detecting Changes in Time

Nathan M. Muncy
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Family, Life Course, and Society Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/8951

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Mnemonic Discrimination: Correcting False Memories
and Detecting Changes in Time

Nathan M. Muncy

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

C. Brock Kirwan, Chair
Shawn D. Gale
Derin Cobia
Jared A. Nielsen

Department of Psychology
Brigham Young University

Copyright © 2020 Nathan M. Muncy
All Rights Reserved
ABSTRACT

Mnemonic Discrimination: Correcting False Memories and Detecting Changes in Time

Nathan M. Muncey
Department of Psychology, BYU
Doctor of Philosophy

Two projects are presented in this dissertation. First, we investigated the impact of false memories on the original trace and whether recovery of the original is possible. Second, we tested whether mnemonic discrimination for temporal duration is possible. Both projects incorporated fMRI techniques in order to implicate any potential neural correlates of these memory behaviors.

Project One. To elicit false memories and study a potential recovery therefrom, thirty-five healthy, young adults performed multiple recognition memory tests where they were induced to make errors in a first test and then participants were tested again in a surprise, second test. These two tests allowed us to determine which memory error would be corrected, if any. Further, fMRI signal associated with the encoding and retrieval processes during the experiment allowed us to implicate the regions associated with false memory correction. We found that false memories do not overwrite the original trace in all instances, as recovery of the original information was possible. Critically, we determined that recovery of the original information was dependent on activity in regions associated with retrieval, saliency attention, and bottom-up attention during the formation of the false memory, and not on processing at the time of encoding or the second test episode.

Project Two. We developed a novel paradigm to test episodic memory for temporal duration. Thirty-five healthy, young adults completed a temporal discrimination task that consisted of a continuous-recognition paradigm in which visual objects were presented one at a time for either 1 or 1.5 seconds. Certain items repeated (Targets and Lures) where Targets were presented for the same duration while the duration of Lures was altered by ±0.5 seconds. Participants were asked to identify whether the stimulus duration changed. Whole-brain high-resolution fMRI data were acquired. Behavioral results indicate that participants were sensitive to both increases and decreases in duration. Further, fMRI analyses revealed that the left entorhinal and perirhinal cortices were differentially involved in encoding and retrieval, respectively, of correct duration representations. These findings support the notion of the entorhinal cortex supporting temporal representations in memory as well as the perirhinal cortex representing the conjunction of item and context.

Keywords: episodic memory, mnemonic discrimination, false memory, time, MRI
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Brock Kirwan, for his mentoring, patience, and enthusiasm throughout these projects. I would also like to thank my wife, Ariana Hedges-Muncy, for the personal and statistical support that is always charitably given. This finished work is not my accomplishment, but our accomplishment.
# TABLE OF CONTENTS

Mnemonic Discrimination: Correcting False Memories and Detecting Changes in Time........... 1

Chapter 1 Correcting False Memories ................................................................. 5

1   INTRODUCTION ......................................................................................... 5

2   METHODS ............................................................................................... 7
   2.1   Participants ..................................................................................... 7
   2.2   Mnemonic Discrimination Task ......................................................... 8
   2.3   Statistical Analyses of Behavioral Data ............................................. 11
   2.4   MRI Acquisition ............................................................................. 12
   2.5   MRI Data Pre-Processing ............................................................... 12

3   RESULTS ............................................................................................... 19
   3.1   Behavioral Analyses and Descriptive Statistics ............................... 19
   3.2   Functional MRI Analyses .............................................................. 22

4   DISCUSSION ......................................................................................... 38

Chapter 2 Discriminating Temporal Information .............................................. 46

1   INTRODUCTION ..................................................................................... 46

2   METHODS – Behavioral Experiment 1 .................................................... 49
   2.1   Participants ..................................................................................... 49
   2.2   Temporal Discrimination Task ......................................................... 49
   2.3   Statistical Analyses ....................................................................... 52

3   RESULTS – Behavioral Experiment 1 ...................................................... 54

4   METHODS – Behavioral Experiment 2 .................................................... 55

5   RESULTS – Behavioral Experiment 2 ...................................................... 56

6   METHODS – fMRI Experiment .............................................................. 56
   6.1   Participants ..................................................................................... 56
   6.2   Temporal Discrimination Task and Statistics ................................ 57
   6.3   MRI Acquisition ............................................................................. 57
   6.4   MRI Data Pre-Processing ............................................................... 58

7   RESULTS – fMRI Experiment ............................................................... 61
   7.1   Behavioral Results ....................................................................... 61
   7.2   FMRI Results - Encoding .............................................................. 61
   7.3   FMRI Results – Retrieval ............................................................... 64

8   DISCUSSION ......................................................................................... 64

References .................................................................................................... 72
LIST OF FIGURES

Figure 1-1: Mnemonic discrimination task ................................................................. 10
Figure 1-2: Hippocampal subregion masks .............................................................. 13
Figure 1-3: NeuroSynth encoding masks ................................................................. 17
Figure 1-4: NeuroSynth retrieval masks ................................................................. 18
Figure 1-5: Boxplots of response proportions ........................................................ 20
Figure 1-6: Schematic of analyses performed ....................................................... 23
Figure 1-7: SpT1 whole-brain results ................................................................. 25
Figure 1-8: Test1 hippocampal subfield results .................................................... 28
Figure 1-9: Test1 NeuroSynth results ................................................................. 30
Figure 1-10: Test1 whole-brain results ................................................................. 32
Figure 1-11: T1pT2 NeuroSynth results ................................................................. 34
Figure 1-12: T1pT2 whole-brain results ................................................................. 35
Figure 1-13: T2fT1 NeuroSynth results ................................................................. 38
Figure 2-1: Temporal discrimination paradigm .................................................... 51
Figure 2-2: Boxplot of d' scores .......................................................................... 55
Figure 2-3: Regions engaged in duration detection ............................................. 63
LIST OF TABLES

Table 1-1: Cluster sizes and test statistics for the exploratory whole-brain analyses ................. 25
Table 1-2: NeuroSynth ROI ANOVA statistics ........................................................................ 29
Table 1-3: NeuroSynth ROI post-hoc t-tests corresponding to Table 1-2 ................................. 29
Table 2-1: Response behavior classifications ......................................................................... 53
Table 2-2: Cluster descriptions of small-volume corrected analyses ........................................ 63
Mnemonic Discrimination: Correcting False Memories and Detecting Changes in Time

Episodic memory is a subset of declarative memory by which the retrieval of personal events is possible (Baddeley et al., 2009). For instance, the retrieval of a college lecture includes specific information about the student’s history (who gave the lecture, where it occurred, what was covered, etc.) at a specific time (when in the semester). Such information organization in memory systems allows for the purposeful recall of relevant information to guide future behaviors like exam performance. An individual’s day-to-day interactions with the environment are, however, on average rather similar: a student attends lectures from the same professors on similar material at the same time of day while seated next to the same classmates, and accordingly the episodic memory of one lecture may be quite similar to that of another. While it may not be necessary or even reasonable to form hyper-accessible episodic memories of every minute detail encountered throughout the day, certain pieces of important yet highly-similar information need to be represented separately and accurately, such as where the car was parked today versus yesterday or whether the iron was unplugged before leaving the house. Given that we are capable of retrieving highly similar but separate memories, a mechanism must exist that serves to write similar information in unique ways for future use.

Stemming from the work by Eccles et al. (1967) on cerebellar architecture, Marr (1969) proposed that such organization of the granule and Purkinje cells performed the function of “amplify[ing] discrepancies between patterns that are rather similar” (p. 440), a process which he termed “pattern separator”. Accordingly, nearly identical afferent information could result in rather distinct neural patterns of activity via such a pattern-separating mechanism, or as Marr
describes, “similar inputs have markedly less similar codons” (Marr, 1969, p. 444). It was not long before such a mechanism was applied to memory (Marr, 1971). Building upon the work of Marr (1969, 1971) and others (Rumelhart et al., 1986; Wigström, 1977), Rolls (1989) proposed a computational model of hippocampal pattern separation based on the exceptionally low probability of any one CA3 neuron being innervated by a dentate gyrus (DG) mossy fiber; such cellular circuitry demonstrated the computational ability to orthogonalize similar patterns of information. Subsequent work in hippocampal tissue and modelling (Rolls, 1991, 1996; Rolls & Treves, 1994; Treves & Rolls, 1992, 1994) indeed demonstrated that pattern separation processes occur within the hippocampus, where information is sparsely encoded by DG input into CA3 via the mossy fibers (Hunsaker & Kesner, 2013; Rolls, 2010, 2013; Rolls & Kesner, 2016). In this way, a neural-computation mechanism has been established by which two highly similar pieces of information (A, A’) may be stored separately in memory systems and independently reactivated.

While work on hippocampal pattern separation processes has largely been conducted with rodents and non-human primates using spatial- and object-based tasks, memory paradigms have been developed for work with human participants (Bakker et al., 2008; Kirwan & Stark, 2007). From such mnemonic paradigms, it was correspondingly demonstrated that the DG/CA3 region was integral to memory performance associated with pattern separation processes. Subsequent work not only established the importance of these medial temporal regions for the behavioral correlates of pattern separation (Lacy et al., 2011; Yassa & Stark, 2011), here termed “mnemonic discrimination”, but also demonstrated that mnemonic discrimination performance deteriorated in proportion to hippocampal and DG degradation (Kirwan et al., 2012). Interestingly, mnemonic discrimination performance also decreases as a function of mild
cognitive impairment (Stark et al., 2013), which implies that while pattern separation is
necessary for mnemonic discrimination, it is not sufficient. Indeed, networks associated with
visual processing, attention, and goal or task orientation also impact how an individual will
perform in memory tasks (Cabeza, 2013; Cabeza et al., 2008; Geib et al., 2017).

Typically, mnemonic discrimination studies present a series of visual images of common
items, with some stimuli repeating while others having a similar counterpart presented, and then
ask the participant to identify whether or not they have been previously exposed to the stimulus
(Stark et al., 2019). If an object has previously been presented, it is called a “Target” with a
correct response resulting in a “Hit” while an incorrect response results in a “Miss”. Finally, if an
object is similar but different it is called a “Lure” with the corresponding correct and incorrect
responses being “Correct Rejection” and “False Alarm”, respectively. Research on mnemonic
discrimination has focused largely on Correct Rejections, where a Lure is correctly identified,
and such a behavior is thought to rely on pattern separation processes. Previous literature has
demonstrated that the DG/CA3 region is essential to this behavior (Bakker et al., 2008; Yassa &
Stark, 2011). However, False Alarms (FA), which are largely unaddressed in the literature, are
also interesting because of their potential impact on the original memory trace; it is possible that
the FA response, which results from the failure to detect how the salient information in the Lure
differs from the original stimulus, could write a memory representation separate from that of the
original stimulus rather than overwriting the original memory trace. That is, incorrectly
remembering where the car is parked could result in a false memory that is separate from the
original memory representation holding the true location of the car. Accordingly, the first aim of
these experiments is to investigate the effects of a FA response on the original memory trace, and
whether or not recovery of the original memory representation is possible. Further, the stimuli
used in mnemonic discrimination tasks have traditionally been either visual or visuospatial in
nature (e.g., Reagh et al., 2014; Roberts et al., 2014; Stark et al., 2013; Yassa & Stark, 2011), but
given that mnemonic discrimination is based on pattern separation processes, which are likely
information agnostic (as they are known to occur outside of the hippocampus; see Marr, 1969;
Wigström, 1977), participants should be able to discriminate in explicit memory other
information than spatial and/or visual. Embedded in episodic memory is temporal information;
for instance, the memory of a college lecture has information coding for both when the lecture
took place as well as the duration of the lecture. This temporal information should be accessible
to the participant just as the other information modalities are available, and indeed it is easy to
recall if a lecture was too long or that it takes longer to drive to one destination than another.
Therefore, the second aim of these experiments is to investigate whether participants can
discriminate for temporal information and what cerebral processes are involved in such a
behavior.

What follows are two studies totaling for three experiments examining the neural
correlates of the FA response and subsequent attempt to retrieve the original memory trace, and
the mnemonic processing and discrimination of temporal information, respectively. In the first
study, termed “Correcting False Memories”, data have been collected and analyzed in order to
address the effects of the FA response as well as how it is that an individual can recover from a
false memory. In the second study, termed “Discriminating Temporal Information”, two
experiments examine memory for temporal duration. In the first experiment, a novel paradigm
was developed to determine whether participants are sensitive to changes of temporal duration in
episodic memory, and the second experiment investigated the capacity of such sensitivities using
MRI techniques.
Chapter 1 Correcting False Memories

1 INTRODUCTION

The field of study on mnemonic discrimination in episodic memory has largely focused on a single aspect—the capacity of the participant to correctly identify lure stimuli. In animal models, where pattern separation processes can be targeted, previous work has demonstrated that the successful identification of lures is dependent on the orthogonalization processes of the dentate gyrus in the hippocampus as well as an intact hippocampal trisynaptic pathway (Gilbert et al., 1998, 2001; Gilbert & Kesner, 2006; Gold & Kesner, 2005; Hunsaker et al., 2008; Kesner, 2007a, 2007b; Leutgeb et al., 2007). Human studies have implicated the same processes for successful mnemonic discrimination performance: participants with healthy hippocampi perform well on tests of mnemonic discrimination while participants with hippocampal atrophy perform poorly and participants with hippocampal lesions perform worst of all (Bakker et al., 2008; Kirwan et al., 2012; Mueller et al., 2011; Small et al., 2011). Additionally, hippocampal CA3/DG volumes in humans are known to positively correlate with performance on tests of mnemonic discrimination (Doxey & Kirwan, 2015; Grady & Ryan, 2017; Leal & Yassa, 2014; Marshall et al., 2016; Reagh et al., 2014; Roberts et al., 2014; Sheppard et al., 2016). That said, while mnemonic discrimination processes appear to be pattern separation dependent, higher order processes such as semantic labeling, visual spatial processing, attention, and decisiveness (Goebel & Vincze, 2007; Hunsaker & Kesner, 2013; Lacy et al., 2011; Leal & Yassa, 2014; Pidgeon & Morcom, 2016; Reagh et al., 2016; Reagh & Yassa, 2014; Rolls & Kesner, 2016; Steemers et al., 2016; Wais et al., 2017) integrate with hippocampal pattern separation output to coordinate the overall behavior. But while the roles of the medial temporal lobe structures in
mnemonic discrimination are well understood, higher-order processes are relatively understudied in this context.

In addition to these gaps in the field of mnemonic discrimination, the failure in detecting lures is also relatively unaddressed. The detection of lure stimuli in a recognition memory test is the result of both encoding and retrieval processes: the lure’s representation must be successfully orthogonalized from the target’s, and the target must be successfully retrieved. At the point of successful lure identification, then, two distinct representations exist (Target, Lure) between which the participant is able to discriminate. This process is sometimes known as “recall to reject” (J. Kim & Yassa, 2013; Kirwan & Stark, 2007; Lacy et al., 2011; van den Honert et al., 2016). A failure in lure detection, then, could have multiple antecedents: a failure to encode the lure’s salient details, a failure to retrieve the salient target information, and/or lure interference may lead to mnemonic generalization rather than discrimination. Such failures may result from underperforming lower-order hippocampal processes, but may also be the result of deficient attention, semantic labelling, and/or decision making. Further, and perhaps more interestingly, the consequence of incorrectly identifying the Lure as Target (a False Alarm, FA) is unknown: as studies of mnemonic discrimination typically use a single study/test presentation, the effect of the FA response on the original memory trace has not been established. It is possible that either the lure representation could overwrite the target memory trace, or that a second representation could exist concurrently with the original. If the latter is true and multiple memory traces result from a FA response, then the two memory traces (one of the Target, and one of the Lure) may be dissociable in subsequent testing. And while certain studies, particularly in the false-memory literature, have specifically targeted the FA responses (rather than record it as a task failure in the mnemonic discrimination task) and the underlying mechanism at both the encoding and test
phase of the trial (Cabeza, 2013; Hanczakowski & Mazzoni, 2011; H. Kim, 2011; Kubota et al., 2006; Moritz et al., 2006; Okado & Stark, 2005), the differential processing that could potentially be involved with correcting versus perpetuating FA behavior has not been investigated.

In order to address these gaps in the literature, we developed a paradigm to investigate both the impact of a false alarm on the original memory trace as well as the relative contributions of hippocampal and cortical processes in promoting and recovering from a failed lure detection. We hypothesized that (a) FAs would result in multiple memory traces that exist concurrently, and as such, we predicted that participants would be able to distinguish target from lure when both were presented simultaneously. Further, we hypothesized that subsequent target-trace recovery would be predicted by (b) increased BOLD signal in CA3/DG during the orienting task and false alarm, and (c) increased BOLD signal in a priori regions associated with encoding and retrieval. Finally, we performed exploratory whole-brain analyses with the prediction that (d) we would observe regions associated with higher-order cognitive processes, such as attention and content processing, that were involved in subsequent target-trace recovery. In short, we predicted that not all FA responses are equal, but that some would occur during better processing of the Lure stimulus which would then precede a retrieval of the Target memory trace during a subsequent test.

2 METHODS

2.1 Participants

Thirty-five participants (15 female, mean age = 23.1, SD = 2.2) were recruited from the University and local community and were compensated for their participation with either payment or a 3D-printed, quarter-scale model of their brain. The university’s Institutional
Review Board approved the research, and all participants gave written informed consent prior to participation. Inclusion criteria consisted of speaking English as a native language, being right-hand dominant, and having normal or corrected-to-normal vision. Exclusion criteria consisted of a history of psychologic disorder or brain injury, head dimensions that were too large for the 32-channel coil, the existence of non-surgical metal in the participant’s body (including non-removable piercings), and color blindness. Participants were excluded from MRI analyses separately for excess motion within a particular phase of the experiment as noted in the MRI Data Preprocessing section below.

2.2 Mnemonic Discrimination Task

Stimuli for the task consisted of 200 pairs of perceptually and semantically similar Target-Lure pairs that were randomly selected from those hosted at https://github.com/celstark/MST. Each stimulus depicted an everyday object in the center of the image with a white background, and one member of each set was randomly selected to be a Target while the other was selected to be a Lure for each participant. Previous to this study, approximately 900 Target-Lure pairs were rated for similarity by 35 individuals (who were independent of the current study) on a 7-point Likert scale where a response of “1” indicated “extremely dissimilar” and “7” indicated “extremely similar”. Only items that received an averaged similarity rating of 6.0 or higher were used in this task in order to ensure a sufficiently large number of FA responses for functional analyses.

The behavioral task consisted of three phases (Study, Test1, and Test2), all of which were performed while the participant was scanned using fMRI. Stimuli were displayed by means of an MR-compatible LCD monitor placed at the head of the scanner and viewed using a mirror mounted on the head coil. Responses were collected via a fiber optic button box. E-prime (v 2.0)
was used to control stimulus display and record behavioral responses. While in the scanner but
dprior to engaging in the task, participants viewed a task instruction video; for consistency, the
same researcher (NMM) screened and debriefed all participants, responded to any questions
about the task, and operated the scanner. Three different instruction videos were shown, one
immediately prior to each respective phase of the task, in which explicit directions and examples
were given to the participant.

Study phase: In the initial encoding phase, participants were shown 200 images of
everyday objects and were instructed to indicate via button press whether each object is typically
encountered indoors or outdoors (Figure 1-1, left). Each stimulus was preceded by a fixation
cross for 500 ms, then presented for 3500 ms across one continuous block. For all phases of the
experiment, 16 blank trials were added in a pseudo-random fashion to create jitter in stimulus
timing, with care taken to ensure that each block contained an equal number of blank trials
(when applicable). These blank trials, in addition to the inter-stimulus fixation crosses, served as
the baseline condition in the single subject regression analysis (see MRI Data Pre-Processing).

Test phase: Participants were shown 100 Target stimuli (items identical to the original
stimuli in the Study phase), and 100 Lure stimuli (items which were similar to, but differed
visually from, the original items) one at a time in the center of the screen. Participants were
instructed to indicate via button press whether each stimulus was the “same” (an exact repeat, or
Target) or “different” (a related Lure) to the study item (Figure 1-1, middle). Stimuli were
presented in a randomized order for both type (Target/Lure) and for level of Target-Lure
similarity. The Test phase was conducted in two blocks of 100 stimuli each (50 Target, 50
Lure), where each stimulus was presented for 3500 ms and preceded by a fixation cross for 500
ms.
Test2 phase: In this phase, all stimuli from the Study phase (Targets), accompanied with the similar item (Lure), were presented side-by-side to the participant in a two-alternative forced-choice format. The stimuli were counterbalanced for whether the Target appeared on the Left or Right (Figure 1-1, right). Participants were asked to indicate which of the two items they had seen during the Study phase (Left or Right) and their confidence in their decision (Very or Somewhat) via one of four potential response options (left very confident; left somewhat confident; right somewhat confident; right very confident). This phase was conducted in four blocks of 50 stimuli in an attempt to combat participant fatigue. Each stimulus was presented for 5000 ms and preceded by a fixation cross for 500 ms. During all phases of this experiment, participants were instructed to respond while the stimulus was presented on the screen, and no feedback was given relating to participant performance.

**Figure 1-1. Mnemonic discrimination task.** Left, during the study phase the participants made indoor/outdoor judgements about a series of objects. Middle, during the first test phase, participants were presented with either the same (Target) stimulus or one that was similar (Lure) and were asked to decide whether the stimulus was the same or similar. Right, during the second test phase, participants were presented with both Target and Lure stimuli simultaneously and asked to indicate which item was originally presented during the Study phase. Additionally, participants indicated the level of confidence (Very or Somewhat) in their decision and made both their target and confidence decision with a single button press. For instance, if they decided that the left stimulus was the Target and they were very confident, they pressed “1” whereas if they were only somewhat confident the right stimulus was the Target, they pressed “3”.
2.3 Statistical Analyses of Behavioral Data

Participant responses during the Study phase of the experiment were not analyzed for indoor-outdoor accuracy as this question was merely meant to orient participants’ attention to the stimuli. During the second phase of the experiment, Test1, participants saw either Target or Lure stimuli and made a forced-dichotomous judgment of whether the item was “Same” or “Different”. Correctly identifying the Target and Lure is termed Hit and Correct Rejection (CR), respectively, and the incorrect identification is termed Miss and False Alarm (FA), respectively. Here, we operationalize a false memory as the FA response to the lure stimulus. The third phase of the experiment, Test2, was a two-alternative forced-choice where both Target and Lure stimuli were presented and the participant was asked to identify the Target; consequently, only Hits and FAs were possible.

Behavioral analyses were conducted on response proportions, considering the number of items to which each participant responded. D-prime (d') scores were calculated for both Test1 and Test2 phases as an indicator of each participant’s sensitivity to the stimuli. Test1 d' scores were calculated the standard way (Equation 1; Wickens, 2002) where \( Z = z\text{-score}, \ pHit = \frac{\text{proportion of Hits}}{\text{number of Targets}}, \ pFA = \frac{\text{proportion of FAs}}{\text{number of Lures}} \). As the Test2 phase presented both stimuli simultaneously, a correction was added to the d' calculation (Equation 2; Wickelgren, 1968).

\[
\begin{align*}
d'_{T1} &= z(pHit) - z(pFA) \\
d'_{T2} &= \frac{z(pHit) - z(pFA)}{\sqrt{2}}
\end{align*}
\]

Equations (1, 2)

Further, one-sample proportion testing using a continuity correction were used assess whether participants were more likely to respond one way versus another (e.g., if participants were more likely to Hit or Miss when presented with a Target, or more likely to respond Very versus Somewhat confident during CRs). Outliers were detected according to the \( 1.5 \times \text{IQR} \).
method. Paired and one-sample \( t \)-testing was used to compare the d' scores to each other and against 0, respectively, utilizing an FDR correction to account for multiple comparisons. When the exclusion of outliers differentially impacted the statistics, tests including and excluding outliers are reported, otherwise reporting represents the exclusion of outliers.

2.4 MRI Acquisition

Functional and structural images were acquired on a Siemens TIM Trio 3T MRI scanner utilizing the 32-channel head coil. Participants contributed a T1-weighted scan, a high-resolution T2-weighted scan, and a series of functional T2*-weighted scans as described below. Standard-resolution structural images were acquired using a T1-weighted magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence with the following parameters: TR = 1900 ms, TE = 2.26 ms, flip angle = 9°, FoV = 218 × 250, voxel size = 0.97 × 0.97 × 1 mm, slices = 176 interleaved. High-resolution structural images were acquired using a T2-weighted sequence with the following parameters: TR = 8020 ms, TE = 80 ms, flip angle = 150°, FoV = 150 × 150 mm, voxel size = 0.4 × 0.4 × 2 mm, slices = 30 interleaved. The T2 data were acquired perpendicular to the long axis of the hippocampus. High-resolution multi-band echo-planar images were acquired with a T2*-weighted pulse sequence with the following parameters: Multi-band factor = 8, TR = 875 ms, TE = 43.6 ms, flip angle = 55°, FoV = 180 × 180 mm, voxel size = 1.8 mm\(^3\), slices = 72 interleaved. This sequence was aligned parallel with the long axis of the hippocampus, and the first 11 TRs were discarded in order to allow for T1 equilibration.

2.5 MRI Data Pre-Processing

The pre-processing and visualization of structural and functional MRI data were accomplished with the following software: dcm2nii (Li et al., 2016), Convert 3D Medical Image Processing Tool (c3d) and ITK-SNAP (Yushkevich et al., 2006), Analysis of Functional
NeuroImages (AFNI; version 18.3.15; Cox, 2012), Advanced Normalization Tools software (ANTs; Avants et al., 2008, 2011), Automatic Segmentation of Hippocampal Subfields (ASHS; Yushkevich et al., 2015), and Multi-image Analysis GUI (Mango; Lancaster et al., 2012).

T1-weighted structural files were converted into 3D NIfTI files, rotated into functional space (referencing the volume with the minimum calculated outlier voxels) using a rigid transformation with a Local Pearson Correlation cost function. The rotated scan was then skull-stripped and warped into MNI space via a non-linear diffeomorphic transformation.

High-resolution T2-weighted files were converted to 3D NIfTI files, and both T1w and T2w files, in native space, were used to segment hippocampal subregions via ASHS, referencing the UPenn atlas (Yushkevich et al., 2015). Hippocampal subregion masks from each participant were then used to construct template priors via Joint Label Fusion (Figure 1-2; H. Wang et al., 2013), and study-specific masks for the subiculum, CA1, and the combined CA2-CA3-DG regions where constructed and resampled into functional dimensions. Voxels associated with overlapping subregions were excluded to combat partial-voluming effects when down-sampling, and the CA2, CA3, and DG masks were combined into a single mask due to functional resolution constraints.

Figure 1-2. Hippocampal subregion masks. Left, a participant’s T2-weighted scan segmented for the medial temporal lobe via ASHS. Middle, template medial temporal lobe masks constructed in MNI space via Joint Label Fusion. Right, the resampled masks, where overlapping labels were excluded.
T2*-weighted DICOMs were first converted to 4D NIfTI files, and the volume from the experiment with the smallest percentage of outlier voxels was extracted to serve as a volume registration base. The calculations for moving each volume into the same space as the volume registration base were then performed, using cubic polynomial interpolations, and the T1-weighted file was also aligned with the volume registration base using a rigid transformation that utilized a Local Pearson Correlation cost function. This rotated structural scan was then used to produce normalization calculations via a non-linear transformation into MNI space. Movement of the functional data into template space was then accomplished by concatenating the volume registration calculations with the MNI transformation calculations, thereby moving all volumes from native space to MNI space via a single transformation, resulting in less partial-voluming and blurring of the functional data than if each transformation was applied separately. A series of masks were then generated in order to remove volumes with missing data as well as determine the intersection between functional and structural data, and the functional data were then scaled by the mean signal. Finally, centered (demeaned) motion files for six degrees of freedom were generated for each run of the experiment.

Behavioral data during each of the three phases (Study, Test1, and Test2) were coded according to behavioral responses during that and subsequent phases of the experiment. Trials from the Study phase were coded according to subsequent Test1 performance (i.e., Study predicting Test1 activation or SpT1): subsequent Hit, subsequent FA, subsequent Miss, or subsequent CR. Additionally, Study-phase trials were coded according to behavioral responses in Test1 and Test2 (SpT1pT2). For example, a stimulus presented during Study might subsequently receive a FA in Test1 and a Hit in Test2 (subsequent FA-Hit). Coding Study trials in this way allows us to potentially localize differences during the encoding process that result in Test1 FAs
that are perpetuated into Test2 (subsequent FA-FA) compared to FAs that are corrected in Test2 (subsequent FA-Hit). It should be noted, however, that a low number of certain behaviors (particularly Test1 Misses preceding Test2 behaviors) necessitated collapsing across CR and Miss response bins in the deconvolution phase in order to avoid producing vectors for some participants consisting entirely of zeros, vectors which are incompatible with deconvolution. Such behavioral bins were not included in group-level analyses. Test1 trials were coded according to four outcomes (Hit, FA, CR, and Miss). Responses during Test1 that preceded Test2 (T1pT2) produced 4×2 possible outcomes since Test2 had two possible behaviors (Hit, FA). Test2 responses that followed Test1 (T2fT1) also produced 4×2 possible outcomes, and Test2 had two outcomes. Dividing all Test2 behavioral data according to confidence interval as well as response type resulted in bins with trial counts that were too small for analyses, and as such the confidence interval was dropped from functional analyses. Finally, trials where participants failed to respond were coded separately from each of the above.

Vectors coding for the above behavioral outcomes were used in single-subject regression models (Generalized Least Square) in which the canonical hemodynamic response was convolved with a boxcar function with a duration equal to stimulus presentation. Modelling included polynomial regressors accounting for scanner drift and scan run and regressors coding for movement. The baseline condition consisted of collapsing across inter-stimulus intervals and jittered blank stimuli. Volumes consisting of movement greater than 0.3° and/or more than 0.3 mm in any direction relative to the previous volume were discarded along with the previous volume. Each deconvolution was assessed for the number of volumes that necessitated removal and any participant with more than 10% of volumes dropped and 1.5 × IQR of volumes dropped were excluded from subsequent analyses: one subject was removed from the Study phase, two
subjects were removed from the Test1 phase, and three subjects were removed from the Test2 phase. As such 34 participants were used to investigate Study performance, 33 for Test1, and 32 for Test2 performance.

Region of interest (ROI) specific masks were constructed to test a priori hypotheses. In addition to the hippocampal subregion masks described above, clusters associated with memory tasks were constructed by using the key words “Memory Encoding” and “Memory Retrieval” as search terms in the NeuroSynth.org algorithm, which conducts a meta-analysis of PubMed articles reporting the same key words and constructs a set of clusters derived from the reported coordinates in said papers. The search term “Memory Encoding” yielded 124 articles, while “Memory Retrieval” yielded 183. Association clusters (corrected for meta-analysis multiple comparisons at the p = .01 level) were downloaded and thresholded at k=80 for the encoding clusters, and k=100 for the retrieval clusters, which produced a set of binary masks that were then resampled into functional dimensions. The “encoding” set contained four clusters, consisting of the left hippocampus, right collateral sulcus, left inferotemporal gyrus, and right amygdala (Figure 1-3). The “retrieval” set contained 10 clusters, consisting of the left retrosplenial, angular gyrus, hippocampus, medial PFC, temporal-parietal junction, dorsal medial PFC, and the right hippocampus, parietal-occipital sulcus, and posterior hippocampus (Figure 1-4). NeuroSynth masks were then rotated and resampled into functional space and used to extract region-of-interest (ROI) specific mean β-coefficients, using “encoding” masks to investigate volumes from the Study phase, and “retrieval” masks for the Test1 and Test2 phases. Two-factor repeated measures ANOVAs were used to investigate an ROI × behavior interaction for hippocampal subregion and NeuroSynth masks, individually, where a Greenhouse-Geisser (GG) correction was conducted for violations of sphericity. Any significant interaction surviving
FDR multiple-comparison adjustments were investigated via a one-way repeated measures ANOVA to determine which ROIs had differential activation for the various behaviors, and subsequent pairwise $t$-testing to determine how the $\beta$-coefficients for the behaviors differed within the ROI.

**Figure 1-3. NeuroSynth encoding masks.** A left hippocampus, B left inferotemporal gyrus, C right collateral sulcus, D right posterior amygdala.
In preparation for exploratory group-level analyses, a gray matter mask was constructed in template space using Atropos priors (Tustison et al., 2014), which was multiplied by the intersection mask to produce a gray matter, intersection inclusion mask. Exploratory group-level analyses were then conducted on all voxels within this inclusion mask for each phase of the experiment using the Equitable Thresholding and Clustering method (ETAC; Cox, 2019) using the p-values 0.01, 0.005, and 0.001, blurs of 4, 6, and 8mm, nearest neighbors 1, and two-sided thresholding. Any surviving clusters were used to extract mean β-coefficients from participant data. All scripts used in this experiment can be found at [https://github.com/nmuncy/STT](https://github.com/nmuncy/STT).
3 RESULTS

3.1 Behavioral Analyses and Descriptive Statistics

Analyses of Test1 performance yield a mean \( d'_{T1} \) score of 0.66, indicating that successfully distinguishing between Targets and Lures was difficult, as the design of the experiment intended. However, participants were successful at detecting Targets and Lures in Test1, as the \( d'_{T1} \) differed significantly from zero \((t(34) = 11.3, p_{FDR} < .0001, 95\% \text{ CI } [0.54, 0.78])\). Response proportions were tested against chance (0.5) for each stimulus type; participants were more likely to Hit than Miss on Targets \((\text{estimate} = .79, \chi^2 = 32.18, p_{FDR} < .0001, 95\% \text{ CI } [.69, .86])\) while being equally likely to CR and FA on Lures \((\text{estimate} = .43, \chi^2 = 1.43, p_{FDR} = .41, 95\% \text{ CI } [.33, .54]; \text{Figure 1-5, left})\). On average, participants had 76 Hits, 54 FAs, 42 CRs, 20 Misses, and 6 non-responses in this phase. The 54 FA responses are of particular interest to the study: we designed the difficulty of the experiment to elicit approximately 60 FA responses in Test1 as two responses (Hit, FA) were possible in Test2, so subdividing Test1 FAs bins according to subsequent Test2 responses would yield sufficient trials to test statistically and model the HRF.
Performance at chance results in a proportion of 0.5. Test1 (T1), Test2 (T2), and Test1 preceding Test2 (T1 p T2). “Hit-Very Rate” = likelihood of responding Very Confident versus Somewhat Confident during Hits. “T1 Hit p T2 Hit Rate” = proportion of Test1 Hits that precede a Test2 Hit versus Test2 FA. Outliers are included.

Analyses of Test2 performance yielded a $d'_{T2}$ of 0.45, which also differed significantly from zero ($t(34) = 9.12, p_{FDR} < .0001, 95\% \text{ CI } [.35, .55]$), and indicated that participants were more likely to Hit than FA (estimate = .62, $\chi^2 = 11.43, p_{FDR} = .002, 95\% \text{ CI } [.55, .69]$; Figure 1-5, middle). In terms of confidence, participants were equally likely to indicate Very and Somewhat confident during Test2 Hits (estimate = .53, $\chi^2 = 0.2, p_{FDR} = .7, 95\% \text{ CI } [.43, .62]$) and during Test2 FAs (estimate = .38, $\chi^2 = 4.1, p_{FDR} = .09, 95\% \text{ CI } [.27, .49]$). Finally, testing the $d'$ scores for Test1 versus Test2 indicated that participants performed better during Test1
(t(34) = 4.37, \( p_{FDR} = .001 \), 95% CI [.11, .31]). On average, participants had 121 Hits, 73 FAs, and 5 non-responses.

Proportion testing of Test1 Hits and FAs which precede Test2 responses (T1pT2) reveal that a Test1 Hit was more likely to result in a subsequent Test2 Hit than a FA (estimate = .72, \( \chi^2 = 13.8, p_{FDR} < .001 \), 95% CI [.60, .82]), but that Test1 FAs were equally likely to result in either a subsequent Test2 Hit or FA (estimate = .5, \( \chi^2 \approx 0 \), \( p_{FDR} \approx 1 \), 95% CI [.37, .64]; Figure 1-5, right). These results, when viewed with respect to the Test2 results (d’ which differs from 0 and a tendency to detect the target), suggest that recovery from a Test1 FA is possible but not guaranteed. It should be noted, though, that Test2 Target detection (particularly following a Test1 Hit) may be driving some of the behavioral findings. Analyses of the BOLD response during Test1 FAs that precede Hits versus FAs in Test2 time will help determine why some Test1 FAs are perpetuated into Test2 while others are corrected, and if participants are indeed performing differently from chance.

Test1 responses that preceded Test2 confidence responses consist of permutations of Test1 behaviors, Test2 behaviors, and Test2 confidence responses. It was originally thought that Test1-Test2 FA-Hit responses would yield higher confidence responses than FA-FA, which could show interesting activation in the BOLD response. Paired \( t \)-testing revealed, however, an equal number of “Very Confident” responses for both Hits and FAs in Test2 following a Test1 FA (\( t(34) = 0.1, p = .9 \), 95% CI [-2.3, 2.5]), indeed, participants were equally likely to select “Very” and “Somewhat” during Test2 Hits and FAs following Test1 FAs (estimate = .4, \( \chi^2 = 0.7, p_{FDR} = .53 \), 95% CI [.23, .61], estimate = .4, \( \chi^2 = 0.8, p_{FDR} = .53 \), 95% CI [.22, .59], respectively). Finally, parsing the data in such a fashion yielded bins that were too small to reasonably model the hemodynamic response (e.g. Test1 FA preceding Test2 Hit-Very = 11 ±6
responses on average). As such, all subsequent analyses were collapsed across confidence ratings.

3.2 Functional MRI Analyses

To address our research hypotheses, we conducted a priori analyses on both hippocampal subregions and NeuroSynth memory regions, separately, as well as a posteriori exploratory, whole-brain analyses. First, analyses investigated the effects of encoding on subsequent test performance. The first analysis examined signal during the Study phase coded for subsequent Test1 behavior, termed “SpT1” or “Study preceding Test1” (Figure 1-6, blue), while the second analysis investigated the Study phase coded instead for trials which resulted in either a Test1 FA that was then corrected (Test2 Hit) or perpetuated (Test2 FA), termed “SpT1pT2” (Figure 1-6, purple). One participant was excluded from these analyses due to excessive movement during the Study phase (n=34). Next, two analyses investigated the Test1 phase of the experiment. One analysis investigated signal during Test1 associated with the various Test1 behaviors, termed “T1”, and the next analysis investigated Test1 FAs according to whether or not they would be corrected in Test2, termed “T1pT2” (Figure 1-6, red). Two participants were excluded from these analyses due to excessive movement in the Test1 phase (n=33). The final set of analyses involved signal from the Test2 phase. First, we investigated signal during Test2 associated with Test2 behaviors, termed “T2”. The final analysis coded trials according to whether a Test2 Hit or FA followed a Test1 FA, here termed “T2fT1” or “Test2 following Test1” (Figure 1-6, green). Three participants were excluded from the last set of analyses due to excessive movement during Test2 (n=32). In this way, the signal from each respective phase of the experiment is investigated, but trials are coded for the current phase (T1 and T2), for future responses to current stimuli (SpT1, SpT1pT2, T1pT2), or with respect to previous responses (T2fT1), thereby
allowing for the investigation of potential encoding and retrieval effects that differentially impact the memory trace and attempted recovery thereof.

<table>
<thead>
<tr>
<th>Study</th>
<th>Test1</th>
<th>Test2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoding</td>
<td>Hit</td>
<td>Hit</td>
</tr>
<tr>
<td>SpT1</td>
<td>FA</td>
<td>FA</td>
</tr>
<tr>
<td>SpT1pT2</td>
<td>T1pT2</td>
<td>T2fT1</td>
</tr>
<tr>
<td>CR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1-6. Schematic of analyses performed.** The experiment consists of three phases, Study, Test1, Test2. Study trials were coded for which stimuli preceded Test1 responses (SpT1, blue) as well as the resulting response in Test2 following a Test1 response (SpT1pT2, purple). Test1 trials were analyzed both for the behavior which occurred in Test1 (not colored) as well as for responses which preceded a Test2 response (T1pT2, red). Similarly, Test2 trials were analyzed for Test2 behaviors (not colored) as well as for which Test1 behavior the Test2 response followed (T2fT1).

### 3.2.1 Study trials preceding Test1 responses (SpT1).

Study stimulus responses were coded for subsequent Test1 behavior. For example, one stimulus in Study could elicit a Hit in Test1 while a different stimulus could elicit a FA response.

For *a priori* analyses, a two-factor repeated measures MANOVA, where hippocampal subregion ROIs (Left and Right Subiculum, CA1, and CA2-CA3-DG) served as within-subject factor A and subsequent Test1 behavior (Hit, FA, CR) served as within-subject factor B, failed to detect any significant interaction between the factors ($F(10, 330) = .89$, $\eta^2_p < .001$, $GG_\epsilon = .45$, $p(GG)_{FDR} = .47$). A similar analysis using NeuroSynth ROIs associated with encoding (Left hippocampus, inferotemporal gyrus, and right amygdala, and collateral sulcus; see Figure 1-3).
also failed to detect a significant interaction between the factors ($F(6, 198) = 1.4, \eta_p^2 < .001, GG\epsilon = .38, p(GG)_{FDR} = .24$).

Exploratory whole-brain analyses revealed seven clusters (Figure 1-7, Table 1-1) which significantly differed in Study volumes for subsequent Test1 Hits versus FAs (two subsequent behaviors were investigated given that ETAC only supports $t$-test equivalent analyses). Three clusters were more active preceding a Hit than a FA (left and right visual stream [Figure 1-7, A, B], right intraparietal sulcus [C]) which are consistent with visual and attention networks. Four clusters demonstrated the opposite pattern of activation (left and right angular gyrus [D, E], left precuneus [F], and left dorsal medial PFC [G]), where the angular gyri and precuneus regions are consistent with the default mode network.
Figure 1-7. SpT1 whole-brain results. Regions which demonstrated differential signal during SpT1. Hits versus FAs during an exploratory whole-brain analysis. A left visual stream, B right visual stream, C right intraparietal sulcus, D left angular gyrus, E right angular gyrus, F left precuneus, G left dorsal medial PFC.

Table 1-1

<table>
<thead>
<tr>
<th>Phase</th>
<th>ROI</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>T(33)</th>
<th>Sig</th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpT1</td>
<td>LVS</td>
<td>1649</td>
<td>-42.6</td>
<td>-74.3</td>
<td>6.8</td>
<td>***</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>RVS</td>
<td>2054</td>
<td>42.1</td>
<td>-72.4</td>
<td>7.1</td>
<td>***</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>RIPS</td>
<td>561</td>
<td>30.1</td>
<td>-75.2</td>
<td>33.7</td>
<td>5.3</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>LAG</td>
<td>825</td>
<td>-47.8</td>
<td>-64.7</td>
<td>39.1</td>
<td>-6.1</td>
<td>-0.11</td>
<td>-0.05</td>
</tr>
</tbody>
</table>
Cluster sizes and test statistics for the exploratory whole-brain analyses. ROI = region of interest, K = cluster size, X-Z = coordinate (peak), T(xx) = t-test(df), Sig = significant, *** = p < .0001, LB = 95% confidence interval lower bound, UB = upper bound. See text for ROI abbreviations.

### 3.2.2 Study trials preceding Test1 FAs preceding Test2 responses (SpT1pT2)

Study trials were coded for whether or not an encoding stimulus which resulted in a Test1 FA was then corrected in Test2. For example, while two stimuli both may have resulted in a Test1 FA, the participant may have successfully identified one stimulus as the Target but failed to do so for the other stimulus in Test2. Such coding allowed us to investigate the potential role of encoding in recovering from interference in Test2. On average there were an equal number of encoding trials (26) that resulted in a T1 FA that were and were not corrected in Test2.

For the a priori analyses, individual two-factor repeated measures MANOVAs were used to investigate the interaction of ROI and behaviors for hippocampal subregion as well as NeuroSynth ROIs associated with encoding. ROIs were used as factor A in the model and factor B consisted of whether a FA would be corrected (Test1 FA, Test2 Hit) or perpetuated (Test1 FA, Test2 FA). No ROI × behavior interaction was detected in this phase of the experiment for either the hippocampal or NeuroSynth ROIs ($F(5, 165) = 1.4$, $\eta^2_G < .001$, GG$\epsilon$ = .72, $p = .23$; $F(3, 99) = 26$)
Further, exploratory whole-brain analyses failed to detect any regions during the encoding phase that discriminated for whether a subsequent Test1 FA would be corrected or perpetuated in Test2. Taken together, our data suggest that while the neural activation during encoding trials had a significant impact on subsequent Test1 performance, the activation during the encoding phase did not predict a recovery from interference, as indicated by a Test1 FA preceding a Test2 Hit.

**3.2.3 Test1 responses.** Test1 trials were coded for the relative success of target and lure detection, using the labels Hit, FA, CR, and Miss. A two-factor repeated measures MANOVA detected a significant interaction of hippocampal subregions (left and right subiculum, CA1, and CA2-CA3-DG) and Test1 behaviors (Hit, FA, CR; \( F(10, 320) = 2.7, \eta_G^2 = .001, \varepsilon = .58, p(GG)_{FDR} = .09 \), respectively). One-way ANOVA testing revealed that left CA2-CA3-DG and CA1 were driving this effect \( F(2, 64) = 8.28, \eta_G^2 = .01, \varepsilon = .96, p(GG) < .001; F(2, 64) = 3.64, \eta_G^2 < .01, \varepsilon = .86, p(GG) < .05 \), respectively. Further, post-hoc pairwise \( t \)-testing indicated that the left CA2-CA3-DG was significantly more active during Hits than FAs and CRs \( t(32) = 2.16, p < .05, 95\% CI [.004, .12]; t(32) = 4.47, p < .001, 95\% CI [.06, .17], \) respectively and that the left CA1 was also more active during Hits than CRs \( t(32) = 3.06, p < .01, 95\% CI [.01, .06]; \) Figure 1-8).
Figure 1-8. Test1 hippocampal subfield results. The CA2-CA3-DG region is more active during Hits than FAs or CRs. The CA1 is more active during Hits than CRs.

Using the same analysis as above to investigate NeuroSynth masks associated with retrieval (left angular gyrus, dorsal medial PFC, hippocampus, middle frontal gyrus, medial PFC, retrosplenial cortex, temporo-parietal junction, and right hippocampus, parieto-occipital sulcus, and posterior hippocampus; see Figure 1-4) revealed a significant mask \times behavior interaction was detected  \( (F(18, 576) = 13.15, \eta_g^2 = .01, GG_\text{FDR} = .36, p(GG)_{\text{FDR}} < .0001) \). Subsequent one-factor ANOVA testing revealed that five regions (Table 1-2, T1) were driving this effect: the left temporo-parietal junction (LTPJ), ventral medial PFC (LVMPFC), dorsal medial PFC (LDMPFC), middle frontal gyrus (LMFG), and angular gyrus (LAG). Post-hoc t-testing was conducted for these regions in order to determine for which behaviors a difference in BOLD signals occurred (Table 1-3, T1), which revealed two patterns of activity (Figure 1-9). First, the...
LTPJ and LVMPFC were associated with greater BOLD signal during Hits than CRs; with such a pattern of activity it appears that these two regions are involved in recognition aspects of memory retrieval. Second, activity in the LDMPFC, LMFG, and LAG regions during CRs demonstrated greater BOLD signal than either Hits or FAs, consistent with novelty detection in memory systems. In sum, the effect found in these regions is driven largely by the differential BOLD response during CRs versus Hits and FAs (Table 1-3, T1), and the effects appear to be associated with recognition memory and novelty detection.

Table 1-2

<table>
<thead>
<tr>
<th>ROI</th>
<th>F(2,64)</th>
<th>η²</th>
<th>GGe</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTPJ</td>
<td>6.9</td>
<td>0.006</td>
<td>0.82</td>
<td>**</td>
</tr>
<tr>
<td>LVMPFC</td>
<td>7.3</td>
<td>0.016</td>
<td>0.81</td>
<td>**</td>
</tr>
<tr>
<td>LDMPFC</td>
<td>22.5</td>
<td>0.029</td>
<td>0.72</td>
<td>***</td>
</tr>
<tr>
<td>LMFG</td>
<td>10.9</td>
<td>0.011</td>
<td>0.80</td>
<td>***</td>
</tr>
<tr>
<td>LAG</td>
<td>13.9</td>
<td>0.012</td>
<td>0.81</td>
<td>***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ROI</th>
<th>F(2,64)</th>
<th>η²</th>
<th>GGe</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1pT2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDMPFC</td>
<td>20.4</td>
<td>0.025</td>
<td>0.7</td>
<td>***</td>
</tr>
<tr>
<td>LAG</td>
<td>17.9</td>
<td>0.02</td>
<td>0.9</td>
<td>***</td>
</tr>
<tr>
<td>LMFG</td>
<td>9.6</td>
<td>0.01</td>
<td>0.86</td>
<td>***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ROI</th>
<th>F(2,62)</th>
<th>η²</th>
<th>GGe</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2fT1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG</td>
<td>16.2</td>
<td>0.023</td>
<td>0.96</td>
<td>***</td>
</tr>
<tr>
<td>LDMPFC</td>
<td>12.2</td>
<td>0.019</td>
<td>0.95</td>
<td>***</td>
</tr>
<tr>
<td>LMFG</td>
<td>10.7</td>
<td>0.017</td>
<td>0.91</td>
<td>**</td>
</tr>
<tr>
<td>LPCU</td>
<td>11.8</td>
<td>0.019</td>
<td>0.88</td>
<td>***</td>
</tr>
<tr>
<td>LTPJ</td>
<td>15.1</td>
<td>0.022</td>
<td>0.87</td>
<td>***</td>
</tr>
<tr>
<td>RPOS</td>
<td>13.9</td>
<td>0.023</td>
<td>0.87</td>
<td>***</td>
</tr>
</tbody>
</table>

NeuroSynth ROI ANOVA statistics. T1 = Test1, T1pT2 = Test1 preceding Test2, T2fT1 = Test2 following Test1. GGe = Greenhouse-Geisser epsilon. ** = p < .01, *** = p < .001.

Table 1-3

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Hit-FA</th>
<th>Hit-CR</th>
<th>FA-CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROI</td>
<td>T(32)</td>
<td>Sig</td>
<td>LB</td>
</tr>
<tr>
<td>T1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTPJ</td>
<td>3.65</td>
<td>***</td>
<td>0.03</td>
</tr>
<tr>
<td>LVMPFC</td>
<td>-0.21</td>
<td>n.s.</td>
<td>-0.04</td>
</tr>
<tr>
<td>LDMPFC</td>
<td>-0.14</td>
<td>n.s.</td>
<td>-0.02</td>
</tr>
<tr>
<td>LMFG</td>
<td>-1.22</td>
<td>n.s.</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast</th>
<th>FpH-FpF</th>
<th>FpH-CpH</th>
<th>FpF-CpH</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1pT2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI</td>
<td>T(32)</td>
<td>Sig</td>
<td>LB</td>
</tr>
<tr>
<td>LDMPFC</td>
<td>2.10</td>
<td>*</td>
<td>0.00</td>
</tr>
<tr>
<td>LAG</td>
<td>0.99</td>
<td>n.s.</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

29
NeuroSynth ROI post-hoc t-tests corresponding to Table 1-2. * = p < .05, n.s. = p > .05.

<table>
<thead>
<tr>
<th>T2fT1</th>
<th>Contrast</th>
<th>Hf-FfF</th>
<th>Hf-HfC</th>
<th>Ff-FhC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG</td>
<td>-1.00</td>
<td>n.s.  -0.08</td>
<td>0.03</td>
<td>-5.84</td>
</tr>
<tr>
<td>LDMPFC</td>
<td>-0.07</td>
<td>n.s.  -0.03</td>
<td>0.03</td>
<td>-5.42</td>
</tr>
<tr>
<td>LMFG</td>
<td>0.42</td>
<td>n.s.  -0.03</td>
<td>0.04</td>
<td>-4.14</td>
</tr>
<tr>
<td>LPCU</td>
<td>0.50</td>
<td>n.s.  -0.04</td>
<td>0.06</td>
<td>-4.91</td>
</tr>
<tr>
<td>LTPJ</td>
<td>0.86</td>
<td>n.s.  -0.03</td>
<td>0.07</td>
<td>-3.87</td>
</tr>
<tr>
<td>RPOS</td>
<td>-0.40</td>
<td>n.s.  -0.06</td>
<td>0.04</td>
<td>-5.02</td>
</tr>
</tbody>
</table>

**Figure 1-9. Test1 NeuroSynth results.** Five regions (left temporo-parietal junction [LTPJ], ventral medial prefrontal cortex [LVMPFC], dorsal medial prefrontal cortex [LDMPFC], middle frontal gyrus [LMFG], angular gyrus [LAG]) differentiate between Hits, FAs, and CRs. The LTPJ and LVMPFC demonstrate significantly greater activity during Hits than CRs, while the LDMPFC, LMFG, and LAG show the opposite pattern of activity.

Exploratory analyses detected nine clusters (Figure 1-10) which demonstrated differential activation according to whether a Hit versus a FA was occurring (Table 1-1, Figure 1-10). Five regions had a greater BOLD response during Hits all of which were located in the right
hemisphere: the putamen (RPUT, A), medial PFC (RMPFC, B), angular gyrus (RAG, C),
anterior middle temporal gyrus (RAMTG, D), and precuneus (RPCU, E). These findings are
largely consistent with known patterns of activity in the default mode network associated with
successful retrieval during episodic memory tasks (Figure 1-10, B-E). Four regions demonstrated
the opposite pattern of BOLD response: left dorsal medial PFC (LDMPFC, F), insular cortex
(LINS, G), right intraparietal sulcus (RIPS, H), and inferior frontal gyrus (RIFG, I). These
regions have been associated with dorsal and ventral networks which attend to self-referential
and saliency processing, respectively, in episodic memory tasks (Figure 1-10, F-H) as well as
attention networks (Figure 1-10, I).
Figure 1-10. Test1 whole-brain results. Regions which demonstrated differential signal during Test1 Hits versus FAs during an exploratory whole-brain analysis. A right putamen, B right medial PFC, C right angular gyrus, D right anterior middle temporal gyrus, E right precuneus, F left dorsal medial PFC, G left insula, H right intraparietal sulcus, I right inferior frontal gyrus.
3.2.4 Test1 FA responses preceding Test2 responses (T1pT2). Test1 FA responses were coded according to whether or not they would be corrected in Test2 (Hit versus FA), resulting in the same number of trials per bin as in Section 3.2.2. This was done in order to investigate any potential processes occurring during the FA that may allow for subsequent recovery from such interference. Additionally, Test1 CR trials which preceded a Hit were included in the analysis in order to investigate processes which may differ during Test1 Lure trials (FA versus CR) that precede the same Test2 performance (Hit).

A two-factor repeated measures MANOVA failed to detect a significant interaction of hippocampal subregions and beta-coefficients ($F(10, 320) = 0.97$, $\eta^2_G < .001$, $GG \varepsilon = .69$, $p(GG)_{FDR} = .45$). However, an investigation into NeuroSynth regions associated with retrieval showed a significant interaction ($F(18, 576) = 7.28$, $\eta^2_G = .008$, $GG \varepsilon = .42$, $p(GG)_{FDR} < .0001$). Subsequent one-factor ANOVA testing revealed that three ROIs were driving this effect (Table 1-2, T1pT2; Figure 1-11): the left dorsal medial PFC (LDMPFC), angular gyrus (LAG), and middle frontal gyrus (LMFG). The LDMPFC differed significantly in signal across all behaviors, and of particular interest had stronger signal during a Test1 FA that preceded a Test2 Hit than a Test2 FA (Table 1-3, T1pT2). The LAG and LMFG did not differ among the Test1 FA responses but showed stronger signal for Test1 CR preceding a Hit than either Test1 FA response. These results suggest that while all three regions are particularly important for Test1 performance in terms of CRs versus FAs, only the LDMPFC had differential activity during Test1 FAs that predicted subsequent Test2 performance.
**Figure 1-11. T1pT2 NeuroSynth results.** Three regions (left dorsal medial prefrontal cortex [LDMPFC], angular gyrus [LAG], middle frontal gyrus [LMFG]) demonstrated differential activity during Test1 behaviors (FA, CR) preceding Test2 Hits and FAs. FpH = FA preceding Hit, FpF = FA preceding FA, CpH = CR preceding Hit.

Exploratory whole-brain analyses of Test1 FA trials that precede a Test2 Hit versus FA implicated two regions as being differentially engaged (Figure 1-12): the left insular cortex (LINS) and right supramarginal gryus (RSMG). Both regions had greater BOLD signal preceding a Hit than a FA (Figure 1-12, Table 1-3).
Figure 1-12. T1pT2 whole-brain results. Clusters which demonstrate differential signal during Test1 FAs that precede a Test2 Hit versus FA. A left insular cortex, B right supramarginal gyrus. FpH = FA preceding Hit, FpF = FA preceding FA.

3.2.5 Test2 responses. A two-factor repeated measures MANOVA failed to detect a significant interaction of hippocampal subregions and Test2 behaviors (Hit, FA; $F(5, 155) =$
1.04, $\eta^2_g < .001$, GGE = .75, $p(GG)_{FDR} = .39$) or among NeuroSynth ROIs associated with retrieval ($F(9, 279) = .19$, $\eta^2_g < .001$, GGE = .51, $p(GG)_{FDR} = .99$). Further, and unexpectedly, an exploratory whole-brain analysis failed to detect any regions which differentially engaged in Hits versus FAs.

### 3.2.6 Test2 responses following Test1 FAs (T2fT1)

Test2 responses (Hit, FA) following a Test1 FA were coded in order to investigate any potential processes during the second testing phase involved in overcoming interference resulting from a Test1 FA. Further, Test2 Hits following a Test1 CRs were also investigated in order to determine whether any processes were differentially involved in Test2 Hits with respect to whether the Hit was preceded by an incorrect or correct Test1 Lure identification (FA, CR). On average there were 27 Hits following a CR.

A two-factor repeated measures MANOVA failed to detect a significant interaction of hippocampal subregions and behavioral outcomes during the behavioral coding described above ($F(10, 310) = 1.1$, $\eta^2_g = .001$, GGE = .53, $p(GG)_{FDR} = .39$). An investigation in NeuroSynth masks associated with Retrieval, however, showed a significant interaction in the same analysis ($F(18, 558) = 4.68$, $\eta^2_g = .004$, GGE = .48, $p(GG)_{FDR} < .0001$), and subsequent one-way ANOVA analyses indicated that six regions were driving this effect (left angular gyrus, dorsal medial PFC, middle frontal gyrus, precuneus, temporoparietal junction, and right parietooccipital sulcus; Table 1-3, T2fT1). Greater signal was detected during Test2 in all regions for Hits following a Test1 CR than for Hits following a Test1 FA (Figure 1-13), while no difference was detected in the same ROIs for Hits versus FAs following a Test1 FA (Table 1-3, T2fT1). As such, we do not detect any compensatory processes in Test2 in these ROIs associated with overcoming interference resulting from a Test1 FA, as would be implicated by differential signal in a region...
during Test2 Hits versus FAs following a Test1 FA. Further, we find the difference between Test2 Hits according to whether they were preceded by a Test1 FA or CR to be interesting; as Test2 presents both the Target and Lure simultaneously, increased activity during a Hit following a CR but not a FA may be indicative of processes associated with selecting between two salient memory representations. Indeed, in Test2 Hits following a Test1 CR, the participant (a) encoded the Target sufficiently to be able to (b) correctly identify the Lure. When both are presented simultaneously and the participant is tasked with identifying the Target, the participant may have to resolve interference resulting from two competing representations. Finally, exploratory whole-brain analyses did not detect any regions that differed significantly between their BOLD signal associated with Test2 Hits versus FAs following a Test1 FA.
Figure 1-13. *T2fT1 NeuroSynth results.* Six regions (left angular gyrus [LAG], dorsal medial prefrontal cortex [LDMPFC], middle frontal gyrus [LMFG], precuneus [LPCU], temporoparietal junction [LTPJ], right parietooccipital sulcus [RPOS]) demonstrated differential activity during Test2 behaviors (Hit, FA) following Test1 responses (FA, CR). HfF = Hit following FA, FfF = FA following FA, HfC = Hit following CR.

4 DISCUSSION

This experiment set out to investigate the impact of a false memory response on the original memory representation, and to determine what, if any, neural processes were associated with recovering the original memory trace. To this end, we employed a Study-Test1-Test2 paradigm within an MRI scanner, in which participants were trained on a series of images (Study), tasked with identifying either Targets or similar Lures presented individually (Test1), and then tasked with identifying the Target when both Targets and Lures were presented simultaneously (Test2). False Alarm (FA) responses in Test1 were operationalized as false
memories and were used to investigate whether participants could correctly identify the Target (Hit) in Test2 following a Test1 FA. Briefly, we hypothesized (a) that a FA would write a memory trace of the Lure without necessarily overwriting the Target’s trace such that the two representations would be dissociable, and that (b) the CA3/DG regions of the hippocampus as well as (c) encoding and retrieval nodes and (d) higher-order executive processes would be integral to the formation of a second, dissociable representation.

As the paradigm anticipated, participants made a large number of FA responses in Test1, and proportion testing revealed that participants were equally likely to correct (Hit) and perpetuate (FA) the mistake in Test2 (Figure 1-5). That is, a FA does not unambiguously overwrite the Target memory trace, as would be indicated by a greater likelihood to perpetuate the FA response in Test2, nor does the FA necessarily write multiple representations from which the Target is dissociable. Given that participants were indeed successful in Test2 (Section 3.1), the behavioral data is ambiguous as to the effect of the FA on the Target trace, and so from the behavioral analyses alone we are unable to reject the null of hypothesis (a). The functional data, however, shed some light on why some FA responses are corrected and others perpetuated.

Functional analyses both replicate and reveal several aspects of memory. First, consistent with literature on subsequent memory and forgetting (Cabeza et al., 2008; H. Kim, 2011; Shrager et al., 2008), seven regions were differentially involved during the encoding phase according to test performance (SpT1). Three regions, the bilateral visual stream and right intraparietal sulcus (Figure 1-7, A-C), showed greater BOLD signal preceding a Test1 Hit versus a Test1 FA, indicative of increased visual processing and attention, respectively. Conversely, three regions, the bilateral angular gyri and left precuneus (Figure 1-7, D-F), had the opposite pattern of activity, which is consistent with increased default mode network activity during encoding.
preceding an erroneous response during test. The final region, the dorsal medial PFC region (Figure 1-7, G), also demonstrated greater BOLD signal preceding a FA versus Hit and appears to be associated with recognition memory (Phillips et al., 2009; C. Suzuki et al., 2009); as our stimulus set consists of everyday objects, increased familiarity with certain objects may result in top-down interference thereby resulting in poorer encoding. As such, suppressed signal in this region during encoding trials for stimuli that precede a Hit may be indicative of decreased familiarity with the stimulus which allowed for greater encoding and subsequent successful target identification.

Second, in addition to confirming that hippocampal subfields and NeuroSynth retrieval clusters demonstrated differential signal during this retrieval task (Figure 1-8, Figure 1-9), multiple regions were detected in an exploratory analysis of the Test1 phase of the experiment that were consistent with the retrieval literature: four regions in the right hemisphere, consisting of the medial PFC, angular gyrus, anterior middle temporal gyrus, and precuneus (Figure 1-10, B-E), all of which are again consistent with the default mode network, were differentially engaged in the task such that they demonstrated greater BOLD signal during Hits relative to FAs. These findings replicate previous results (H. Kim, 2010; Sestieri et al., 2011) which have demonstrated that the default mode network is involved in self-referential processing, an aspect of which is episodic memory retrieval. Additionally, regions associated with saliency processing and attention networks (left anterior insula, right intraparietal sulcus, right inferior frontal gyrus; Figure 1-10, G-I) had greater signal during FA responses relative to Hits. It appears that such patterns of activation are the result of the unsuccessful attempt to overcome significant interference resulting from the highly similar Lures used in this task. Indeed, the right intraparietal sulcus is particularly interesting as it is implicated in post-retrieval processes, where
 hippocampal information is integrated with executive orchestration (Cabeza et al., 2008; Gilmore et al., 2015; Hao et al., 2018; Sestieri et al., 2011; Uncapher & Wagner, 2009).

Third, and importantly, differential signal was detected during Test1 FAs according to whether or not the FA was perpetuated into Test2 (T1pT2). Three regions, the left dorsomedial PFC, left anterior insula, and right supramarginal gyrus (Table 1-3, Figure 1-11, Figure 1-12), demonstrated significantly greater BOLD signal during Test1 FAs the preceded a Test2 Hit versus a Test2 FA. The dorsomedial PFC region was derived from a NeuroSynth search term of “Memory Retrieval” and is often implicated in the retrieval of episodic memory while the left anterior insular cortex and right supramarginal gyrus are themselves associated with saliency and bottom-up attention networks (Cabeza et al., 2008; Euston et al., 2012; Menon & Uddin, 2010; Seeley et al., 2007; Shirer et al., 2012). These findings, then, suggest that increased activity in these regions during a Test1 FA may be associated with better retrieval of the original memory trace (LDMPFC) and attention to the relevant details of the stimulus (LINS, RSMG) which, while subthreshold for Lure detection, nevertheless led to the formation of a Lure representation such that the Target and Lure memory traces were subsequently dissociable in Test2.

Correspondingly, no dissociable pattern of activity was detected during the encoding phase that predicted which Test1 FAs would be corrected in Test2 (SpT1pT2), nor is any such pattern detected in Test2 when comparing Hits to FAs for stimuli following a Test1 FA (T2fT1) as would be expected if such resolution of Target-Lure interference was only a function of the quality of Target encoding (SpT1pT2) or retrieval processes used to dissociate Target and Lure representations (T2fT1). Instead, our data suggest that the capacity to recover the Target trace following a FA is entirely dependent on processes occurring during the FA response, and not before or after, where greater attention and retrieval processes result in a dissociable Lure trace.
As such, and with respect to the behavioral data, the rejection of hypothesis (a) may or may not be warranted; it appears that multiple representations can occur during a FA but we do not demonstrated that multiple representations will occur, as the probability appears to dependent on retrieval and attention processes during Test1. We do, however, reject the null of hypotheses (c) and (d).

Some results of this study were surprising. First, participants were equally likely to Hit or FA in Test2 to stimuli that had previously resulted in a Test1 FA (Figure 1-5, right), which is to say that their Test2 behavior for these trials was not significantly different from chance. While the d’ score of Test2 differed significantly from 0, this was likely driven by the large number of Test2 Hits following a Test1 Hit. Analyses of confidence ratings indicate that participants may have been less confident in the Test2 FA responses (Figure 1-5, middle); note, though, that while the confidence intervals of the proportion test do not include 0.5, indicating a difference from chance, the FDR-corrected p-value is not significant at the predetermined $\alpha$. Having confidence intervals that do not include 0.5 gives some indication that participants were not responding with FAs and Hits in Test2 at random. Regardless of the responses at varying confidences, however, an equal likelihood to Hit or FA in Test2 following a Test1 Hit undermines subsequent fMRI analyses in that any differential signal detected may simply be the capitalization on noise and confirmation bias. That said, such a result may very well be a function of the experimental design itself: a large number of difficult Target-Lure pairs were utilized in order to elicit a sufficient number of Test1 FA responses such that dividing the number of Test1 FA responses by subsequent behavior would still have sufficient trials to model the HRF, resulting in a task that lasted 1.5 hours. Such an ambiguous result may very well therefore be the result of participant fatigue in a long, arduous task. It is interesting, however, that the response proportion did not
approach 0, which would be indicative of a greater tendency for the Test1 FA to overwrite the original representation. It is our interpretation, then, that the participants were indeed engaged in the task and not performing at random given the Test2 d' and a decreased confidence in Test2-FA responses. As such, a false alarm may or may not overwrite the original representation, according to processes occurring during the initial false alarm (Section 3.2.4). Somewhat relatedly, another minor limitation is the fact that participants performed better during the Test1 phase than Test2, even though it has been documented that a two-alternative forced-choice task is easier than a Yes/No (Jou et al., 2016). We interpret this fact as evidence of participant fatigue.

A second unexpected result was the lack of findings within the hippocampal subfields in all analyses save those of Test1 (Section 3.2.3). While we demonstrate the differential roles of left CA1 and CA3-CA3-DG in this task with respect to Hits and CRs, in which the hippocampus appeared to be biased towards Target detection in this task, the lack of subfield findings in other analyses of the task was surprising to us given the wealth of literature that has well established the roles of these regions in similar tasks (Bakker et al., 2008; J. Kim & Yassa, 2013; Kirwan & Stark, 2007; Lacy et al., 2011; Rolls & Kesner, 2016). Accordingly, we cannot reject the null of hypothesis (b) as we did not detect differential subregion processing during the FA that predicted subsequent performance. These null results are perhaps a type II or false negative error which resulted from extracting a single parameter estimate for the entire ROI, thereby collapsing across a significant amount of noise as well as signal, possibly washing out the effect. Indeed, it has been noted that the long-axis of the hippocampus contains a functional gradient and discrete domains (Strange et al., 2014), so collapsing across all variance along the anterior-posterior axis of the hippocampal subregions may be inappropriate. Similarly, our contrast of Test1 FA signal coded for subsequent Test2 performance may have simply failed to detect a difference between
those two conditions in terms of subregion BOLD signal, but this does not necessarily imply that
the hippocampus is not differentially processing the lures such that subsequent behavior results
in a Hit versus FA. Rather, the processes involved in such a computation could be rather similar
to one another and so a comparison of the mean subfield BOLD signal associated with each
process may again be simply too insensitive.

Third, we were surprised by the lack of significant differences for the fMRI data of the
Test2 phase. The behavioral data indicate that successful Target detection was possible, and that
participants were more likely to Hit in Test2 following a Test1 Hit, so performance was not at
chance. The lack of findings may not be unwarranted, however, as there are very few studies
reporting positive results from two-alternative, forced choice, recognition memory paradigms
using fMRI. One notable exception by Olsen and colleagues (2009) is somewhat similar to this
experiment in terms of participants and power. Olsen et al. (2009) demonstrated that the medial
temporal region was active while maintaining the memory of a set of faces during a delay period
prior to a two-alternative forced choice test, an area similar to where we expected activation to
occur. The lack of findings in the present experiment may be due to differences between the
experiments in fMRI processing methods. One stark difference is in the methods modelling both
the noise and auto-correlation function at the group-analysis step employed in this experiment, as
this improvement is more conservative than previous random-field based models in terms of
false discovery rates but was only recently developed (Cox, 2018, 2019; Eklund et al., 2016).
Furthermore, while significant differences were not detected within the fMRI data at the group
level, one should not assume that “[i]f the result is not statistically significant, then it proves that
no effect or difference exists” (Chen et al., 2017, p. 955). Indeed, as Chen et al. (2017) points
out, null results may very well be the result of sub-threshold signal and/or a lack of differential signal between our chosen contrasts at our resolution.

In conclusion, our data suggest that a false alarm does not necessarily overwrite the original memory trace of the target, but that subsequent detection of the target is possible, suggesting that a false alarm results in multiple memory representations in some instances. Further, the successful recovery of the target trace following a false alarm appears to be dependent not on the initial encoding trial or on retrieval processes following the false alarm, but on processes which occurred during the false alarm itself. Differential activation was detected in the left dorsal medial prefrontal cortex, left insular cortex, and right supramarginal gyrus during the false alarm that predicted whether or not the target would subsequently be retrieved, regions which are associated with episodic memory, saliency, and bottom-up attention, respectively.
Chapter 2 Discriminating Temporal Information

1 INTRODUCTION

Temporal information is critical for episodic memory. For example, the memory of a college lecture contains both when the lecture occurred as well as the length of the lecture, in addition to the visual, spatial, and auditory properties. Temporal information is unique in the sense that the temporal properties do not directly map to any receptor in the way that other sensory modalities do. Rather, time is a “hidden dimension” (Howard et al., 2014); the information is inferred through intermediate neural representations rather than actual interactions with the environment. Recent work suggests that it is precisely due to such intermediate representations that disparate information is able to be associated (Howard et al., 2014; Howard & Eichenbaum, 2013; MacDonald et al., 2011). For example, differing stimulus modalities in the example above (visual presentations, auditory explanations) arrive at the central nervous system with discrete but meaningful temporal associations which are then integrated into a neural representation of time, thereby allowing for the disparate information to be associated into a unique and unitized (Clewett & Davachi, 2017) memory representation.

The hippocampus receives multi-modal input from various cortical regions, including the visual stream as well as parietal, prefrontal, and temporal cortices (Aggleton, 2012; Baldassano et al., 2016; Cabeza et al., 2008; Eichenbaum et al., 2012; Mayes et al., 2007; Qin et al., 2016; Rockland & Van Hoesen, 1999; W. Suzuki & Naya, 2014; Thierry et al., 2000). As such, the hippocampus is ideally situated to form associations among disparate sensory modalities, stitching them together into a single representation that will subsequently be retrievable (Backus et al., 2016; DuBrow & Davachi, 2014; Howard & Eichenbaum, 2013; Thavabalasingam et al., 2018). Recent work has demonstrated that the hippocampus also generates a representation of
time: work by Pastalkova et al. (2008) demonstrated that the hippocampus contained a self-organized internal mechanism by which temporal information was maintained. Further, more recent modelling (Howard & Eichenbaum, 2013) has demonstrated that the hippocampus is capable of producing scale-invariant representations of time. It is through these representations that the recovery of contextual temporal information in episodic memory is possible, thereby allowing for the “backward jump in time” (Howard & Eichenbaum, 2013). MacDonald et al. (2011) has also shown that certain ensembles of hippocampal neurons, termed “time cells”, had peak activity at different and sequential periods, effectively serving to “bridge the delay period” and maintain information through a representation of time. Further subsequent work by Howard et al. (2014) argued that varying time scales may also be represented by differing neural ensembles coding for different time constants. Such time constants are then thought to be combined with unimodal input from various cortical regions, thereby forming “holistic representation that captures the relationships between stimuli separated by time and space” (Howard et al., 2014, p. 4703; Staresina & Davachi, 2009). As such, not only does the hippocampus serve to stitch together multiple information modalities in time, it also constructs the temporal foundation upon which such associations are formed.

Given the intimate relationship between time and information in episodic memory, it follows that the hippocampus would be sensitive to event duration at different time scales. In addition to the varying scales formed via neural ensembles (Howard et al., 2014), the hippocampus has been shown to be sensitive to time scales both on the order of seconds (Barnett et al., 2014; Thavabalasingam et al., 2018) as well as on the order of minutes (Montchal et al., 2019). Given that pattern separation processes have been detected within hippocampal circuitry (Gilbert et al., 1998, 2001; Gilbert & Kesner, 2006; Hunsaker & Kesner, 2013), that pattern
separation processes are information agnostic (Huffman & Stark, 2014; Yassa & Stark, 2011), and hippocampal sensitivity to temporal information, we would expect the hippocampus to be capable of discriminating between similar but different temporal properties within episodic memory. Many have studied temporal pattern separation or mnemonic discrimination (Gilbert et al., 2001; Hunsaker & Kesner, 2013; Kesner et al., 2002; Montchal et al., 2019), but these paradigms have largely tested the sensitivity of hippocampal processes to order of stimuli. While the order of events is important for episodic representations, event duration is also critical. Importantly, while order contains temporal components, order itself is not inherently temporal, given that “[o]rdinal information can be extracted from a temporal representation, but the converse is not true” (Howard & Eichenbaum, 2013, p. 1212). This idea was addressed by Montchal et al. (2019) when discussing the importance of considering the relationship of sequence and duration, concluding “[i]t is likely that both types of information are important for making temporal judgments” (p. 287). However, the literature on the ability of human subjects to discriminate duration differences at short (1-2 second) time scales is sparse. Further, although implicated by previous studies, there are no demonstrations that the human hippocampus performs pattern separation for event duration. We therefore designed a set of experiments to (1) determine whether participants can distinguish between highly similar temporal durations and (2) implicate the important neural regions associated with detecting changes in temporal information. In the first experiment we hypothesized that participants would be (a) sensitive to changes in temporal duration on the order of 0.5 seconds, (b) more sensitive when fewer stimuli separate the first presentation from the Target or Lure, and (c) more sensitive to increases than decreases in stimulus duration. In the second experiment, we hypothesized that (d) the medial temporal regions, and in particular the CA1, would be involved in successful Lure detection,
given that this region is largely implicated by extant research (Azab et al., 2014; Gilbert et al., 2001; Howard et al., 2014; Kitamura et al., 2014; Montchal et al., 2019; Schapiro et al., 2016; Thavabalasingam et al., 2018); we expected an increased BOLD response in the CA1 that corresponded with successful Lure detection relative to an error. Further, we expected that (e) a gradient would exist in which the anterior hippocampus (head) would show greater sensitivity to temporal information (DuBrow & Davachi, 2014; Montchal et al., 2019) than more posterior regions of the hippocampus.

2 METHODS – Behavioral Experiment 1

2.1 Participants

Forty healthy, young-adult participants were recruited from the university student population. This number was determined via a priori G*Power (Bruin, n.d.) calculations, where the test family = F tests, statistical test = ANOVA: repeated measures, within factors, Effect size = 0.25, Power = 0.95, Number of groups = 1, Number of measurements = 4, Correlation = 0.5, and Nonsphericity correction = 1; such input yielded a required sample size of 36. The inclusion criteria consisted of speaking English as a native language and having normal or corrected-to-normal vision. Exclusion criteria consisted of a history of psychologic disorder or brain injury, and color blindness. The university’s Institutional Review Board approved the research, and all participants gave written informed consent prior to participation. Participants were compensated for their time via course credit.

2.2 Temporal Discrimination Task

Task description. Each block followed a continuous-recognition paradigm in which a continuous stream of trials was presented (Figure 2-1). Each trial consisted of a fixation cross (a black cross on a white background) for 0.5s, the stimulus (a colored image of an everyday
object) for a variable duration (see below), a colored masking image to combat iconic memory presented for 0.5s, and a response screen displayed for 2.5s. If the stimulus was novel (a Foil or the first presentation of either the Target or Lure), one of three potential response screens was presented at random, prompting the participant to indicate a judgment about a visual property of the object via a button response. These response options were a) whether the object was Smooth, Rough, or Sharp, b) whether the object was made of Metal, Plastic, or Other, and c) whether the item’s weight was Heavy, Medium, or Light. Each response option was presented an equal number of times. The aim of these response options was to maintain engagement in the task as the perception of time is related to attention (Coull et al., 2004; Radua et al., 2014). If the stimulus had been presented before, (i.e., a Target or Lure), the response screen prompted the participant to decide whether the stimulus was presented for either a Longer, Same Time, or Shorter duration than the first presentation of the same stimulus. Targets were trials where the same stimulus was presented with the same duration as the initial presentation. Lures were trials where the same stimulus was presented but for a different duration: 1.5s if the first counterpart was presented for 1s, and vice versa if the first was presented for 1.5s. In total, the earliest a second presentation could occur, assuming Lag = 4 (see below) and all intermediate durations = 1s, was 17.2 seconds, and the latest was 57.6 seconds (Lag = 12, Duration = 1.5s). Note that while it is possible that the second presentation could occur within a timeframe governed by working memory, it was expected that the cognitive demand elicited by the first-presentation response prompts would overwhelm the processes such that second-presentation responses are reflective of episodic memory processes. Prior to the first block, the participant was instructed in the task via a video with explicit instructions as well as a practice period, and no feedback on performance was given during the task. Stimuli were presented in PsychoPy (v2.0, n.d.).
Figure 2-1. Temporal discrimination paradigm. A trial started with a fixation cross (0.5s, not shown) and then a stimulus was presented for either 1 or 1.5s. Following the stimulus, a colored mask was presented for 0.3s, and then a response screen (RS) was presented for 2.5s. Four RS were possible: for first presentations one of three memory prompts for visual properties were presented, and for second presentations a memory prompt for stimulus duration was presented. Targets and Lures (second presentations) followed their first presentation counterpart by either 4 or 12 trials. Targets had the same duration as the first presentations, whereas Lures were presented for 1.5s if the first counterpart was presented for 1s, and vice versa if the first was presented for 1.5s.

Stimuli, duration, and lag. Stimuli for the task consisted of randomly sampling 324 images from the 570 stimuli used in the Mnemonic Similarity Task (Stark et al., 2019) stimuli sets C-E, type “a”, for each participant. The 324 stimuli were then assigned as either Target (120), Lure (120), or Foil (84) trials. Target and Foil trials were presented for either 1 or 1.5 seconds (if the first presentation of the Target had a duration of 1 second, then the second presentation also had a duration of 1 second). Lures were first presented for 1 or 1.5 seconds and then the duration of the second presentation was adjusted by 0.5s, where if the first presentation had a duration of 1s then the second presentation had a duration of 1.5s, and vice versa; in this way all stimuli (Target, Lure, Foil) were presented for either 1 or 1.5s but Lure trials either
increased or decreased by 0.5s. Targets and Lures repeated after either 4 or 12 stimuli (Lag = 4, 12). With 324 stimuli, 240 of which repeated once, the total number of trials was 564. Duration and Lag were counterbalanced such that (a) an equal number of stimuli were presented for both durations of 1 and 1.5s (282 each), (b) an equal number of Targets were presented for both durations (60 1s, 60 1.5s), (c) an equal number of Lures increased and decreased in stimulus duration (60 increase, 60 decrease), (d) an equal number of Foils were presented for both durations (42 each), and (e) an equal number of Targets and Lures repeated at Lag = 4 or 12 (60 Lag = 4, 60 Lag = 12). Further, each Lag (e.g., 4) had an equal number of trial durations for Targets (30 1s, 30 1.5s) and Lures (30 increase, 30 decrease). The stimuli were presented in 6 blocks of 94 trials, where each block was counterbalanced for stimulus type (Target, Lure, Foil), duration, and lag. Stimuli were presented in a pseudorandom order with the constraints of Lag and number of trials. Each block contained 20 Targets (5 for each combination of Lag and Duration), 20 Lures (5 for each combination of Lag and In/Decrease), and 14 Foils (7 of each Duration).

2.3 Statistical Analyses

As there were two possibilities to incorrectly identify Targets and Lures in this experiment, an adjusted signal-detection naming convention was adopted (Table 2-1). A correct identification of Targets was termed “Hit-Same” (Hitsa) and incorrect identification was called either “Miss-Long” (MissL) or “Miss-Short” (MissS) according to whether the participant indicated if the Target duration increased or decreased with respect to the first presentation. For Lures which decreased in duration (first presentation = 1.5s, second = 1s), correctly identifying the stimulus was termed “Correct Rejection-Short” (CRS), and incorrectly identifying the stimulus was termed either “False Alarm-Long” or “False Alarm-Same” (FA1, FASa), according
to whether the participant indicated that the Lure was presented for a longer or same duration as
the first presentation, respectively. For Lures which increase in duration (first = 1s, second =
1.5s), correct identification was called “Correct Rejection-Long” (CR_L). Incorrect identification
of the increasing Lure was termed either FA_Sa or “False Alarm-Short” (FA_S). Note that while
some response types received the same classification, such as “Hit_Sa” resulting from correctly
identifying a Target of either 1 or 1.5s, individual classifications were maintained separate
during analyses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Longer</td>
</tr>
<tr>
<td>Target (1s)</td>
<td>Miss_L</td>
</tr>
<tr>
<td>Target (1.5s)</td>
<td>Miss_L</td>
</tr>
<tr>
<td>Lure, Decrease (1s)</td>
<td>FA_L</td>
</tr>
<tr>
<td>Lure, Increase (1.5s)</td>
<td>CR_L</td>
</tr>
</tbody>
</table>

Response behavior classifications. Stimuli were presented for either 1 or 1.5s. Second
presentations were either Targets or Lures, where a Lure stimulus duration differed from the first
presentation by 0.5s. “Lure, decrease (1s)” corresponds to “Target (1.5s)”, and “Lure, increase
(1.5s)” corresponds to “Target (1s)”. Participants responded whether the second presentation was
presented for Longer (L), Same Time (Sa), or Shorter (S) than the first presentation.

A separate d’ calculation was conducted for the 1-second and 1.5 second Target durations
in order to identify the response criteria, according to the adjusted formulas

$$d'_{1.5s} = z(p_{Hit}) - z(p_{FA})$$  \hspace{1cm} \text{Equation 3}$$

where

$$p_{Hit} = \frac{(Hit_{Sa} + Miss_L)}{\# \ 1.5s \ Targets}$$ \hspace{1cm} \text{and} \hspace{1cm}$$p_{FA} = \frac{(FA_L + FA_{Sa})}{\# \ 1s \ Lures}$$

and

$$d'_{1s} = z(p_{Hit}) - z(p_{FA})$$  \hspace{1cm} \text{Equation 4}$$
where
\[ p_{Hit} = \frac{(Hit_{sa} + Miss_s)}{\text{# 1s Targets}} \quad \text{and} \quad p_{FA} = \frac{(FA_s + FA_{sa})}{\text{# 1.5s Lures}} \]

Note that while Hit_{sa} and FA_{sa} appear in both formulas, these response bins are kept separate in response to their respective stimulus duration.

D-prime scores were then used as the dependent variable to investigate whether or not a Lag, Duration, or interaction effect exists using a two-factor within-subject ANOVA utilizing a Greenhouse-Geisser correction for any violations of sphericity, and d' scores were individually tested against zero using the Student’s t to determine if participant behavior differed from chance. A d' score significantly different from 0 would indicate that participants are sensitive to changes of stimulus duration on the order of 0.5 seconds (as in hypothesis [a]). A test of whether the d' for stimuli of Lag=4 that is significantly greater than the d' associated with Lag=12 tested hypothesis (b), and finally, a test of whether d' scores are greater for decreases in duration relative to increases tested hypothesis (c). The FDR method for correcting for multiple comparisons was used, when applicable.

3 RESULTS – Behavioral Experiment 1

Four participants were excluded due to computer malfunction, leaving a total n=36. Outliers were detected according to the 1.5 × IQR method; replacing outlier values with the median value had no effect on the significance of test statistics and so all statistics reported reflect the datasets where outliers were included. A two-factor within-subject ANOVA failed to detect any effect of Duration, Lag, or interaction thereof on d' scores (all p-values > 0.05), while one sample t-testing revealed that d'_{1.5s} and d'_{1s} scores differed significantly from zero (t(34) = 14.07, p_{FDR} < .001, 95% CI [1.13, 1.52]; t(34) = 13.83, p_{FDR} < .001, 95% CI [1.21, 1.63], respectively). Further, these d' scores did not differ from each other according to paired t-testing.
(pFDR > .05; Figure 2-2, Experiment 1). Together, the evidence suggests that participant performance differs significantly from chance and that participants are equally sensitive regardless of whether the Lure increases, decreases, or appears after Lag 4 or 12.

**Figure 2-2. Boxplot of d' scores.** 1.5s = 1.5 second Target, 1 second Lure; 1s = 1 second Target, 1.5 second Lure. *** = p < .001, * = p < .05.

4 METHODS – Behavioral Experiment 2

As no effect of duration, lag, or interaction thereof was detected in the previous analyses, and given that the experiment lasted for approximately 45 minutes, we elected to reduce the experiment length: the number of trials in each category was reduced by exactly one-half, yielding 60 Targets, 60 Lures, and 42 Foils. The resulting experiment was still counterbalanced
for duration and lag as well as Target and Lure types but no longer powered to investigate interactive effects of duration and lag, and had a duration of approximately 25 minutes. In order to make sure that such a reduction would not affect the outcome, 34 more participants (16 Female, Age = 20.1±1.7) were recruited to participate in the shortened version of the experiment. Participants gave written informed consent prior to participation. All methodologies and analyses conducted, save the reduction of trial items, remain identical to those reported above.

5 RESULTS – Behavioral Experiment 2

In the reduced experiment replacing outlier values with the group median again had no effect on test statistics and so all statistics reported include outliers. One-sample t-testing on the d'1.5s and d'1s scores demonstrated that participants performed above chance (t(33) = 14.85, pFDR < .001, 95% CI [1.08, 1.43]; t(33) = 16.44, pFDR < .001, 95% CI [1.27, 1.64], respectively), but this time we detected that participants were more sensitive in d'1s than d'1.5s according to paired t-testing (t(33) = -2.22, pFDR < .05, 95% CI [-0.37, -0.02]; Figure 2-2, Experiment 2).

6 METHODS – fMRI Experiment

6.1 Participants

Forty healthy, young-adult participants were recruited from the Brigham Young University student population in order to replicate the procedures of section 4 (the shortened experiment) and investigate any potential neural correlates of discrimination for temporal duration. All participants were naïve to the task. In addition to the inclusion and exclusion criteria defined above, exclusion criteria also consisted of incompatibilities with the magnet or head coil, colorblindness and left-handedness, and the capability to remain still for an hour. The university’s Institutional Review Board approved the research, and all participants gave written
informed consent prior to participation. Participants were compensated for their time via course credit, money ($20 USD), or a 3d-printed quarter-scale model of their brain.

6.2 Temporal Discrimination Task and Statistics

The task was identical to that in section 4. Correspondingly, analyses of behavioral data were conducted in an identical fashion to section 4.

6.3 MRI Acquisition

Functional and structural data were collected on a Siemens Tim Trio 3T MRI scanner at the Brigham Young University MRI Research Facility, using a 32-channel head coil. T1-weighted, T2-weighted, and T2*-weighted scans were collected from each participant. The T1-weighted structural scan was acquired using a magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence, according to the following parameters: TR = 1900 ms, TE = 2.26 ms, flip angle = 9°, FoV = 218 × 250, voxel size = 0.97 × 0.97 × 1 mm, slices = 176 interleaved. A high-resolution T2-weighted scan was used to image hippocampal subfields according to the following parameters: TR = 8020 ms, TE = 80 ms, flip angle = 150°, FoV = 150 × 150 mm, voxel size = 0.4 × 0.4 × 2 mm, slices = 30 interleaved; the T2 data were acquired perpendicular to the long axis of the hippocampus. High-resolution multi-band echo-planar scans were collected from each participant according to the following parameters: Multi-band factor = 4, TR = 2200 ms, TE = 45.2 ms, flip angle = 90°, acquisition matrix = 100 × 100 mm, voxel size = 1.8 mm³, slices = 72 interleaved; the first 11 TRs were discarded to allow for T1 equilibration. Finally, a reverse blip echo-planar scan was collected using the same parameters as the previous protocol save that direction of acquisition was opposite (P>>A) and only 10 volumes were collected.
6.4 MRI Data Pre-Processing

All pre-processing was conducted using the software packages dcm2niix (Li et al., 2016), Analysis of Functional NeuroImages (AFNI; Cox, 1996), convert 3D (c3d; Yushkevich et al., 2006) and Automatic Segmentation of Hippocampal Subfields (ASHS; Yushkevich et al., 2015). AFNI and Mango (Lancaster et al., 2012) were used for visualizations.

T2*-weighted DICOMs were converted into NIfTI files. Field distortion was then corrected by incorporating information from the reverse blip scan. Briefly, the median input of all voxels from concatenated blip and experiment functional scans were individually calculated and then the nonlinear transformation between the blip and experimental files were calculated thereby producing warping and reverse warping vectors. These reverse warping files were then used to “unwarp” the experimental files thereby recovering some signal in regions of high signal distortion. The experimental volume with the minimum of voxel outliers was then extracted to use as a volume registration base. To move data into template (MNI 152) space, the calculations of (a) rigidly aligning the structural to the registration base (using an lpc+ZZ cost function), the (b) non-linear warp of the structural file in original space to template space, and (c) the volume registration (rigid transformation) of each volume to the registration base were concatenated and applied to the functional data in native space such that all volumes were aligned with the template utilizing a diffeomorphic transformation via a single interpolation. A binary extension mask corresponding to regions within the domain with sufficient minimum signal was constructed; data within this region were scaled by the mean signal.

Single-subject regression used a generalized-least-squares fit utilizing a nonlinear residual maximum likelihood estimate to determine the best fitting autoregressive-moving-average model, where the timeseries corresponding to an eroded white matter mask was utilized.
as a nuisance regressor. Centered motion regressors for six degrees of freedom were included, and volumes that were (a) outliers in terms of total volume signal (>10% of the voxels in the volume had outlier signal) and (b) either preceded or were involved in a motion event (>0.3° rotation) were censored. Regressors for the various behaviors (described below) were modelled with a duration modulated block function. Two single-subject regressions were conducted, where the BOLD signal associated with behavioral responses was modelled using a boxcar convolution with the canonical hemodynamic response. First, test trials (when participants made their response) were modeled for correct and incorrect responses to Targets (Hits and Misses, respectively) and Lures (Correct Rejections and False Alarms, respectively); only the time between the start of the response screen and participant response was modelled. Second, responses to encoding the stimuli (when the stimulus was onscreen) were modeled for subsequent behavioral responses; the entire stimulus duration was modelled. Finally, participants that had more than 10% of their total volumes censored would have been excluded from all subsequent group-level analyses but no participant reach this criterion.

For a priori analyses, behavioral parameter estimates (β-coefficients) were extracted from anterior and posterior hippocampal regions of interests (ROIs) from each deconvolution utilizing masks that were previously generated in template space. Briefly, ASHS was used to segment the medial temporal regions of high-resolution T2-weighted scans for a number of participants. Joint Label Fusion (H. Wang et al., 2013) then utilized this set of segmentations to generate template hippocampal subfield priors. For this experiment, template MTL masks were split along the anterior-posterior axis at X=-20, the CA2, CA3, and DG masks were combined, and then masks were resampled into functional dimensions. Any voxels labeled by more than one mask were excluded, thereby producing anterior and posterior, right and left Subiculum,
CA1, and CA2-3-DG masks. A two-factor within-subject MANOVA was then used to investigate the role of hippocampal structures in the modeled behaviors, where factor A was the ROI, factor B the behavioral responses, and the parameter estimate was the dependent variable. ROIs in the left hemisphere were treated as independent from the right, and a separate analysis was conducted on the anterior and posterior masks for both the encoding and response estimates. The FDR method was used to correct for multiple comparisons.

For exploratory analyses, an intersection mask was constructed, which indicated where meaningful signal exists for both T1- and T2*-weighted scans across all participants, and multiplied with a gray matter mask that had been generated in template space, thereby producing an intersection-gray-matter mask that was used to constrain the number of voxels on which analyses were conducted. Further, a small volume mask was generated in a similar fashion but only using masks associated with the gray matter of bilateral hippocampi, entorhinal cortex, perirhinal cortex, and parahippocampal cortex. Noise was modelled by calculating the autocorrelation function of single-subject regression residuals from voxels which lay within the generated masks, and these parameter estimates were then used in Monte Carlo simulations to generate multiple-comparison correction criteria. A multivariate model was used to investigate any regions that show differential signal associated with the various behaviors modelled where the behaviors were modeled as within-subject factors. Clusters which showed differential activity were then used to extract mean parameter estimates in order to report the (a) difference in parameter estimates and (b) statistics. All scripts used for running these experiments and analyzing the data are available at https://github.com/nmuncy/MST_Temporal.
7 RESULTS – fMRI Experiment

7.1 Behavioral Results

From the 40 participants recruited, five were excluded from analyses due to equipment problems and incompatibilities (n=3), responding with the wrong button presses (n=1), and “excessive eye watering” during the task which was reported to be “very distracting” and resulted in at-chance performance (n=1). The group size reported below is therefore n=35, 19 female, age=22.9±2.7.

As above, statistics were conducted both with and without outliers. Replacing outlier values with median values had no effect on test statistics and so all analyses below include outliers. Performance was very similar to the above experiments (Figure 2-2, Experiment 3). Participants d' scores differed significantly from 0 according to one-sample t-testing for both increases and decreases in duration (t(34)=13.9, pFDR<.001, 95% CI [1.35, 1.81]; t(34)=16.4, pFDR<.001, 95% CI [1.32, 1.69], respectively). In addition, and similarly to the results of experiment 1 but not experiment 2, no difference was detected between sensitivity to changes in increases versus decreases according to two-sample t-testing (t(34)=0.65, pFDR=.52, 95% CI [-.16, .31]).

7.2 FMRI Results - Encoding

As no effect of Lag, Duration, or an interaction thereof was detected in behavioral performance (Section 3) and participants performed equally well regardless of increasing or decreasing the Lure duration (Section 7.1), responses were collapsed across Lag and Duration, and further, types of Misses, Correct Rejections, and False Alarms (Table 2-1) were collapsed as well. Collapsing resulted in 28 Hits, 31 Misses, 31 Correct Rejections, and 23 False Alarms, on average.
The effect of encoding on task performance was assessed by investigating the BOLD response to stimulus presentation that preceded a task response (Hit, CR, etc.). For the a priori ROIs, no effect was determined via a two-factor within-subject MANOVA for either ROI, Behavior, or an interaction thereof (all p-values > .05 once correcting for sphericity via the Greenhouse-Geisser method) for both the anterior and posterior regions of the hippocampus.

An exploratory whole-brain analysis also failed to detect any clusters which showed differential signal during trials preceding task responses (Trials preceding Hit, FA, CR, Miss) when using the thresholding criteria of cluster size $k = 20$, $NN = 1$, two-sided comparison, $p < .001$, $FWE = .05$. A small-volume correction, using masks corresponding to the MTL (hippocampus, entorhinal, and parahippocampal cortex) did, however, reveal two clusters with differential activity using the thresholding criteria of $k = 6$, $NN = 1$, two-sided comparison, $p < .001$, $FWE = .05$. Neuroanatomic classification of the clusters for these and subsequent analyses were determined by the criteria established by Frankó et al. (2014). The left entorhinal cortex was more active during encoding trials preceding a Hit versus a Miss (LEC, Figure 2-3 Left, Table 2-2) while the right hippocampus was more active during encoding trials preceding a FA than a CR (RHC, Figure 2-3 Middle, Table 2-2).
Figure 2-3. Regions engaged in duration detection. Left and Middle: left entorhinal cortex (LEC) and right hippocampus (RHC) are differentially engaged during encoding trials preceding task performance. Right: Left perirhinal cortex (LPRC) is differentially involved in Target detection. “Tp” = trial preceding a test response.

Table 2-2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Cluster</th>
<th>K</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>T(34)</th>
<th>Sig</th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoding</td>
<td>LEC</td>
<td>9</td>
<td>-16</td>
<td>-13</td>
<td>-25</td>
<td>4.13</td>
<td>***</td>
<td>.21</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td>RHC</td>
<td>7</td>
<td>27</td>
<td>-25</td>
<td>-9</td>
<td>-4.55</td>
<td>***</td>
<td>-.81</td>
<td>-.31</td>
</tr>
<tr>
<td>Retrieval</td>
<td>LPRC</td>
<td>7</td>
<td>-18</td>
<td>-23</td>
<td>-25</td>
<td>-3.15</td>
<td>**</td>
<td>-.61</td>
<td>-.13</td>
</tr>
</tbody>
</table>

Cluster descriptions of small-volume corrected analyses. K = cluster size, X-Z = peak MNI coordinate. T(34) = paired t-test with 34 df, LB UB = lower and upper bound of 95% confidence interval. LEC = left entorhinal cortex, RHC = right hippocampus, LPRC = left perirhinal cortex. *** = p < .001, ** = p < .01.
7.3 FMRI Results – Retrieval

To investigate processes associated with task performance the HRF was modeled for the time in which participants responded to Target and Lure stimuli. *A priori* investigations into the subregions of the medial temporal lobe failed to detect a significant interaction between region and behavior in either portion of the hippocampus. Correspondingly, an exploratory whole-brain analysis using the same thresholding criteria as was reported above failed to detect any clusters that demonstrated differential activity for the various task behaviors. A small-volume correction, however, detected a cluster in the left perirhinal cortex that was differentially engaged in correct responses (LPRC, Figure 2-3 Right, Table 2-2). That is, this cluster was more active during the correct identification of Targets (Hits) than during the correct identification of Lures (CRs).

8 DISCUSSION

This set of experiments investigated whether mnemonic discrimination was possible for stimulus duration and identified functional neuroanatomic correlates thereof. The task (Figure 2-1) involved serially presenting images of everyday items in a continuous-recognition paradigm for either 1 or 1.5 seconds. Following the first presentation of a stimulus the participants performed an orienting task wherein they made decisions from memory about physical properties of the stimulus. If a stimulus repeated, it was either presented for the same duration (Targets) or for a difference of 0.5 seconds (Lures); in these cases, the participants were explicitly instructed to attend the presentation duration and subsequently tasked with deciding whether a change of stimulus duration occurred. Such a small change in stimulus duration was intentionally selected in order to elicit a sufficient number of Misses and FAs for functional BOLD analyses. The paradigm was conducted three times in naïve groups of participants, experiments 2 and 3 serving
to replicate the behavioral findings of experiment 1 and bringing the total number of participants that performed the task to n=69.

These experiments had several findings. First, behavioral evidence indicated that participants were successful at remembering and detecting changes of stimulus duration on the order of 0.5 seconds and that they were equally sensitive regardless of the direction of duration change (Section 3). With respect to these behavioral findings, we reject the null of hypothesis (a), given that participants had d' scores which differed significantly from zero, but fail to reject the null of hypotheses (b) and (c); participants were equally sensitive at Lag 4 as at Lag 12 and were equally sensitive to increases and decreases in duration. These subsecond findings replicate those of Palombo et al. (2019), in which amnestic patients were tasked detecting whether test spinwheels spun for a different duration than study pinwheels. Palombo et al. (2019) noted, as we do here, that participants had strong d' scores and were sensitive to both increases and decreases in duration. Further, in the present experiment, no effect of lag was detected; equal performance at either lag suggests that the distracting task was sufficiently difficult to overwhelm working memory such that behavioral results here are reflective of episodic memory processes. Were this not the case, we would expect to see an interaction between d' scores and lag due to Lag 4 performance being governed more by short-term memory and Lag 12 by episodic. Corresponding with this line of reasoning, the Palombo et al. (2019) experiment, who tested participants within the domain of short-term memory, presented d' scores that where two standard deviations higher than those presented here. Finally, the behavioral evidence suggests that the detected effect is robust as the d' scores were stable across three different groups of participants.
Second, two clusters showed differential activation during the encoding of the stimuli which was predictive of subsequent test performance. A cluster in the right hippocampus (Figure 2-3, middle) was differentially engaged in encoding such that the cluster was more active preceding the incorrect detection of Lures (FA) compared to correct detection (CR). This pattern of activity is consistent with the deleterious effect that the default mode network has on subsequent test performance; we interpret this cluster as being indicative of participants failing to attend the training stimulus such that subsequent performance was adversely affected (H. Kim, 2011). It should be noted, though, that a second a posteriori whole-brain analysis did not identify any additional clusters to those detected in the small-volume correction. As such, the lack of additional default-mode findings weakens the interpretation of this hippocampal cluster.

Next, the left medial entorhinal cortex (mEC) demonstrated differential activity which predicted subsequent Target-trial performance, in that the cluster was more active preceding a Hit relative to preceding a Miss (Figure 2-3, left). This cluster appears to be critically involved in the encoding the temporal properties of the stimuli, as visual information remained identical between study and test, a finding that is consistent with the previous literature: numerous studies have detected cells with a selectivity for temporal information within the mEC. For instance, Kraus et al. (2015) demonstrated mEC grid cell activity which signaled for elapsed time in addition to distance traveled, showing an integration of time and distance, and Heys and Dombeck (2018) demonstrated both that while the mEC is sensitive to elapsed time, it can also remap its sensitivity according to contextual task demands (also, see Sugar & Moser, 2019).

Robinson et al. (2017) showed that the mEC helps support the hippocampal CA1 representation of time, and further, Suh et al. (2011) demonstrated that the monosynaptic side of the perforant pathway, the connection between the mEC and CA1, is integral for temporal associations. This
hippocampal structure, the CA1, is downstream from the mEC and has been strongly implicated in the temporal aspects of memory (Allen et al., 2016; Cai et al., 2016; Dimsdale-Zucker et al., 2018; MacDonald et al., 2011; Mankin et al., 2012; Salz et al., 2016; Suh et al., 2011; Y. Wang et al., 2015). Finally, inactivation of the mEC is also associated with disruptions in temporal processing (Heys & Dombeck, 2018; Robinson et al., 2017; Sugar & Moser, 2019). As such, the mEC is proposed to bind information of different modalities (i.e. item, spatial, temporal) as well as providing a scaffold of spatiotemporal information to the hippocampus (Sugar & Moser, 2019).

Further, Maass et al. (2015) and Schröder et al. (2015) demonstrate that the mediolateral divisions of the rodent entorhinal cortex do not translate to the human entorhinal cortex, but that the human entorhinal cortex is functionally divided along an anterolateral-posteromedial axis such that the human homologue of the rodent mEC is the posteromedial entorhinal cortex (pmEC). Maass et al. (2015) shows the PHC is functionally coupled with the pmEC, implicating this region in processing contextual information as opposed to item information which tends to be sent from the PRC to the lateral anterior EC (Aminoff et al., 2007; Diana et al., 2013; Hsieh et al., 2014; Jenkins & Ranganath, 2010; Naya et al., 2017; F. Wang & Diana, 2017). Critically, the cluster detected in this experiment during the encoding phase of the experiment entirely corresponds with the pmEC. Further still, Schröder et al. (2015) demonstrated that pmEC forms part of the posterior-medial system, a system which carries spatial information. The rodent and human literature in conjunction therefore implicate the pmEC in processing not just spatial but contextual information which varies in a dimension (i.e. temporal information) according to task demands, a notion that is reminiscent of Salz et al. (2016), who noted that temporal processing in
the hippocampus mirrors that of spatial. In the current experiment, then, the detected pmEC cluster is believed to be processing the temporal context information associated with each item.

Third, a cluster within the left perirhinal cortex (PRC) was differentially engaged during the test phase such that it was more active during Hits than Correct Rejections (Figure 2-3, right). Similarly, recent work implicated the PRC in processing temporal information during the test phase of an experiment: Montchal et al. (2019) had participants first view a television episode and then tested participants by having them place still-frames from the episode on a timeline, a task designed to test memory for temporal order. While the paper emphasized the roles of the anterior-lateral EC and subregions of the HC, Montchal et al. (2019) nevertheless demonstrated that the PRC was differentially engaged in high versus low accuracy for temporal information. In a separate but related vein, the Binding of Item and Context theory (Diana et al., 2013) has proposed a functional dissociation in the medial temporal lobe where the parahippocampus processes contextual information, the PRC item information, and the hippocampus serving to bind together item and context; the role of the PRC in such a model has recently been supported (Diana et al., 2013; Hsieh et al., 2014; F. Wang & Diana, 2017). Together then, the PRC has been implicated in processing both temporal and item information. Such a conjunction of item and time is supported by F. Wang et al. (2017), who noted PRC activation associated with the retrieval of contextual temporal information, perhaps due to the ‘unitization’ of the item and context (Ford et al., 2010; Giovanello et al., 2006). More explicitly, work using single-unit recordings in non-human primates demonstrated that PRC cells fired for both item and temporal information (Naya et al., 2017). Thus, the PRC activation detected in this experiment is consistent with the conjunction of item-temporal information processing, in which the PRC differentially aides in the successful retrieval of item and temporal duration information. Indeed,
increased signal from this region is associated with Hits, i.e., correctly retrieving the duration information (as the visual information was identical), while suppressed signal is associated with a mismatch detection between the Lure duration and that of the original.

A number of limitations exist for this experiment. First, while only Targets (same duration) and Lures (different duration) were tested in this experiment, participants had three response options necessitating an adjustment to the d' calculations. This adjustment resulted in collapsing across some response types such that these adjusted-d' scores do not represent a bias-corrected deflection point but rather indicate sensitivity to whether a change occurred. Indeed, in calculating d'_{1.5s} (when the Target was 1.5s and the Lure was 1s), \( p_{Hit} \) combined both responses that did not involve a “shorter” judgment for Targets, and \( p_{FA} \) did the same. With such a calculation, a higher adjusted-d' score therefore indicates that a participant is less likely to respond “shorter” on longer Target trials and more likely to respond the same on Lure trials, a response pattern that is related to, but separate from, accuracy in judging stimulus duration in episodic memory. Second, and related, several response categories were collapsed across in the functional analyses: as no main effect of Lag or Duration was detected, the same responses from different stimulus types (such as Hit_{Sa} for both a 1.5s and 1s Target for both Lags) were combined into the same bin. This method ensured a proper number of trials per behavioral bin existed such that the shape of the hemodynamic response function could reasonably be modeled, but it nevertheless may have contributed noise to the analyses. Indeed, if certain regions of tissue were involved in only detecting increases, decreases, or no change in stimulus duration then such a combination would likely have introduced unnecessary noise to any existing signal. As such, our functional findings do not illustrate signal associated with correctly identifying increases or decreases of duration in Targets and Lures, but rather demonstrate signal associated with
detecting temporal change between the study and test stimuli. Third, while the hippocampal subregion CA1 has been shown to be critically involved in temporal processing for memory tasks (Allen et al., 2016; Cai et al., 2016; Dimsdale-Zucker et al., 2018; MacDonald et al., 2011; Mankin et al., 2012; Montchal et al., 2019; Robinson et al., 2017; Salz et al., 2016; Y. Wang et al., 2015), our results did not replicate these findings. This is likely due to the fact that the rodent studies which have implicated the CA1 have largely involved requiring the rodent to maintain information over time (e.g., Kraus et al., 2015; Pastalkova et al., 2008), information which is not necessarily temporal and is commonly a previous decision. Our experiment instead involved the retrieval of temporal duration information, and likewise it is unsurprising that we detected signal associated with tissue in a different region. With respect to previous studies involving humans (Aminoff et al., 2007; Clewett & Davachi, 2017; Diana et al., 2013; El-Kalliny et al., 2019; Hsieh et al., 2014; Jenkins & Ranganath, 2010, 2016; Montchal et al., 2019; F. Wang & Diana, 2017), our results are similar in nature and any differences are likely due to the fact the we here tested for episodic memory of temporal duration rather than temporal order or sequence, or working memory for temporal duration (Palombo et al., 2019). As such, we partially reject hypothesis (d), as we did find differential activation in the MTL associated with the task, but not in the hypothesized region of CA1. Further, we reject hypothesis (e), as no differential signal was detected in the anterior versus posterior HC. It is possible that we failed to detect differential signal in these regions due to collapsing across too much variance, or that the selected contrasts did not reflect differences in the signal, or perhaps even due to the explicit nature of our task. Either way, one must take caution in interpreting null results (Chen et al., 2017) and note that we were unable to falsify these hypotheses.
In summary, we presented participants with stimuli of different duration and tasked them with detecting a change in duration. We found that participants were sensitive on the order of 0.5 seconds, a finding which was demonstrated in three different groups of participants. We also demonstrate that the left posteromedial entorhinal cortex is differentially involved during the encoding of items such that greater activity in this region is predictive of subsequent Hit versus subsequent Miss. Finally, the left perirhinal cortex was noted to be differentially engaged during successful test performance, being more active during Hits than Correct Rejections, activation which is consistent with previous work demonstrating that this region is sensitive to the conjugation of item and temporal context.
References


https://doi.org/10.1126/science.1152882


https://doi.org/10.1016/j.neuroimage.2016.09.066


https://doi.org/10.1002/hipo.22321


https://doi.org/10.1016/j.actpsy.2016.03.014


https://doi.org/10.1002/hbm.22627