

HHS Public Access

Author manuscript *Neuroimage*. Author manuscript; available in PMC 2015 June 18.

Published in final edited form as: *Neuroimage*. 2011 June 1; 56(3): 1726–1734. doi:10.1016/j.neuroimage.2011.02.048.

Neural mechanisms of semantic interference and false recognition in short-term memory

Alexandra S. Atkins1,2 and **Patricia A. Reuter-Lorenz**¹

¹ Department of Psychology, University of Michigan

² Center for Cognitive Neuroscience, Duke University

Abstract

Decades of research using the Deese-Roediger-McDermot (DRM) paradigm have demonstrated that episodic memory is vulnerable to semantic distortion, and neuroimaging investigations of this phenomenon have shown dissociations between the neural mechanisms subserving true and false retrieval from long-term memory. Recently, false short-term memories have also been demonstrated, with false recognition of items related in meaning to memoranda encoded less than 5 seconds earlier. Semantic interference is also evident in short-term memory, such that correct rejection of related lures is slowed relative to correct rejection of unrelated lures. The present research constitutes the first fMRI investigation of false recognition and semantic interference in short-term memory using a short-term DRM paradigm in which participants retained 4 semantic associates over a short 4 second filled retention interval. Results showed increased activation in the left mid-ventrolateral prefrontal cortex (BA45) associated with semantic interference, and significant correlations between these increases and behavioral measures of interference across subjects. Furthermore, increases in dorsolateral PFC occurred when related lures were correctly rejected versus falsely remembered. Compared with false recognition, true recognition was associated with increases in left fusiform gyrus, a finding consistent with the notion that increased perceptual processing may distinguish true from false recognition over both short and long retention intervals. Findings are discussed in relation to current models of interference resolution in short-term memory, and suggest that false short-term recognition occurs as a consequence of the failure of frontally-mediated cognitive control processes which adjudicate semantic familiarity in support of accurate mnemonic retrieval.

Keywords

false memory; short-term memory; interference; cognitive control; VLPFC; DLPFC

Correspondence should be addressed to: Alexandra S. Atkins, LSRC Building, B203, Duke University, Box 90999, Durham, NC 27708-0999, (919) 668-2299, alex.atkins@duke.edu or Patricia A. Reuter-Lorenz, 530 Church Street, University of Michigan, Ann Arbor, MI 48109-1043, (734)-764-6577, parl@umich.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Distortions of memory have been a subject of interest for cognitive psychology since its inception. One reason for this interest is that examination of the circumstances under which our memories fail us can illuminate our understanding of how memory is organized. In the last two decades, the term *false memory* has come to describe instances in which memories become distorted, leading to false recognition and false recall of previously unstudied items. In the Deese-Roediger-McDermott (DRM; 1959; Roediger & Mcdermott, 1995) paradigm, participants study lists of 12 –15 words which are all related in meaning to a common unstudied theme word, or related lure. At test, participants are required to either recognize studied items from a list of probes that includes related lures, or to recall studied items in free report. Investigations using variants of this procedure have shown that participants consistently and confidently falsely recognize related lures, and even produce these items in free recall (see Gallo, 2006, for review).

Although initial investigations of this false memory phenomena were limited to paradigms that included long study lists and retention intervals that varied from several seconds to many hours, there is recent evidence that false memories are produced rapidly, within the time and load constraints of traditionally defined short-term memory tasks (Atkins & Reuter-Lorenz, 2005, 2008; Coane, McBride, Raulerson et al., 2007; Flegal, Atkins, & Reuter-Lorenz, 2010). For example, using a short-term variation of the DRM (ST-DRM) paradigm with 4-item lists, we recently demonstrated reliable false recognition and recall of unstudied lures only 4 seconds following encoding (Atkins & Reuter-Lorenz, 2008). Furthermore, in the recognition version of our task, we found strong evidence that the semantic relationship between related lures and memoranda induced interference even when related lures were not falsely recognized. Specifically, participants took longer to correctly reject a related lure compared to one that was unrelated to the memoranda. We refer to this response time (RT) difference as semantic interference (SI).

The increased time required to reject related lures is consistent with the notion that correct rejection of these items requires a control process that resolves interference induced by the semantic familiarity of these items. When interference is resolved successfully, the related lure can be correctly rejected. False recognition of these items could indicate either a failure of this control process, or a failure to engage it at all.

Interestingly, false recognition has not been widely investigated as a failure of cognitive control. One reason for this is the paucity of crosstalk between research on false long-term memory, and investigations of cognitive control in short-term memory. The control of proactive interference in short-term memory is an executive process that has been studied extensively using the recent probes (RP) task (Jonides, Smith, Marshuetz et al., 1998; Monsell, 1978) and may be relevant to controlling false recognition as well. In the RP task, participants study a set of 4 memoranda, typically letters or words. Following a brief retention interval, a probe item is presented that requires a yes/no recognition response. Generally, 4 probes types are employed. Probes that require a No" response include *recent negative* (RN) probes that are not present on the current trial, but were members of the memory set on the trial immediately preceding the current one, and *non-recent negative*

(NRN) probes that are not members of the current set and have not appeared for the last several trials. A Yes response is required to standard *positive* (POS) probes which are members of the current memory set, and *recent positive* (RPOS) that appeared as memoranda on both the current and immediately preceding trial.

Participants are markedly slower at rejecting RN relative to NRN probes, suggesting that temporal familiarity makes RN probes more difficult to reject. This slowing has been attributed to the engagement of an interference control process that adjudicates this familiarity in the service of accurate recognition memory. Numerous neuroimaging studies indicate an important role for left mid-ventrolateral prefrontal cortex (L VLPFC, in the region of Brodmann s area 45) in the circuitry subserving this resolution process. Activation increases for RN relative to NRN probes have been shown and replicated across a variety of RP tasks (see Jonides & Nee, 2006 for review). Furthermore, behavioral indices of proactive interference (PI), calculated as the RT difference between correct responses to RN relative to NRN probes, show positive correlations with these increases L VLPFC activity (Badre & Wagner, 2005, 2007; Jonides & Nee, 2006).

The importance of L VLPFC is also demonstrated by patient studies demonstrating that focal lesions to this region are associated with behavioral increases in PI (Hamilton $\&$ Martin, 2005; Thompson-Schill, D'Esposito, Aguirre et al., 1997). Furthermore, increased activity in L VLPFC regions has been linked to semantic elaboration during episodic retrieval (Raposo, Han, & Dobbins, 2009), and to selection between semantic competitors (Hirshorn, Aguirre, & Thompson-Schill, 2005; Thompson-Schill et al., 1997; see also Gold & Buckner, 2002; Oztekin, Curtis, & McElree, 2009; Poldrack, Wagner, Prull et al., 1999). In the Verb Generate task, for instance, participants are asked to generate a verb corresponding to a noun, which is presented to them in the scanner. For example, given the noun SCISSORS, a participant may generate the response CUT. Both L VLPFC activity and RT have been found to increase when verbs must be generated in response to nouns that have many associated verbs (for example, BALL) than those that have few (for example SCISSORS; Nelson, Reuter-Lorenz, Persson et al., 2009; Persson, Sylvester, Nelson et al., 2004; Thompson-Schill et al., 1997). Under these conditions left VLPFC has been thought to contribute to a post-semantic retrieval process that selects between semantic competitors (Badre & Wagner, 2007; Kan & Thompson-Schill, 2004).

Thus, converging evidence suggests that regions of L VLPFC may be involved in mediating interference from temporally familiar or semantically related competitors. Further, a recent investigation comparing the RP and Verb Generate tasks demonstrated overlap in the L VLPFC activations associated with interference in both tasks (Nelson et al., 2009). In the present study, we test the hypothesis that L VLPFC may also play a role in controlling SI and vulnerability to false memory in the ST-DRM task. In this task, correct rejection of related lures requires participants to overcome interference induced by semantic, rather than temporal familiarity, as in the RP task. Therefore a primary goal of the present study is to investigate the role of L VLPFC in SI. Whereas one prior study (Paz-Alonso, Ghetti, Donohue et al., 2008) of false long-term memory has implicated more anterior regions of L VLPFC (Brodmann s area 47) in semantic elaboration during retrieval, we were specifically interested in mechanisms associated with the cognitive control of semantic interference.

An additional goal of the present study is to more precisely examine the role of L VLPFC in cognitive control of interference. Although correlations between behavioral measures of interference and increases L VLPFC activity are consistent with the interpretation that this region is primary involved in the resolution of interference (and therefore must work harder to resolve interference as it increases), such correlations are also consistent with the notion that L VLPFC does not play an active role in the resolution of interference per say, but rather provides an index of interference that is passed on to other regions within frontal cortex to support accurate task performance. Given high levels of task performance in the RP and Verb Generate tasks, research using these paradigms has focused almost exclusively on correct trials. Thus, little currently known regarding the distinction between the neural mechanisms which respond to the presence of interference and those associated with successful versus unsuccessful resolution of this interference. By comparing the neural activity associated with correct rejection versus false recognition of lure items in the ST-DRM paradigm, the present study will directly assess this question. If L VLPFC plays a direct role in the resolution of semantic interference in our paradigm, we would expect L VLPFC to distinguish between these two trial types, demonstrating increased activity for correct rejections versus false recognitions. Alternatively, if L VLPFC provides an index of interference that is passed along to other regions which mediate the resolution process more directly, we would expect a positive relationship between increase in L VLPFC activity and behavioral measures of SI, but expect no difference in this region s response to correctly rejected versus falsely recognized lure probes.

A final important goal of this work is to compare true and false recognition in the short-term memory domain. Several studies have examined neural regions that distinguish true and false retrieval from long-term memory (e.g. Abe, Okuda, Suzuki et al., 2008; Cabeza, Rao, Wagner et al., 2001; Garoff-Eaton, Slotnick, & Schacter, 2006; Johnson, Nolde, Mather et al., 1997; Kim & Cabeza, 2007; Okado & Stark, 2003; Slotnick & Schacter, 2004). In one such investigation, Cabeza et al. (2001) demonstrated similar activation of the hippocampus during true and false recognition, but showed increased parahippocampal activation for true memories, this demonstrating a dissociation between the neural mechanism subserving true and false recognition. Additional work has highlighted remarkable overlap in the frontal, parietal and medial temporal regions subserving true and false recognition (see Schacter & Slotnick, 2004 for review). More recently, interactions between memory veracity and confidence have also been demonstrated, with medial temporal lobe (MTL) regions showing increased activity during confident veridical recognition, and frontal-parietal regions show increased activity during confident false recognition (Kim & Cabeza, 2007).

It is currently unknown whether distinctions such as these will carry over from the long-term to the short-term memory domain, or whether similar mnemonic signatures are available within seconds of stimulus encoding. The present study will address these questions, and will examine the relationship between neural regions supporting resolution of semantic interference and successful vs. distorted remembering over short delays.

2. Method

2.1 Participants

Twenty participants (12 females; mean age=20) were recruited from the University of Michigan. All participants gave informed consent as reviewed by the university s Institutional Review Board. Participants were paid \$20 per hour for their participation.

2.2 Task and Procedure

Participants completed the ST-DRM task (see Figure 1; Atkins & Reuter-Lorenz, 2008) during 6 task runs. At the beginning of each trial, a blinking red fixation appeared for 500 ms to warn the participant the trial was beginning. This was followed by a memory set consisting of 4 semantically related items, all associated with a common theme word. The memory set appeared for 1200ms. Five hundred ms following the offset of the memory items, a dual-operation math equation appeared at the center of the screen for 3000ms. This equation was solved either correctly for example, $(4 \text{ X } 3) - 2 = 10$?, or incorrectly, and participants made a left-handed response to indicate whether the math was correct or incorrect.¹ Five hundred ms following the offset of the math problem, a memory probe appeared and participants made a right-handed yes or no response indicating whether or not the probe was a member of the memory set.

During this task, theme words served as the probes on all the trials. There were two variations of "No" trials. The first were *unrelated lure* (UL) trials, in which the probe consisted of the theme word associated with a nonpresented list. The second were related *lure* (RL) trials, in which the probe consisted of the (unstudied) theme associated with the present memory set. On *positive* (POS) trials, the associated theme was embedded in the memory set, and served as the positive probe.

With the exception of positive probes, which by definition occurred twice within the same trial, no participant was exposed to any theme or memoranda more than once during the experiment. Backward associative strength (BAS), a measure of the degree of association between theme words and memoranda (see Roediger, Watson, McDermott et al., 2001; Hicks & Hancock, 2002), was equated across memory lists associated with each probe type, and probe type was counterbalanced with lists across participants. This procedure ensured that participants encountered the same probes, all theme words, but in different contexts, as related lures, unrelated lures, or positive probes. Trials were presented in random order for each participant.

Participants completed 102 ST-DRM trials that were distributed across 6 task runs. Trials were equally distributed across all three probe types in each run. We used a 16sec ITI to allow for the hemodynamic response to return to baseline between trials (Glover, 1999). Participants completed 2 practice runs prior to entering the scanner, in order to become familiar with task and response requirements.

¹Atkins and Reuter-Lorenz (2008) demonstrated false recognition and SI effects both with and without the math verification task during the retention interval. In the present work, we included the math task to insure an adequate number of false memory errors.

Neuroimage. Author manuscript; available in PMC 2015 June 18.

2.3 fMRI Data Acquisition and Preprocessing

Our data were collected using a 3 Tesla GE whole-body scanner equipped with a standard quadrature headcoil. Head movement during scanning was minimized with the used of foam padding. Experimental stimuli were presented using E-Prime software.

Functional T2* blood oxygenation level-dependent (BOLD) images were collected using a spiral sequence with 40 contiguous slices of $3.44 \times 3.44 \times 3$ mm voxels (repetition time $(TR) = 2000$ ms, echo time $(TE) = 30$, flip angle = 90, and field of view $(FOV) = 22$ cm). T1-weighted gradient echo (GRE) anatomical image was also acquired in the same FOV and slices as were used in the functional data collection (TR=250, TE=5.7, and flip angle=90). A high-resolution (106 slice) set of anatomical images was acquired via spoiled gradientrecalled acquisition in steady state (SPGR) imaging (TR=10.5, TE=3.4, flip angle -25 , FOV=24, slice thickness = 1.5mm). SPGR images were corrected for signal inhomogeneity (G. Glover and K. Kristoff, [http://www.psych.standford.edu/_kalina/SPM99/Tools/](http://www.psych.standford.edu/_kalina/SPM99/Tools/vol_homocor.html) [vol_homocor.html](http://www.psych.standford.edu/_kalina/SPM99/Tools/vol_homocor.html)) and skull-stripped using the Brain Extraction Tool provided by FSL (Smith, Jenkinson, Woolrich et al., 2004). These images were then normalized to the Montreal Neurological Institute (MNI) template (avg152t1.img) using SPM5 (Wellcome Department of Cognitive Neurology, London, UK). Functional images were corrected for slice-time differences using 4-point sinc interpolations (Oppenheim, Schafer, & Buck, 1999), and were corrected for head movement using MCLFIRT (Jenkinson, Bannister, Brady et al., 2002). In order to reduce the effect of spike artifacts, functional images were winsorized on a voxel by voxel basis to ensure that no voxel had a signal more than 3.5 standard deviations greater than the mean of the current run (Lazar, Eddy, Genovese et al., 2001). Functional images were then normalized to MNI space using transformations from the normalization of structural images, and were smoothed using an 8mm Guasian kernel. All analyses included a 128s high-pass filter and AR(1) modeling to correct for temporal autocorrelation. For all analyses, each image was scaled to a global mean intensity of 100.

2.4 fMRI Data Analysis

Neuroimaging analyses were conducted using the General Linear Model implemented in SPM5 (Wellcome Department of Cognitive Neurology, London, UK) with separate regressors for each trial type in each run. Event-related activity to probes was modeled by convolving probe onsets with the canonical HRF. Statistical models examined probe-related activations associated with correct recognition of positive probes (hits), correct rejection of unrelated lures, correct rejection of related lures, and false alarms to related lures (false recognitions). Statistical models were estimated for each participant. The number of observations per condition depended on participant performance and therefore varied by participant. All but one participant produced a sufficient number of observations to estimate probe-related activity in each condition, including an average of 29.6 hits, 31.2 correct rejection of unrelated lures, 26.9 correct rejection of related lures, and 7.5 false alarms to related lures. The single participant who produced no false recognition responses, was excluded from analyses that included this condition. All other participants exceeded a minimum criterion of 4 observations per condition. For each comparison of interest described below, contrast maps for each participant were submitted to random effects comparisons.

3. Results

3.1 Behavior

Behavioral findings replicate the SI and false memory effects demonstrated previously (Atkins & Reuter-Lorenz, 2008). Mean accuracy and response time (RT; correct trials only) measures were submitted to a repeated measures analysis of variance (ANOVA). There were main effects of probe type (positive, unrelated, related) on both accuracy, F=16.98, p<.001, $n/2 = .47$ and RT, F=37.22, p<.001, $n/2 = .66$.

Post-hoc tests were conducted to examine false memory and semantic interference effects, and were submitted to a Bonferonni correction for multiple comparisons. Table 1 shows the proportion of POS, UL, and RL items that received a yes response during item recognition, We found a reliable false memory effect, with participants falsely recognizing related lures at a rate over four times that for unrelated lures, $t=6.01$, $p<0.01$, $d=1.54$. Mean RTs for accurate responses were 900ms (SE=29) for POS, 905ms (SE=40) for UL, and 1062ms (SE=41) for RL items. Mean RT for false alarms to RL items was 1132ms (SE=52). With respect to accurate trials, participants were reliably slower to correctly reject related lures compared to unrelated lures, $t=8.10$, $p<.001$, $d=.86$. Our SI index, measured as the difference in RT for correct rejections of related lure vs. unrelated lures had a mean of 156.92ms (SE=19.36).

Paired t-tests comparing false recognition RTs (false alarms to related lures) to true recognition RTs (hits to positive probes) indicate that false recognition of related lures was reliably slower than true recognition (t=4.40, $p < .001$, $d=1.02$). Furthermore, RTs associated with false recognition vs. correct rejections of related lures did not differ reliably, indicating that false recognition did not result from fast responding.

Mean accuracy on the math verification task was .80 (SE=.02). Overall recognition performance did not vary as a function of incorrect or correct responding on the math task (p>.6). Furthermore, post-hoc tests examining each probe type separately revealed no significant differences in the accuracy of responses to positive, unrelated lure or related lure items following correct vs. incorrect math judgments (p>.3 for all).

3.2 Neuroimaging Results

3.2.1 Whole brain analyses—Results from our whole-brain analyses are presented in Table 2, and summarized below. Unless otherwise stated all comparisons reported were significant at p<.005, uncorrected, with threshold requirement of 20 or more contiguous voxels (Forman, Cohen, Fitzgerald et al., 1995; Lieberman & Cunningham, 2009).

In order to examine the neural mechanisms associated with SI, we identified regions that showed increased probe-related activity in response to correctly rejected related lures, relative to correctly rejected unrelated lures. In both cases, correct No responses were made to unstudied items, but SI is present only for related lures. This comparison is thus directly analogous to the recent vs. non-recent negative probe comparisons used in investigations of PI in the RP task.

Figure 2 displays regions with greater activity for correct rejections of related lures vs. unrelated lures. These included a large cluster of voxels in left mid-VLPFC, with a single peak in BA45. This suggests, consistent with predictions, that L VLPFC is recruited when there is interference from unstudied items that are semantically associated with items in memory. Smaller clusters of activation within the bilateral anterior cingulate (ACC, BA32) and bilateral inferior parietal cortex (BA 40/7) also distinguished related lures from unrelated lures.

We examined neural mechanisms of true and false memory by first identifying regions associated with true and false recognition separately. For true recognition, we compared correct recognition of positive probes to correct rejections of unrelated lures (Pos. Hit > Unrelated CR). For false recognition, we compared false alarms to related lures and correct rejections of unrelated lures (Related FA> Unrelated CR).

True recognition was associated with increased activity in a network of fronto-parietal regions consistently associated with verbal short-term memory (Bedwell, Horner, Yamanaka et al., 2005; Cappell, Gmeindl, & Reuter-Lorenz, 2010; Chein & Fiez, 2001; Cohen, Perlstein, Braver et al., 1997; D'Esposito, Postle, Jonides et al., 1998; Rypma & D'Esposito, 1999). Most notably, these included large increases of activation in left anterior prefrontal/ dorsolateral prefrontal cortex (BA 10/46) and bilateral inferior parietal cortices (BA 40). False recognition showed a similar pattern of fronto-parietal activation, as well as a large cluster of activation in left posterior cingulate/retrosplenial cortex, a region previously linked to phenomenological feelings of remembering that may be independent of retrieval accuracy (Wagner, Shannon, Kahn et al., 2005).

In order to examine regions common to both true and false memory, we conducted a conjunction analyses of our true and false memory contrasts. For each contrast, we utilized a threshold of $p<01$ with a cluster extent of 10 or more contiguous voxels. The conjoint probability estimate for the conjunction thus approached p=.0001 (Lazar, Luna, Sweeney et al., 2002), but this value should be considered with caution given the non-independence of the two contrasts. Results (Figure 3A) showed left frontal and bilateral parietal activations common to both true and false recognition. We next examined regions that distinguished true from false recognition by directly contrasting activation associated with correct recognition of positive probes and false recognition of related lures (Pos. Hit > Related FA). Results are displayed in Figure 3B. Compared with false recognition, true recognition was associated with increased activity in the left putamen/parahippocampal gyus (PHG), the left fusiform gyrus, and the right VLPFC.

In order to more directly examine regions mediating successful rejection of interferenceinducing related lures, we identified regions that distinguished correct rejection vs. false recognition of these items. Unlike our SI contrast, which examined differences between correct rejection of related and unrelated lures, this contrast directly compared neural activity associated with correct rejection vs. false recognition of related lures (Related CR > Related FA). Results from this contrast revealed increased activation in left dorsolateral prefrontal cortex (L DLPFC, BA 46/9; see supplementary Figure 1), as well as the fusiform gyrus and putamen. This finding suggests that while L VLPFC becomes activated in the

present of SI, L DLPFC may play a larger role in supporting correct rejection of interfering items.

3.2.2. ROI analyses

VLPFC: Results from our whole brain analysis supported our hypothesis that L VLPFC would distinguish between unrelated lures and related lures that induce SI. In order to investigate the relationship between VLPFC activity in the present task and that observed in studies of proactive interference, we conducted an ROI analysis to determine whether activation in the L VLPFC region associated with PI is also relevant for the control of SI. An ROI was formed by creating a 10mm sphere surrounding peak activation reported previously for the Recent Negative>Non-Recent Negative contrast in the Repeated Probes task (MNI peak: −51 21 11; Jonides et al., 1998; Nelson et al., 2003).

Figure 4A plots mean percent signal change in L VLPFC for hits to positive probes, correct rejections of unrelated lures, and both correct rejections and false alarms to related lures. Activity in this ROI distinguished correctly rejected related lures and unrelated lures, t=2.5, p<.05, *d*=.45. Furthermore, change in L VLPFC activity was strongly and positively correlated with individual differences in SI, $r=.67$, $p<.01$, (Figure 3.5B). A positive correlation between L VLPFC activity and PI has also been reported in the recent probes task (Jonides & Nee, 2006)².

Compared to correct rejections of unrelated lures, L VLPFC activation showed similar numerical increases in activity during both correct rejection and false recognition of related lures (see Fig. 4A), although the VLPFC increase during false recognition did not reach statistical significance, a finding most likely due to increased variability in the VLPFC response during false recognition. Direct comparisons of ROI activity during both correct rejection and false recognition of related lures also showed no difference between the two $(t=.144, p<.8)$. Although this is a null finding, this result coupled with results from our direct comparison of whole-brain activity during correct rejection versus false recognition of related lures suggests that L DLPFC, rather than L VLPFC, may play an important role in mediating whether or not semantic relatedness leads to false recognition.

Another possibility is that L VLPFC might simply respond to the presence of semantic familiarity rather than interference per say. We tested this possibility by examining activation associated with positive probes. Because positive probes were studied items, they should have been both semantically and temporally familiar. However, this familiarity should not have induced interference because it was consistent with veridical recognition. If L VLPFC responded to familiarity only, activation should increase to positive probes as well as related lures. This was not the case, however, as probe-related activity did not differ for positive and unrelated negative probes in this region (t=.186, p>.8).

Neuroimage. Author manuscript; available in PMC 2015 June 18.

²We also examined the correlation between individual variations in SI and changes in L VLPFC activity in a 10mm sphere surrounding our whole brain peak for the semantic interference contrast (−48 21 21). For this ROI, we found a similar, though slightly weaker, positive correlation between increased activity for lure vs. negative probes and increases in our behavioral SI measure, r=.51, p<.05.

MTL: Given previous findings indicating a role for MTL in distinguishing true and false long-term memories, we were particularly interested in examining differences between true and false short-term memory in this region. Although, our whole brain analysis revealed a large cluster of activation extending from the left putamen into the L parahippocampal gyrus (PHG), no other MTL activations distinguished true from false memory. In order to increase our sensitivity to detect MTL differences, we conducted an exploratory analysis by examining true vs. false differences in bilateral MTL at reduced threshold of $p<0.05$. Our MTL ROI was defined as the bilateral hippocampus, PHG and amygdala using the Pick Atlas (Wake Forest University; <http://www.fmri.wfubmc.edu>). Results showed a single cluster of 12 voxels in L PHG (−28, −3, −12), which distinguished between true and false. This cluster was contiguous with the putamen/PHG cluster observed in our whole brain analysis. No other portions of MTL distinguished between true and false recognition.

3. Discussion

The present study investigated the neural mechanisms of semantic interference and false recognition in a short-term memory version of the DRM task (Atkins & Reuter-Lorenz, 2008; Deese, 1959; Roediger & Mcdermott, 1995). We examined the neural mechanisms of SI by identifying regions of increased activation during the correct rejection of probes related in meaning to current memoranda (i.e., related lures) as compared to unrelated probes (i.e. unrelated lures). Past research indicating a role for L VLPFC in resolving interference induced by temporal familiarity in the recent probes task (Jonides & Nee, 2006), led us to predict this region would also be involved in interference induced by the semantic familiarity or similarity of related lure items. Consistent with this prediction, we found increased L VLPFC (BA45) activity associated with the correct rejection of related vs. unrelated lures (Figure 2).

ROI analyses revealed that across individuals there was a strong positive correlation between the magnitude of SI and L VLPFC activity in response to related vs. unrelated lures (see Figure 4). This finding demonstrates that the positive relationship between behavioral indices of interference and L VLPFC activity is not unique to temporal familiarity, and suggests that common neural substrates are engaged in response to interference induced by either temporal or semantic familiarity.

Positive correlations between behavioral measures of interference and increased activity in L VLPFC could be interpreted as evidence that this region is either (a) the site of interference resolution and therefore must work harder to resolve interference as it increases, or (b) a site that provides an index of interference for each trial that is used by other cortical regions to support accurate memory performance. We attempted to distinguish between these possibilities by comparing L VLPFC activity for related lures that were ultimately rejected to those that were falsely recognized. Results showed a similar increase in probe-related activity for both correct rejections of and false alarms to related lures, indicating that the L VLPFC activity alone does not distinguish between semantic interference that is resolved correctly or not. In contrast, whole brain comparisons showed L DLPFC activation does distinguish between correct rejections and false alarms to related lures (see Supplementary Figure 1), indicating that this region may play an important role in determining the extent to

which interference indexed by LVLPFC can be mitigated to reduce false memories. This interpretation is consistent with previous work linking L DLPFC to post-retrieval source monitoring (e.g. Achim & Lepage, 2005).

In addition to L VLPFC, our semantic interference comparison of correct rejections of related vs. unrelated lures also showed increased activation in the bilateral ACC. Given strong evidence associating similar increases in ACC (BA 32/24) activity with responselevel conflict across a variety of tasks (e.g. Milham & Banich, 2005; Milham, Banich, Webb et al., 2001; Nelson, Reuter-Lorenz, Sylvester et al., 2003), involvement of this region most likely reflects the need to forgo a yes response to a probe that is familiar in favor of a correct no response.

Our examination of neural activity common to true and false short-term recognition suggests some similarity between the neural mechanism supporting short and long-term retrieval True and false recognition were both associated with increased activity with the left anterior PFC and bilateral PPC, both regions which have been previously associated with retrieval effort and monitoring that may be independent of the success of retrieval from long-term memory (e.g. Cabeza, 2008; Cabeza, Ciaramelli, Olson et al., 2008; Kompus, Eichele, Hugdahl et al., in press). These findings converge with previous reports of common prefrontal activations during short and long-term retrieval (Ranganath, Johnson, & D'Esposito, 2003) and contribute to a growing body of literature highlighting similarities between the neural mechanism supporting retrieval from short-term and long-term memory (e.g. Cabeza, Dolcos, Graham et al., 2002;Ranganath, Cohen, & Brozinsky, 2005; Ranganath et al., 2003). The common recruitment of bilateral precuneus during both true and false recognition is also found in long-term memory tasks, and has been interpreted as reflecting the *perceived* oldness that may occur independently of memory accuracy (Cabeza, 2008; Cabeza et al., 2008). Replication of this finding in the current task suggests common neural representations of perceived oldness may be active during true and false recognition that occurs over short and long retention intervals.

Our comparisons of neural activity associated with true and false short-term memory also suggest some overlap between the neural mechanisms that distinguish true and false recognition over the short- and long-term. Compared with false recognition, true recognition was associated with increased activation in the left fusiform gyrus, a finding consistent with the notion that increased perceptual processing may serve as a signature distinguishing true from false recognition (Slotnick & Schacter, 2004). Increased left fusiform activity is also consistent with previous work which has indicated a potential role for this region in semantic processing by showing repetition priming effects that are selective for semantically meaningful stimuli (see Vuilleumier, Henson, Driver et al., 2002, Simons, Koutstaal, Prince et al., 2003).

Whole-brain and ROI analyses also showed increases left PHG activation associated with true versus false recognition, a result which is consistent with previous work in the longterm memory domain (Cabeza et al., 2001) and which indicates a role for left PHG in distinguishing veridical and false memories following short or long-term retention intervals. True recognition was also associated with increased activation in a region of right VLPFC

(BA44) consistently implicated in inhibitory control across a variety of task contexts (Chikazoe, Jimura, Hirose et al., 2009; Chikazoe, Konishi, Asari et al., 2007; Garavan, Ross, & Stein, 1999). Taken together, these findings suggest increased perceptual processing, as well as the need to exert inhibitory control or increased task monitoring in order to support correct recognition of studied items within a task context which includes a high degree of interference.

In summary, the present work extends our understanding of the neural mechanisms supporting the cognitive control of interference and veridical short-term memory. False alarm rates in RP tasks used to investigate the neural mechanisms of PI are normally quite low. As such, these investigations have focused almost exclusively on the successful resolution of interference, and have generally interpreted L VLPFC increases in this context. Our findings are consistent with the interpretation that L VLPFC responds to selection demands associated with multiple semantic competitors (Badre & Wagner, 2007; Thompson-Schill et al., 1997) but suggest that the region may not directly distinguish between interference that is successfully mitigated in service of accurate task performance, and that which is not. Furthermore, results suggest that when interference is sufficient to produce source confusion regarding the old or new status of a memory probe, monitoring operations mediated by the adjacent L DLPFC may be critical for supporting accurate task performance. Finally, findings indicate that increased sensory and PHG activity may serve as neural signatures that distinguish true and false recognition even when memory is tested only seconds following initial learning.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank John Jonides, Cindy Lustig and Douglas Noll for their comments on earlier drafts of this paper, as well as Halle Zucker for her assistance with data collection.

This research was supported by NIA Grant R01AG18286 (to P.A.R.-L.) and NIMH Grant F31MH079536 (to A.S.A.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Abe N, Okuda J, Suzuki M, Sasaki H, Matsuda T, Mori E, et al. Neural Correlates of True Memory, False Memory, and Deception. Cerebral Cortex. 2008; 18:2811–2819. [PubMed: 18372290]
- Achim AM, Lepage M. Dorsolateral prefrontal cortex involvement in memory post-retrieval monitoring revealed in both item and associative recognition tests. Neuroimage. 2005; 24:1113– 1121. [PubMed: 15670688]
- Atkins, AS.; Reuter-Lorenz, PA. 'False' working memories: Memory distortion in a mere 4 seconds. Paper presented at the Paper presented at the Psychonomic Society 46th Annual Meeting; Toronto, ON. 2005.
- Atkins AS, Reuter-Lorenz PA. False working memories? Semantic distortion in a mere 4 seconds. Memory & Cognition. 2008; 36:74–81. [PubMed: 18323064]
- Badre D, Wagner AD. Frontal lobe mechanisms that resolve proactive interference. Cerebral Cortex. 2005; 15:2003–2012. [PubMed: 15788702]

- Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia. 2007; 45:2883–2901. [PubMed: 17675110]
- Bedwell JS, Horner MD, Yamanaka K, Li XB, Myrick H, Nahas Z, et al. Functional neuroanatomy of subcomponent cognitive processes involved in verbal working memory. International Journal of Neuroscience. 2005; 115:1017–1032. [PubMed: 16051547]
- Cabeza R. Role of parietal regions in episodic memory retrieval: The dual attentional processes hypothesis. Neuropsychologia. 2008; 46:1813–1827. [PubMed: 18439631]
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. The parietal cortex and episodic memory: an attentional account. Nature Reviews Neuroscience. 2008; 9:613–625.
- Cabeza R, Dolcos F, Graham R, Nyberg L. Similarities and differences in the neural correlates of episodic memory retrieval and working memory. Neuroimage. 2002; 16:317–330. [PubMed: 12030819]
- Cabeza R, Rao SM, Wagner AD, Mayer AR, Schacter DL. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:4805–4810. [PubMed: 11287664]
- Cappell KA, Gmeindl L, Reuter-Lorenz PA. Age differences in prefontal recruitment during verbal working memory maintenance depend on memory load. Cortex. 2010; 46:462–473. [PubMed: 20097332]
- Chein JM, Fiez JA. Dissociation of verbal working memory system components using a delayed serial recall task. Cerebral Cortex. 2001; 11:1003–1014. [PubMed: 11590110]
- Chikazoe J, Jimura K, Hirose S, Yamashita K, Miyashita Y, Konishi S. Preparation to Inhibit a Response Complements Response Inhibition during Performance of a Stop-Signal Task. Journal of Neuroscience. 2009; 29:15870–15877. [PubMed: 20016103]
- Chikazoe J, Konishi S, Asari T, Jimura K, Miyashita Y. Activation of right inferior frontal gyrus during response inhibition across response modalities. Journal of Cognitive Neuroscience. 2007; 19:69–80. [PubMed: 17214564]
- Coane JH, McBride DM, Raulerson BA, Jordan JS. False memory in a short-term memory task. Experimental Psychology. 2007; 54:62–70. [PubMed: 17341016]
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, et al. Temporal dynamics of brain activation during a working memory task. Nature. 1997; 386:604–608. [PubMed: 9121583]
- D'Esposito M, Postle BR, Jonides J, Smith EE. Isolating the neural correlates of maintenance and inhibitory processes using event related fMRI. Journal of Cognitive Neuroscience. 1998:86–86.
- Deese J. On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall. Journal of Experimental Psychology. 1959; 58:17–22. [PubMed: 13664879]
- Flegal K, Atkins AS, Reuter-Lorenz PA. False memories seconds later: The rapid and compelling onset of illusory recognition. Journal of Experimental Psychology-Learning Memory and Cognition. 2010; 36:1331–1338.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved Assessment of Significant Activation in Functional Magnetic-Resonance-Imaging (Fmri) - Use of a Cluster-Size Threshold. Magnetic Resonance in Medicine. 1995; 33:636–647. [PubMed: 7596267]
- Gallo, DA. Associative Illusions of Memory. New York: Psychology Press; 2006.
- Garavan H, Ross TJ, Stein EA. Right hemispheric dominance of inhibitory control: An event-related functional MRI study. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96:8301–8306. [PubMed: 10393989]
- Garoff-Eaton RJ, Slotnick SD, Schacter DL. Not all false memories are created equal: The neural basis of false recognition. Cerebral Cortex. 2006; 16:1645–1652. [PubMed: 16400158]
- Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage. 1999; 9:416–429. [PubMed: 10191170]
- Gold BT, Buckner RL. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron. 2002; 35:803–812. [PubMed: 12194878]
- Hamilton AC, Martin RC. Dissociations among tasks involving inhibition: a single-case study. Cogn Affect Behav Neurosci. 2005; 5:1–13. [PubMed: 15913003]

- Hicks JL, Hancock TW. Backward associative strength determines source attributions given to false memories. Psychonomic Bulletin & Review. 2002; 9:807–815. [PubMed: 12613687]
- Hirshorn EA, Aguirre GK, Thompson-Schill SL. Proactive and reactive cognitive control during semantic retrieval. Journal of Cognitive Neuroscience. 2005:123–123.
- Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17:825–841. [PubMed: 12377157]
- Johnson MK, Nolde SF, Mather M, Kounios J, Schacter DL, Curran T. The similarity of brain activity associated with true and false recognition memory depends on test format. Psychological Science. 1997; 8:250–257.
- Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. Neuroscience. 2006; 139:181–193. [PubMed: 16337090]
- Jonides J, Smith EE, Marshuetz C, Koeppe RA, Reuter-Lorenz PA. Inhibition in verbal working memory revealed by brain activation. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95:8410–8413. [PubMed: 9653200]
- Kan IP, Thompson-Schill SL. Selection from perceptual and conceptual representations. Cognitive, Affective, & Behavioral Neuroscience. 2004; 4:466–482.
- Kim H, Cabeza R. Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. Journal of Neuroscience. 2007; 27:12190–12197. [PubMed: 17989285]
- Kompus K, Eichele T, Hugdahl K, Nyberg L. Multimodal imaging of incidental retrieval: the low route to memory. Journal of Cognitive Neuroscience. (in press).
- Lazar NA, Eddy WF, Genovese CR, Welling J. Statistical issues in fMRI for brain imaging. International Statistical Review. 2001; 69:105–127.
- Lazar NA, Luna B, Sweeney JA, Eddy WF. Combining brains: a survey of methods for statistical pooling of information. Neuroimage. 2002; 16:538–550. [PubMed: 12030836]
- Lieberman MD, Cunningham WA. Type I and Type II error concerns in fMRI research: re-balancing the scale. Social, Cognitive, and Affective Neuroscience. 2009; 4:423–428. [PubMed: 20035017]
- Milham MP, Banich MT. Anterior cingulate cortex: An fMRI analysis of conflict specificity and functional differentiation. Human Brain Mapping. 2005; 25:328–335. [PubMed: 15834861]
- Milham MP, Banich MT, Webb A, Barad V, Cohen NJ, Wszalek T, et al. The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. Cognitive Brain Research. 2001; 12:467–473. [PubMed: 11689307]
- Monsell S. Recency, immediate recognition memory, and reaction-time. Cognitive Psychology. 1978; 10:465–501.
- Nelson JK, Reuter-Lorenz PA, Persson J, Sylvester CYC, Jonides J. Mapping interference resolution across task domains: A shared control process in left inferior frontal gyrus. Brain Research. 2009; 1256:92–100. [PubMed: 19111526]
- Nelson JK, Reuter-Lorenz PA, Sylvester CYC, Jonides J, Smith EE. Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:11171–11175. [PubMed: 12958206]
- Okado Y, Stark C. Neural processing associated with true and false memory retrieval. Cogn Affect Behav Neurosci. 2003; 3:323–334. [PubMed: 15040552]
- Oppenheim, AV.; Schafer, RW.; Buck, JR. Discrete-Time Signal Processing. 2. Upper Saddle River, NJ: Prentice Hall; 1999.
- Oztekin I, Curtis CE, McElree B. The Medial Temporal Lobe and the Left Inferior Prefrontal Cortex Jointly Support Interference Resolution in Verbal Working Memory. Journal of Cognitive Neuroscience. 2009; 21:1967–1979. [PubMed: 18855551]
- Paz-Alonso PM, Ghetti S, Donohue SE, Goodman GS, Bunge SA. Neurodevelopmental correlates of true and false recognition. Cerebral Cortex. 2008; 18:2208–2216. [PubMed: 18203693]
- Persson J, Sylvester CYC, Nelson JK, Welsh KM, Jonides J, Reuter-Lorenz PA. Selection requirements during verb generation: differential recruitment in older and younger adults. Neuroimage. 2004; 23:1382–1390. [PubMed: 15589102]

- Ranganath C, Cohen MX, Brozinsky CJ. Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. Journal of Cognitive Neuroscience. 2005; 17:994–1010. [PubMed: 16102232]
- Ranganath C, Johnson MK, D'Esposito M. Prefrontal activity associated with working memory and episodic long-term memory. Neuropsychologia. 2003; 41:378–389. [PubMed: 12457762]
- Raposo A, Han S, Dobbins IG. Ventrolateral prefrontal cortex and self-initiated semantic elaboration during memory retrieval. Neuropsychologia. 2009; 47:2261–2271. [PubMed: 19038275]
- Roediger HL, Mcdermott KB. Creating False Memories Remembering Words Not Presented in Lists. Journal of Experimental Psychology-Learning Memory and Cognition. 1995; 21:803–814.
- Roediger HL, Watson JM, McDermott KB, Gallo DA. Factors that determine false recall: A multiple regression analysis. Psychonomic Bulletin & Review. 2001; 8:385–407. [PubMed: 11700893]
- Rypma B, D'Esposito M. The roles of prefrontal brain regions in components of working memory: Effects of memory load and individual differences. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96:6558–6563. [PubMed: 10339627]
- Schacter DL, Slotnick SD. The cognitive neuroscience of memory distortion. Neuron. 2004; 44:149. [PubMed: 15450167]
- Simons JS, Koutstaal W, Prince S, Wagner AD, Schacter DL. Neural mechanisms of visual object priming: evidence for perceptual and semantic distinctions in fusiform cortex. Neuroimage. 2003; 19:613–626. [PubMed: 12880792]
- Slotnick SD, Schacter DL. A sensory signature that distinguishes true from false memories. Nat Neurosci. 2004; 7:664–672. [PubMed: 15156146]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23:S208–S219. [PubMed: 15501092]
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94:14792–14797. [PubMed: 9405692]
- Vuilleumier P, Henson RN, Driver J, Dolan RJ. Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nature Neuroscience. 2002; 5:491–499.
- Wagner AD, Shannon BJ, Kahn I, Buckner RL. Parietal lobe contributions to episodic memory retrieval. Trends in Cognitive Sciences. 2005; 9:445–453. [PubMed: 16054861]

Figure 1.

Example trials from the ST-DRM task. *Positive* probes are those that did appear in the memory set. *Unrelated lure*s were not in the memory set, and were unrelated in meaning to items in the memory set. *Related lure* probes were also not in the memory set, but were semantically associated theme words related to items in the memory set.

Atkins and Reuter-Lorenz Page 17

Page 17

Figure 2.

Regions showing increased activity for correct rejection of lure vs. unrelated negative probes. As predicted, we found a large cluster of activation in L VLPFC (BA45) associated with semantic interference.

Figure 3.

Regions common to both true and false memory are displayed in Panel B. Regions that distinguished true from false memory are displayed in Panel C (*red* = True Recognition; *blue* = False Recognition). See Section 3 text for details.

Figure 4.

Mean percent signal change (PSC) in our L VLPFC ROI displayed as a function of trial type (A). Individual differences in L VLPFC activity for lure vs. unrelated negative probes was positively correlated with the RT index of semantic interference (B).

Table 1

Mean proportion of positive, unrelated lure, and related lure probes to which participants responded 'Yes'.

Note- A 'Yes' response indicates that the probe was recognized as a member of the current memory set. The proportion recognized therefore represents the hit rate for *positive* probes, and the false recognition rate for *unrelated* and *related* lures. The mean false memory rate, defined as the difference between false recognition for *unrelated* and *related* lures was .10 (SE=.01).

Author Manuscript

Peak voxels for whole-brain analyses (p<.005, uncorrected, 20 contiguous voxels). Peak voxels for whole-brain analyses (p<.005, uncorrected, 20 contiguous voxels).

