Neural Signatures of Test-Depressed Encoding: Dynamic Modulations in the Memory Encoding Network and Anterior Cingulate Cortex

Jian Zhu, Pamela J. LaMontagne, and Reza Habib

Southern Illinois University Carbondale

Author Note

Jian Zhu, Pamela J. LaMontagne, and Reza Habib, Department of Psychology, Southern Illinois University Carbondale, Carbondale, IL.

Jian Zhu is now at Department of Psychiatry, University of Wisconsin Madison, Madison, WI, USA. Pamela J. LaMontagne is now at Department of Radiology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA.

Jian Zhu processed the data, analyzed the results, and prepared the manuscript. Pamela J. LaMontagne collected and processed the data. Reza Habib designed the experiment and supervised the research. All authors contributed to the writing and have read and approved the final manuscript.

Correspondence concerning this article should be addressed to Reza Habib, Department of Psychology, Southern Illinois University Carbondale, Carbondale, IL 62901. Email: rhabib@siu.edu

The authors declared no conflicts of interest.

The authors wish to thank Natalie Bryant and Timothy Callahan for assistance with data collection and Julie Alstat and Gary Etherton for assistance with MRI scanning.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Although previous behavioral studies have demonstrated that restudying already-learned items is mnemonically ineffective, the neural mechanisms underlying this phenomenon remain largely unknown. Here, functional magnetic resonance imaging (fMRI) was used to assess neural changes during a multi-trial associative learning task where participants repeatedly studied and were tested on unrelated English word pairs. The results revealed that the brain network involved in the initial encoding of the word pairs was less activated when participants restudied previously learned ones, supporting a pattern that we refer to as *test-depressed encoding* (TDE). More distinctively, the dorsal anterior cingulate cortex (dACC) showed greater activation when restudying learned vs. unlearned word pairs, suggesting dACC-mediated cognitive control of TDE. Functional connectivity analysis further consistently showed that the memory encoding network had a weakened neural coupling within itself, but an enhanced coupling with dACC under TDE. Together, the present work is the first to demonstrate that these dynamic modulations in regional activity and connectivity may reflect the neural foundation for TDE.

Keywords: encoding, testing, repeated study, subsequent memory, fMRI

Significance Statement

The current work aims to elucidate the largely unknown neural mechanisms underlying the behavioral observation that repeatedly studying previously learned information is ineffective to improve memory retention. The neuroimaging findings from our report may help inform optimal educational practice for teaching and learning.

Neural Signatures of Test-Depressed Encoding: Dynamic Modulations in the Memory Encoding Network and Anterior Cingulate Cortex

In real-world situations, it is not uncommon for learners to engage in repeated studying and testing, perhaps in the form of quizzes, in order to improve their memory for to-be-learned information. A retrieval attempt (i.e. the quiz) in this learning sequence, even if unsuccessful, has been shown to benefit memory encoding on a subsequent study of that information (Arnold & McDermott, 2013a, 2013b), a phenomenon first identified by Izawa (1966) and known as *test-potentiated encoding* (TPE; Vestergren & Nyberg, 2014).

Examining the neural correlates of TPE, X. L. Liu, Liang, Li, and Reder (2014) observed that the left prefrontal cortex and left hippocampus tended to be activated during immediate feedback/restudy of previously unrecalled items if these items were successfully recalled later. Similarly, Vestergren and Nyberg (2014) observed that activity in the anterior insula signified TPE when unrecalled items were successfully recalled after a restudy opportunity. Together these studies indicate that regions typically associated with encoding of information into long-term memory (Kim, 2011)—including left inferior frontal gyrus / insula and hippocampus—demonstrate greater brain activity when previously unrecalled items are learned as a result of subsequent re-exposure (TPE; see Nelson, Arnold, Gilmore, & McDermott, 2013 and later Discussion for a possible role of retrieval processes in TPE).

The flipside of TPE concerns the fate of items that were successfully retrieved on a prior test. Would these already-learned items gain any further benefit from restudy as well? Two separate lines of research converge to suggest that they do not (cf. Jang et al., 2012). First, Pashler, Cepeda, Wixted, and Rohrer (2005) and Fazio, Huelser, Johnson, and Marsh (2010) showed that providing the correct answer as feedback immediately following an incorrect recall

of initially studied materials improved later memory (TPE), but feedback following a correct recall did not convey any additional information to the learners such that final memory performance was the same with or without feedback. Second, Karpicke and Roediger (2008) and Soderstrom, Kerr, and Bjork (2016) systematically varied items in an interpolated study-test learning sequence (i.e, STST…) and demonstrated that once an item had been successfully recalled, repeated studies did not improve its final retention when it continued to be tested. In sum, these results suggest that successfully retrieved items do not benefit from TPE the same way that unrecalled items do.

This lack of appreciable memory improvement observed for a learned item raises a question regarding the encoding processes during its restudy. It is theoretically possible that the learner still engages in active learning during the restudy, but the strength of that item's mnemonic representation has reached a ceiling and could therefore not be enhanced further. On the other hand, the learner may exert less effort and fewer cognitive resources on its restudy and therefore activity related to its encoding is actually decreased given that the item has already been learned (and likely remains available for retrieval on subsequent tests). We refer to the latter scenario as *test-depressed encoding* (TDE). It would be impossible to contrast these alternative explanations with purely behavioral approaches because both accounts would lead to the same behavioral outcome: a null memory improvement for items that have been learned on a previous trial.

Thus, the purpose of the present study was to use fMRI to elucidate the encoding-related neural patterns during the restudy of previously unrecalled and recalled information. To this end, we adopted a multi-trial associative learning paradigm in which word pairs were repeatedly studied and tested in order to accumulate sufficient trials of successfully and unsuccessfully

retrieved items for analysis. Neural activity for these classes of items was then examined on subsequent restudy trials. Based on prior research, we hypothesized that the typical prefronto-temporo-parietal memory encoding network (for reviews, see Moscovitch, Cabeza, Winocur, & Nadel, 2016; Ofen, 2012; Paller & Wagner, 2002; Rugg, Johnson, & Uncapher, 2015) would be activated for items that were newly learned (TPE). More importantly, if less effort and resources are allocated during subsequent restudy episodes when items were previously successfully recalled, (parts of) the same encoding network may not be activated as strongly. To foreshadow the findings, our data supported the pattern of TDE in memory encoding brain networks. Additionally, we found that the engagement of dorsal anterior cingulate cortex (dACC) differentiated between TPE and TDE, and that dACC exerted cognitive control during TDE via dynamic changes in functional connectivity with the frontal and temporal encoding regions. Together, our results provide a neural signature for TDE.

Method

Participants

Eleven healthy right-handed volunteers (aged 21-34 years; $M = 25$; 6 males) each received a $30 gift card as remuneration for participating in the study. All participants were native English-speakers with normal or corrected-to-normal vision and hearing. The study was approved by the Human Subjects Committee of Southern Illinois University Carbondale. Informed consent was obtained from all participants prior to their participation in the study.

Procedure

The experiment consisted of a series of study-test trials. MRI scans were acquired during the *study* phases only. During the study phase, participants were presented with a series of 40 unrelated randomly paired words (e.g., SISTER - ESSAY). Each word pair was shown using 26-

point Arial font in white letters on a black background on the center of an MRI-compatible LCD screen (IFIS-SA; Invivo, Orlando, FL), which was attached to the back of a standard MRI head-coil. Participants viewed the LCD screen via a mirror placed directly above their head. Word pairs were shown sequentially for 2.5 s and were separated by a fixation cross that was shown on the center of the screen for 1.5 s. To introduce jitter in the sampling of the hemodynamic response curve (HRF), 20 null events (a fixation cross, 2.5 s in duration) were randomly intermixed with the word pairs. Participants were instructed to read to themselves each word pair and decide whether both words were shown in upper or lower case letters by pressing one of two buttons on an MRI compatible response pad with their right index and middle fingers. Half of the word pairs were shown in upper case and the other half were shown in lower case. Shallow encoding instructions were adopted here to slow down the rate at which participants learned the word pairs such that sufficient numbers of items could be obtained in each condition for the statistical analyses. Each study phase lasted 4 min and 20 s.

The test phase followed each study trial after a brief (3-5 min) rest and occurred without MRI scanning. Participants were shown the same set of 40 word pairs (in a different random order) as they had seen during the prior study phase. For each pair, the right-hand word was replaced with a question mark (e.g., SISTER - ?). The presentation rate was self-paced and stimuli thus remained on the screen until participants provided a response. A fixation cross was presented for 1.5 s between each stimulus item. For each item, participants were instructed to say aloud the right-hand word that was associated with each left-hand cue. If they could not remember the right-hand word, they were instructed to say “pass”. An MRI-compatible microphone attached to the head coil transmitted participants’ responses to the control room where it was recorded by an experimenter. No feedback was provided to participants. Although

participants were asked to provide a confidence judgment on a 5-point scale regarding the accuracy of their response following a response other than "pass", all of them did not comply. Therefore, this dimension of the data was not considered here.

This study-test procedure was repeated a minimum of 5 times for each participant and a maximum of 7 times until they learned all or the majority (≥ 85%) of the word pairs (see the Behavioral Results for details). The presentation order of the word pairs was randomized on each study and test trial. At the end of the experiment, participants were compensated and debriefed.

Scanning and Analysis

FMRI imaging was performed with a Philips Intera 1.5 T magnet using a standard head coil. Each functional run consisted of 96 contiguous whole brain volumes (T2* single-shot EPI, TR = 2.5 s, TE = 50 ms, flip angle = 90°, FOV = 220 x 220 mm^2, 64 x 64 matrix, 3.44 x 3.44 x 5.5 mm^3 voxels, 26 x 5.5 mm axial slices in an interleaved fashion, 0 mm gap, first 8 images were discarded). Conventional high-resolution T1 weighted 3-D structural images (256 x 256 matrix, 200 slices) were acquired prior to the start of the functional imaging stage.

Data were analyzed with SPM 12 (http://www.fil.ion.ucl.ac.uk/spm/) implemented in Matlab (Mathworks, Natick, MA) environment. Images were 1) slice time corrected for acquisition order, 2) realigned and motion corrected to the first image of the session, 3) normalized to the Montreal Neurological Institute (MNI) space involving a coregistration and unified segmentation of individual participant's structural image, 4) resampled to 2 x 2 x 2 mm^3 voxels, and 5) spatially smoothed with an isotropic Gaussian kernel of 10 mm full width at half maximum (FWHM). A 128 s high-pass filter was applied to each time course in order to eliminate low-frequency noises. Single-participant statistical contrasts were created using the

general linear model (GLM). Conditions of interest (E00, E01, E11, see below) were modeled using a canonical hemodynamic response function (HRF). Regressors for motion parameters were also included in the GLM model. Group comparisons were created using a random effects model. All unmasked contrasts were thresholded at $p_{unc} < .001$ and k (extent) ≥ 10 voxels. For a masking analysis, the mask itself was thresholded at $p_{unc} < .01$, and a threshold of $p_{unc} < .005$ and $k \geq 10$ was used for the contrast. All coordinates were reported in the MNI coordinate system.

An additional functional connectivity analysis was conducted to explore functional coupling among the activated brain regions. For this analysis, a separate GLM model was constructed in which each trial was entered as a separate regressor (together with motion parameters as in the earlier traditional GLM), yielding beta estimates for each individual trial of each condition within each participant (i.e., beta series analysis, see Rissman, Gazzaley, & D'Esposito, 2004). For each condition, the beta series were extracted from and averaged over voxels within identified brain regions separately, and outlier betas (defined as 3 SD away from the mean, in practice less than an average of 3% betas, range 0-8%) were rejected. Then, pairwise correlations were calculated to represent region-region functional connectivity and tested against zero using robust one-sample *t*-tests (Wilcox, 2012). Finally, paired *t*-tests with FDR correction (Storey, 2002) were performed to compare condition difference in corresponding region-region connectivity. All the operations were implemented in an SPM toolbox, BASCO (https://www.nitrc.org/projects/basco/) (Göttlich, Beyer, & Krämer, 2015) with homebrew customization.

With the repeated study-test sequence (…$S_{n-1}T_{n-1}S_nT_n$…), encoding activity (S_n) on study trials 2 through 7 were classified into four categories based on memory performance on consecutive pairs of cued-recall trials 1 through 7 (T_{n-1}, T_n; see Figure 1 for a schematic

illustration and note that only study trials were scanned): 1) E11 items were recalled on both the immediately preceding test trial (T_{n-1}) and the subsequent test trial (T_n), b) E01 items were not recalled on the immediately preceding test trial (T_{n-1}) but were recalled on the subsequent test trial (T_n), c) E10 items were recalled on the immediately preceding test trial (T_{n-1}) but not on the subsequent test trial (T_n), and d) E00 items were recalled on neither test trials (T_{n-1}, T_n). In this designation, E refers to classified encoding items on the study trial S_n, the first 0/1 refers to memory performance on retrieval trial T_{n-1} and the second 0/1 refers to memory performance on retrieval trial T_n, with a 0 indicating that an item was not retrieved and a 1 indicating that the item was successfully retrieved. In practice, few items (an average of 1.45 pairs across all trials per participant, see Figure 2) fell into the E10 category and were consequently not included in the fMRI modeling.

The comparison between E01 and E00, similar to the standard single-trial subsequent memory comparison (Paller & Wagner, 2002; van den Broek et al., 2016), reveals the encoding areas that contribute to the successful recall on the subsequent test (presumably, greater TPE for E01 items than for E00 items induced by previous unsuccessful retrieval attempts). The comparison between E01 and E11 permits an examination of activity during learning that is conditional on whether items have been previously retrieved, with E01 reflecting TPE and E11 reflecting TDE. That is, the comparison between E01 and E11 should identify brain regions/circuits that are activated during encoding for previously unlearned items (E01) or previously learned items (E11).

Results

Behavioral Data

Across the 11 participants, a minimum of 5 and a maximum of 7 study-test trials were presented (M = 6 trials). By the 5th study-test cycle, six participants had learned all 40 word pairs, two participants required a 6th trial to learn all 40 word pairs, and the remaining three participants learned on average 36 of the 40 words-pairs towards their last trial. See Figure 2 also for the dynamic changes in the number of items falling into each of the four operationalized conditions (E00, E01, E11, E10). On average, each participant had 45.82 (±39.57) E00 items, 37.64 (±4.03) E01 items, 115.09 (±34.02) E11 items, and 1.45 (±1.63) E10 items across all trials.

Overall, performance on the cued recall test improved with each subsequent repeated study of the word pairs, reaching an asymptote on the 5th trial (see Figure 2). Bonferroni-corrected paired t-tests were carried out on consecutive study-test trials. Results indicated that significant improvements in performance occurred between consecutive trials up to the 5th trial, at which point performance did not significantly improve with further subsequent presentations: Trial 1 (M = 0.07) vs. Trial 2 (M = 0.39), $t(10) = 6.0$, $p < .001$; Trial 2 (0.39) vs. Trial 3 (0.71), $t(10) = 7.9$, $p < .001$; Trial 3 (0.71) vs. Trial 4 (0.87), $t(10) = 4.3$, $p < .05$; Trial 4 (0.87) vs. Trial 5 (0.92), $t(10) = 3.8$, $p < .05$; Trial 5 (0.92) vs. Trial 6 (0.94), $t(7) = 1.33$, n.s.; Trial 6 (0.94) vs. Trial 7 (0.93), $t(2) = 0.68$, n.s.

FMRI Activation

The standard single-trial subsequent memory analysis involves comparing neural activity measured during encoding between back-sorted items that are subsequently remembered and those that are subsequently forgotten, thereby revealing loci of operations supporting successful encoding (Paller & Wagner, 2002; van den Broek et al., 2016). Likewise, in our multi-trial design, this analysis (E01 > E00) was performed across all repeated (test-potentiated) encoding trials except the very first study (see the Method for details). Many of the resultant regions listed

in Table 1 were consistent with the encoding network identified by standard subsequent memory paradigms (Kim, 2011) and included: left inferior frontal gyrus (LIFG, BA 45/46/47), right dorsal lateral prefrontal cortex (DLPFC, BA 9), left parahippocampal gyrus / fusiform gyrus (BA 19, 35/36), bilateral superior frontal gyrus / premotor cortex (BA 6/4), and bilateral temporal gyri (BA 37, 38). Additional regions included right thalamus, right midbrain, left cerebellum and bilateral occipital lobe (left lingual gyrus / cuneus, BA 17/18; right middle occipital gyrus, BA 18). Unexpectedly, there were higher activities in right medial frontal gyrus (BA 10), bilateral posterior cingulate (BA 23/31) and bilateral precuneus (BA 31/7, see Figure 3), all of which are part of the default mode network (Buckner, Andrews-Hanna, & Schacter, 2008). The right caudate showed a greater activity for subsequently forgotten (vs. remembered) word pairs (E00 > E01).

We next asked whether the activity of regions identified in the above subsequent memory analysis—regions that were involved in successful (test-potentiated) encoding—would be modulated by whether an item had been previously successfully retrieved. To answer this question, the subsequent memory contrast (E01 > E00) was used as an inclusive mask to compare activity during encoding for items that had not been previously recalled (E01) with items that had been previously recalled (E11). The results (Table 1) showed that neural activity in LIFG (BA 45/13), left DLPFC (BA 9), left middle temporal gyrus (LMTG, BA 37), and bilateral fusiform gyri (LFFG and RFFG, BA 37, 19) was greater for items classified as E01 than E11 (see Figure 4). No regions showed increased activity in the E11 condition relative to the E01 condition within the mask. These results demonstrate that relearning-related activity is lower in the encoding regions when items have been successfully retrieved on an earlier trial (E11) compared to when they have not (E01).

By restricting brain activity between the E11 and E01 conditions to regions activated within the subsequent memory (E01 > E00) contrast, regions exhibiting differences in activity between the E01 and E11 had to also differ in the subsequent memory analysis. This strategy was adopted to assess whether prior learning would modulate the activity of regions involved in successful encoding. However, it is possible that activity may differ between the E01 and E11 conditions in regions other than those identified in the subsequent memory contrast. That is, there may be regions that differ between E01 and E11 but are unrelated to successful encoding per se (e.g., regions related to cognitive control in TDE). To discover such regions, the E01 and E11 conditions were contrasted directly without the use of an inclusive mask, using a stricter threshold (see Method). This unmasked contrast revealed a single region that had different brain activity between the E11 and E01 conditions: the right dACC showed greater activity for previously retrieved items (E11) relative to previously unrecalled items (E01; Figure 4).

Functional Connectivity

Motivated by the results that activity in the LMTG, LIFG, left DLPFC, and bilateral FFG was higher under the E01 condition than the E11 condition, whereas activity in the dACC showed the opposite pattern, a functional connectivity analysis was performed to examine how these regions were coupled during repeated studying as a function of learning outcome on the previous test. Specially, pairwise connectivity (see Method) was first calculated among the six above-identified regions in the E01 and E11 conditions separately, and then subjected to paired *t* tests with FDR correction comparing the difference in connectivity strength between the two conditions. The results are shown in Figure 5. The neural coupling between two encoding areas—LIFG and LMTG—was weaker for items that were successfully retrieved on an earlier trial (E11) compared to items that were not successfully retrieved on an earlier trial (E01) (r_{E01} =

.24, $p < .001$; $r_{E11} = .04$, n.s.), $p_{fdr} < .05$. On the other hand, the coupling between dACC and LIFG was stronger under the E11 condition ($r_{E11} = .09$, $p < .05$) than the E01 condition ($r_{E01} = -.06$, n.s.), $p_{fdr} < .05$, suggesting a dACC-mediated cognitive control of memory encoding. None of the other pairwise connectivity correlations varied as a function of condition, ps > .05, FDR corrected (Storey, 2002).

Discussion

While both behavioral (Arnold & McDermott, 2013a, 2013b; Izawa, 1966) and fMRI (X. L. Liu et al., 2014; Vestergren & Nyberg, 2014; Nelson et al., 2013) studies have found that previously unlearned items benefit from subsequent restudy opportunities, a finding known as testing-potentiated encoding (TPE), it is less clear whether previously learned items receive a similar benefit. Although learned items may continue to be strengthened with repeated presentations of that item, previous behavioral studies (Fazio et al., 2010; Karpicke & Roediger, 2008; Pashler et al., 2005; Soderstrom et al., 2016) suggest that additional presentations of a successfully learned item may prove ineffective, and the strength of that item remains essentially unchanged.

The main purpose of the present study was to compare brain activity during encoding for items that were previously successfully remembered with those that were not. On the second and subsequent encoding trials of a multi-trial study-test paradigm, items were classified into three categories: those that were successfully retrieved on both the previous and subsequent trial (E11), those that were not retrieved on the previous trial but were on the subsequent trial (E01), and those that were not retrieved on either the previous or subsequent trial (E00). The comparison between E01 and E00 from trial 2 onwards identified the brain regions involved in successful encoding, similar to a standard subsequent memory analysis (Paller & Wagner, 2002;

van den Broek et al., 2016). More importantly, amongst these regions, we observed that five areas, left inferior frontal gyrus (LIFG), left dorsolateral prefrontal cortex (DLPFC), left middle temporal gyrus (LMTG), and bilateral fusiform gyri (FFG) were less activated, and the connectivity between the LIFG and LMTG was suppressed on encoding trials 2 onwards if items had been successfully retrieved on the prior trial. Thus, once an item is learned, subsequent restudy is not effective, as has been shown numerous times behaviorally, because brain areas that are recruited to learn unlearned items are not as strongly activated or interconnected.

LIFG and FFG are the two prevalent regions associated with memory encoding as identified in a recent meta-analysis (Kim, 2011). The putative function of LIFG in episodic memory encoding is to support retrieval of stored knowledge from semantic memory and goal-directed selection amongst multiple semantic representations (Badre & Wagner, 2007; Habib, Nyberg, & Tulving, 2003; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill, 2003). Consistently, LIFG is more strongly engaged during the encoding of verbal materials than that of non-verbal materials (Golby, Poldrack, Brewer, & Spencer, 2001; Kim, 2011). In contrast, FFG is thought to be more involved in the encoding of pictorial vs. verbal stimuli (Kim, 2011), likely attributable to its differentiation function in high-level visuoperceptual processing (Garoff, Slotnick, & Schacter, 2005; Kanwisher, McDermott, & Chun, 1997). In more direct relation to the shallow encoding of word pairs (i.e., type-case decision) in the current experiment, both LIFG and FFG have also been reported to be engaged in successful encoding of words via similar structural processing (Baker, Sanders, Maccotta, & Buckner, 2001; Otten, Henson, & Rugg, 2001).

While the ventrolateral part of the PFC (VLPFC, i.e., LIFG) contributes to long-term memory through the maintenance, retrieval and selection of information, the dorsolateral part of

the PFC enhances memory via its capability to manipulate and organize multiple pieces of information in working memory (Blumenfeld & Ranganath, 2007). Indeed, Blumenfeld, Parks, Yonelinas, and Ranganath (2011) demonstrated that DLPFC helped form relationships among items to support associative memory, but did not promote memory for item-specific information. The distinction between the roles of VLPFC and DLPFC in memory encoding also fits well with the generic "what/how" axis specified by a recent model of PFC organization (O'Reilly, 2010), proposing that VLPFC determines what features should be processed/extracted depending on the goals whereas DLPFC concerns how operations should be carried on given certain execution rules.

Left (posterior) middle temporal gyrus has also been implicated in encoding and storage of paired-associate information and in verbal language processing. Previously, our lab (Habib & Nyberg, 2007) observed that LMTG showed a graded activation in response to memory of different strengths at both encoding and retrieval, such that accessible items (both recalled and recognized) had the highest brain activity, followed by inaccessible but available items (not recalled but later recognized) and forgotten items (neither recalled and recognized). Interestingly, recent advancement in neurolinguistics suggests that LMTG may subserve retrieval of lexical knowledge from long-term memory (Snijders et al., 2009), thereby supporting verbal learning.

Memory research has long proposed the idea that the frontal and temporal brain regions work together in support of memory (Eichenbaum, 2017). The weakened neural coupling between LIFG and LMTG during study after an item has already been successfully learned highlights the importance of this connectivity to memory encoding. In fact, LIFG showed the same linear response to memory strength as LMTG in our previous study (Habib & Nyberg,

2007). Previous research has also directly demonstrated a functional connection between the LIFG and LMTG during language comprehension at the sentence level (Snijders, Petersson, & Hagoort, 2010), thus suggesting that the prefrontal modulation of posterior cortical representations might be critical to word-word associative learning. Structurally, this left fronto-temporal connection seems to arise from two direct anatomical circuits between them—dorsally via the arcuate fasciculus and ventrally via the extreme capsule (Papoutsi et al., 2011).

Only one region, the dACC, showed increased activity per se, as well as functional connectivity with LIFG, during the restudy of learned items (Figure 4 and 5). This pattern, which was opposite to that of the above encoding regions, therefore differentiated between TPE and TDE. Thus far, a wide spectrum of functions have been linked to dACC, including motivation (Holroyd & Yeung, 2012), reward-based decision-making (Wallis & Kennerley, 2011), and monitor and control (Carter et al., 1998; Sheth et al., 2012). Researchers (Carter & van Veen, 2007; Heilbronner & Hayden, 2016; Mansouri, Egner, & Buckley, 2017; Shenhav, Cohen, & Botvinick, 2016) have attempted to unify these diverse functions under a "comparator" model in which dACC tracks the ongoing conflict/competition and cost/benefit between alternative actions, and generates an outgoing control signal as a result of this evaluation to indicate whether to preserve or switch a behavioral strategy in the best service of current task-relevant goals. In relation to the present results, once an item has been successfully retrieved on an earlier trial, an economical strategy would be to prevent subsequent effort to be expended in re-encoding it. To that end, the enhanced connectivity between the dACC and LIFG in the E11 condition could reflect the neural implementation of such a strategy by which an "interfering" signal from the dACC to the LIFG would prevent additional and "unnecessary" encoding efforts towards already learned items.

Finally, it is interesting to note that our contrast between E01 and E00 revealed higher activities in the regions that are part of the default mode network (DMN, Buckner et al., 2008), including right medial frontal gyrus (BA 10), bilateral posterior cingulate (BA 23/31) and bilateral precuneus (BA 31/7) (Figure 3). The activation of DMN regions at encoding is often predictive of subsequent forgetting, possibly reflecting mind-wandering, a failure to disengage from task-negative cognitive processing, or momentary lapse of attention, all of which could lead to encoding impairment (Kim, 2011; Rugg et al., 2015). How could DMN be involved in the enhanced encoding from TPE (presumably, greater TPE for E01 items than for E00 items induced by previous unsuccessful retrieval attempts)? Recent research (e.g., Kim, Daselaar, & Cabeza, 2010) has demonstrated overlaps between encoding failure and retrieval success activity within these DMN midline areas. Therefore, increased DMN activity in E01 may reflect (covert) retrieval processes during the post-test restudy such that recalling an earlier test experience (e.g, the self-awareness of the unlearned item) could serve as an additional encoding context to improve memory (see Q. Liu, Dong, Chen, & Xue, 2014 for a similar interpretation). Indeed, Nelson et al. (2013) found that neural activity in the left posterior parietal cortex during a post-test restudy predicted the amount of "new learning"; the authors concluded that the engagement of retrieval process associated with the parietal area facilitated restudy encoding.

Conclusion

There are a number of limitations to the present study, such as the small sample size, the unavoidable dissimilarity of memory performance across individuals, and the necessary choice of categorizing items with different learning histories into the same conditions (e.g., the combination of E01, E001, E0001, etc. into the E01 condition) in order to accumulate sufficient trials of successfully and unsuccessfully retrieved items for analysis. Despite these limitations,

however, the ties between our results with those from previous studies of memory encoding and subsequent memory suggest that the activations and connectivity observed in the present study do indeed represent the neural signature of TDE, the flip side of TPE. Future studies are called for to further elucidate the relationship between TPE and TDE.

To summarize, our findings showed that once an item has been learned, a subset of the brain network involved in the initial successful encoding processes, such as retrieval from semantic memory and elaboration supported presumably by LIFG, LMTG and DLPFC, is less activated during subsequent restudy, likely via dACC-mediated cognitive control. Thus, the present work is the first to demonstrate that these dynamic modulations in the memory encoding network and dACC in terms of regional activity and functional connectivity may lay a neural foundation for the behavioral observation that once an item has been learned, further study does not prove beneficial, a pattern we refer to as test-depressed encoding. Reciprocally, our neuroimaging findings have the potential to help formulate new cognitive theories of test-depressed encoding.

References

Arnold, K. M., & McDermott, K. B. (2013a). Free recall enhances subsequent learning. *Psychonomic Bulletin & Review*, *20*(3), 507–513. http://doi.org/10.3758/s13423-012-0370-3

Arnold, K. M., & McDermott, K. B. (2013b). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 940–945. http://doi.org/10.1037/a0029199

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–2901. http://doi.org/10.1016/j.neuropsychologia.2007.06.015

Baker, J. T., Sanders, A. L., Maccotta, L., & Buckner, R. L. (2001). Neural correlates of verbal memory encoding during semantic and structural processing tasks. *Neuroreport*, *12*(6), 1251–1256.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal Cortex and Long-Term Memory Encoding: An Integrative Review of Findings from Neuropsychology and Neuroimaging, *Neuroscientist, 13*(3), 280–291. http://doi.org/10.1177/1073858407299290

Blumenfeld, R. S., Parks, C. M., Yonelinas, A. P., & Ranganath, C. (2011). Putting the pieces together: the role of dorsolateral prefrontal cortex in relational memory encoding. *Journal of Cognitive Neuroscience*, *23*(1), 257–265. http://doi.org/10.1162/jocn.2010.21459

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. http://doi.org/10.1196/annals.1440.011

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 367–379.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*(5364), 747–749.

Eichenbaum, H. (2017). Memory: Organization and Control. *Annual Review of Psychology*, *68*(1), 19–45. http://doi.org/10.1146/annurev-psych-010416-044131

Fazio, L. K., Huelser, B. J., Johnson, A., & Marsh, E. J. (2010). Receiving right/wrong feedback: Consequences for learning. *Memory*, *18*(3), 335–350. http://doi.org/10.1080/09658211003652491

Garoff, R. J., Slotnick, S. D., & Schacter, D. L. (2005). The neural origins of specific and general memory: The role of the fusiform cortex. *Neuropsychologia*, *43*(6), 847–859. http://doi.org/10.1016/j.neuropsychologia.2004.09.014

Golby, A. J., Poldrack, R. A., Brewer, J. B., Spencer, D., Desmond, J. E., Aron, A. P., & Gabrieli, J. D. (2001). Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain*, *124*(9), 1841-1854.

Göttlich, M., Beyer, F., & Krämer, U. M. (2015). BASCO: a toolbox for task-related functional connectivity. *Frontiers in Systems Neuroscience*, *9*, 126. https://dx.doi.org/10.3389%2Ffnsys.2015.00126

Habib, R., & Nyberg, L. (2007). Neural Correlates of Availability and Accessibility in Memory. *Cerebral Cortex*, *18*(7), 1720–1726. http://doi.org/10.1093/cercor/bhm201

Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: the HERA model revisited. *Trends in Cognitive Sciences*, *7*(6), 241–245. http://doi.org/10.1016/S1364-6613(03)00110-4

Heilbronner, S. R., & Hayden, B. Y. (2016). Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annu Rev Neurosci*, *39*(1), 149–170. http://doi.org/10.1146/annurev-neuro-070815-013952

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*(2), 121–127. http://doi.org/10.1016/j.tics.2011.12.008

Izawa, C. (1966). Reinforcement-test sequences in paired-associate learning. *Psychological Reports*, *18*(3), 879-919.

Jang, Y., Wixted, J. T., Pecher, D., Zeelenberg, R., & Huber, D. E. (2012). Decomposing the interaction between retention interval and study/test practice: The role of retrievability. *The Quarterly Journal of Experimental Psychology*, *65*(5), 962-975.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302-4311.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966–968. http://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://doi.org/10.1016/j.neuroimage.2010.09.045

Kim, H., Daselaar, S. M., & Cabeza, R. (2010). Overlapping brain activity between episodic memory encoding and retrieval: Roles of the task-positive and task-negative networks. *NeuroImage*, *49*(1), 1045–1054. http://doi.org/10.1016/j.neuroimage.2009.07.058

Liu, Q., Dong, Q., Chen, C., & Xue, G. (2014). Neural processes during encoding support durable memory. *NeuroImage*, *88*(C), 1–9. http://doi.org/10.1016/j.neuroimage.2013.11.031

Liu, X. L., Liang, P., Li, K., & Reder, L. M. (2014). Uncovering the Neural Mechanisms Underlying Learning from Tests. *PLOS ONE*, *9*(3), e92025–7. http://doi.org/10.1371/journal.pone.0092025

Mansouri, F. A., Egner, T., & Buckley, M. J. (2017). Monitoring Demands for Executive Control: Shared Functions between Human and Nonhuman Primates. *Trends in Neurosciences*, *40*(1), 15–27. http://doi.org/10.1016/j.tins.2016.11.001

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. *Annual Review of Psychology*, *67*(1), 105–134. http://doi.org/10.1146/annurev-psych-113011-143733

Nelson, S. M., Arnold, K. M., Gilmore, A. W., & McDermott, K. B. (2013). Neural Signatures of Test-Potentiated Learning in Parietal Cortex. *Journal of Neuroscience*, *33*(29), 11754–11762. http://doi.org/10.1523/JNEUROSCI.0960-13.2013

Ofen, N. (2012). The development of neural correlates for memory formation. *Neuroscience and Biobehavioral Reviews*, *36*(7), 1708–1717. http://doi.org/10.1016/j.neubiorev.2012.02.016

Otten, L. J., Henson, R. N., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding. *Brain*, *124*(2), 399-412.

O'Reilly, R. C. (2010). The What and How of prefrontal cortical organization. *Trends in Neurosciences*, *33*(8), 355–361. http://doi.org/10.1016/j.tins.2010.05.002

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102.

Papoutsi, M., Stamatakis, E. A., Griffiths, J., Marslen-Wilson, W. D., & Tyler, L. K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *NeuroImage*, *58*(2), 656–664. http://doi.org/10.1016/j.neuroimage.2011.06.036

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When Does Feedback Facilitate Learning of Words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 3–8. http://doi.org/10.1037/0278-7393.31.1.3

Rissman J, Gazzaley A, D'Esposito M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage, 23*(2), 752-763

Rugg, M. D., Johnson, J. D., & Uncapher, M. R. (2015). Encoding and Retrieval in Episodic Memory. In D. R. Addis, M. Barense, & A. Duarte (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Memory* (pp. 84-107). Chichester, UK: John Wiley & Sons. doi: 10.1002/9781118332634.ch5

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291. http://doi.org/10.1038/nn.4384

Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., et al. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*(7410), 218–221. http://doi.org/10.1038/nature11239

Snijders, T. M., Petersson, K. M., & Hagoort, P. (2010). Effective connectivity of cortical and subcortical regions during unification of sentence structure. *NeuroImage*, *52*(4), 1633–1644. http://doi.org/10.1016/j.neuroimage.2010.05.035

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and Unification of Syntactic Structure in Sentence Comprehension: an fMRI Study Using Word-Category Ambiguity. *Cerebral Cortex*, *19*(7), 1493–1503. http://doi.org/10.1093/cercor/bhn187

Soderstrom, N. C., Kerr, T. K., & Bjork, R. A. (2016). The Critical Importance of Retrieval—and Spacing—for Learning. *Psychological Science*, *27*(2), 223–230. http://doi.org/10.1177/0956797615617778

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(3), 479–498. http://doi.org/10.1111/1467-9868.00346

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia*, *41*(3), 280–292. http://doi.org/10.1016/S0028-3932(02)00161-6

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*, *94*(26), 14792-14797.

van den Broek, G., Takashima, A., Wiklund-Hörnqvist, C., Wirebring, L. K., Segers, E., Verhoeven, L., & Nyberg, L. (2016). Neurocognitive mechanisms of the "testing effect": A review. *Trends in Neuroscience and Education*, *5*(2), 52–66. http://doi.org/10.1016/j.tine.2016.05.001

Vestergren, P., & Nyberg, L. (2014). Testing alters brain activity during subsequent restudy: Evidence for test-potentiated encoding. *Trends in Neuroscience and Education*, *3*(2), 69–80. http://doi.org/10.1016/j.tine.2013.11.001

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329–338.

Wallis, J. D., & Kennerley, S. W. (2011). Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Annals of the New York Academy of Sciences*, *1239*(1), 33–42. http://doi.org/10.1111/j.1749-6632.2011.06277.x

Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing*. London: Academic.

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLOS ONE*, *8*(7), e68910.

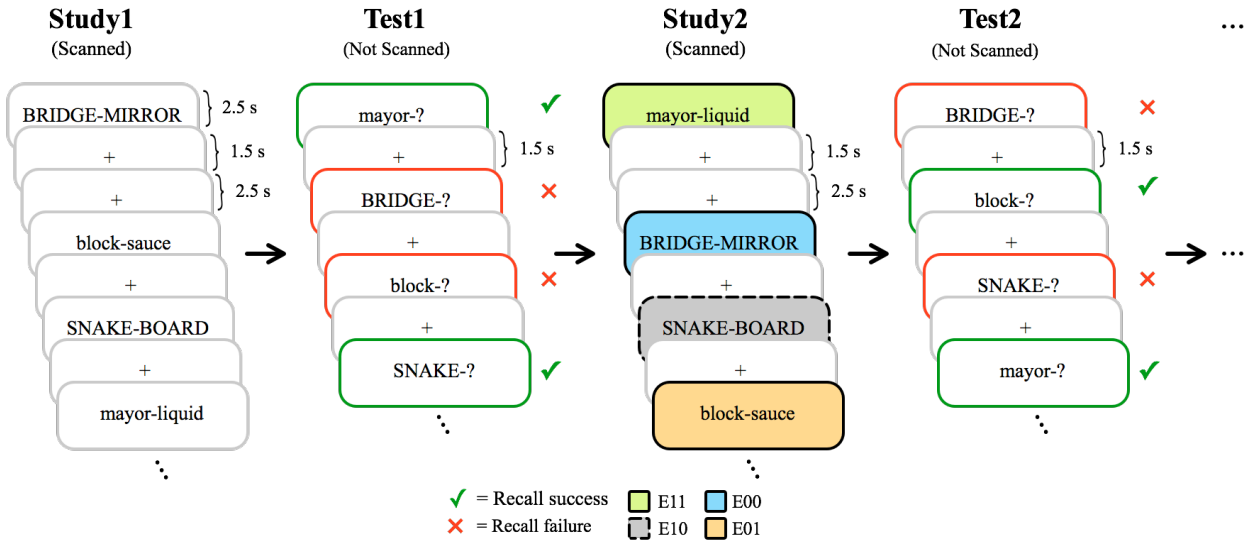

Figure 1. Schematic overview of the experimental procedure and design. In an MRI scanner, participants first studied 40 unrelated word pairs (half lower case, half upper case) with brain image scanning and then had a cue-recall test (via microphone) without scanning. During study, word pairs were shown for 2.5 s followed by a 1.5 s crosshair and a randomly jittered 2.5 s null event; during test, word pairs were presented in a self-paced fashion followed by a 1.5 s crosshair. This study-test trial was repeated 5-7 times for each participant until they learned all or the majority (≥ 85%) of the word pairs. Memory encoding during study on trial 2-7 were classified into four categories based on subsequent memory *and* test performance on trial 1-6 respectively: a) E11 if both recalled on the previous test and on the subsequent test (e.g., mayor-liquid), b) E10 if recalled on the previous test but not on the subsequent test (SNAKE-BOARD), c) E00 if recalled on neither test (BRIDGE-MIRROR), and d) if not recall on the previous test but recalled on the subsequent test (block-sauce).

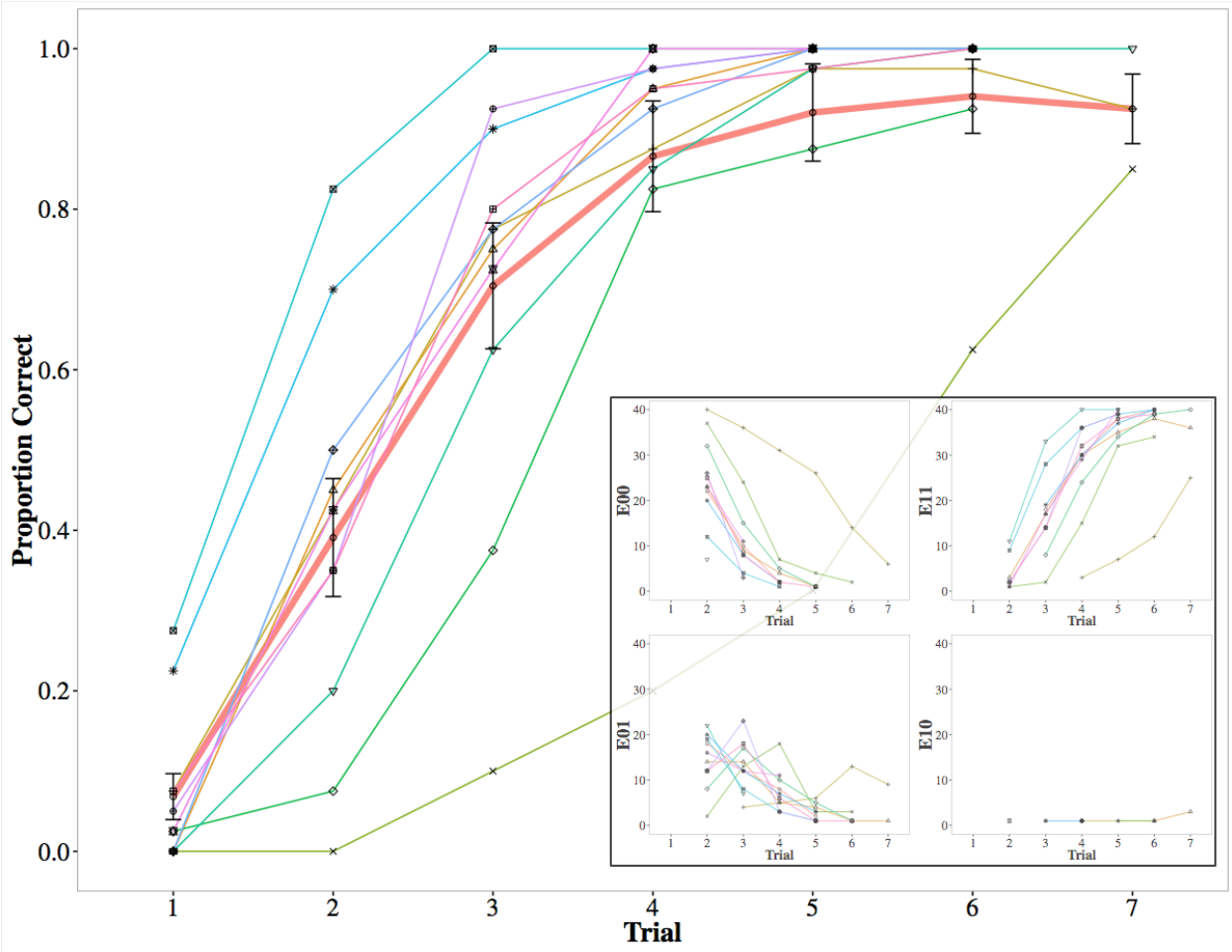

Figure 2. Individual and averaged (n = 11) cued-recall performance as a function of study-test trials. Each thin line with a different marker represents one participant. Mean performance is indicated by the thick line. Error bars represent standard errors of the means. See text for detailed statistics. The transparent inset shows individual dynamic changes in the number of items falling into the four categories/conditions: a) items remained unlearned (E00) decreased as learning progressed, b) items remained learned (E11) increased as learning progressed, c) items that became learned increased first and then decreased as learning progressed, d) items that became learned first but unlearned next occurred infrequently were therefore not included in the fMRI analysis.

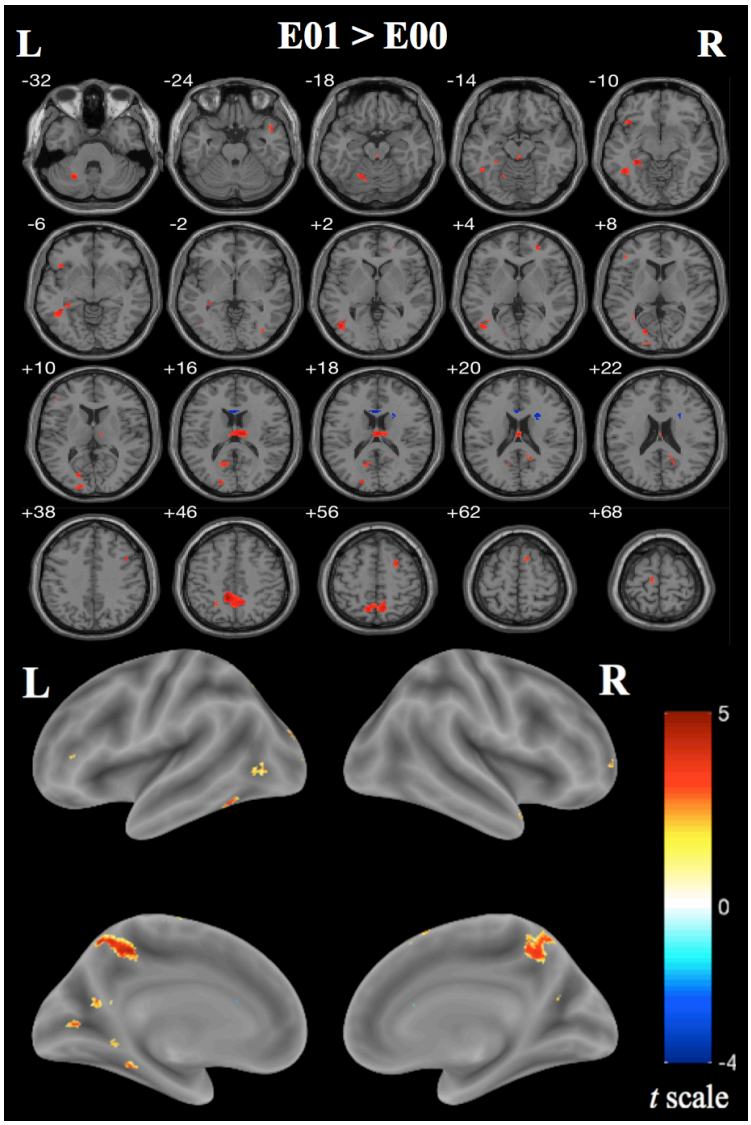

Figure 3. Top: Axial view of significant clusters showing greater brain activity across repeated studies for E01 condition than for E00 condition. The number on the top left corner of each slice denotes its z-axis coordinate. Bottom: The same clusters rendered on the cortical surface (lateral and medial views). All maps were generated using an SPM toolbox, bspmview (http://www.bobspunt.com/bspmview/). Color bar represents *t* statics (*df* = 30). See Table 1 for detailed cluster information.

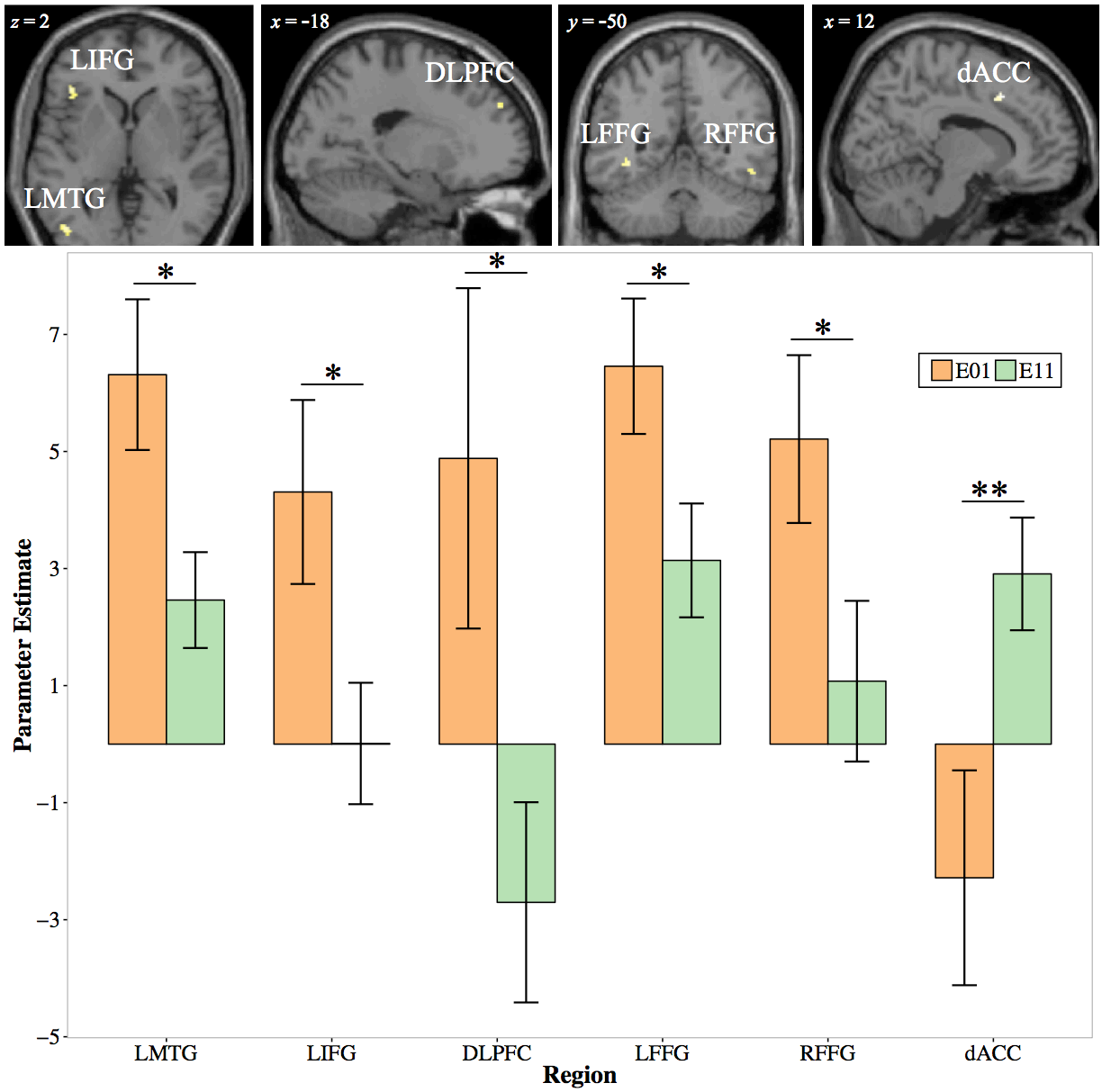

Figure 4. When an item was successfully (vs. unsuccessfully) learned previously (i.e., E11 vs. E01), brain activities in LMTG, LIFG, and DLPFC decreased (with the subsequent memory contrast [E01 > E00] as an inclusive mask), whereas activity in the dACC increased (unmasked). The top panel shows the slice views of significant clusters. The Y-axis in the bottom panel represents the parameter estimate (beta weight) extracted from and averaged over voxels within each cluster. See Table 1 for detailed cluster information. LMTG = left middle temporal gyrus, LIFG = left inferior frontal gyrus, DLPFC = dorsolateral prefrontal cortex, LFFG = left fusiform gyrus, RFFG = right fusiform gyrus, dACC = dorsal anterior cingulate cortex. *$p < .05$, **$p < .01$.

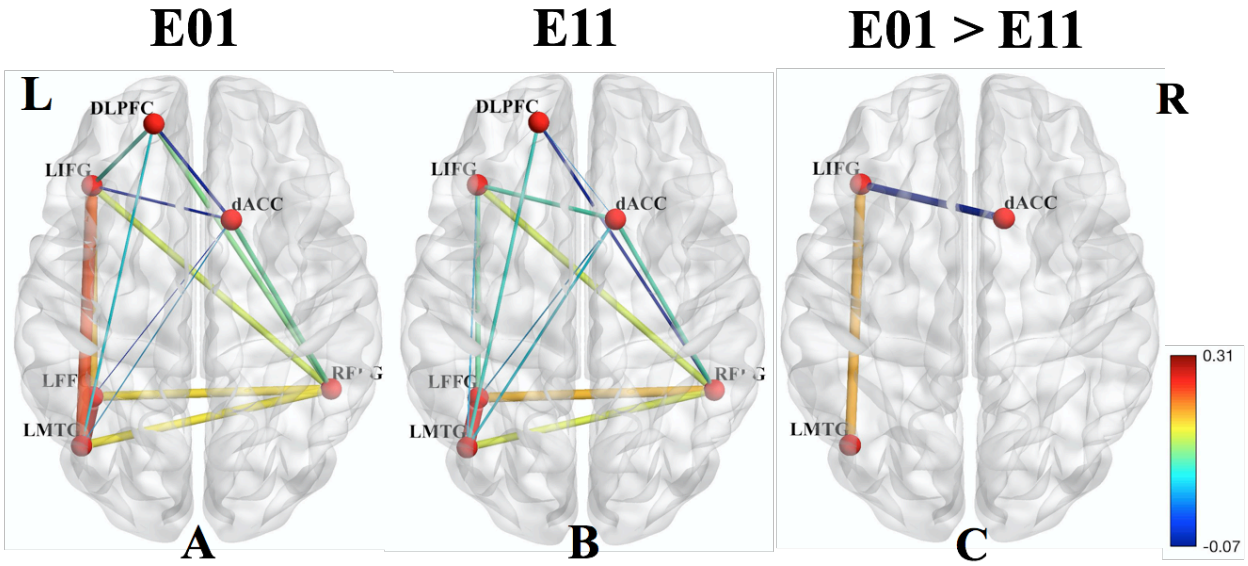

Figure 5. Pairwise functional connectivity among the six identified TDE-sensitive regions. The left image shows the connectivity under E01 condition (A), middle under E11 condition (B), and the right image (C) shows the significant difference in connectivity strength between the two conditions (E01 – E11). The neural coupling between LIFG and LMTF was strongly enhanced under E01 condition than E11 condition (r_{E01} = .24, p < .001; r_{E11} = .04, n.s.), p_{fdr} < .05, whereas the coupling between dACC and LIFG was strengthened under E11 condition (r_{E11} = .09, p < .05) than E01 condition (r_{E01} = -.06, n.s.), p_{fdr} < .05. All other connectivity did not vary as a function of condition, ps > .05, FDR corrected (Storey, 2002). Warm colors indicate positive connectivity or connectivity difference, and cold colors indicate negative. The radius/thickness of the line reflects the absolute value of connectivity or connectivity difference. Color bar represents the overall range of connectivity (difference) values. Brain networks were visualized with the BrainNet Viewer (http://www.nitrc.org/projects/bnv/) (Xia et al., 2013). DLPFC = dorsolateral prefrontal cortex, LIFG = left inferior frontal gyrus, dACC = dorsal anterior cingulate cortex, LFFG = left fusiform gyrus, RFFG = right fusiform gyrus, LMTG = left middle temporal gyrus. L = left, R = right.

Table 1

Brain regions showing greater activity across repeated study dependent on conditions

Anatomy	BA	L/R	Peak MNI *x*	*y*	*z*	*Z* score	*k* (voxels)
E01 > E00							
Medial Frontal Gyrus	10	R	24	48	4	3.48	13
Middle Frontal Gyrus / Dorsolateral Prefrontal Cortex	9	R	42	16	38	3.67	22
Inferior Frontal Gyrus	45/46	L	-44	34	8	3.38	10
Inferior Frontal Gyrus	47	L	-36	22	-10	3.87	37
Superior Frontal Gyrus / Premotor Cortex	6/4	L	-10	-18	68	3.43	13
Superior Frontal Gyrus	6	R	24	6	56	3.55	23
Superior Frontal Gyrus	6	R	12	16	62	3.28	15
Posterior Cingulate	23/31	L	-12	-60	16	3.76	41
Precuneus	31/7	L/R	-8	-44	46	4.59	712
Posterior Cingulate / Precuneus	31	R	18	-56	22	3.47	18
Thalamus	-	R	10	-14	16	4.08	146
Midbrain	-	R	2	-32	-14	3.33	21
Parahippocampal Gyrus / Fusiform Gyrus	19	L	-38	-44	-6	3.87	115
Parahippocampal Gyrus	35/36	L	-28	-38	-10	4.27	74
Superior Temporal Gyrus	38	R	50	12	-24	3.58	14
Middle/Inferior Temporal Gyrus	37	L	-44	-68	2	3.97	71
Lingual Gyrus / Cuneus	17/18	L	-18	-76	10	3.91	33
Cuneus	17/18	L	-18	-86	16	3.73	61
Middle Occipital Gyrus	18	R	36	-76	-2	3.33	10
Cerebellum	-	L	-24	-62	-32	3.88	28
Cerebellum	-	L	-16	-62	-18	3.74	50
E00 > E01							
Caudate	-	R	22	14	20	3.40	23
E01 > E11 (masked by E01 > E00)							
Superior Frontal Gyrus / Dorsolateral Prefrontal Cortex	9	L	-18	50	34	2.82	12
Inferior Frontal Gyrus	45/13	L	-42	32	4	2.99	15
Middle Temporal Gyrus	37	L	-44	-70	4	3.22	48
Fusiform Gyrus	19	L	-38	-50	-8	2.90	17
Fusiform Gyrus	37	R	50	-46	-14	3.02	16
E11 > E01 (masked by E01 > E00)							
None							

E01 > E11

None

E11 > E01

Dorsal Anterior Cingulate	32	R	12	16	42	3.43	13

Note. BA, Brodmann area; L/R hemisphere. Peak coordinates are in MNI space. All contrasts were unmasked unless otherwise stated.