

## Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect



Erik A. Wing<sup>a,b,\*</sup>, Elizabeth J. Marsh<sup>a</sup>, Roberto Cabeza<sup>a,b</sup>

<sup>a</sup> Department of Psychology & Neuroscience, Duke University, Box 90086, 417 Chapel Drive, Duke University, Durham, NC 27708, United States

<sup>b</sup> Center for Cognitive Neuroscience, Duke University, Box 90999, B203 Levine Science Research Center, Durham, NC 27708, United States

### ARTICLE INFO

Available online 19 April 2013

#### Keywords:

Testing effect  
Retrieval practice  
Episodic memory  
fMRI  
Hippocampus  
Medial PFC

### ABSTRACT

Restudying material is a common method for learning new information, but not necessarily an effective one. Research on the *testing effect* shows that practice involving retrieval from memory can facilitate later memory in contrast to passive restudy. Despite extensive behavioral work, the brain processes that make retrieval an effective learning strategy remain unclear. In the present experiment, we explored how initially retrieving items affected memory a day later as compared to a condition involving traditional restudy. In contrast to restudy, initial testing that contributed to future memory success was associated with engagement of several regions including the anterior hippocampus, lateral temporal cortices, and medial prefrontal cortex (PFC). Additionally, testing enhanced hippocampal connectivity with ventrolateral PFC and midline regions. These findings indicate that the testing effect may be contingent on processes that are typically thought to support memory success at encoding (e.g. relational binding, selection and elaboration of semantically-related information) in addition to those more often associated with retrieval (e.g. memory search).

© 2013 Elsevier Ltd. All rights reserved.

### 1. Introduction

In the learning literature, there is abundant evidence that practice involving retrieval is more effective than practice involving passive review (Karpicke & Roediger, 2008; Roediger & Butler, 2011). In a typical memory experiment demonstrating this effect, to-be-learned items are first studied under uniform encoding conditions (*Study trials*) and are then practiced either through additional study (*Restudy trials*) or through retrieval from memory (*Test trials*). The reliable finding of this paradigm is that when memory is later assessed on a final memory test, items practiced in Test trials are remembered better than those practiced in Restudy trials (Roediger & Karpicke, 2006; Toppino & Cohen, 2009). The mnemonic advantage of the Test condition over the Restudy condition—known as the *testing effect*—illustrates the powerful role that retrieval can play during the course of learning. Research on the testing effect has demonstrated this difference across a wide range of stimuli from foreign-language vocabulary (Carrier & Pashler, 1992) to visuospatial information (Carpenter & Kelly, 2012; Carpenter & Pashler, 2007), and is increasingly directed towards improving learning in real-world settings (Larsen, Butler, & Roediger, 2009; McDaniel, Roediger, & McDermott, 2007).

Behavioral evidence suggests that the testing effect is mediated by the enhancement of cognitive processes typically associated with encoding and/or retrieval. As an example of a process associated with encoding, there is evidence that the testing effect is mediated by semantic elaboration. For instance, in word pair learning, the testing effect is larger for weakly rather than strongly related word pairs, as would be expected if pairs with lower intrinsic relatedness benefited to a greater extent by semantic binding during testing (Carpenter, 2009). Consistent with this idea, testing enhances subsequent memory not only for cue and target words, but also for semantic mediators linking cues and targets (Carpenter, 2011; Pyc & Rawson, 2010). As an example of a process associated with retrieval, there is evidence that the testing effect is enhanced by memory search during Test trials. For instance, there is substantial evidence that the testing effect is larger when Test trials involve recall rather than recognition (Carpenter, Pashler, & Vul, 2006; Glover, 1989; Kang, McDermott, & Roediger, 2007). This advantage holds even when the final test uses a different format (e.g. recognition, Carpenter & DeLosh, 2006; Glover, 1989), suggesting that recall-related processes like memory search may help explain the testing advantage, rather than the mere congruence between initial and final tests (Morris, Bransford, & Franks, 1977).

Despite a wealth of information from behavioral research, little work has been done to directly link the benefits of testing memory to brain function. This is particularly surprising given that functional neuroimaging techniques, such as functional MRI (fMRI), have successfully identified neural mechanisms associated with

\* Corresponding author at: Center for Cognitive Neuroscience, Duke University, Box 90999, B203 Levine Science Research Center, Durham, NC 27708, United States. Tel.: +1 919 668 2299.

E-mail address: [erik.wing@duke.edu](mailto:erik.wing@duke.edu) (E.A. Wing).

memory success at encoding and retrieval. A particularly powerful event-related fMRI method for investigating encoding is the *subsequent memory paradigm* (for a review see [Paller & Wagner, 2002](#)), which identifies greater encoding-phase activity for items that were remembered rather than forgotten on a later memory test. This difference in activity is known as the *subsequent memory effect* (SME) and is assumed to reflect successful encoding processes. A recent meta-analysis described a consistent set of regions that exhibit SMEs, including the medial temporal lobes (MTL), left prefrontal cortex (PFC), and superior parietal cortex ([Kim, 2011](#)). Within the MTL, SMEs in the hippocampus have been attributed to the storage of new contextual or semantic associations, which allow later recollection ([Davachi, 2006](#); [Eichenbaum, 2004](#); [Norman & O'Reilly, 2003](#); [Prince, Daselaar, & Cabeza, 2005](#)). Within left PFC, SMEs in ventrolateral regions are assumed to reflect processing and evaluation of semantic features ([Otten, Henson, & Rugg, 2001](#); [Wagner et al., 1998](#)).

Event-related fMRI can also be used to identify brain regions involved in successful retrieval operations by comparing activity for remembered and forgotten items during the test. Across numerous studies ([Spaniol et al., 2009](#)), retrieval success has been associated with activations in the hippocampus, left PFC, ventral parietal cortex, and posterior midline regions (e.g. posterior cingulate). Thus, MTL and left PFC regions are associated with both encoding and retrieval success. A study that directly compared word pair encoding vs. retrieval ([Prince et al., 2005](#)) found that SMEs were stronger in anterior MTL regions, and retrieval success in posterior MTL regions, consistent with other reports in the literature ([Lepage, Habib, & Tulving, 1998](#); [Saykin et al., 1999](#); [Strange, Fletcher, Henson, Friston, & Dolan, 1999](#)). Within left PFC, encoding success was associated with ventrolateral regions, and retrieval success with dorsolateral regions ([Prince et al., 2005](#)). In contrast to the MTL and left PFC, ventral parietal cortex and posterior midline cortex are regions very rarely associated with encoding success. In fact, these regions often show “negative SMEs” by displaying greater activity for subsequently forgotten than remembered items ([Daselaar et al., 2009](#); [Huijbers et al., 2012](#); [Uncapher & Wagner, 2009](#)).

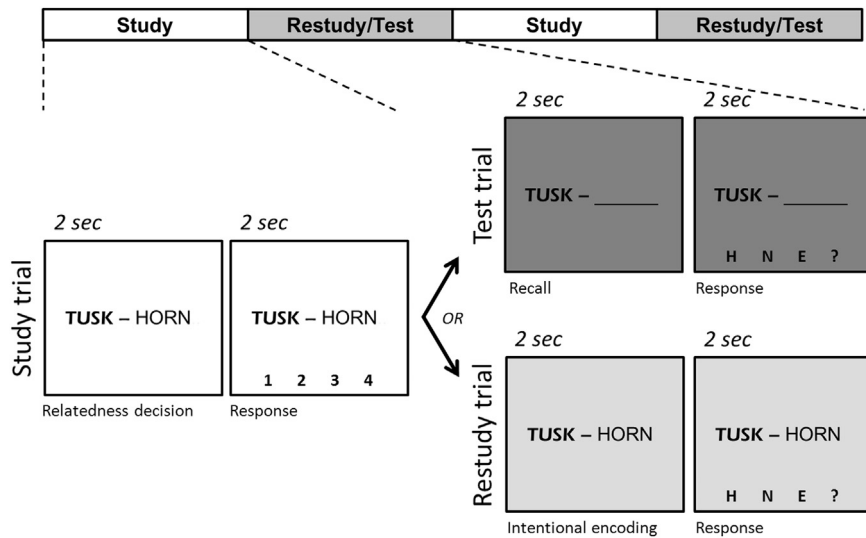
A few fMRI studies have investigated encoding processes occurring during retrieval by applying the subsequent memory paradigm to new (distractor) items presented during old/new recognition tasks. This method involves testing memory for the distractors through a surprise memory test after scanning and using performance on this second test to backsort the distractors in the first test as subsequently remembered or forgotten. In these studies, SMEs for the distractors were found in typical encoding regions including MTL ([Stark & Okado, 2003](#)) and left ventrolateral PFC ([Buckner, Wheeler, & Sheridan, 2001](#); [Huijbers, Pennartz, Cabeza, & Daselaar, 2009](#)). Given that these regions facilitate incidental encoding during retrieval, they may also be important candidate regions involved in strengthening retrieved representations during testing. Nonetheless, these findings have only limited applicability to the testing effect, which necessitates a direct comparison of items from the Test and Restudy conditions. Unlike contrasts using a set of new items during recognition, an ideal testing comparison would examine a single Restudy or Test exposure for a set of items that have been initially exposed under uniform conditions (Study). This format would minimize potential novelty signals elicited by recognition lures and would help ensure that subsequent memory success could more easily be attributed to the single Test or Restudy exposure.

Research on other aspects of memory also offers some insight into questions surrounding the testing effect, including how repeated exposure to an item influences retention. Repeatedly presented items typically produce a reduced neural response in sensory cortex, as compared with their initial presentation (for a

review, see [Grill-Spector, Henson, & Martin, 2006](#); [Henson & Rugg, 2003](#)), and the extent of reductions has been tied to subsequent memory success or failure in a number of studies ([Turk-Browne, Yi, & Chun, 2006](#); [Wagner, Maril, & Schacter, 2000](#); [Xue et al., 2011](#)). This pattern has also been found in the hippocampus ([Vannini, Hedden, Sullivan, & Sperling, 2012](#)), which exhibits activity reductions corresponding to subsequent gist-based memory rather than recollection ([Manelis, Paynter, Wheeler, & Reder, 2013](#)). Related findings have been described in continuous recognition tasks, where activity in the hippocampus decreases across successive presentations but increases during successful recollection ([Suzuki, Johnson, & Rugg, 2011a, 2011b](#)), with potential significance for the long-term retention of recollected items.

Neuroimaging studies of retrieval induced forgetting (e.g. [Kuhl, Dudukovic, Kahn, & Wagner, 2007](#)) also provide information about repeated processing of a stimulus in differing contexts. In one study, experimental blocks containing selective retrieval of word pair associates were compared against blocks in which items were passively rehearsed ([Wimber, Rutschmann, Greenlee, & Bäuml, 2009](#)). The resulting contrast showed increased activity for retrieval blocks in several areas including the hippocampus, ventrolateral PFC, and lateral parietal cortex, even though a final test showed no condition-related memory differences. However, activity in such block-related contrasts may reflect broad differences in task demands between retrieval and restudy conditions that do not relate to the enhancement of later memory—a question better suited to trial-specific subsequent memory analyses. Another recent study directly examined the influence of retrieval on subsequent memory, finding that activity in dorsal anterior cingulate cortex (ACC) increased with the success of previous retrieval attempts, and also predicted across-subject memory performance on a later test ([Eriksson, Kalpouzos, & Nyberg, 2011](#)). While this finding suggests the importance of the ACC in test-related memory retention, the absence of a Restudy comparison complicates the attempt to connect this result with retrieval-specific processes underlying the testing effect. The presence of repeated retrieval practice prior to testing in the scanner also means that memory enhancements cannot be as easily linked to neural activity within the specific test trials that were examined.

The current experiment was designed to address such issues by directly comparing the processes through which retrieval benefits subsequent memory in contrast to simple restudy (see [Fig. 1](#)). We used pairs of weakly related English words that were both conducive to a cued-recall format and enabled testing over a large number of items after a substantial delay. Each fMRI run consisted of alternating Study and Practice blocks. During Study blocks, participants intentionally encoded a set of word pairs and also rated the relatedness between the left word (cue) and right word (target) in each pair. Study blocks were followed by Practice blocks, in which the previously-shown pairs were evenly split into Restudy and Test trials. In Restudy trials, word pairs were presented intact, and participants intentionally re-encoded the entire pair. In Test trials, only cue words were presented, and participants covertly retrieved the associated targets. Retrieval accuracy was gauged by requiring participants to select the last letter of the recalled target word from several possible options that appeared in the final half of Test trials ([Fig. 1](#)). In order to equate visual and motor components, this procedure was also included in the final half of Restudy trials, following intentional re-encoding. Unlike testing effect studies in which the same items are practiced several times (e.g. [Karpicke & Roediger, 2007](#)), in our study each item was practiced only once (as Restudy or Test) in order to better isolate testing effects within a single learning trial. One day after the fMRI session, participants were given a final cued-recall test outside the scanner. Performance during this final cued-recall phase was used to backsort both Restudy and Test



**Fig. 1.** Schematic of one run from the fMRI portion of experiment. Each Study and Practice block contained 16 trials. Twenty-four hours after the scan session, memory for all initial word pairs was assessed in a final cued-recall test (not shown). Performance on this final test was used to backsort Practice (i.e., Test/Restudy) trials as either subsequently remembered or forgotten.

trials as either subsequently remembered or subsequently forgotten. The resulting  $2 \times 2$  factorial design allowed us to compare the size of SMEs (subsequently remembered vs. forgotten) for word pairs presented in Test trials vs. Restudy trials.

In the context of this approach, the mnemonic processes contributing to the testing effect advantage should be reflected in areas showing greater SMEs (subsequently remembered vs. forgotten) for Test than Restudy trials (i.e., a subsequent memory  $\times$  condition interaction). We investigated this question by measuring not only activation levels but also differences in functional connectivity. Given that the hippocampus has been associated with successful relational memory during both encoding and retrieval (Davachi & Wagner, 2002; Giovanello, Schnyer, & Verfaellie, 2004; Hannula & Ranganath, 2008; Prince et al., 2005), we reasoned that this region might be particularly involved in strengthening word pair associations during Test trials. We also expected that Test-related memory enhancement might involve areas of left ventrolateral PFC, particularly if semantic elaboration during Test trials contributes to subsequent memory as suggested by behavioral research. Additionally, we were interested in whether regions more uniquely associated with retrieval than encoding success, such as ventral parietal and posterior midline regions, would contribute to encoding success in the special context of Test trials.

## 2. Methods

### 2.1. Participants

Twenty-four right-handed, college-aged participants took part in the study. Participants were healthy, native English speakers with no reported history of neurological or psychiatric episodes. All participants gave written informed consent in accordance with a protocol approved by the Duke University Institutional Review Board. One participant was excluded from analysis due to high final cued-recall performance that yielded an insufficient number of miss trials ( $< 10$ ) for the key hit vs. miss contrast, and another participant was excluded after failing to follow instructions during the final cued-recall portion of the experiment. All behavioral and fMRI analyses were conducted on the remaining 22 participants.

### 2.2. Stimuli

Stimuli consisted of 192 weakly related English word pairs drawn from a database of free association norms (Nelson, McEnvoy, & Schreiber, 1998). The mean forward associate strength (i.e., the probability of producing the target word given the cue) for the set was 0.043. The pairs were divided into two sets of 96 pairs, which were assigned to Restudy or Test conditions (counter-balanced across participants).

### 2.3. Scanned paradigm

The scan session consisted of six identically-structured runs (Fig. 1), each with four blocks (Study, Practice, Study, Practice). Both types of blocks started with an instruction screen (4 s) and consisted of 16 trials (4 s each) separated by jittered fixation periods with a mean of 3 s. During each Study block, participants intentionally encoded 16 new word pairs, while rating the relatedness between the words in each pair (1 = moderately related, 4 = highly related). In the next Practice block, half of the 16 pairs were practiced in Test trials, and half in Restudy trials. Separate background colors were used for each of the Study, Test, and Restudy trial types. In each Test trial, only cue words were displayed, while a blank space appeared in place of the target. During the first 2 s of Test trials, participants were instructed to covertly recall the target word, and during the last 2 s, they had to indicate the last letter of the invisible recalled word from three letter options at the bottom of the screen. In each Restudy trial, word pairs were displayed again in full. During the first 2 s, participants were instructed to use these trials as an additional opportunity to re-encode the pairs, and during the last 2 s they had to press a key corresponding to the last letter of the visible target word. Thus, the response component of Test and Restudy trials was identical, and hence, it was subtracted out in direct contrast between these two types of trials. In both Restudy and Test trials, response options appeared during the final half of the trial to ensure participants focused on either restudying or recalling during the first half of the trial. Last letters were used instead of first letters to discourage a strategy in which letter options could be used as retrieval cues, and a post-scan questionnaire confirmed that participants did not rely on such a strategy to successfully recall targets. If participants could not recall the

target, or the final letter for the word they retrieved was not among the three letter options provided in the trial, they selected a question mark option. No feedback was given to avoid confounding activity related to recall with activity related to feedback processing. In both Test and Restudy trials, the last letter response options changed across trials. To reduce working memory contributions, pairs presented in the first and second halves of a Study block were re-presented in the first and second halves of the Practice block, respectively (order randomized within each half). Thus, the minimum lag was between the Study and Practice trials for a pair was eight trials.

### 2.3.1. Post-scan session

Participants returned 24 h after the scan for a surprise cued-recall test that assessed memory for all 192 word pairs. During this session, which took place in a computer laboratory, each cue word from the scan session was presented for 7 s on a white background while participants attempted to recall the previously-associated target word. Verbal responses – captured via microphone – were scored as either correct or incorrect, and then used to backsort both Test and Restudy trials from the fMRI session as either subsequently remembered or forgotten. Following the final cued-recall test, participants completed a short reading span task after which they were compensated and debriefed. Even though the correspondence in format between initial and final test is not thought to contribute substantially to the testing effect (Carpenter, 2009; Glover, 1989), we note several elements of the final test that help distinguish it from the initial Test trials, including response format, trial duration, background color, and physical location/context. Nonetheless, it is difficult to completely rule out the potential influence of cued-recall format similarity across the initial and final tests.

## 3. fMRI methods

All MRI data acquisition was conducted with a 3-T GE scanner. Scanner noise was reduced with ear plugs, and head motion was minimized with foam pads. Behavioral responses were recorded with a 4-key fiber-optic response box (Resonance Technology, Inc.), and when necessary, vision was corrected using MRI-compatible lenses that matched the distance prescription used by the participant. High-resolution structural images were collected using a 3D, T1-weighted FSPGR sequence (256 × 256 matrix, 166 slices, 1 mm slice thickness). Functional images were acquired using a SENSE spiral sequence (64 × 64 matrix, TR=2000 ms, TE=27 ms, FOV=24 cm, flip angle=60). Thirty-four contiguous slices were acquired in an interleaved fashion. Slice thickness was 3.8 mm, resulting in 3.75 × 3.75 × 3.8 mm voxels.

## 4. fMRI analyses

Preprocessing was performed using SPM8 software implemented in MATLAB ([www.fil.ion.ucl.ac.uk/spm/](http://www.fil.ion.ucl.ac.uk/spm/)). Segmented tissue probability maps were generated from anatomical volumes, and the VBM8 toolbox was then used to generate deformation fields for each participant based on the DARTEL template brain. After discarding the first six volumes of each run, functional images were corrected for slice time acquisition and motion. The corrected images were then coregistered to native space grey-matter tissue maps and normalized to MNI space using the deformations generated during normalization of the anatomical images.

Statistical fMRI analysis at both the subject and group level was performed in SPM5. Data were high-pass filtered using a cutoff of 128 s. For each subject, evoked hemodynamic responses to event

types were modeled with a delta (stick) function corresponding to the onset of stimulus presentation convolved with a canonical hemodynamic response function and temporal derivative in the context of the general linear model (GLM). Separate trial type regressors were defined for subsequently remembered and forgotten Test and Restudy trials, with additional regressors corresponding to the initial Study trials, instructional screens, and trials with no response. Regressors for session means and motion parameters were also included in the model.

Separate contrasts for subsequently remembered and forgotten Test and Restudy trials were generated for each subject and then submitted to a random effects analysis. For group analyses, a 2 × 2 repeated measures ANOVA was performed, with factors of practice condition (Test vs. Restudy) and subsequent memory (Remembered vs. Forgotten). The main effects of condition and memory were evaluated, along with the testing effect interaction that identified regions associated with memory success for the Test condition in contrast to Restudy (all contrasts were evaluated at  $p < .001$ , extent threshold=5 voxels).

An additional functional connectivity analysis was conducted to explore functional coupling between the hippocampus – identified in the initial memory by condition interaction above – and other regions of the brain. For this analysis, a separate model was constructed in which each trial was entered as a separate regressor, yielding estimates for each individual trial within each participant (for details of this method, see Daselaar, Fleck, & Cabeza, 2006; Daselaar, Fleck, Prince, & Cabeza, 2006; Rissman, Gazzaley, & D'Esposito, 2004). Within each trial type, the mean activity from the hippocampal ROI (suprathreshold voxels from the above interaction contrast within an anatomical mask of the hippocampus) was correlated with every other voxel to produce separate correlation volumes. The single subject correlation volumes for subsequently remembered and forgotten Test and Restudy were then entered into a 2 × 2 repeated measures ANOVA, allowing for group level comparisons of hippocampal coupling across conditions. A connectivity interaction contrast was then generated to detect regions that covaried with the hippocampus during successful Test more than Restudy trials ( $p < .001$ ,  $k=5$ , as with the corresponding activity ANOVA, the interaction effect was further inclusively masked with the Test subsequent memory contrast ( $p < .01$ ) to ensure effects were not merely attributable to a reverse memory effect within Restudy trials).

## 5. Results

### 5.1. Behavioral

Confirming that participants were able to successfully recall target words during Test trials, the last letter option was correctly selected at a rate of 91.3%, while Restudy trials showed a predictable ceiling effect of 99.6%. Final cued-recall demonstrated a clear testing effect. Memory was higher for words that had been practiced in the Test condition ( $M=0.63$ ,  $SD=0.14$ ) than those that had been shown again in the Restudy condition ( $M=0.51$ ,  $SD=0.14$ ;  $t(21)=9.25$ ,  $p < .0001$ ). As illustrated by Fig. 2, this testing effect difference ( $M=0.12$ ,  $SD=0.06$ ) was evident in all but one participant, with a range of 0.00 to 0.23.

### 5.2. Activations

To investigate the brain activity associated with retrieval-based memory enhancement, we conducted a 2 (subsequent memory: remembered vs. forgotten) × 2 (condition: Test vs. Restudy) ANOVA. As noted before, the critical comparison regarding the testing effect is the interaction between subsequent memory and



condition, which reveals areas where SMEs are greater for Test than Restudy trials. The main effects of condition are not directly related to the testing effect because they reflect the particular features of the Test and Restudy conditions that may not play a role in promoting later memory. Regions showing a main effect of condition are listed in Table 1. Compared to Test trials, Restudy trials were associated with greater activity in inferior parietal, lateral temporal, and dorsal PFC. Conversely, stronger activations for Test than Restudy trials were found in dorsal ACC, bilateral ventrolateral PFC extending into anterior insula, and midbrain. A

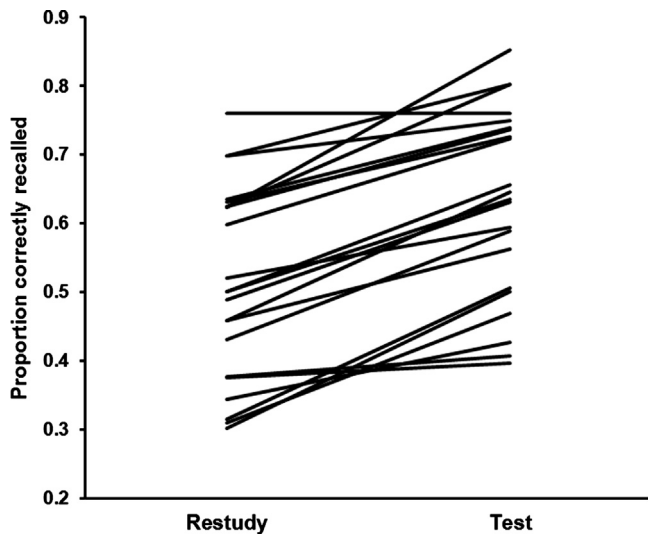


Fig. 2. Behavioral testing effect. Single subject behavioral performance showing proportion correct recall during final cued-recall test as a function of scan-session practice condition (Test/Restudy).

Table 1  
Main effect of condition.

Region	Hem	BA	x	y	z	Z	vox
<i>Test &gt; Restudy</i>							
Inferior frontal gyrus	R	47	34	19	-8	> 8	1442
	L	47	-30	23	-8	> 8	
	L	45	-53	34	4	6.18	
Globus pallidus	R	-	11	8	-8	5.90	
Anterior cingulate gyrus	-	32/24	0	23	34	7.59	756
Inferior temporal gyrus	L	37	-49	-49	-11	3.76	7
Parahippocampal gyrus	L	-	-23	-41	0	3.50	6
Precuneus	L	7	-26	-68	38	4.55	58
Posterior cingulate gyrus	-	23	0	-26	19	4.19	37
Middle occipital gyrus	L	19	-34	-83	15	3.44	9
Cerebellum	-	-	0	-60	-34	4.28	26
Cerebellum	-	-	15	-41	-46	3.81	5
<i>Restudy &gt; Test</i>							
Middle/medial frontal gyrus	R	8	23	19	42	6.45	756
Middle/superior frontal gyrus	L	9	-30	23	38	5.05	125
Medial frontal gyrus	R	10	11	60	8	4.02	37
Middle temporal gyrus	R	21	64	-49	-4	4.64	37
	R	21	56	-19	-23	5.49	112
	L	20	-56	-4	-30	3.61	37
	L	21	-45	4	-34	3.55	14
Superior temporal/Postcentral gyrus	L	22/40	-49	-11	-4	4.94	161
IPL/Angular gyrus	R	39/40/7	49	-68	30	6.81	329
Supramarginal gyrus	L	40	-56	-56	27	6.28	172
Angular gyrus	L	39	-49	-71	34	4.92	172
Precuneus	R	31	8	-53	34	3.68	15
Postcentral gyrus	R	40	53	-30	19	3.96	39
	R	40	23	-41	57	3.51	5
Lingual gyrus	L	18	-23	-83	-8	5.77	264
Cuneus	L/R	18/19	8	-83	19	5.19	96
Lingual gyrus	R	18	11	-71	-8	5.24	111

Up to five local maxima set 16 mm apart are reported for each cluster. BA, Brodmann area; Hem, hemisphere. Coordinates are in MNI space.

main effect of subsequent memory was found in a single area in dorsomedial PFC (MNI: 8, 45, 42), which was more active for subsequently remembered than forgotten trials in both Test and Restudy conditions. The relative lack of SMEs shared by Test and Restudy trials is consistent with previous evidence that SMEs are task dependent (Fletcher, Stephenson, Carpenter, Donovan, & Bullmore, 2003; Otten & Rugg, 2001a), as well as with substantial SME differences between Test and Restudy conditions, as revealed by the interaction findings below.

The main focus of the study was to identify regions showing greater SMEs for Test than Restudy conditions, which were isolated by memory  $\times$  condition interactions. As listed in Table 2, interaction effects were found in the bilateral hippocampus, lateral temporal cortex, medial prefrontal cortex, and left striatum. The hippocampal interaction occurred in the anterior hippocampus, with somewhat stronger differences evident in left hemisphere (Fig. 3B, C). The effects in lateral temporal cortex were also stronger in the left hemisphere, where they were localized to middle and inferior temporal gyri in contrast to superior temporal gyrus in the right hemisphere (Fig. 3A). In addition to showing higher activity for subsequently remembered vs. forgotten Test trials, some regions from the interaction also showed an inverted pattern on Restudy trials, with higher activity for subsequently forgotten than remembered trials. No significant SMEs were found for the reversed contrast (Restudy > Test).

### 5.3. Functional connectivity

Motivated by the pattern of findings from the interaction contrast, a functional connectivity analysis was performed to examine the potential for differential coupling between the hippocampus and other brain regions. This analysis addressed the possibility that regions insensitive to condition differences

might nonetheless show different patterns in coactivity with a region directly identified as important to later memory. Consequently, a second  $2 \times 2$  repeated measures ANOVA with factors of condition and memory success was constructed using individual subject maps reflecting condition-specific hippocampal correlation instead of activity. The corresponding interaction revealed regions where differential coactivation with the hippocampus predicted subsequent memory success to a greater extent for Test than Restudy trials (Fig. 4, Table 3). This pattern was found in the posterior cingulate cortex (Fig. 4A), medial PFC (Fig. 4B), and left ventrolateral PFC (Fig. 4C).

## 6. Discussion

The current study explored the neural correlates of the testing effect. The capacity of retrieval to promote subsequent memory was examined, as well as the how mnemonic differences during testing differ from simple restudy. On a final memory test one day after the scan, target words that had been practiced through retrieval were remembered at a higher rate than those practiced via restudy, confirming a behavioral testing effect. The critical interaction between practice condition and subsequent memory showed that testing enhanced subsequent memory effects (SMEs) in the hippocampus, left temporal cortex, and medial PFC. Functional connectivity analyses identified regions whose interactions with the hippocampus predicted subsequent memory primarily for the Test condition, including left ventrolateral PFC, medial PFC, and posterior cingulate. Below we discuss these findings in the context of relational encoding processes, which may allow for

relevant semantic information to be incorporated into coherent representations through testing. We also discuss the potential contribution of retrieval processes like memory search, and connect the present findings to related research on how the integration of information during retrieval influences memory consolidation.

### 6.1. Hippocampus

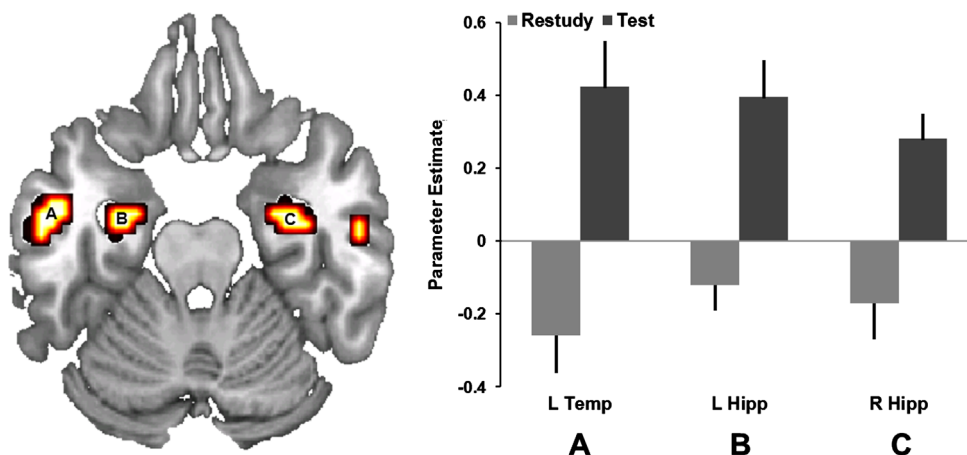
In contrast to Restudy trials, bilateral anterior hippocampus activity was greater for Test trials that were subsequently remembered than those that were forgotten on the final test. While the hippocampus figures centrally in many mnemonic operations, it is thought to play a particularly important role in binding disparate information into coherent representations at encoding (Eichenbaum, Yonelinas, & Ranganath, 2007; Yassa & Stark, 2011). At retrieval, the hippocampus may coordinate reinstatement of initially formed associations through pattern completion processes that operate in response to partially overlapping cues (Norman & O'Reilly, 2003). Differential engagement of relational memory processes suggests one way in which retrieval practice may enhance memory over traditional restudy. In the present study, Test trials might be expected to induce processing of previously formed cue-target associations more often than Restudy trials, wherein such associations are of less strategic importance. The element of memory search during target retrieval may also produce novel associations between word pairs and related information (e.g. retrieval candidates). When incorporated into an updated representation, such related information – whether formed during initial encoding or during the process of retrieval – may provide additional cues during final retrieval, improving retention and facilitating recollection-based memory.

Support for such an account of the present hippocampal findings is evident in both the neuroimaging literature on relational memory and in testing effects research. Consistent with the localization of observed SME effects, several studies of relational encoding have found activity in the anterior hippocampus that tracks subsequent memory success (Jackson & Schacter, 2004; Prince et al., 2005; Sperling et al., 2003). Attempts to link brain function with recollective experience have also emphasized the critical role of the hippocampus. Numerous studies using verbal material have reported hippocampal activity for encoding that led to subsequent source or relational memory (Davachi, Mitchell, & Wagner, 2003; Ranganath et al., 2004; Uncapher, Otten, & Rugg, 2006), with a similar pattern evident for analogous comparisons

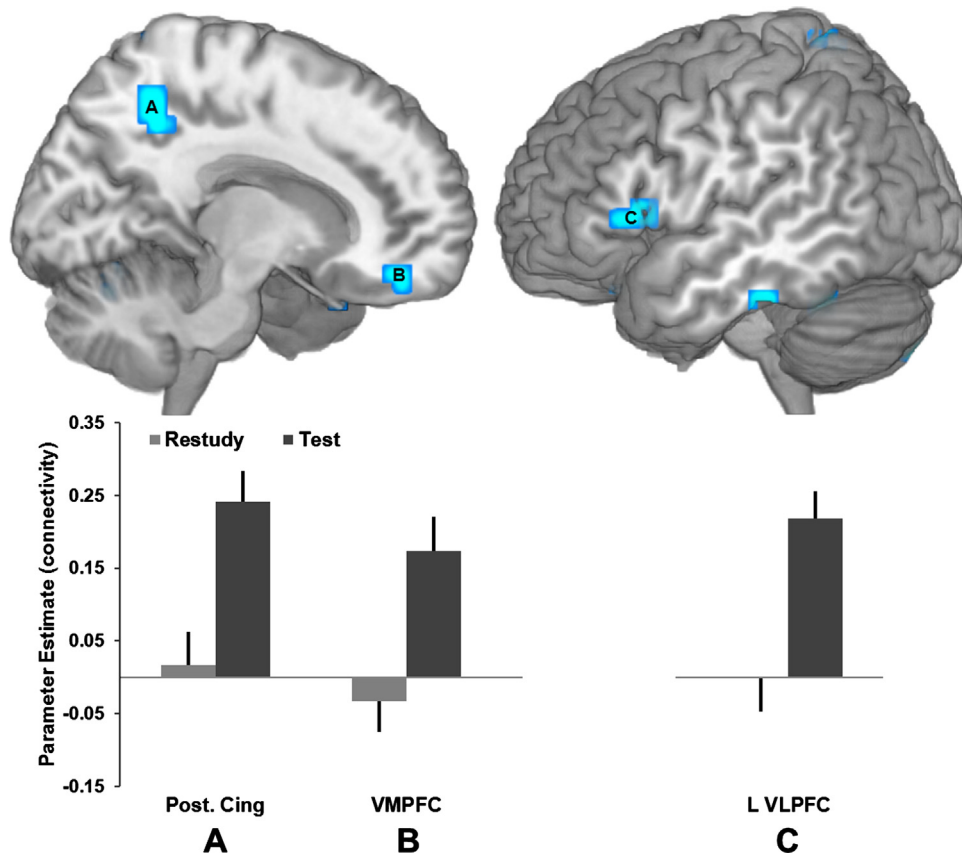
**Table 2**  
Condition  $\times$  memory interaction.

Region	Hem	BA	x	y	z	Z	vox
<i>Test &gt; Restudy</i>							
Anterior cingulate gyrus	L	32	-15	30	0	3.99	13
	L	32	-15	41	-4	3.60	13
Inferior temporal gyrus	L	20	-49	-8	-23	4.41	18
Superior temporal gyrus	R	22	53	-8	-11	4.29	23
Hippocampus	R	-	30	-8	-23	4.1	11
Hippocampus/parahippocampal gyrus	L	-	-30	-11	-27	3.67	12
Insula	R	13	38	-23	15	4.03	23
	R	13	45	-30	19	3.65	9
Insula/claustrum	L	-	-34	-4	4	4.02	50

Up to three local maxima set 8 mm apart are reported for each cluster. BA, Brodmann area; Hem, hemisphere. Coordinates are in MNI space.



**Fig. 3.** Testing effect interaction. Brain regions showing a condition (Test/Restudy) by memory (subsequently remembered/subsequently forgotten) interaction. Interaction effects are evident in left middle/inferior temporal gyri (A) and bilateral anterior hippocampus (B, C). Bars reflect SME (difference in activity between subsequently remembered and forgotten). Error bars denote standard error.



**Fig. 4.** Hippocampal connectivity interaction. Brain regions showing hippocampal connectivity differences by condition and subsequent memory. Regions in both posterior (posterior cingulate/precuneus, A) and anterior (medial PFC, B) midline showed this effect, along with left ventrolateral PFC (C). Bars reflect difference in hippocampal connectivity between subsequently remembered and forgotten. Error bars denote standard error.

**Table 3**  
Hippocampal connectivity interaction.

Region	Hem	BA	x	y	z	Z	vox
<i>Test &gt; Restudy interaction</i>							
Inferior frontal gyrus	L	47	-23	19	-19	4.63	20
	L	44	-60	15	8	3.98	13
	L	45	-49	26	4	3.49	13
	R	47	53	26	-4	3.81	7
Middle frontal gyrus	R	11	23	38	-19	3.74	5
Medial frontal gyrus	R	11	11	41	-15	3.69	6
Superior temporal gyrus	L	22	-64	-38	4	3.52	12
	R	21	49	4	-19	4.44	17
Insula/Superior temporal gyrus		13/22	45	-15	0	3.63	18
Middle temporal gyrus	R	37	56	-45	-23	3.76	8
Inferior temporal gyrus	L	20	-56	-26	-27	4.04	5
Parahippocampal gyrus	L	36	-34	-23	-23	4.31	5
Supramarginal gyrus	R	40	60	-53	34	4.25	21
Postcentral gyrus	L	7	-19	-53	72	4.14	7
	R	5/40	26	-49	68	3.66	13
Precuneus	R	7	15	-49	46	3.93	14
Cerebellum	R	-	4	-68	-23	3.72	7
Cerebellum	R	-	23	-86	-42	4.33	11
Cerebellum	R	-	45	-56	-30	3.81	14
Cerebellum	L	-	-41	-53	-30	3.59	6

Up to three local maxima set 8 mm apart are reported for each cluster. BA, Brodmann area; Hem, hemisphere. Coordinates are in MNI space.

during the retrieval phase (Giovanello et al., 2004; Yonelinas, Otten, Shaw, & Rugg, 2005). Hippocampal activity has also been observed during successful encoding (Fernández et al., 1998; Habib & Nyberg, 2008; Staresina & Davachi, 2006) and retrieval (Meltzer & Constable, 2005; de Zubicaray et al., 2007) of

information during cued or free recall tests, which are thought to diminish the opportunity for familiarity-based memory responses.

Recent behavioral findings also implicate relational processing in retrieval practice effects. In comparison to traditional restudy, testing has been shown to increase memory for semantic mediators (words that link cue-target pairs). This is the case both for mediators that are explicitly generated to aid retrieval, (Pyc & Rawson, 2010) and for mediators that are never explicitly produced but that contain a strong semantic connection between cues and targets (Carpenter, 2011). Carpenter (2011) found that when word pairs (e.g. mother–child) were practiced through testing as opposed to restudy, participants were later more likely to false alarm to related semantic mediators (e.g. father). Initial testing also made these mediators more effective substitute cues from which to retrieve targets on a modified final test. The availability of mediating information to serve as additional memory cues may also help explain why retrieval practice also promotes subsequent recollection, as observed in studies exploring the testing effect in a dual process framework (Chan & McDermott, 2007; Verkoeijen, Tabbers, & Verhage, 2011). Although multiple factors likely contribute to the testing effect, the pattern of hippocampal findings observed in the present study provides further support for the involvement of relational processing, which should be particularly beneficial when centered on stimulus-related semantic associations, as detailed in the above behavioral studies.

## 6.2. Lateral temporal cortex

The integration of related stimulus associations requires not only a binding mechanism, but also access to pre-existing

knowledge. Such knowledge is necessary for assessing the relatedness of word pairs during the initial Study trials and may be tapped to varying degrees during Restudy and Test trials. Whereas Restudy trials contain both components of the to-be-remembered pair, Test trials require retrieval of previously stored information that may be facilitated by initially formed associations. Semantic knowledge would also be necessary during the process of memory search, as semantically relevant concepts are brought to mind and evaluated in the attempt to produce the specific target.

Findings from the interaction of condition and memory success provide some indication that retrieval practice may benefit from differential processing of semantic information. Most notably, activity in left middle temporal cortex was associated with stronger SMEs during Test than during Restudy trials. Left lateral temporal cortex has been linked to both semantic and episodic memory systems, which are thought to interact closely (Tulving, 1972, 2002). Neuroimaging studies of semantic memory retrieval often report activity in left lateral and anterior temporal cortices (Simmons & Martin, 2009; Thompson-Schill, 2003), and degeneration of the left temporal lobe contributes to memory deficits in semantic dementia (Hodges & Patterson, 2007; Snowden, Griffiths, & Neary, 1996). In episodic memory studies, lateral temporal cortex has been associated with successful memory during both encoding and retrieval (Spaniol et al., 2009). Exploring the relationship between semantic and episodic memory, Menon et al. (2002) found that activity in left lateral temporal cortex (BA 21/22) corresponded to successful retrieval both for the semantic properties of words, and for their mnemonic status in an episodic memory test (Menon, Boyett-Anderson, Schatzberg, & Reiss, 2002). The authors suggest that this finding reflects the use of previously encoded semantic information during episodic retrieval. Such an account fits well with the pattern of current lateral temporal SMEs for the Test condition. Presumably, not all relational information active during Test helps enhance subsequent memory. Associations that contain more specific or detailed connections to target concepts are likely to serve as better subsequent cues. Accordingly, processes mediating selection and controlled retrieval of unique associations should also be important in producing coherent and memorable representations.

### 6.3. Left ventrolateral PFC

Left ventrolateral PFC effects were observed in an analysis performed to help clarify how hippocampal connectivity during retrieval practice contributes to subsequent memory. Differences in hippocampal connectivity were examined across Test and Restudy trials that varied in final test memory outcome. This comparison produced an expanded and spatially distinct set of regions from those observed in the standard testing effect contrast. Stronger connectivity with ventrolateral PFC was found for Test trials that were effective in producing successful memory on the final recall test. This result indicates that the beneficial aspects of retrieval practice may arise through differential coupling between the hippocampus and other cortical regions that do not differ in overall activity.

The involvement of left ventrolateral PFC in the memory-related cognitive control operations (e.g., see Badre & Wagner, 2007) is particularly relevant to test-enhanced learning in the current design. This region has been linked to retrieval of knowledge-based representations in a number of studies comparing semantic vs. non-semantic conditions (e.g. Devlin, Matthews, & Rushworth, 2003; McDermott, Petersen, Watson, & Ojemann, 2003). Damage to ventrolateral PFC has also been shown to impair performance on semantic tasks (Martin & Cheng, 2006; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). The nature of current findings in the ventrolateral PFC indicates a more

complex relationship when considering how Restudy or Test trials influence subsequent memory. While this region may operate in a similar capacity to facilitate immediate cued recall during initial Test trials (i.e., a main effect of Test > Restudy), it may not invariably strengthen retrieved items such that they are better remembered in the future. Instead, information selected by the ventrolateral PFC may enhance later memory only when it can be effectively integrated into active representations via the hippocampus. In this scenario, the present connectivity findings would reflect a hippocampally-mediated elaboration effect in which the behavioral advantages of Test draw on the increased likelihood of collaboration between frontal regions, involved in controlled semantic processing, and the hippocampus, which integrates associated content into a more durable representation.

### 6.4. Medial PFC

The medial PFC was another region where hippocampal connectivity predicted subsequent memory for Test more than Restudy trials. The medial PFC contains considerable anatomical connections with limbic regions including the hippocampal formation (Cavada, Compañy, Tejedor, Cruz-Rizzolo, & Reinoso-Suárez, 2000), and memory studies have also reported functional associations between these regions. Functional connectivity between anterior midline and hippocampus has been found during episodic encoding in the context of working memory (Ranganath, Heller, Cohen, Brozinsky, & Rissman, 2005) and schema formation (van Kesteren, Fernández, Norris, & Hermans, 2010). Related areas of memory research have explored how retrieval affects pre-existing mnemonic associations. For example, during retrieval-mediated associative learning, newly acquired memories are integrated with (or influenced by) distinct but overlapping representations that have already been encoded (Zeithamova, Dominick, & Preston, 2012). Studies in both humans and rodents have emphasized the importance of the hippocampus and medial PFC during this process, which may enhance consolidation of newly encoded memories (Tse et al., 2007, 2011).

Consolidation processes may also contribute to retrieval practice advantages, as the magnitude of the testing effect typically increases with the temporal separation between retrieval practice and final test (i.e., the retention interval) (Roediger & Butler, 2011; Toppino & Cohen, 2009). Research on the testing effect has only recently begun to consider the role of consolidation, (Eriksson et al., 2011; Finn & Roediger, 2011). While not directly assessing Test/Restudy differences, a recent study found Test-related activity in a more dorsal ACC region that tracked the success of previous retrieval attempts while predicting retention across subjects (Eriksson et al., 2011). In the present study, medial PFC activity was evident in the key testing effect interaction, but this region also demonstrated differential connectivity with the hippocampus during Test and Restudy trials. This behaviorally-selective functional relationship provides a further connection to the research on retrieval-mediated learning for distinct but related material, and supports the notion that testing may strengthen memory through the retrieval and elaboration of initially encoded associative information.

### 6.5. Posterior cingulate cortex

The functional connectivity difference in posterior midline is interesting because this region is known to play a very different role during retrieval than at encoding (Spaniol et al., 2009). Whereas this region is associated with retrieval success, it is rarely associated with encoding success. Examination of encoding effects in posterior midline and other default mode network (DMN) regions including ventral parietal cortex have even shown an inverted success pattern, with higher activity for subsequently



forgotten than remembered items (Daselaar, Prince, & Cabeza, 2004; Otten & Rugg, 2001b; Wagner & Davachi, 2001). In posterior midline, reverse SMEs have been shown to overlap with retrieval success activations both within subjects and across a variety of memory studies, underscoring the reliability and pervasiveness of these phase-related differences (Daselaar et al., 2009; Huijbers, Pennartz, Cabeza, & Daselaar, 2011). Furthermore, a study comparing task-related hippocampal activation with resting state default mode activity found greater coupling between the hippocampus and DMN during retrieval, but divergent activity profiles during encoding (Huijbers et al., 2011). These findings illustrate the complex relationship between activity and connectivity in the present results, where the hippocampus appears to act in concert with traditional retrieval regions during Test trials that promote later memory. While many of the present activity findings for Test/Restudy interactions appear to reflect the enhancement of traditional encoding regions (hippocampus, left temporal cortex, left PFC, etc.), patterns of hippocampal connectivity suggest that the testing effect is also associated with recruitment of regions consistently tied to retrieval operations (e.g. posterior cingulate). Further research is necessary to disentangle these two possible mechanisms.

## 7. Conclusion

Following the extensive behavioral research on the testing effect, we used event-related fMRI to investigate the processes through which retrieval promotes subsequent memory in contrast to traditional restudy. Consistent with past findings, final test memory performance was higher for items that had been previously retrieved vs. restudied during scanning. The critical interaction between practice condition (Test or Restudy) and final memory outcome showed that increased activity in the hippocampus and lateral temporal cortex predicted later memory during testing but not restudy. Additionally, a connectivity analysis revealed increased coupling between the hippocampus and ventrolateral PFC, medial PFC, and posterior cingulate cortex when retrieval practice was effective at producing subsequent memory. These results suggest that the power of retrieval to promote memory stems in part from relational memory processes that operate on selected semantic associations during testing. The hippocampus may be involved in integrating relevant information into updated representations, supporting memory search, and interfacing with other regions involved in consolidation processes to produce more durable memories.

## Acknowledgements

This work was supported by the National Institute on Aging [Grant number R56 AG23770]. We thank Henry L. Roediger III for his advice on experimental design and theoretical issues, James E. Kragel and Jared Stokes for their assistance with programming of the paradigm, and Kerry Townsend for her assistance with data collection and analysis.

## References

- Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, 45(13), 2883–2901, <http://dx.doi.org/10.1016/j.neuropsychologia.2007.06.015>.
- Buckner, R. L., Wheeler, M. E., & Sheridan, M. A. (2001). Encoding processes during retrieval tasks. *Journal of Cognitive Neuroscience*, 13(3), 406–415, <http://dx.doi.org/10.1162/08989290151137430>.
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1563–1569, <http://dx.doi.org/10.1037/a0017021>.
- Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1547–1552, <http://dx.doi.org/10.1037/a0024140>.
- Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34(2), 268–276.
- Carpenter, S. K., & Kelly, J. W. (2012). Tests enhance retention and transfer of spatial learning. *Psychonomic Bulletin & Review*, 19(3), 443–448, <http://dx.doi.org/10.3758/s13423-012-0221-2>.
- Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*, 14(3), 474–478.
- Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cue recall test? *Psychonomic Bulletin & Review*, 13(5), 826–830.
- Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, 20(6), 633–642.
- Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral Cortex*, 10(3), 220–242.
- Chan, J. C. K., & McDermott, K. B. (2007). The testing effect in recognition memory: a dual process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 431–437, <http://dx.doi.org/10.1037/0278-7393.33.2.431>.
- Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. *Journal of Neurophysiology*, 96(4), 1902–1911, <http://dx.doi.org/10.1152/jn.01029.2005>.
- Daselaar, S. M., Fleck, M. S., Prince, S. E., & Cabeza, R. (2006). The medial temporal lobe distinguishes old from new independently of consciousness. *The Journal of Neuroscience*, 26(21), 5835–5839, <http://dx.doi.org/10.1523/JNEUROSCI.0258-06.2006>.
- Daselaar, S. M., Prince, S. E., & Cabeza, R. (2004). When less means more: deactivations during encoding that predict subsequent memory. *NeuroImage*, 23(3), 921–927, <http://dx.doi.org/10.1016/j.neuroimage.2004.07.031>.
- Daselaar, S. M., Prince, S. E., Dennis, N. A., Hayes, S. M., Kim, H., & Cabeza, R. (2009). Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Frontiers in Human Neuroscience*, 3, 13, <http://dx.doi.org/10.3389/neuro.09.013.2009>.
- Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, 16(6), 693–700, <http://dx.doi.org/10.1016/j.conb.2006.10.012>.
- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, 100(4), 2157–2162, <http://dx.doi.org/10.1073/pnas.0337195100>.
- Davachi, L., & Wagner, A. D. (2002). Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *Journal of Neurophysiology*, 88(2), 982–990.
- Devlin, J. T., Matthews, P. M., & Rushworth, M. F. S. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of cognitive neuroscience*, 15(1), 71–84, <http://dx.doi.org/10.1162/089892903321107837>.
- Eichenbaum, H. B. (2004). Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron*, 44(1), 109–120, <http://dx.doi.org/10.1016/j.neuron.2004.08.028>.
- Eichenbaum, H. B., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30(1), 123–152, <http://dx.doi.org/10.1146/annurev.neuro.30.051606.094328>.
- Eriksson, J., Kalpouzos, G., & Nyberg, L. (2011). Rewiring the brain with repeated retrieval: a parametric fMRI study of the testing effect. *Neuroscience Letters*, 505(1), 36–40, <http://dx.doi.org/10.1016/j.neulet.2011.08.061>.
- Fernández, G., Weyerts, H., Schrader-Bölsche, M., Tendolkar, I., Smid, H. G., Tempelmann, C., Hinrichs, H., et al. (1998). Successful verbal encoding into episodic memory engages the posterior hippocampus: a parametrically analyzed functional magnetic resonance imaging study. *The Journal of Neuroscience*, 18(5), 1841–1847.
- Finn, B., & Roediger, H. L. (2011). Enhancing retention through reconsolidation: negative emotional arousal following retrieval enhances later recall. *Psychological Science*, 22(6), 781–786, <http://dx.doi.org/10.1177/0956797611407932>.
- Fletcher, P. C., Stephenson, C. M. E., Carpenter, T. A., Donovan, T., & Bullmore, E. T. (2003). Regional brain activations predicting subsequent memory success: an event-related fMRI study of the influence of encoding tasks. *Cortex*, 39(4–5), 1009–1026.
- Giovanello, K. S., Schnyer, D. M., & Verfaellie, M. (2004). A critical role for the anterior hippocampus in relational memory: evidence from an fMRI study comparing associative and item recognition. *Hippocampus*, 14(1), 5–8, <http://dx.doi.org/10.1002/hipo.10182>.
- Glover, J. A. (1989). The “Testing” phenomenon: not gone but nearly forgotten. *Journal of Educational Psychology*, 81(3), 392–399.
- Grill-Spector, K., Henson, R. N. A., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23, <http://dx.doi.org/10.1016/j.tics.2005.11.006>.
- Habib, R., & Nyberg, L. (2008). Neural correlates of availability and accessibility in memory. *Cerebral Cortex*, 18(7), 1720–1726, <http://dx.doi.org/10.1093/cercor/bhm201>.
- Hannula, D. E., & Ranganath, C. (2008). Medial temporal lobe activity predicts successful relational memory binding. *The Journal of Neuroscience*, 28(1), 116–124, <http://dx.doi.org/10.1523/JNEUROSCI.3086-07.2008>.

- Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, 41(3), 263–270.
- Hodges, J. R., & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurology*, 6(11), 1004–1014, [http://dx.doi.org/10.1016/S1474-4422\(07\)70266-1](http://dx.doi.org/10.1016/S1474-4422(07)70266-1).
- Huijbers, W., Pennartz, C. M. A., Cabeza, R., & Daselaar, S. M. (2009). When learning and remembering compete: a functional MRI study. *PLoS biology*, 7(1), e11, <http://dx.doi.org/10.1371/journal.pbio.1000011>.
- Huijbers, W., Pennartz, C. M. A., Cabeza, R., & Daselaar, S. M. (2011). The hippocampus is coupled with the default network during memory retrieval but not during memory encoding. In: G. Chapouthier (Ed.), 6 (p. e17463). *PLoS One*, <http://dx.doi.org/10.1371/journal.pone.0017463>.
- Huijbers, W., Vannini, P., Sperling, R. A., Pennartz, C. M. A., Cabeza, R., & Daselaar, S. M. (2012). Explaining the encoding/retrieval flip: memory-related deactivations and activations in the posteromedial cortex. *Neuropsychologia*, 1–11, <http://dx.doi.org/10.1016/j.neuropsychologia.2012.08.021>.
- Jackson, O., & Schacter, D. L. (2004). Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. *NeuroImage*, 21(1), 456–462, <http://dx.doi.org/10.1016/j.neuroimage.2003.09.050>.
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4–5), 528–558, <http://dx.doi.org/10.1080/09541-440601056620>.
- Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57(2), 151–162.
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968, <http://dx.doi.org/10.1126/science.1152408>.
- Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *NeuroImage*, 54(3), 2446–2461, <http://dx.doi.org/10.1016/j.neuroimage.2010.09.045>.
- Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience*, 10(7), 908–914, <http://dx.doi.org/10.1038/nn1918>.
- Larsen, D. P., Butler, A. C., & Roediger, H. L. (2009). Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Medical Education*, 43(12), 1174–1181, <http://dx.doi.org/10.1111/j.1365-2923.2009.03518.x>.
- Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET activations of memory encoding and retrieval: the HIPER model. *Hippocampus*, 8(4), 313–322, [http://dx.doi.org/10.1002/\(SICI\)1098-1063\(1998\)8:4<313::AID-HIPO1>3.0.CO;2-I](http://dx.doi.org/10.1002/(SICI)1098-1063(1998)8:4<313::AID-HIPO1>3.0.CO;2-I).
- Manelis, A., Paynter, C. a, Wheeler, M. E., & Reder, L. M. (2013). Repetition related changes in activation and functional connectivity in hippocampus predict subsequent memory. *Hippocampus*, 23(1), 53–65, <http://dx.doi.org/10.1002/hipo.22053>.
- Martin, R. C., & Cheng, Y. (2006). Selection demands versus association strength in the verb generation task. *Psychonomic Bulletin & Review*, 13(3), 396–401.
- McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14(2), 200–206 doi:VL – 14.
- McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia*, 41(3), 293–303, [http://dx.doi.org/10.1016/S0028-3932\(02\)00162-8](http://dx.doi.org/10.1016/S0028-3932(02)00162-8).
- Meltzer, J. a, & Constable, R. T. (2005). Activation of human hippocampal formation reflects success in both encoding and cued recall of paired associates. *NeuroImage*, 24(2), 384–397, <http://dx.doi.org/10.1016/j.neuroimage.2004.09.001>.
- Menon, V., Boyett-Anderson, J. M., Schatzberg, A. F., & Reiss, A. L. (2002). Relating semantic and episodic memory systems. *Cognitive Brain Research*, 13(2), 261–265.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519–533, [http://dx.doi.org/10.1016/S0022-5371\(77\)80016-9](http://dx.doi.org/10.1016/S0022-5371(77)80016-9).
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida Word Association, Rhyme, and Word Fragment Norms.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*, 110(4), 611–646, <http://dx.doi.org/10.1037/0033-295X.110.4.611>.
- Otten, L. J., Henson, R. N., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. *Brain*, 124(Pt 2), 399–412.
- Otten, L. J., & Rugg, M. D. (2001a). Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cerebral Cortex*, 11(12), 1150–1160.
- Otten, L. J., & Rugg, M. D. (2001b). When more means less: neural activity related to unsuccessful memory encoding. *Current Biology*, 11(19), 1528–1530, [http://dx.doi.org/10.1016/S0960-9822\(01\)00454-7](http://dx.doi.org/10.1016/S0960-9822(01)00454-7).
- Paller, K. a, & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, 6(2), 93–102.
- Prince, S. E., Daselaar, S. M., & Cabeza, R. (2005). Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. *The Journal of Neuroscience*, 25(5), 1203–1210, <http://dx.doi.org/10.1523/JNEUROSCI.2540-04.2005>.
- Pyc, M. a, & Rawson, K. a. (2010). Why testing improves memory: mediator effectiveness hypothesis. *Science*, 330(6002), 335, <http://dx.doi.org/10.1126/science.1191465>.
- Ranganath, C., Heller, A., Cohen, M. X., Brozinsky, C. J., & Rissman, J. (2005). Functional connectivity with the hippocampus during successful memory formation. *Hippocampus*, 15(8), 997–1005.
- Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*, 42(1), 2–13.
- Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, 23(2), 752–763, <http://dx.doi.org/10.1016/j.neuroimage.2004.06.035>.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27, <http://dx.doi.org/10.1016/j.tics.2010.09.003>.
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255, <http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x>.
- Saykin, A. J., Johnson, S. C., Flashman, L. A., McAllister, T. W., Sparling, M., Dorcey, T. M., Moritz, C. H., et al. (1999). Functional differentiation of medial temporal and frontal regions involved in processing novel and familiar words: an fMRI study. *Brain*, 122 (Pt 1), 1963–1971.
- Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, 15(5), 645–649, <http://dx.doi.org/10.1017/S155617709990348>.
- Snowden, J. S., Griffiths, H. L., & Neary, D. (1996). Semantic-episodic memory interactions in semantic dementia: implications for retrograde memory function. *Cognitive Neuropsychology*, 13(8), 1101–1139, <http://dx.doi.org/10.1080/026432996381674>.
- Spaniol, J., Davidson, P. S. R., Kim, A. S. N., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47(8–9), 1765–1779, <http://dx.doi.org/10.1016/j.neuropsychologia.2009.02.028>.
- Sperling, R. A., Chua, E. F., Cocchiarella, A., Rand-Giovannetti, E., Poldrack, R. A., Schacter, D. L., & Albert, M. (2003). Putting names to faces: successful encoding of associative memories activates the anterior hippocampal formation. *NeuroImage*, 20(2), 1400–1410, [http://dx.doi.org/10.1016/S1053-8119\(03\)00391-4](http://dx.doi.org/10.1016/S1053-8119(03)00391-4).
- Staresina, B. P., & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *The Journal of Neuroscience*, 26(36), 9162–9172, <http://dx.doi.org/10.1523/JNEUROSCI.2877-06.2006>.
- Stark, C. E. L., & Okado, Y. (2003). Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. *The Journal of Neuroscience*, 23(17), 6748–6753.
- Strange, B. A., Fletcher, P. C., Henson, R. N. A., Friston, K. J., & Dolan, R. J. (1999). Segregating the functions of human hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, 96(7), 4034–4039.
- Suzuki, M., Johnson, J. D., & Rugg, M. D. (2011a). Recollection-related hippocampal activity during continuous recognition: a high-resolution fMRI study. *Hippocampus*, 21(6), 575–583, <http://dx.doi.org/10.1002/hipo.20781>.
- Suzuki, M., Johnson, J. D., & Rugg, M. D. (2011b). Delements in hippocampal activity with item repetition during continuous recognition: an fMRI study. *Journal of Cognitive Neuroscience*, 23(6), 1522–1532, <http://dx.doi.org/10.1162/jocn.2010.21535>.
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where”. *Neuropsychologia*, 41(3), 280–292.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 94(26), 14792–14797.
- Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval: questions and answers. *Experimental Psychology*, 56(4), 252–257, <http://dx.doi.org/10.1027/1618-3169.56.4.252>.
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. a, Wood, E. R., Witter, M. P., et al. (2007). Schemas and memory consolidation. *Science*, 316(5821), 76–82, <http://dx.doi.org/10.1126/science.1135935>.
- Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., et al. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, 333(6044), 891–895, <http://dx.doi.org/10.1126/science.1205274>.
- Tulving, E. (1972). Episodic and semantic memory. In: E. Tulving, & W. Donaldson (Eds.), *Organization of Memory*, Vol. 1 (pp. 381–403). Academic Press, <http://dx.doi.org/10.1017/S0140525x00047257>.
- Tulving, Endel. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology*, 53(1), 1–25, <http://dx.doi.org/10.1146/annurev.psych.53.100901.135114>.
- Turk-Browne, N. B., Yi, D.-J., & Chun, M. M. (2006). Linking implicit and explicit memory: common encoding factors and shared representations. *Neuron*, 49(6), 917–927, <http://dx.doi.org/10.1016/j.neuron.2006.01.030>.
- Uncapher, M. R., Otten, L. J., & Rugg, M. D. (2006). Episodic encoding is more than the sum of its parts: an fMRI investigation of multifaceted contextual encoding. *Neuron*, 52(3), 547–556, <http://dx.doi.org/10.1016/j.neuron.2006.08.011>.
- Uncapher, M. R., & Wagner, A. D. (2009). Posterior parietal cortex and episodic encoding: insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiology of Learning and Memory*, 91(2), 139–154, <http://dx.doi.org/10.1016/j.nlm.2008.10.011>.

- Vannini, P., Hedden, T., Sullivan, C., & Sperling, R. a. (2012). Differential functional response in the posteromedial cortices and hippocampus to stimulus repetition during successful memory encoding. *Human Brain Mapping, 00*, <http://dx.doi.org/10.1002/hbm.22011> (January 2011).
- Verkoeijen, P. P. J. L., Tabbens, H. K., & Verhage, M. L. (2011). Comparing the effects of testing and restudying on recollection in recognition memory. *Experimental Psychology, 58*(6), 490–498.
- Wagner, A. D., & Davachi, L. (2001). Cognitive neuroscience: forgetting of things past. *Current Biology, 11*(23), R964–R967.
- Wagner, A. D., Maril, A., & Schacter, D. L. (2000). Interactions between forms of memory: when priming hinders new episodic learning. *Journal of Cognitive Neuroscience, 12*(Suppl. 2), 52–60, <http://dx.doi.org/10.1162/089892900564064>.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., et al. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*(5380), 1188–1191.
- Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K.-H. (2009). Retrieval from episodic memory: neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience, 21*(3), 538–549, <http://dx.doi.org/10.1162/jocn.2009.21043>.
- Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience, 23*(7), 1624–1633, <http://dx.doi.org/10.1162/jocn.2010.21532>.
- Yassa, M. a., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34*(10), 515–525, <http://dx.doi.org/10.1016/j.tins.2011.06.006>.
- Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience, 25*(11), 3002–3008, <http://dx.doi.org/10.1523/JNEUROSCI.5295-04.2005>.
- Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron, 75*(1), 168–179, <http://dx.doi.org/10.1016/j.neuron.2012.05.010>.
- de Zubicaray, G., McMahon, K., Eastburn, M., Pringle, A. J., Lorenz, L., & Humphreys, M. S. (2007). Support for an auto-associative model of spoken cued recall: evidence from fMRI. *Neuropsychologia, 45*(4), 824–835, <http://dx.doi.org/10.1016/j.neuropsychologia.2006.08.013>.
- van Kesteren, M. T. R., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences of the United States of America, 107*(16), 7550–7555, <http://dx.doi.org/10.1073/pnas.0914892107>.