

# Neural Signatures of Test-Potentiated Learning in Parietal Cortex

Steven M. Nelson,<sup>1</sup> Kathleen M. Arnold,<sup>1</sup> Adrian W. Gilmore,<sup>1</sup> and Kathleen B. McDermott<sup>1,2</sup>

<sup>1</sup>Department of Psychology, Washington University, St. Louis, Missouri 63130, and <sup>2</sup>Department of Radiology, Washington University School of Medicine, St. Louis, Missouri 63110

Testing, or retrieval practice, is beneficial for long-term memory both directly, by enhancing performance on tested information, and indirectly, by facilitating learning from subsequent encounters with the information. Although a wealth of behavioral research has examined the “testing effect,” neuroimaging has provided little insight regarding the potential mechanisms that underlie the benefits of retrieval practice. Here, fMRI was used to examine the effects of retrieval practice on later study trials. Human subjects studied pairs of associated words, which were then tested, restudied, or neither tested nor restudied. All pairs were then studied once more in expectation of a final test. We asked how this Final Study episode was affected by prior history (whether the pair had been previously tested, restudied, or neither). The data revealed striking similarities between responses in lateral parietal cortex in the present study and those in a host of studies explicitly tapping recognition memory processes. Moreover, activity in lateral parietal cortex during Final Study was correlated with a behavioral index of test-potentiated learning. We conclude that retrieval practice may enhance learning by promoting the recruitment of retrieval mechanisms during subsequent study opportunities.

## Introduction

Taking a test on recently studied information increases the likelihood that the information will be retained in the long-term (“the testing effect”; Abbott, 1909; Gates, 1917; Spitzer, 1939; Tulving, 1967; Glover, 1989; Carrier and Pashler, 1992; Roediger and Butler, 2011). The benefits of testing, or retrieval practice, often exceed those achieved by simply studying (Karpicke and Roediger, 2008). These benefits have been demonstrated not only in laboratory settings, but also in classrooms (Larsen et al., 2008; McDaniel et al., 2011). Therefore, retrieval practice not only provides a window to elucidate human memory function, but also has applications for educational settings.

Related studies have demonstrated that testing recently studied information also enhances the effectiveness of a subsequent study opportunity (“test-potentiated learning”; Izawa, 1971; Rohrer et al., 2010; Arnold and McDermott, 2013a); that is, a retrieval attempt occurring between two study phases can enhance the amount of information gleaned from the second study phase (Arnold and McDermott, 2013b). How does retrieval prac-

tice augment the later study phase? This question is difficult to answer because memory encoding does not lend itself to direct behavioral observation. However, fMRI is particularly well suited for the study of memory encoding because encoding processes can be indexed without the need to infer such processes solely from later retrieval performance (Brewer et al., 1998; Wagner et al., 1998). Therefore, fMRI can help us better understand why memory encoding is especially effective when a prior retrieval attempt has occurred.

In the present study, we adopted a straightforward paradigm to measure the BOLD response while subjects studied a list of word pairs twice. Between these study episodes, the word pairs were tested, restudied, or neither tested nor restudied. We used a data-driven approach to identify regions in which activity differed between Initial Study and Final Study as a result of intermediate testing and differentiated between tested and untested items at Final Study.

One possible outcome is that the facilitation that accrues from testing might manifest as diminished activation during Final Study. This pattern would be similar to findings in the implicit memory literature, in which repeated processing of items is accompanied by “repetition suppression,” sometimes also referred to as “neural priming” (Buckner et al., 2000; Schacter et al., 2007). Indeed, previous research has shown that a greater degree of neural priming can result in better remembering (Turk-Browne et al., 2006; but see Xue et al., 2011). A second possibility is that testing allows for more elaborative encoding during Final Study. Encoding-related neural activity may be greater at Final Study than Initial Study (“repetition enhancement”) because the tested items have been processed in multiple contexts. Individuals may therefore draw on a richer set of representations when studying the material a second time ( Craik and Tulving, 1975; Bradshaw

Received March 2, 2013; revised May 27, 2013; accepted June 8, 2013.

Author contributions: S.M.N., K.M.A., and K.B.M. designed research; S.M.N., K.M.A., and A.W.G. performed research; S.M.N. analyzed data; S.M.N., A.W.G., and K.B.M. wrote the paper.

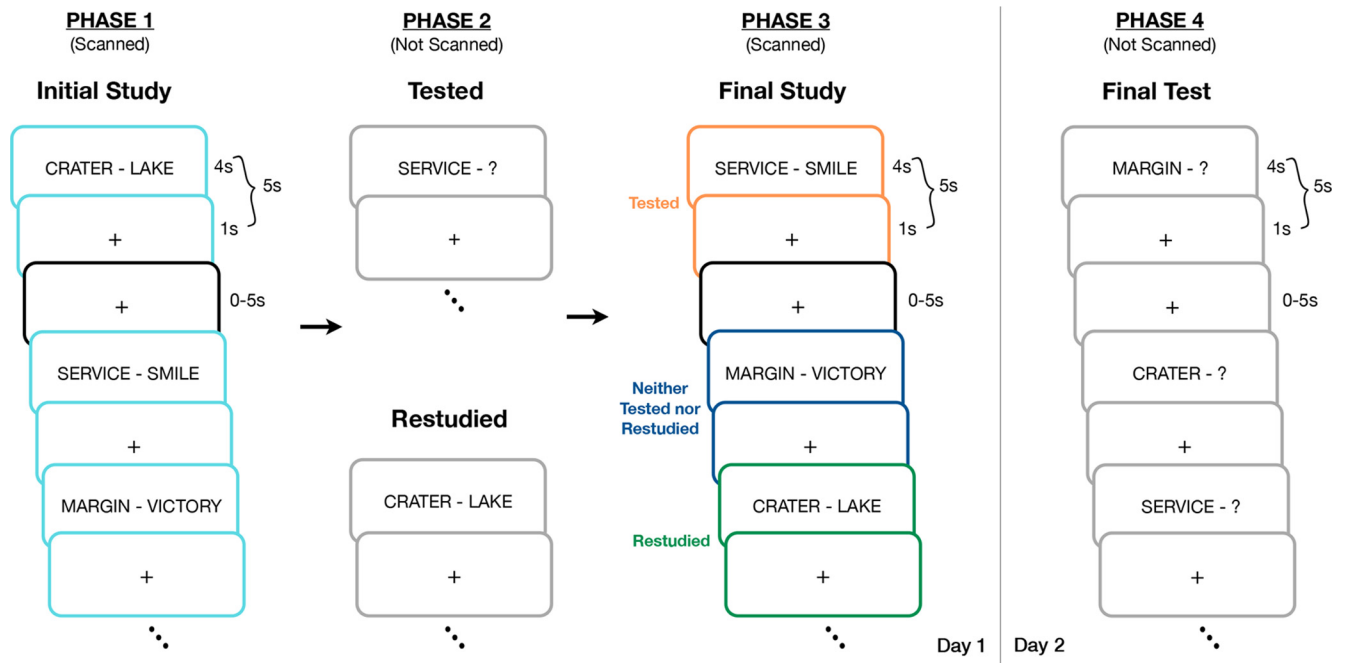
This work was supported by the McDonnell Center for Systems Neuroscience at Washington University, a James S. McDonnell Foundation 21st Century Science Initiative Grant: Bridging Brain, Mind and Behavior/Collaborative Award, and the National Science Foundation Graduate Research Fellowship (Grant #DGE-1143954). We thank Laura Najjar and Bridgid Finn for help with data collection and analysis, and Roddy Roediger, Dave Balota, Steve Petersen, Brad Schlaggar, Chris Fetsch, Gagan Wig, Jonathan Power, Mark Wheeler, Nico Dosenbach, Maital Neta, and Chris Wahlheim for discussions and helpful comments.

The authors declare no competing financial interests.

Correspondence should be addressed to Steven M. Nelson, Washington University, 1 Brookings Drive, Campus Box 1125, St. Louis, MO 63130. E-mail: stevenelson@wustl.edu.

DOI:10.1523/JNEUROSCI.0960-13.2013

Copyright © 2013 the authors 0270-6474/13/3311754-09\$15.00/0



**Figure 1.** In Phase 1, subjects were presented with pairs of words (cue-target) while being scanned during an Initial Study session (teal). Word pairs were shown for 4 s followed by a 1 s presentation of a centrally located crosshair and 0, 2.5, or 5 s of “fixation” in which a crosshair was also present (black). Word pairs were then tested by presenting the subjects with the cue and asking them to speak the target word, presented for restudy, or neither tested nor restudied in Phase 2. Finally, in Phase 3 (Final Study), subjects were simply shown all of the word pairs a final time while being scanned. The procedure for Phase 3 was identical to that of Phase 1, with the only difference emerging as a result of the Phase 2 manipulation (orange, tested; green, restudied; dark blue, neither tested nor restudied). A day later, participants took a final cued recall test on all 126 word pairs (Phase 4).

and Anderson, 1982). After distinguishing between these alternatives, we provide a framework for test-potentiated learning, incorporating into our results brain-behavior correlations from the present study as well as data from a recent meta-analysis of retrieval-related activity (Nelson et al., 2010).

### Materials and Methods

**Subjects.** Twenty-six subjects participated in the experiment and were recruited from Washington University and the St. Louis area. Two subjects were excluded for failing to comply with task instructions. For the remaining 24 subjects (12 female), ages ranged from 21 to 30 years (mean, 24.5 ± 2.54). All participants were right-handed native speakers of English with normal or corrected-to-normal vision. All participants were neurologically healthy with no reported history of psychiatric illness. Participants were consented in accordance with the guidelines set forth by Washington University’s Human Research Protection Office and were compensated for their time at a rate of \$25/h.

**Stimuli.** Stimuli consisted of 126 pairs of weakly associated words (e.g., disc-laser) drawn from the Nelson et al. (2004) norms. These pairs were divided into four lists that were equated for forward associative strength (range, 0.010–0.049; mean, 0.016) calculated as the proportion of subjects reported in Nelson et al. (2004) who produced a specific target (e.g., laser) given a cue (e.g., disc), cue-word frequency (range, 0.00–315; mean, 38.60) and target-word frequency (range, 0.00–967; mean, 85.08), both of which were determined by the Kuçera and Francis (1967) norms. Lists were also equated for cue-word length (range, 3–14; mean, 6.29), target-word length (range, 3–11; mean, 5.54), cue syllable count (range, 1–6; mean, 1.92), and target syllable count (range, 1–3; mean, 1.61). The order of these lists was counterbalanced across subjects and item order was randomized within each list for each participant. All stimuli were presented in black, 48-point Arial font on a white background.

**Procedure.** The experiment took place over 2 d and consisted of 4 phases. Phases 1–3 (Fig. 1) took place during the first day. In Phase 1, subjects studied the 126 weakly related paired associates with instructions to learn them for a later test. Specifically, subjects were presented with each word pair for 4 s and asked to attempt to learn each so that if

given the cue (e.g., “crater”), they could generate the target word (e.g., “lake”). Word pairs were separated by a jittered interstimulus interval of 1–6 s. During this phase (Phase 1) of the experiment, we measured the BOLD response as subjects studied each cue-target pair, all of which were novel within the context of the experiment. Subjects studied the 126 pairs across 2 consecutive scanning runs, each of which contained 63 pairs.

During Phase 2, the 126 studied pairs of items were arranged into 3 blocks of 42 pairs each (i.e., each block contained 1/3 of the word pairs from Phase 1). For one block, subjects were given a cued recall test in which they received the first item from each pair (e.g., “crater”) and asked to verbally recall its paired associate (e.g., “lake”). Subjects were not provided with feedback during this test period. For a separate block, participants were shown each word pair again with the instruction to restudy these items. The final block of items was neither tested nor restudied (i.e., items in this block were not presented to participants during this phase of the experiment). Therefore, of the original 126 word pairs, participants were tested on 42 word pairs, restudied 42 word pairs, and did not see 42 word pairs during Phase 2 of the experiment. The order of the testing and restudy blocks was counterbalanced across participants. Subjects remained in the scanner for this phase of the experiment, but no imaging data were collected.

During Phase 3, subjects restudied all 126 pairs of words once more. Word pairs were studied in a different order in this phase than during Phase 1, but the instructions provided to participants were identical. As with Phase 1, the BOLD response was measured for each cue-target pair. Word pairs were again presented to subjects in two scanning runs containing 63 pairs each. Upon completing Phase 3, participants exited the scanner and were reminded that they would return the following day.

Phase 4 occurred approximately 1 d after scanning (mean delay, 20.97 ± 3.06 h). During this phase, participants were given a final cued recall test in which they were presented with the first word of all 126 pairs of items (i.e., the cue) and 42 new words and asked to either recall the target word or to respond “new” if they had not seen the word previously. No feedback was given during the test period and the order of items on this test was random. After completing the cued-recall test, participants

were compensated for their time and debriefed in accordance with standard University policies.

**fMRI data acquisition.** Images were acquired in adherence to a standard protocol. To help stabilize head position, subjects were provided with a foam pillow and were fitted with a thermoplastic mask fastened to the head coil. All images were obtained with a Siemens MAGNETOM Tim Trio 3.0T Scanner and a Siemens 12-channel Matrix Head Coil. A T1-weighted sagittal MPRAGE structural image was obtained (TE = 3.08 ms, TR partition = 2.4 s, TI = 1000 ms, flip angle = 8 degrees, 176 slices with  $1 \times 1 \times 1$  mm voxels; Mugler and Brookeman, 1990). A T2-weighted turbo spin echo structural image (TE = 84 ms, TR = 6.8 s, 32 slices with  $2 \times 1 \times 4$  mm voxels) in the same anatomical plane as the BOLD images was also obtained to improve alignment to an atlas. An auto align pulse sequence protocol provided in the Siemens software was used to align the acquisition slices of the functional scans parallel to the anterior commissure–posterior commissure (AC-PC) plane and centered on the brain. This plane is parallel to the slices in the Talairach atlas (Talairach and Tournoux, 1988), which was used for subsequent data analysis. Functional imaging was performed using a BOLD contrast-sensitive gradient echo echoplanar sequence (TE = 27 ms, flip angle = 90°, in-plane resolution =  $4 \times 4$  mm). Whole-brain EPI volumes (MR frames) of 32 interleaved, 4-mm-thick axial slices were obtained every 2.5 s. The first four image acquisitions were discarded to allow net magnetization to reach steady state.

Headphones dampened scanner noise and enabled communication with participants. An Apple iMac computer and PsyScope software (Cohen et al., 1993) were used for display of visual stimuli. An LCD projector (model PG-C20XU; Sharp) was used to project stimuli onto a MRI-compatible rear-projection screen (CinePlex) at the head of the bore, which the participants viewed through a mirror attached to the coil.

**Preprocessing.** Imaging data from each subject were preprocessed to remove noise and artifacts, including: (1) correction for movement within and across runs using a rigid-body rotation and translation algorithm (Snyder, 1996); (2) whole-brain normalization to a common mode of 1000 to allow for comparisons across subjects (Ojemann et al., 1997); and (3) temporal realignment of all slices to the temporal midpoint of the first slice using sinc interpolation to account for differences in slice time acquisition. Functional data were then resampled into 3 mm isotropic voxels and transformed into stereotaxic atlas space (Talairach and Tournoux, 1988). Atlas registration involved aligning each subject's T1-weighted image to a custom atlas-transformed (Lancaster et al., 1995), target T1-weighted template (711–2B) using a series of affine transforms (Michelon et al., 2003).

**fMRI data analysis using the general linear model.** Preprocessed data were analyzed at the voxel level using a general linear model (GLM) approach (Friston et al., 1994; Miezin et al., 2000). Details of this procedure have been described previously (Ollinger et al., 2001). Briefly, the model treats the data at each time point (in each voxel) as the sum of all effects present at that time point. Effects can be produced by events in the model and by error. Estimates of the time course of effects were derived from the model for each response category by coding time points as a set of delta functions immediately after onset of the coded event (Ollinger et al., 2001).

Data from each subject consisted of 4 separate runs of 167 frames each (after discarding the first 4 frames to allow for T1 equilibration) that were concatenated into a single time series for functional analysis. Runs 1 and 2 contained the BOLD data corresponding to the Initial Study portion of the experiment (Phase 1), and Runs 3 and 4 included the BOLD data from Final Study (Phase 3). The GLM for each participant therefore consisted of a time series of 668 frames and this number of frames did not differ between participants (i.e., no runs were lost to movement, noise, etc., for any of the 24 subjects included the analysis).

Within the GLM, 8 conditions, each having eight time points (TR = 2.5 s), were modeled, for a total of 64 regressors. Specifically, for items that were tested in Phase 2, we wanted to examine activation during Initial Study (Phase 1) and Final Study (Phase 3) as a function of whether the item was indeed recalled in Phase 2 (see “Behavioral results” and Fig. 5). Therefore, the 8 conditions of interest were: (1) Initial Study for items both tested and recalled in Phase 2; (2) Initial Study for items tested but not recalled in Phase 2; (3) Initial Study for items restudied in Phase 2; (4)

Initial Study for items neither tested nor restudied in Phase 2; (5) Final Study for items that had been recalled in Phase 2; (6) Final Study for items tested but not recalled in Phase 2; (7) Final Study for items that had been restudied in Phase 2; and (8) Final Study for items neither tested nor restudied in Phase 2. This approach allowed us to examine activation for tested items during Phases 1 and 3 as a function of whether the item was correctly recalled in Phase 2. This same set of GLMs was used for all analyses; in cases in which we collapsed conditions (e.g., examined Initial Study activation regardless of subsequent condition), we simply averaged the relevant conditions (e.g., conditions 1 and 2 above for Fig. 3).

In addition to the regressors described above, over each run, a trend term accounted for linear changes in signal and a constant term modeled the baseline signal. Therefore, there were a total of 72 columns in the design matrix. Event-related effects are described in terms of percent signal change, defined as signal magnitude divided by a constant term. This approach makes no assumptions about the shape of the BOLD response, but does assume that all events included in a category are associated with the same BOLD response (Ollinger et al., 2001). Therefore, we could extract time courses without placing constraints on their shape. Image processing and analyses were performed using in-house software written in IDL (Research Systems).

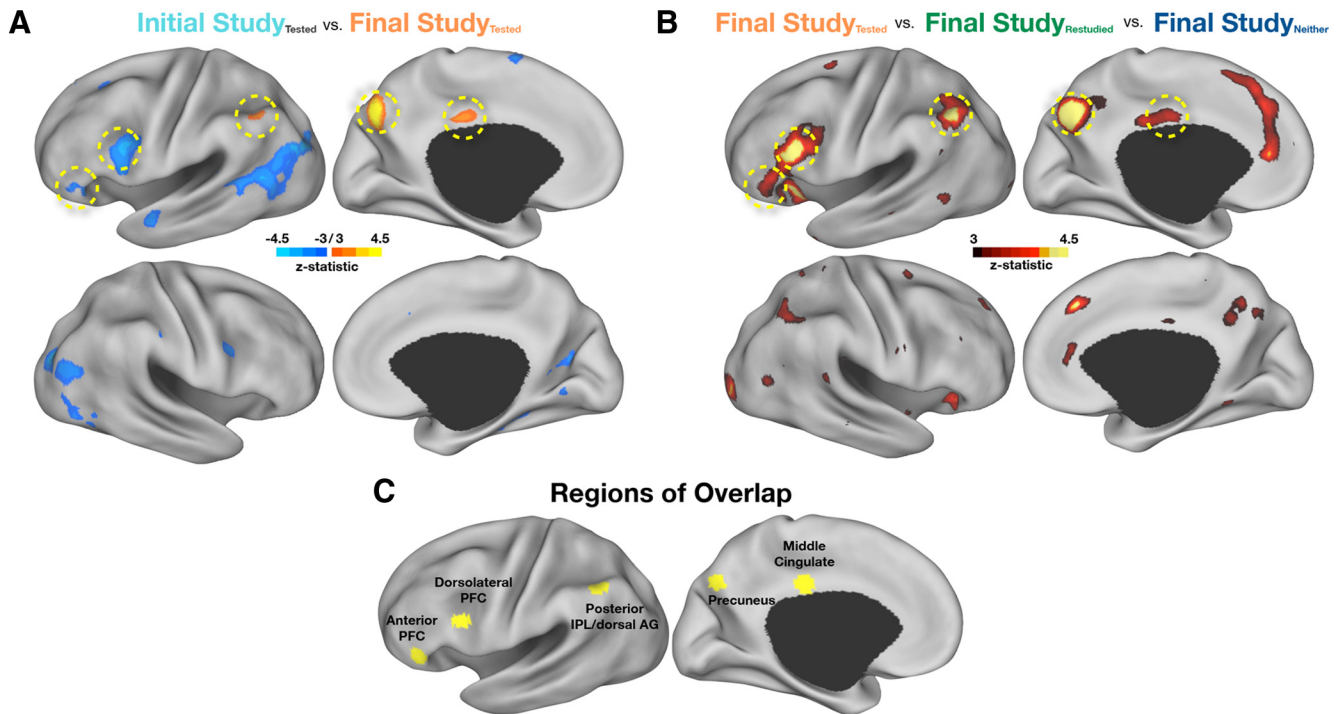
**Whole-brain voxelwise analysis and region of interest definition.** We conducted whole-brain voxelwise analyses to generate two separate images, the conjunction of which formed the basis for region of interest (ROI) definition. All statistical tests were conducted on cross-correlation magnitudes calculated for each voxel. Magnitudes were computed as the inner product of the estimated time course of the BOLD response and a vector of contrast weights modeling a  $\gamma$  function with a delay of 2 s and a time contrast of 1.25 s (Boynton et al., 1996). Three additional delays of 1 s accounted for variations in the onset of the hemodynamic response.

A region that is sensitive to retrieval practice that occurs between two study epochs should exhibit two characteristics: (1) differential activity between the study epochs (Initial and Final Study) for the tested items, and (2) differences within the second study epoch (Final Study) as a function of the Phase 2 manipulation (tested, restudied, or neither tested nor restudied). Therefore, ROIs were defined on the basis of these two contrasts. Although we were specifically interested in differential activity between Initial and Final Study for the tested items, we also examined statistical maps comparing Initial and Final Study for the restudied items and the same contrast for those items neither restudied nor tested. These contrast maps elicited a subset of regions emerging from the contrast of Initial and Final Study for the tested items, with the exception of a single additional region in left dorsolateral prefrontal cortex (left dlPFC; seen as more active for Initial Study than Final Study for the items restudied in Phase 2). Permitting these additional contrasts to contribute to region definition did not alter any of the conclusions.

First, we compared activity between tested items at Initial Study (Phase 1) and Final Study (Phase 3) using a paired *t* test (Fig. 2A). The image was Monte Carlo corrected at a *z*-score of 3.00 with at least 13 contiguous voxels (McAvoy et al., 2001). Next, we compared activity for tested items, restudied items, and neither tested nor restudied items at Final Study (Phase 3) using repeated-measures ANOVA (Fig. 2B). This image was Monte Carlo corrected at a *z*-score of 3.00 with at least 17 contiguous voxels (McAvoy et al., 2001). Note the difference in the number of contiguous voxels as a function of statistical test (*t* test vs ANOVA; McAvoy et al., 2001).

Finally, we created a binary mask for each image where statistically significant voxels were given a value of 1 and all other voxels were given a value of 0. Summing the statistical images yielded voxel values of 0, 1, or 2. ROIs were then defined using a peak-finding algorithm that searched for locations where voxel values were 2 and, after smoothing the data with a 4 mm blurring kernel, contained at least 13 contiguous voxels. Spherical regions of 10 mm diameter were then created around the peak locations derived from the search algorithm. Five ROIs emerged from this analysis, including regions in left lateral prefrontal cortex, left lateral parietal cortex (LLPC), and medial parietal cortex (Fig. 2C).

**fMRI meta-analysis of studies contrasting hits versus correct rejections.** We used a meta-analysis previously reported in Nelson et al. (2010) as a means of placing our findings in a broader context. Briefly, studies from the meta-analysis included a total of 140 neuro-



**Figure 2.** **A**, *t* test reveals voxels that were statistically significant ( $-3.0 > z > 3.0$ , multiple-comparison corrected, see Materials and Methods) when contrasting tested items at Initial Study with tested items at Final Study. Regions in blue showed greater activity at Initial Study than Final Study, whereas regions in orange showed greater activity at Final Study than Initial Study. Dotted yellow circles indicate regions of overlap in **C**. **B**, ANOVA reveals voxels that were statistically significantly different when comparing tested items, restudied items, and neither tested nor restudied items at Final Study ( $z > 3.0$ ). Dotted yellow circles indicate regions of overlap in **C**. **C**, In yellow are five regions defined from voxels that appeared in both of the above contrasts. All data are shown on inflated cortical surface renderings using Caret software (Van Essen et al., 2001).

logically normal adults between the ages of 18 and 35 who were recruited from both the Washington University and the University of Pittsburgh communities. Data were collected on either a 1.5 T Siemens MAGNETOM Vision Scanner (at Washington University in St. Louis) or a 3 T Siemens Allegra Scanner (at the University of Pittsburgh). Studies included a variety of different tasks in which judgments about item status were embedded within various source attribution, remember/know, or basic old/new decisions (Velanova et al., 2003; Wheeler and Buckner, 2003; Wheeler and Buckner, 2004; Phillips et al., 2009; Donaldson et al., 2010). In addition, the encoding tasks contained either visual or auditory stimuli that were either presented once or many times to enhance retrieval success.

Time courses were extracted from GLMs that were processed in the same manner as described above, in which each time point was separately estimated as the sum of all effects present at that time point with no assumptions made about the shape of the hemodynamic response. Time courses for “hits” and “correct rejections” were extracted separately for each condition in each study and averaged across conditions.

*Creation of conjunction image from eight ‘retrieval success’ conditions.* We constructed the conjunction image from all 8 conditions by thresholding each image from each condition at  $p < 0.05$  (uncorrected) and creating a binary mask in which voxels significant at the  $p < 0.05$  level were given a value of 1 and all other voxels given a value of 0. All voxels from the images were then summed so that the value of any given voxel could range from 0 (not significant at  $p < 0.05$  in any condition) to 8 (significant at  $p < 0.05$  in all conditions). Therefore, this image is indicative of the reliability with which the effect of interest (i.e., hits vs correct rejections) was present.

**Results**

**Behavioral results**

On the Initial Test (Phase 2), subjects correctly recalled an average of 40% of all possible targets (16.83/42) when prompted with the cue (Table 1). This relatively low level of performance was by

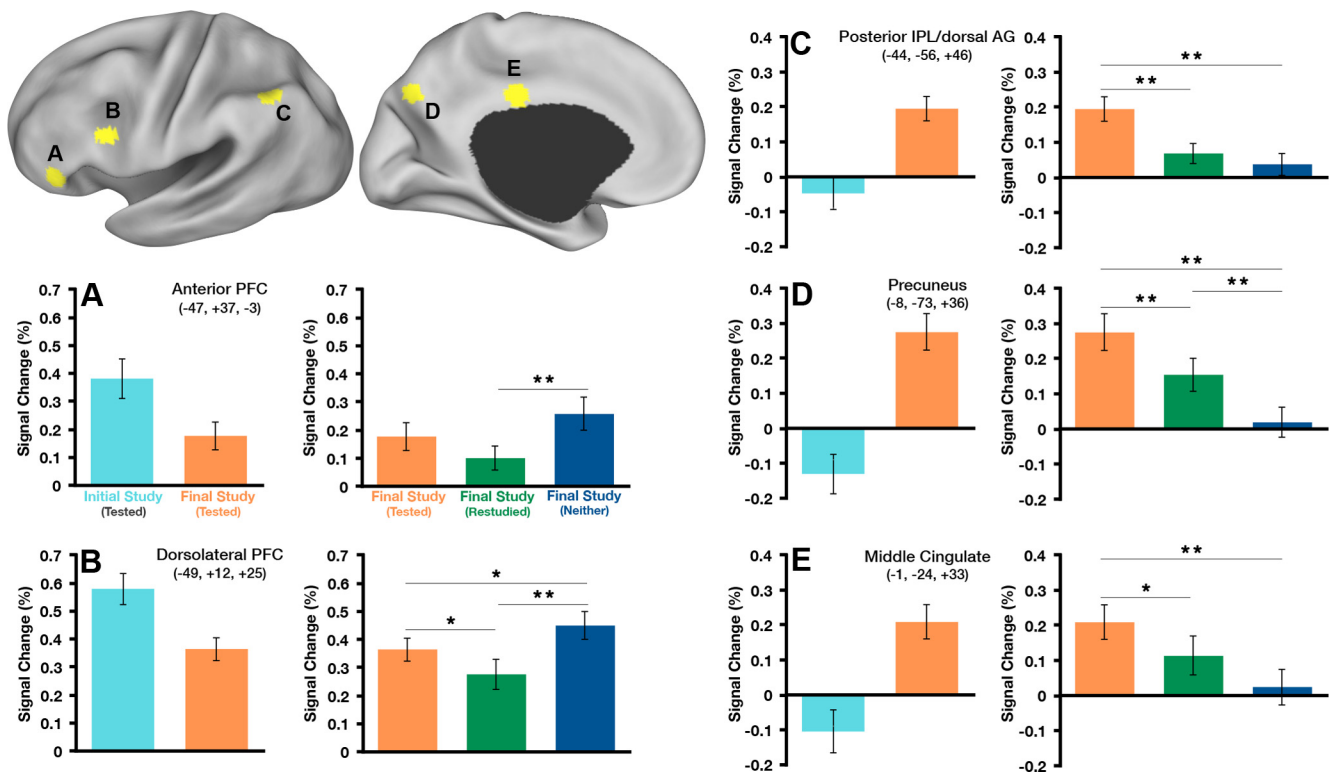
**Table 1. Behavioral results**

	P recall (SD)
Phase 2	0.40 (0.17)
Phase 4	
Tested	0.53 (0.19)
Restudied	0.53 (0.23)
Neither	0.35 (0.20)

Show are the proportions of items recalled during Phase 2 and Phase 4 separated by condition.

design (on the basis of pilot studies) so that subjects would benefit from the subsequent restudy episode. Of the 60% of items that subjects did not recall correctly, subjects responded with the word “Pass” (41%), did not respond (9%), or responded with an incorrect target word (10%).

The likelihood of recall on the Final Test for items initially tested during Phase 2 was 53% (22.08/42; Table 1). Final Test performance for restudied items was also 53% (22.04/42), whereas that for items neither tested nor restudied was 35% (14.50/42). Repeated-measures ANOVA showed a statistically significant difference across conditions ( $F_{(1,23)} = 41.25, p < 0.001$ ). *Post hoc* comparisons revealed significant differences for the Tested versus Neither Tested nor Restudied comparison ( $t_{(23)} = 9.67, p < 0.001$ ), as well as the Restudied versus Neither Tested nor Restudied comparison ( $t_{(23)} = 7.67, p < 0.001$ ), but no difference between performance for Tested and Restudied items ( $t_{(23)} = 0.38, p = 0.97$ ). Although many studies have shown a difference at Final Test for Tested versus Restudied items (Karpicke and Roediger, 2008), the failure to find such a difference here is not surprising given the low accuracy or retrievability during initial testing (Kang et al., 2007; Jang et al., 2012), coupled with the relatively short re-



**Figure 3.** Top left: The five ROIs that showed both a difference between Initial Study and Final Study for tested items and a difference in tested, restudied, and neither tested nor restudied items at Final Study are shown on inflated cortical surface renderings using Caret software (Van Essen et al., 2001). **A–E**, Average magnitude of the BOLD response is plotted for each condition of interest in each region (Initial Study<sub>(Tested)</sub> vs Final Study<sub>(Tested)</sub>, and Final Study<sub>(Tested)</sub> vs Final Study<sub>(Restudied)</sub> vs Final Study<sub>(Neither)</sub>). Stereotaxic coordinates are in MNI space. \* $p < 0.05$ , \*\* $p < 0.01$  ( $p$ -values are not indicated for the Initial Study<sub>(Tested)</sub> vs Final Study<sub>(Tested)</sub> contrast because the difference between conditions is captured in the voxelwise  $t$  test [Fig. 2A]; therefore, indicating  $p$ -values would imply circularity).

tention interval of  $\sim 24$  h (Roediger and Karpicke, 2006). Our intention was to make the Initial Test difficult enough so that a sufficient number of initially forgotten items might be subsequently recalled at Final Test (Phase 4). This allowed us to calculate an index of “New Learning” across subjects, which was defined as the proportion of items that were recalled at final test that had not initially been recalled [ $P(\text{Final Test}_{\text{Correctly Recalled}} | \text{Initial Test}_{\text{Not Recalled}})$ ] during Phase 2. For example, if a subject initially failed to recall 20 items of a possible 42 and subsequently recalled 10 of those items at final test, that subject’s New Learning value would be  $10/20 = 0.5$  or 50%. During Phase 4, subjects correctly produced an average of 28.2% (range, 5.8–55%) of all items that were not correctly recalled during the Phase 2 test.

### fMRI results

#### Regions in prefrontal cortex are sensitive to repetition, whereas regions in parietal cortex are specifically sensitive to retrieval practice

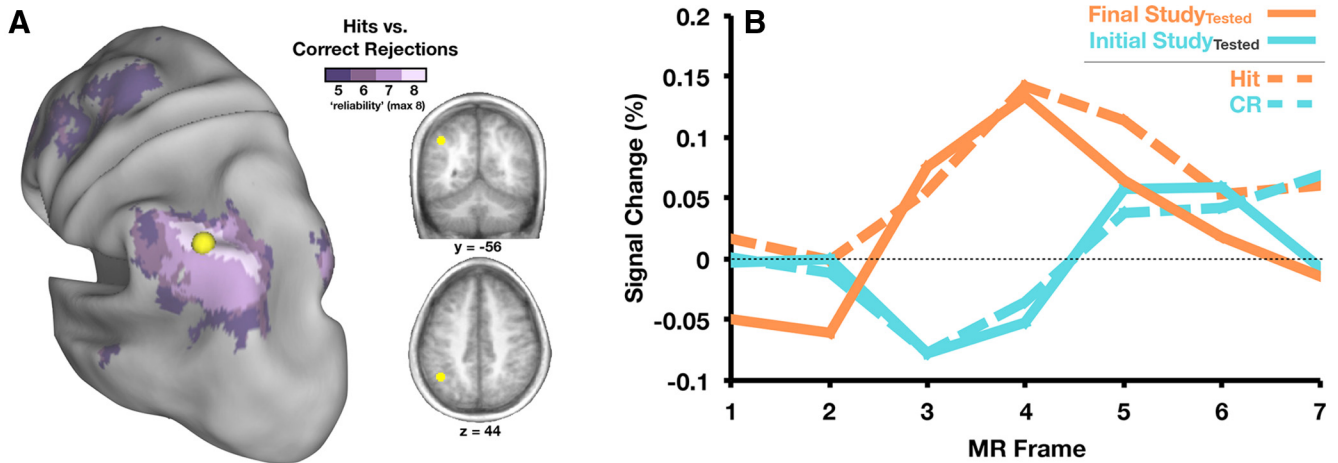
We extracted response magnitudes for each ROI defined in the conjunction map and found two basic patterns of results. The first pattern was a decrease in activity from Initial Study to Final Study (Fig. 3A,B); this pattern occurred in left dlPFC and anterior PFC. The magnitudes within the Final Study conditions exhibited graded activity such that the response to items that had been neither tested nor restudied was greater than the response to items that had been restudied, whereas tested items showed intermediate levels of activity (Table 2). This general pattern of neural activity is reminiscent of the decrease in activity typically observed in repetition priming experiments. Considering that the cue and target are fully absent, fully present, or

**Table 2. Statistical comparison (paired  $t$  test) of activity at final study between each of the Phase 2 conditions (tested, restudied, or neither tested nor restudied)**

Region	$t$ -statistic	$z$ -score	$p$ -value
<b>Tested versus restudied</b>			
Anterior PFC	1.98	1.88	0.06
Dorsolateral PFC	2.55	2.36	0.02
PIPL/dorsal AG	3.90	3.38	<0.001
Precuneus	3.73	3.27	0.001
Middle cingulate	2.87	2.62	0.01
<b>Tested versus neither tested nor restudied</b>			
Anterior PFC	−1.87	−1.79	0.07
Dorsolateral PFC	−2.95	−2.69	0.01
PIPL/dorsal AG	4.63	3.85	<0.001
Precuneus	6.44	4.82	<0.001
Middle cingulate	3.67	3.22	0.001
<b>Restudied versus neither tested nor restudied</b>			
Anterior PFC	−3.61	−3.19	0.001
Dorsolateral PFC	−4.57	−3.82	<0.001
PIPL/dorsal AG	1.65	1.59	0.11
Precuneus	3.66	3.21	0.001
Middle cingulate	2.01	1.92	0.06

Corresponding data are plotted in Figure 3. Region coordinates ( $x, y, z$ ) in MNI space: anterior PFC (−47, 37, −3); dorsolateral PFC (−49, 12, 25); PIPL/dorsal AG (−44, −56, 46); precuneus (−8, −73, 36); middle cingulate (−1, −24, 33).

partially present (cue only at test) during Phase 2 in each of these respective conditions, a repetition suppression account fits the data quite well. Even though activity was statistically significantly different between all item types at Final Study in the left dlPFC (Fig. 3A), activity for the tested items was intermediate to the items that were restudied and those that were



**Figure 4.** *A*, Conjunction map displaying voxels that show a difference between hits and correct rejections (CR) in between 5 (dark purple) and 8 (light purple) recognition memory studies are shown on an inflated cortical surface rendering using Caret software (Van Essen et al., 2001). Also shown in yellow is a 10-mm-diameter sphere representing the region defined in the current experiment as the left pIPL/dorsal AG. The region is also shown in coronal ( $y = -56$ ) and transverse ( $z = 44$ ) slices of an average anatomical volume. *B*, Time courses of activity in left pIPL/dorsal AG are shown for the tested items at Final Study and Initial Study (solid lines), as well as the average response for hits and correct rejections from the meta-analysis (dotted lines, see Materials and Methods).

neither tested nor restudied. Therefore, the pattern of activity in left dlPFC appears to be a function of repetition and not retrieval practice per se.

The second pattern showed greater activity at Final Study relative to Initial Study (Fig. 3C–E) and occurred in left lateral and medial parietal cortex. That is, whereas prefrontal regions showed repetition suppression, parietal regions exhibited a pattern of repetition enhancement. These parietal regions were also most active during Final Study for items that had been tested compared with those that had been restudied and neither tested nor restudied (Table 2). Therefore, unlike the regions in left prefrontal cortex, the parietal regions all exhibited a pattern of activity that suggests specific sensitivity to prior retrieval practice.

#### Overlap of time course activity in LLPC with studies of recognition memory

Thus far, we have identified regions in parietal cortex that are sensitive to retrieval practice. What can we determine about the contribution of these regions in the context of test-potentiated learning? To answer this question, we go beyond the current experiment and consider memory retrieval studies that have focused on the role of parietal cortex. In particular, LLPC has featured prominently in recognition memory, in that it consistently activates more robustly when subjects correctly recognize old (or previously studied) items than when they correctly identify new items as such (i.e., greater activity for hits than correctly rejected lures). A recent meta-analysis (Nelson et al., 2010) identified a specific region in LLPC, referred to as left posterior inferior parietal lobule (pIPL)/dorsal angular gyrus (AG), as among the most reliable showing retrieval-related activity (McDermott et al., 2009). Interestingly, the region identified in the current experiment was located within this area of high reliability (Fig. 4A). Therefore, a tentative hypothesis is that similar types of cognitive processes occur during Final Study and during explicit recognition decisions. More specifically, might subjects be reminded of the prior test episode for items that had been previously tested and this covert reminder is the source of a similar activation pattern?

Of course, the presence of anatomic overlap, no matter how precise, is only suggestive. Stronger claims may be made by looking more closely at the activity patterns exhibited by this region in the current experiment and across recognition memory experi-

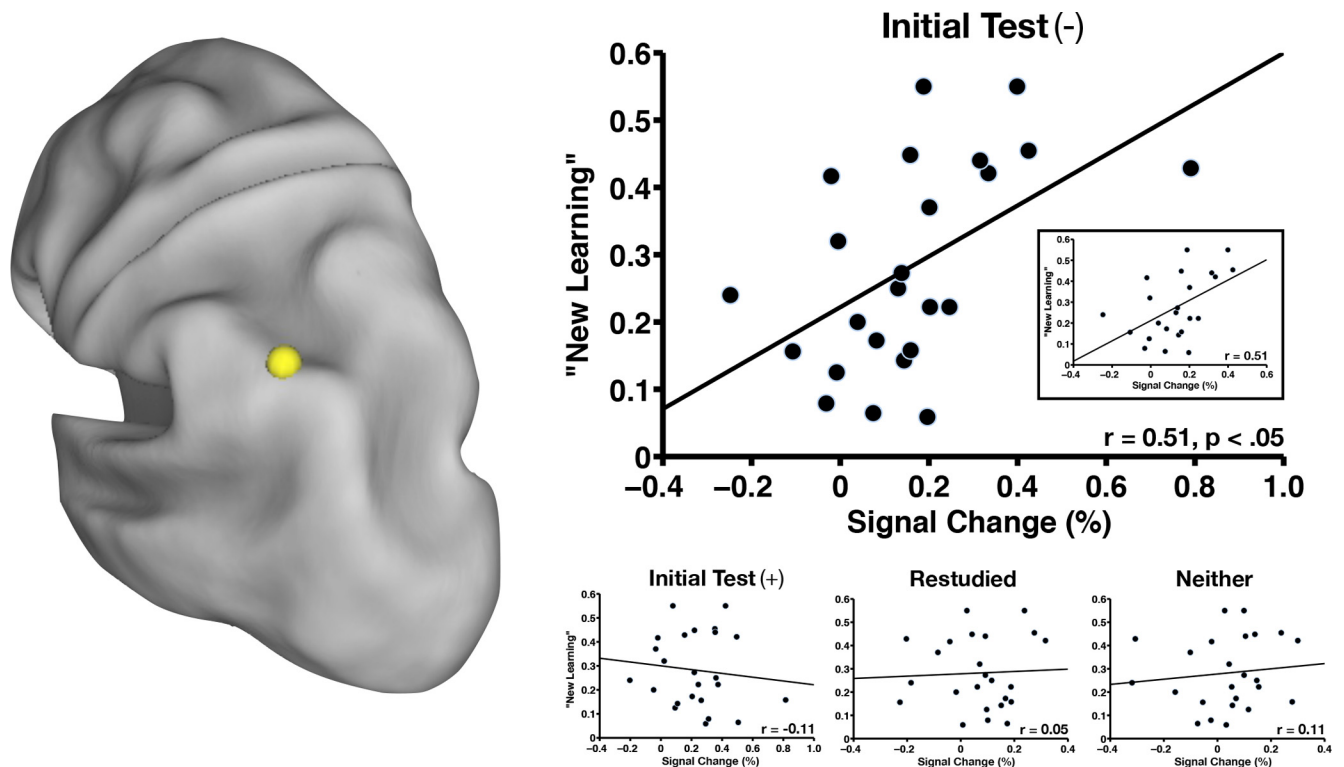
ments. When we compare time course profiles within left pIPL/dorsal AG in both the current study and the eight recognition studies that comprised the meta-analysis, they are virtually identical to one another (Fig. 4B). Items correctly identified as being “old” (hits) in recognition memory experiments produce a strikingly similar time course of activity as items that are being studied after testing. In the latter case, subjects are simply told to study the items for an upcoming test and are in no way encouraged to retrieve information about prior study or test episodes. In addition, the time course of activity for items correctly identified as “new” (correct rejections) looks identical to the time course shown during Initial Study. “New” items are similar to word pairs presented at Initial Study because the item had not yet been encountered and thus there is no episodic content to retrieve.

#### Neural activity in left pIPL/dorsal AG is correlated with ‘new learning’

The left pIPL/dorsal AG region appears to be sensitive to retrieval practice and activity present there at Final Study may reflect the engagement of retrieval processes without explicit task demands. However, although testing may have potentiated activity in this region, we have not yet shown evidence of learning as indexed by behavior. To determine whether left pIPL/dorsal AG may play a role in test-potentiated learning, we calculated the relationship between amount of New Learning (see “Behavioral results”) and activity in left pIPL/dorsal AG during Final Study for items that had not been recalled on the prior test and found a significant relationship (Fig. 5A;  $r = 0.51$ ,  $p < 0.05$ ). Therefore, subjects who learned a greater proportion of items after the Initial Test showed greater activity in left pIPL/dorsal AG for items not initially recalled. This relationship was not present for items that had been recalled during the Initial Test ( $r = -0.11$ ,  $p = \text{NS}$ ), items that were restudied ( $r = 0.05$ ,  $p = \text{NS}$ ), or items that were neither tested nor restudied ( $r = 0.11$ ,  $p = \text{NS}$ ; Fig. 5A), suggesting that this correlation is not a global feature of activity in this region and recall on the final test. Rather, it seems specific to New Learning after an unsuccessful retrieval attempt.

#### Discussion

The present study investigated the neural correlates of the effects of testing on subsequent study and found that a region in left



**Figure 5.** The region in left pIPL/dorsal AG (yellow sphere) is shown along with a scatterplot of New Learning as a function of percent signal change across subjects ( $n = 24$ ) for items that were *not* initially recalled (Initial Test (-)). A statistically significant relationship exists between the two variables ( $r = 0.51$ ,  $p < 0.05$ ). The inset shows the same plot with removal of an outlier in which the percentage signal change was greater than 3 SDs from the mean (far right data point on main scatterplot). The correlation coefficient does not change with removal of the outlier ( $r = 0.51$ ,  $p < 0.05$ ). New Learning is also plotted as a function of percent signal change across subjects ( $n = 24$ ) for items that were initially recalled (Initial Test (+)), items that were restudied, and items that were neither tested nor restudied. There is no statistically significant relationship between the two variables in any of these three plots (bottom right).

pIPL/dorsal AG may play an important role in test-potentiated learning. These results relate to theories in both cognitive psychology and cognitive neuroscience. We start by discussing a potential role for “reminders” (Hintzman, 1974) in test-potentiated learning. We then discuss the implications of our findings for theories that attempt to explain parietal lobe contributions to memory retrieval. Finally, we end by mentioning the ways in which reverse inference can benefit our understanding of brain and behavior.

#### Reminders framework and the effects of testing on subsequent study

The imaging data presented here, specifically in left pIPL/dorsal AG, suggest a mechanistic account of the effects of retrieval practice on subsequent encoding. How does taking a test alter the way in which material is studied? At least a partial answer to this question appears to arise from the concept of reminders or “study-phase retrieval” (Thios and D’Agostino, 1976; Greene, 1989; Benjamin and Tullis, 2010; Hintzman, 2011; Wahlheim and Jacoby, 2013), which has been used to explain the spacing effect (Hintzman, 1974)—the finding that items are better remembered when study trials are spaced in time rather than massed. Subsequent spaced presentations of items may enhance long-term retention because they “allow active retrieval of old information stored during the initial presentation” (Thios and D’Agostino, 1976). It is important to note that this form of active retrieval need not be the result of an explicit intent to retrieve before stimulus onset, but rather may happen in a more involuntary or “bottom-up” manner when the stimulus is presented (Greene, 1989). Therefore, the present data fit nicely with the

theory because subjects are not likely adopting a retrieval “task set” during Final Study or engaging in some form of retrieval mode as a result of the mixed presentation of the stimuli with regard to the different Phase 2 conditions. Instead, retrieval occurs as the subject studies the word pair and is reminded of the previous retrieval attempt. The subject is then able to integrate the current study episode with the prior test episode, perhaps determining a more optimal strategy to learn the item if that item was not correctly recalled or reinforcing the context that was present for correctly recalled items during Phase 2. Although the left pIPL/dorsal AG showed greater activity for tested items at Final Study than either class of untested items, it also showed significantly greater activity at Final Study than Initial Study for untested items ( $p < 0.05$  for both restudied items and items that were neither tested nor restudied). Therefore, if activity in left pIPL/dorsal AG is in some way indexing the degree to which the subject is reminded of a prior encounter with an item, this reminding is still present in the absence of testing, but we would argue to a much lesser degree. This is important but not necessarily surprising given that the untested items are being studied in a spaced manner, which, as noted previously, can encourage reminders.

#### Theories of parietal cortex and memory retrieval

The pattern of activity in left pIPL/dorsal AG is consistent with a recent model proposed by Jaeger et al. (2013) suggesting that this region may play a role in involuntary orienting when a stimulus is unexpectedly familiar. The “memory orienting model” (Jaeger et al., 2013) is the most recent in a set of theories that propose to explain how the parietal cortex contributes to memory retrieval

(Wagner et al., 2005). In addition to the convergence of the memory orienting model with our data at a process level (i.e., orienting), the specific region of LLPC is in an almost identical location to the region we defined here in left pIPL/dorsal AG. Therefore, there is functional and anatomic overlap, both of which are crucial if data and the explanations that arise from those data are going to mutually inform one another. A previous model proposed by Cabeza et al. (2008) dubbed the “attention to memory model” has some conceptual similarities to the memory orienting model in that it posits a role for parietal cortex in “bottom-up” attention. However, the locations within ventral parietal cortex that Cabeza et al. (2008) refer to, most notably the supramarginal gyrus and AG, are distinct from left pIPL/dorsal AG as emerging from the current dataset, as well as the region defined by Jaeger et al. (2013) and Nelson et al. (2010). Therefore, the attention to memory model is not anatomically relevant to the data presented here.

### Utility of reverse inference

In recent years, the topic of reverse inference has garnered considerable attention in cognitive neuroscience (Poldrack, 2006, 2011; Moran and Zaki, 2013). Although the presence of activity in a particular brain region cannot be taken as definitive evidence for the engagement of a specific cognitive process, there are certain boundary conditions in which reverse inference is a reasonable means of hypothesis generation. We posit that the data presented here, which led us to a reminders framework for understanding test-potentiated learning, represent an instance in which arguing from cognitive function (i.e., memory retrieval) based on the presence of activity in the brain (i.e., left pIPL/dorsal AG) is profitable. What are the conditions in the current experiment that make reverse inference viable? The selectivity of the region in LLPC and its established distinctiveness from more ventral regions in AG and more dorsal regions in the intraparietal sulcus is one reason. From an anatomical standpoint, the region is present in approximately the same location (~7 mm apart in Euclidean distance) as a putative area defined in Nelson et al. (2010) in a small strip of cortex that shows the most reliable retrieval-related activity in LLPC (Fig. 4A). In addition, the similarity between the left pIPL/dorsal AG time courses in the present study and those from the meta-analysis provides considerable leverage. Indeed, had the effect of Final Study versus Initial Study been in the same direction as that for hits and correct rejections with very different responses both temporally and with respect to baseline, our conclusion would be much weaker or untenable. In the end, given the paucity of research that has used fMRI to understand behavioral phenomena related to testing, this is a case in which neuroimaging can be used to generate hypotheses in an emerging subfield.

### Future directions

Going forward, it will be important to address whether other tasks interposed between multiple study opportunities might give rise to a similar increase in activity in left pIPL/dorsal AG. For example, if subjects engaged in a different task during Phase 2 (e.g., a standard semantic task such as judging whether the cue and target were both living or nonliving), would this also encourage reminders over and above restudying? If so, then perhaps engaging the items in various contexts (intentional encoding during study and incidental encoding during a semantic judgment) is why testing results in greater activity in parietal cortex. In other words, testing may be one example of such an alternate context. However, if tested items still showed greater activity dur-

ing Final Study than items for which subjects made a semantic judgment, then “context change” cannot be the only explanation for the effect of testing on subsequent study.

### Conclusions

Broadly speaking, the data presented here converge on the conclusion that testing can indeed affect the way in which information is processed during a subsequent study episode. In a situation with no explicit behavioral index, such as restudying, fMRI can be used to extract signals that reflect how the brain is processing information. A critical point, however, is that we were not only able to identify activity related to the effects of retrieval practice on subsequent study, but we could also localize that activity to a specific region in left pIPL/dorsal AG. Because this region features prominently in theories of recognition memory, we could then leverage understanding of this region to inform how test-potentiated learning may emerge. The data presented here suggest that a mechanism by which retrieval practice facilitates subsequent encoding is via engagement of retrieval processes during the subsequent study phase. In addition, although the results from our experiment are not immediately translatable into effective practices for learning in educational settings, we think they provide an entry into how the nascent field of neuroeducation (Varma et al., 2008; Carew and Magsamen, 2010) might one day help to create more effective teaching strategies by informing theories of cognitive psychology. However, we echo Bruer’s sentiment (Bruer, 1997) that any cross talk between neuroscience and education must be done through the language and frameworks developed in cognitive psychology. In the present study, the concept of reminders provides such a background.

### References

- Abbott EE (1909) On the analysis of the factors of recall in the learning process. *Psychological Monographs* 11:159–177. [CrossRef](#)
- Arnold KM, McDermott KB (2013a) Free recall enhances subsequent learning. *Psychon Bull Rev* 20:507–513. [CrossRef Medline](#)
- Arnold KM, McDermott KB (2013b) Test-potentiated learning: distinguishing between direct and indirect effects of tests. *J Exp Psychol Learn Mem Cogn* 39:940–945. [CrossRef Medline](#)
- Benjamin AS, Tullis J (2010) What makes distributed practice effective? *Cogn Psychol* 61:228–247. [CrossRef Medline](#)
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221. [Medline](#)
- Bradshaw G, Anderson J (1982) Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Memory* 21:165–174. [CrossRef](#)
- Brewer JB, Zhao Z, Desmond JE, Glover GH, Gabrieli JD (1998) Making memories: Brain activity that predicts how well visual experience will be remembered. *Science* 281:1185–1187. [CrossRef Medline](#)
- Bruer J (1997) Education and the brain: a bridge too far. *Educational Researcher* 26:4–16.
- Buckner RL, Koutstaal W, Schacter DL, Rosen BR (2000) Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain* 123:620–640. [CrossRef Medline](#)
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 9:613–625. [CrossRef Medline](#)
- Carew TJ, Magsamen SH (2010) Neuroscience and education: an ideal partnership for producing evidence-based solutions to guide 21st century learning. *Neuron* 67:685–688. [CrossRef Medline](#)
- Carrier M, Pashler H (1992) The influence of retrieval on retention. *Mem Cognit* 20:633–642. [CrossRef Medline](#)
- Cohen JD, MacWhinney B, Flatt M, Provost J (1993) PsyScope: a new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments and Computers* 25:257–271. [CrossRef](#)
- Craik FMI, Tulving E (1975) Depth of processing and the retention of words



- in recognition memory. *Journal of Experimental Psychology: General* 82:472–481.
- Donaldson DI, Wheeler ME, Petersen SE (2010) Remember the source: dissociating frontal and parietal contributions to episodic memory. *J Cogn Neurosci* 22:377–391. [CrossRef Medline](#)
- Friston K, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1:153–171. [CrossRef](#)
- Gates AI (1917) Recitation as a factor in memorizing. *Archives of Psychology* 6:1–104.
- Glover JA (1989) The “testing” phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology* 81:392–399. [CrossRef](#)
- Greene RL (1989) Spacing effects in memory: evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:371–377. [CrossRef](#)
- Hintzman D (2011) Research strategy in the study of memory: fads, fallacies, and the search for the “coordinates of truth.” *Perspect Psychol Sci* 6:253–271. [CrossRef](#)
- Hintzman DL (1974) Theoretical implications of the spacing effect. In: *Theories in cognitive psychology, The Loyola symposium* (Solso RL, ed), pp 77–99. Hillsdale, NJ: Erlbaum.
- Izawa C (1971) The test trial potentiating model. *Journal of Mathematical Psychology* 8:200–224. [CrossRef](#)
- Jaeger A, Konkel A, Dobbins IG (2013) Unexpected novelty and familiarity orienting responses in lateral parietal cortex during recognition judgment. *Neuropsychologia* 51:1061–1076. [CrossRef Medline](#)
- Jang Y, Wixted J, Pecher D, Zeelenberg R, Huber D (2012) Decomposing the interaction between retention interval and study/test practice: the role of retrievability. *Q J Exp Psychol (Hove)* 65:962–975. [CrossRef Medline](#)
- Kang SHK, McDermott KB, Roediger HL III (2007) Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology* 19:528–558. [CrossRef](#)
- Karpicke JD, Roediger HL 3rd (2008) The critical importance of retrieval for learning. *Science* 319:966–968. [CrossRef Medline](#)
- Kučera H, Francis W (1967) *Computational analysis of present-day American English*. Providence, RI: Brown UP.
- Lancaster JL, Glass TG, Lankipalli BR, Downs H, Mayberg H, Fox PT (1995) A modality-independent approach to spatial normalization of tomographic images of the human brain. *Human Brain Mapping* 3:209–223. [CrossRef](#)
- Larsen DP, Butler AC, Roediger HL 3rd (2008) Test-enhanced learning in medical education. *Med Educ* 42:959–966. [CrossRef Medline](#)
- McAvoy MP, Ollinger JM, Buckner RL (2001) Cluster size thresholds for assessment of significant activation in fMRI. *Neuroimage* 13:S198.
- McDaniel MA, Agarwal PK, Huelser BJ, McDermott KB, Roediger HL (2011) Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology* 103:399–414. [CrossRef](#)
- McDermott KB, Szpunar KK, Christ SE (2009) Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia* 47:2290–2298. [CrossRef Medline](#)
- Michelon P, Snyder AZ, Buckner RL, McAvoy M, Zacks JM (2003) Neural correlates of incongruous visual information. An event-related fMRI study. *Neuroimage* 19:1612–1626. [CrossRef Medline](#)
- Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000) Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11:735–759. [CrossRef Medline](#)
- Moran JM, Zaki J (2013) Functional neuroimaging and psychology: what have you done for me lately? *J Cogn Neurosci* 25:834–842. [CrossRef Medline](#)
- Mugler JP 3rd, Brookeman JR (1990) Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn Reson Med* 15:152–157. [CrossRef Medline](#)
- Nelson DL, McEvoy CL, Schreiber TA (2004) The University of South Florida free association, rhyme, and word fragment norms. *Behav Res Methods Instrum Comput* 36:402–407. [CrossRef Medline](#)
- Nelson SM, Cohen AL, Power JD, Wig GS, Miezin FM, Wheeler ME, Velanova K, Donaldson DI, Phillips JS, Schlaggar BL, Petersen SE (2010) A parcellation scheme for human left lateral parietal cortex. *Neuron* 67:156–170. [CrossRef Medline](#)
- Ojemann JG, Akbudak E, Snyder AZ, McKinstry RC, Raichle ME, Conturo TE (1997) Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage* 6:156–167. [CrossRef Medline](#)
- Ollinger JM, Shulman GL, Corbetta M (2001) Separating processes within a trial in event-related functional MRI. *Neuroimage* 13:210–217. [CrossRef Medline](#)
- Phillips JS, Velanova K, Wolk DA, Wheeler ME (2009) Left posterior parietal cortex participates in both task preparation and episodic retrieval. *Neuroimage* 46:1209–1221. [CrossRef Medline](#)
- Poldrack RA (2006) Can cognitive processes be inferred from neuroimaging data? *Trends in Cogn Sci* 10:59–63. [CrossRef Medline](#)
- Poldrack RA (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* 72:692–697. [CrossRef Medline](#)
- Roediger HL 3rd, Butler AC (2011) The critical role of retrieval practice in long-term retention. *Trends Cogn Sci* 15:20–27. [CrossRef Medline](#)
- Roediger HL, Karpicke JD (2006) Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci* 17:249–255. [CrossRef Medline](#)
- Rohrer D, Taylor K, Sholar B (2010) Tests enhance the transfer of learning. *J Exp Psychol Learn Mem Cogn* 36:233–239. [CrossRef Medline](#)
- Schacter DL, Wig GS, Stevens WD (2007) Reductions in cortical activity during priming. *Curr Opin Neurobiol* 17:171–176. [CrossRef Medline](#)
- Snyder AZ (1996) Difference image vs. ratio image error function forms in PET-PET realignment. In: *Quantification of brain function using PET* (Myer R, Cunningham VJ, Bailey DL, Jones T, eds), pp 131–137. San Diego: Academic.
- Spitzer HF (1939) Studies in retention. *Journal of Educational Psychology* 30:641–656. [CrossRef](#)
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.
- Thios SJ, D’Agostino PR (1976) Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior* 15:529–536. [CrossRef](#)
- Tulving E (1967) The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior* 6:175–184. [CrossRef](#)
- Turk-Browne NB, Yi DJ, Chun MM (2006) Linking implicit and explicit memory: common encoding factors and shared representations. *Neuron* 49:917–927. [CrossRef Medline](#)
- Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001) An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc* 8:443–459. [CrossRef Medline](#)
- Varma S, McCandliss BD, Schwartz DL (2008) Scientific and pragmatic challenges for bridging education and neuroscience. *Educational Researcher* 37:140–152. [CrossRef](#)
- Velanova K, Jacoby LL, Wheeler ME, McAvoy MP, Petersen SE, Buckner RL (2003) Functional-anatomic correlates of sustained and transient processing components engaged during controlled retrieval. *J Neurosci* 23:8460–8470. [Medline](#)
- Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998) Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281:1188–1191. [CrossRef Medline](#)
- Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005) Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* 9:445–453. [CrossRef Medline](#)
- Wahlheim CN, Jacoby LL (2013) Remembering change: the critical role of recursive reminders in proactive effects of memory. *Mem Cognit* 41:1–15. [CrossRef Medline](#)
- Wheeler ME, Buckner RL (2003) Functional dissociation among components of remembering: control, perceived oldness, and content. *J Neurosci* 23:3869–3880. [Medline](#)
- Wheeler ME, Buckner RL (2004) Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337–1349. [CrossRef Medline](#)
- Xue G, Mei L, Chen C, Lu ZL, Poldrack R, Dong Q (2011) Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *J Cogn Neurosci* 23:1624–1633. [CrossRef Medline](#)