Dissociable correlates of two classes of retrieval processing in prefrontal cortex

Charan Ranganath1,2, Aaron S. Heller1, and Edward L. Wilding3
1 Center for Neuroscience, University of California at Davis
2 Department of Psychology, University of California at Davis
3 School of Psychology, Cardiff University, Wales, UK

Abstract

Although substantial evidence suggests that the prefrontal cortex (PFC) implements processes that are critical for accurate episodic memory judgments, the specific roles of different PFC subregions remain unclear. Here, we used event-related functional magnetic resonance imaging to distinguish between prefrontal activity related to operations that (1) influence processing of retrieval cues based on current task demands, or (2) are involved in monitoring the outputs of retrieval. Fourteen participants studied auditory words spoken by a male or female speaker and completed memory tests in which the stimuli were unstudied foil words and studied words spoken by either the same speaker at study, or the alternate speaker. On “general” test trials, participants were to determine whether each word was studied, regardless of the voice of the speaker, whereas on “specific” test trials, participants were to additionally distinguish between studied words that were spoken in the same voice or a different voice at study. Thus, on specific test trials, participants were explicitly required to attend to voice information in order to evaluate each test item. Anterior (right BA 10), dorsolateral prefrontal (right BA 46), and inferior frontal (bilateral BA 47/12) regions were more active during specific than during general trials. Activation in anterior and dorsolateral PFC was enhanced during specific test trials even in response to unstudied items, suggesting that activation in these regions was related to the differential processing of retrieval cues in the two tasks. In contrast, differences between specific and general test trials in inferior frontal regions (bilateral BA 47/12) were seen only for studied items, suggesting a role for these regions in post-retrieval monitoring processes. Results from this study are consistent with the idea that different PFC subregions implement distinct, but complementary processes that collectively support accurate episodic memory judgments.

Keywords

Memory; Recognition; Prefrontal; Anterior; Lateral; Frontopolar; Cortex; Frontal; Inferior; Dorsolateral; Source; Monitoring; declarative; executive; auditory; control; fMRI; neuroimaging; event-related
Evidence from neuropsychological (Ranganath and Knight, 2003; Shimamura, 1996), neuroimaging (Fletcher and Henson, 2001; Ranganath and Knight, 2003), and electrophysiological (Friedman and Johnson, 2000; Herron and Wilding, 2004; Wilding and Herron, in press) studies suggests that the lateral prefrontal cortex (PFC) implements control processes that support accurate episodic memory judgments. For example, prefrontal damage does not induce an amnesic syndrome, but it impairs the use of strategies that facilitate retrieval when there are few retrieval cues (Gershberg and Shimamura, 1995; Incisa della Rochetta and Milner, 1993; Ranganath and Blumenfeld, in press).

Complementing the neuropsychological evidence described above, results from neuroimaging studies have led to a growing appreciation of the fact that different PFC subregions make different contributions to episodic retrieval processing (Buckner, 2003; Cabeza and Nyberg, 2000; Fletcher and Henson, 2001; Ranganath, 2004). Based on anatomical considerations and patterns of activation in neuroimaging studies, many researchers have distinguished between caudal prefrontal regions lying along the inferior frontal gyrus (at or near Brodmann’s areas [BA] 44, 45, and 47), and more rostral regions in the middle and superior frontal gyri (at or near BA 9, 10, and 46). Initial attempts to characterize the roles of these prefrontal regions in episodic retrieval focused on descriptive distinctions, such as “retrieval success” (Konishi et al., 2000; Rugg et al., 1996) and “retrieval effort” (Buckner et al., 1998a; Buckner et al., 1998b; Kapur et al., 1995). Researchers have also attempted, however, to link neural correlates of memory retrieval in prefrontal regions with processes suggested by psychological models and by event-related potential (ERP) studies of memory retrieval processing (Buckner, 2003; Rugg and Wilding, 2000).

For example, according to strength-based models, memory attributions are made by assessing the memory strength of a given item relative to a decision criterion. When memory strength falls close to a criterion for distinguishing between different classes of items, “retrieval monitoring” operations may be implemented in order to guide decisions (Atkinson and Juola, 1974). Broadly consistent with this idea, ERP studies of source memory for voice information have revealed late-onsetting (e.g., emerging 800ms or later) ERP modulations that differentiate between old and new items (Senkfor and Van Petten, 1998; Wilding, 1999; Wilding and Rugg, 1996). Typically, these late-occurring positive ERP shifts are most apparent over frontal scalp sites, lending support to speculations that prefrontal regions contribute to post-retrieval processing (Buckner et al., 1998a; Rugg et al., 1996; Schacter et al., 1997).

Another kind of memory control process may involve modulations of the processing of retrieval cues in service of task demands (Johnson, 1992; Johnson et al., 1993; Ranganath, 2004; Ranganath et al., 2000). For example, behavioral and ERP research has supported the idea that people can constrain retrieval “at the front end,” by selecting an appropriate task set—a retrieval orientation (Rugg and Wilding, 2000). Adoption of a specific retrieval orientation might influence the initial processing of a cue (e.g., focusing on the auditory features of each word, rather than the meaning) in order to maximize the likelihood that task-relevant information will be recovered. Consistent with this idea, studies of memory for pictures (Ranganath and Paller, 1999, 2000) or visual details of words (Werkle-Bergner et al., 2005) have shown that frontal ERPs to studied and unstudied items are modulated by the degree to which one must retrieve relatively specific details about studied items. Even when the specificity of a memory decision is controlled, ERPs to new items (Dzulkifli et al., 2006; Herron and Wilding, 2006; Hornberger et al., 2004; Hornberger et al., 2006; Johnson et al., 1997; Wilding, 1999) or even to instruction cues (Herron and Wilding, 2004) can differ according to the kind of information that is targeted for retrieval (e.g., retrieving information about an imagery- or a semantically-based encoding task). Critically, ERP correlates of a retrieval orientation can emerge before neural correlates of successful retrieval are apparent (as early as 200ms post-stimulus) (Ranganath and Paller, 1999; Werkle-Bergner et al., 2005). These findings demonstrate that
people can effectively engage in pre-retrieval sets that induce changes in the way the brain evaluates potential retrieval cues (Rugg and Wilding, 2000).

Numerous functional magnetic resonance imaging (fMRI) studies have attempted to specify the processes that might be implemented by different prefrontal regions by contrasting activation between trials that vary in terms of the amount or quality of information that is retrieved. For example, some imaging studies have reported that activation in dorsolateral PFC (particularly right BA 46) is enhanced during processing of studied items that are recognized with low confidence, as compared with studied items that are confidently recognized or new items that are confidently rejected (Henson et al., 2000; Henson et al., 1999a; but see Yonelinas et al., 2005, for a conflicting finding). These results have been interpreted to suggest that right dorsolateral PFC contributes to post-retrieval monitoring, based on the assumption that monitoring processes are disproportionately engaged on low confidence trials. However, activation differences related to response confidence could reflect any number of processes, and it is possible that right dorsolateral PFC and other prefrontal regions contribute more generally by modulating retrieval cue processing based on task demands (Dobbins et al., 2002; Dobbins et al., 2003; Ranganath et al., 2000).

An alternative approach to disentangling neural correlates of retrieval processes is to separately examine neural responses to studied and unstudied (“new”) items in test conditions that vary in cue processing requirements. This approach has been adopted in a number of ERP studies of memory retrieval processing (Dobbins et al., 2003; Johnson et al., 1997; Nolde et al., 1998a; Ranganath et al., 2000; Ranganath and Paller, 1999, 2000), and in some fMRI studies (Hornberger et al., 2004; Nolde et al., 1998a; Ranganath et al., 2000; Rugg et al., 2003). In the present study, we developed a paradigm based on previous ERP studies of memory retrieval (Ranganath and Paller, 1999; Wilding and Rugg, 1997), and used event-related fMRI to identify prefrontal activity patterns related to engagement of retrieval monitoring processes and adoption of a retrieval orientation.

Participants studied a series of words spoken by either a male or female speaker, and were scanned during two different retrieval test conditions. Both test conditions included studied words spoken by the same speaker (“old/same”), studied words spoken by a different speaker (“old/different”), and new words spoken by a male or a female. On “general” test trials, participants were asked to indicate whether each word was on the study list, irrespective of the speaker, whereas on “specific” test trials, participants were asked to indicate whether the word was on the study list and spoken by the same speaker as at test. Thus, both specific and general test trials required participants to differentiate familiar and unfamiliar words, but specific trials additionally required participants to make an attribution about the voice in which each word was spoken at study. The test conditions were blocked in order to encourage participants to adopt a particular retrieval orientation during each type of test, but within each test block, old/same, old/different, and new items were intermixed randomly.

With this design, we were able to examine the degree to which different lateral prefrontal regions exhibited responses consistent with post-retrieval monitoring or task-specific processes that are engaged when retrieval cues are encountered. To identify candidate regions that might be involved in post-retrieval monitoring, we compared activation during processing of old/different items (words previously studied but presented in a different voice at test) between the specific and general tests (see Methods). To identify strategic changes in the processing of a test item, we contrasted activation during processing of new words between the specific and general test conditions.
Materials and Methods

Participants

Fourteen right-handed, neurologically intact, native English speakers (seven females) 20–32 years of age participated in the study. The volunteers were recruited from the University of California at Davis community and were financially compensated for their participation.

Stimuli

The stimuli were 396 recordings of English words spoken by a male or female. These stimuli were subdivided into 6 word lists that were matched on the Kucera-Francis (Kucera and Francis, 1967) word frequency count ($M=45.41$), number of syllables ($M=2.27$), and imageability ratings ($M=509.72$). Stimuli for each trial condition (old/same, old/different, and new items on specific and general test trials; see below) were drawn from one of the word lists, and the mapping of word list to trial condition was randomly varied across participants.

Procedure

Before scanning, participants performed practice study and test trials to ensure they fully understood the instructions for each task. During scanning, participants first performed a visuomotor response task that was used to estimate a subject-specific hemodynamic response function (HRF, cf. Aguirre et al., 1998; Handwerker et al., 2004). Next, they performed alternating runs of study and test trials. Across the scanning session, six study-test cycles were performed. During these trials, auditory stimuli were presented binaurally through pneumatic headphones (Avotec Inc., Stuart, FL.). On each run of study trials, participants fixated on a cross in the center of the screen, and were presented with a list of 44 words. During presentation of each word, participants were instructed to manually respond as to whether the speaker of the word was male (right hand) or female (left hand), and whether the word was pleasant (index finger) or unpleasant (middle finger). This orienting task ensured that participants attended to both the semantic characteristics and the speaker of each word.

During each run of test trials, participants performed one block of “specific” test trials and one block of “general” test trials. The relative order of specific and general test blocks within a run was varied such that on half of the runs specific tests preceded general tests, and on the other half general tests preceded specific tests. Each block of test trials began with an instruction screen that was presented for 2s, followed by a 4s inter-trial interval. Throughout the remainder of each block, the letter “S” or “G” was shown at the center of the screen (signifying a specific or general test trial) and participants were instructed to listen carefully for test words. Each test block included 33 spoken words: 11 previously studied words that were spoken in the same voice at test (“old/same”), 11 previously studied words that were spoken by a different voice at test (“old/different”), and 11 unstudied (“new”) words. Within each test block, old/same, old/different, and new trials were intermixed in a pseudo-random sequence. Thus, within each test block, participants could not predict which item type would be presented on a given trial. The stimulus onset asynchrony (SOA) was varied from 4 to 20s, with a mean SOA of 5.33s. On general test trials, participants were instructed to respond “old” if the word had been presented in the previous study phase (regardless of whether the speaker was the same) and to respond “new” otherwise. On specific test trials, participants were instructed to respond “old” only for studied words that were spoken in the same voice at test, and to respond “new” to studied words spoken by a different speaker, as well as to unstudied words.

MRI Acquisition and Processing

All imaging was performed using a 1.5T, whole-body, neuro-optimized GE Signa Horizon LX NV/I MRI system (GE Medical Systems, Waukesha, WI) at the UC Davis Imaging Research
Center (http://ucdirc.ucdavis.edu/) in Sacramento, California. Functional images were acquired using a T2*-weighted gradient-echo, echoplanar imaging sequence (TR=2s, TE=40 ms, FOV=240mm, 64×64 Matrix, flip angle=90°). Twenty-four contiguous slices were collected in ascending interleaved order with a voxel size of 3.75 × 3.75 × 5.0 mm. Each scanning run began with 20s of image acquisition without stimuli, in order to achieve gradient stabilization. These images were not included in any of the functional analyses. Next, high-resolution anatomical images were acquired using a T1-weighted sequence with TR=27 ms, TE=7 ms, and FA=90° to collect 124 images of 1 mm thickness.

FMRI data pre-processing was performed with Statistical Parametric Mapping (SPM99) software for all participants. EPI images were sinc-interpolated in time to correct for between-slice timing differences in image acquisition and realigned using a six-parameter, rigid-body transformation algorithm. Next, images were spatially normalized to the EPI template image included in the SPM 99 package (which is registered to the Montreal Neurological Institute 305 template), resliced into 3.5 mm isotropic voxels, and spatially smoothed with an 8mm full-width half-maximum Gaussian kernel.

**Data Analysis**

Event-related blood oxygenation level dependent (BOLD) responses were analyzed using a modified general linear model (GLM, Worsley and Friston, 1995), as implemented in the VoxBo software package (www.voxbo.org). Each GLM incorporated empirically-derived estimates of intrinsic temporal autocorrelation (Zarahn et al., 1997) and filters to attenuate frequencies above .25 Hz and below .006 Hz. To account for scaling differences between runs, each voxel’s time series was normalized relative to its mean signal value across the scanning run.

Separate covariates modeled BOLD responses during old/same, old/different, and new trials within the specific and general test conditions. Each of the six covariates was constructed by first creating a vector of neural impulses corresponding to the onset of each event and then convolving the vector with a subject-specific HRF. In addition to these covariates of interest, nuisance covariates were included in the model, including the global signal (orthogonalized with respect to the design matrix, see Desjardins et al., 2001), covariates to model the mean of each scanning run, and an intercept.

Results from single-subject analyses were used to compute contrast images for each comparison of interest (i.e., linear combinations of beta values from the regression analyses described above). Contrast images for each participant were then entered into a second-level group analysis—a one-sample t-test, in which the mean value across the group for each voxel was tested against zero. Significant regions of activation were identified using an uncorrected threshold of \( p < .001 \) and an extent threshold of 8 voxels. Results from mapwise analyses were used to define regions of interest in the right dorsolateral PFC (BA 46), right anterior PFC (BA 10), and bilateral anterior ventrolateral PFC (BA 47/12). Activations in these three regions also survive corrected thresholds for a mapwise false discovery rate (Genovese et al., 2002) of \( p < .05 \).

**Results**

**Behavioral Results**

Mean test phase accuracy and RT data for each item type (old/same vs. old/different vs. new) and test condition (specific vs. general) are shown in Table 1. Participants were extremely accurate at quickly differentiating between old and new items, but were slower and less accurate at making voice identity decisions for studied items on the specific test than on the general test.
ANOVA on the accuracy data revealed significant main effects of Item Type \(F(2,26) = 16.26; p < .0001\) and Test Condition \(F(1,13) = 27.35; p < .0001\) that were qualified by a significant interaction between these factors \(F(1.25,16.25) = 19.77; p < .0001\). Because participants had to make a more fine-grained decision on specific test trials, as compared with general test trials, we were interested in characterizing behavioral differences for each stimulus class between the two test conditions. Follow-up analyses determined that accuracy was significantly lower in the specific test than the general test for old/same \(t(13) = 4.69, p < .0005\) and for old/different \(t(13) = 4.93, p < .0005\) items. In contrast, no significant accuracy differences were seen for new items \(t(13) < 1\).

Due to an error in linking the response device to the appropriate timer, RT data were not available for one participant. Analyses of RT data from the remaining participants mirrored the accuracy analyses described above. An ANOVA revealed significant effects of Item Type \(F(2,24) = 24.6; p < .0001\) and Test Condition \(F(1,12) = 109.29; p < .0001\). These effects were qualified by a significant interaction between the two factors \(F(2,24) = 43.72; p < .0001\). Follow-up analyses determined that RTs were significantly slower in the specific test than the general test for old/same \(t(12) = 7.85, p < 5 \times 10^{-6}\) and old/different \(t(12) = 9.52, p < 5 \times 10^{-7}\) items. However, no significant RT difference was observed for new items \(t(12) < 1\).

Collectively, these behavioral data indicate that: (1) for previously studied items, participants were faster and more accurate at making item recognition decisions (general test trials) than at making voice identity decisions (specific test trials), and (2) accuracy and RTs for new items were equivalent across the two test conditions. This pattern of results was critical for guiding the selection of fMRI contrasts, as described below.

**fMRI Results**

As noted earlier, the present study allowed us to test the extent to which activity in prefrontal regions is sensitive to adoption of a specific retrieval orientation and/or post-retrieval monitoring. We ran two targeted contrasts to address these questions, the results of which are summarized in Table 2 and Figure 1. The first analysis contrasted activation during processing of new words between the specific and general test conditions. Under the assumption that new items are unlikely to elicit successful retrieval, any activity differences between specific and general test conditions for these items should presumably reflect strategic changes in the processing of a potential retrieval cue (i.e., the consequences of having selected/adopted different retrieval orientations). As shown in Figure 1A, this contrast revealed activation in right anterior PFC (BA 10, see Ongur et al., 2003).

Our second contrast identified candidate regions that might be involved in post-retrieval monitoring by comparing activation during processing of old/different items (words previously studied but presented in a different voice at test) between the specific and general tests. Our reasoning is that participants were extremely accurate at detecting that old/different words were familiar (as evidenced by accuracy data for old/different items on the general test), but they were near criterion at detecting the difference in voice between study and test on specific test trials (as evidenced by the low accuracy and slow RTs for these trials). Accordingly, it is reasonable to suggest that, for old/different items, participants engaged post-retrieval monitoring processes to a greater extent on specific test trials as compared with general test trials (Atkinson and Juola, 1974). As shown in Table 2 and Figure 1B, this contrast revealed significant activation changes in the right dorsolateral PFC (BA 46, see Rajkowska and Goldman-Rakic, 1995) and bilaterally in the anterior ventrolateral PFC (BA 47/12, see Ongur et al., 2003).

The mapwise analyses described above relied on theory-driven assumptions regarding the engagement of memory control operations across the different test conditions. However, it is
possible that the full pattern of activation in regions identified in these contrasts might not correspond to our *a priori* theoretical assumptions. Furthermore, even though right anterior PFC (BA 10) was activated in a different contrast than right dorsolateral (BA 46) and bilateral ventrolateral (BA 47/12) PFC, it is not clear whether these regions exhibited qualitatively different response profiles. Indeed, such a dissociation might emerge artifactually from statistical thresholding, even if the three regions exhibited qualitatively similar responses.

To address these concerns, we conducted a series of region-of-interest (ROI) analyses to fully characterize the nature of activity in these regions. The results from the retrieval orientation contrast were used to define one ROI in right BA 10 and results from the post-retrieval monitoring contrast were used to define ROIs in bilateral BA 47/12 and right BA 46 (see Methods). Parameter estimates were extracted from these ROIs indexing activation for each trial type for each participant (Figure 2), and two types of analyses were conducted. The first set of analyses aimed to specify the pattern of results across all trial types within each ROI, whereas the second set aimed to determine whether the three ROIs exhibited qualitatively different patterns of activation\(^1\).

As noted above, right BA 10 was identified in the contrast designed to target activations associated with retrieval orientation. To determine whether this characterization was accurate, parameter estimates extracted from this ROI were submitted to an ANOVA with Test Condition (Specific vs. General) and Item Type (Old/Same vs. Old/Different vs. New) as factors. Activation in this region was significantly greater during specific than during general \([F(1,13) = 11.23; p = .005]\) test trials, and this effect was qualified by a significant Item Type × Test Condition interaction \([F(1.99,25.8) = 3.71; p = .039]\). Follow-up analyses revealed that activity was generally larger on specific than on general trials, with the differences being significant for old/same \([t(13)=2.77, p=.016]\) and for new \([t(13)=4.95, p<.001]\) items, and approaching significance for old/different \([t(13)=1.97, p=.070]\) items (see Fig. 2). These analyses are consistent with a role for this region in modulating retrieval cue processing, possibly related to the adoption of a retrieval orientation.

Right BA 46 was identified in a contrast designed to target post-retrieval monitoring processes. Further analyses, however, revealed a more complex pattern of activation in this region. An ANOVA revealed a main effect of Item Type \([F(1.23,15.9) = 4.61; p = .041]\) and Test Condition \([F(1,13) = 14.52; p = .002]\), but no significant Item Type × Test Condition interaction was observed \([F(1.79,23.3) = 1.86; p > 0.10]\). Follow-up tests revealed that activity was marginally greater for old/same than for new items \([t(13)=2.05, p=.06]\), but no significant differences were observed between old/different items and new or old/same items [both t’s<1]. Comparisons between test conditions revealed that BA 46 activation was increased during specific test trials, as compared with general test trials, for new \([t(13)=2.27, p=.04]\), old/same \([t(13)=3.69, p=.002]\), and old/different \([t(13)=5.59, p<.001]\) items (see Figure 2). The outcomes of these analyses suggest that activity in right BA 46 was relatively sensitive to task demands, but these modulations did not differ between old and new items. These results are not consistent with the idea that right BA 46 is engaged in post-retrieval monitoring processes.

As noted above, BA 47/12 was also identified in the contrast related to post-retrieval monitoring, but activation in this ROI differed in a number of respects from what was observed in right BA 46. An initial ANOVA revealed significant main effects of Test Condition \([F(1,13) = 5.77; p = .032]\), Item Type \([F(1.77,23) = 5.30; p = .015]\), and a significant Item Type × Test Condition interaction \([F(2,26) = 12.99; p < .001]\). To clarify the patterns of results in each

---

\(^1\)The outcomes of targeted ROI analyses for BA 11 identified in the (new/specific)-(new/general) contrast are analogous to those described for right BA 10. The outcomes of targeted ROI analyses for the region of right BA 9 identified in the (old/different specific)-(old/different general) contrast are analogous to those described for right BA 46. These results can be obtained from the first author on request.
region, we ran follow-up analyses contrasting activation between the two test conditions separately for each item type. As shown in Figure 2, BA 47/12 generally showed specific-general activation differences for studied items [old/same: t(13)=1.87, p=.084; old/different: t (13)=5.13, p<.001] but not for new items (t(13)<1). These results suggest that BA 47/12 may implement processes that are specifically relevant for post-retrieval monitoring.

The analyses described above revealed similar patterns of results for right BA 10 and BA 46, in that activation was generally higher on specific test trials than on general test trials. In contrast, bilateral BA 47/12 only exhibited between-test activation differences for studied items. We next conducted an analysis to determine whether there were qualitatively distinct patterns of retrieval-related activation differences across the three ROIs. In this analysis, the differences between parameter estimates on specific and general test trials were computed for old/same, old/different, and new items separately for each ROI and for each participant. These results were then submitted to ANOVAs with Region and Item Type as factors. This analysis revealed that the pattern of between-test activation differences for each trial type differed qualitatively between right BA 10 and BA 47/12 [F(2,26)=5.64, p=.02]. The analysis comparing right BA 46 and BA 47/12 also revealed a marginal Region × Item Type interaction [F(2,26)=3.10, p=.091]. In contrast, no qualitative differences were observed in the analysis comparing right BA 46 and right BA 10 [F(2, 26)<1].

**Discussion**

In the present study, we examined the neural correlates of memory control operations by comparing prefrontal activation to studied and unstudied items across two retrieval tests that involved more specific or more generalized memory decisions. Our results revealed a dissociation between neural correlates of two different classes of retrieval processing within the PFC. Right anterior (BA 10) and dorsolateral (BA 46) prefrontal areas exhibited more activation during specific test trials than during general test trials, irrespective of whether the retrieval cue was an old or a new item. In contrast, bilateral BA 47/12 exhibited activation differences between the two test conditions only when studied items were presented as retrieval cues. This region showed increased activation during processing of studied items on specific test trials as compared to general test trials, whereas activation during processing of new items did not differ between the two test conditions. We comment on these findings in more detail below.

**Potential limitations of the current paradigm**

The primary motivation for this study was to examine control processes that may influence the likelihood of recovery of source information, or the way that cues are processed in service of task demands. Our paradigm was designed to vary the engagement of these control processes between the two test conditions, but it is important to consider some potential limitations in the current paradigm that might affect interpretation of the FMRI results. One issue, inherent in any comparison of different test conditions, is the degree to which different processes were engaged in the two tests. For instance, it is possible that participants sometimes incidentally retrieved specific voice information on general test trials, even though it was not necessary. Critically, this would be expected to minimize differences in activity between the two test conditions, but it could not drive the significant between-test activation differences that were observed in the PFC.

A second issue to consider is that the engagement of memory control operations might vary between old and new items. For instance, to the extent that participants engaged in a specific retrieval orientation, such operations would likely be maintained during the period in which task-relevant contextual information becomes available. This period is longer for old than for new items, as in the latter case, little if any contextual information will be recovered. This
proposal might explain why BA 46 tended to show increased activation for old items, relative to new items (Fig. 2). A related possibility is that when processing familiar items on general test trials, participants adopted a retrieval orientation towards recovery of voice information. In this event, we might expect differential activation between old and new items in the general test condition, as was observed in right BA 10 (Fig. 2).

Another potential concern is the possibility that activation differences between specific and general trials were secondary to differences in accuracy and/or RTs between the two test conditions. Critically, although accuracy and RTs for old items differed between the two test conditions, behavioral results for new items were similar across the two tests. Nonetheless, activation in BA 10 and BA 46 was increased during specific trials, relative to general trials, for both old and new items (Figure 2). Thus, it is unlikely that, in these regions, differential activation between the test conditions can be attributed to performance differences. In BA 47/12, however, activation only differed between the two test conditions for old items. Furthermore, the activation patterns tracked accuracy and RT data, such that between-test activation differences were largest for old/different items (Figure 2). Thus, it is possible that activation in this region reflected a non-specific process that correlates with task difficulty, rather than monitoring per se. However, engagement of monitoring processes is usually correlated with task difficulty (Fletcher and Henson, 2001), and it is therefore challenging to disentangle the two.

Neural evidence for dissociable retrieval processes in PFC

As noted above, our results revealed a qualitative difference in activation patterns between three regions in PFC. Within anterior ventrolateral PFC (BA 47/12), we observed increased activation during processing of old items on the specific test, as compared with the general test. This activation might have reflected the engagement of monitoring processes, under the assumption that such processes are preferentially engaged when the strength of an item falls close to a criterion for distinguishing between different classes of items (Henson et al., 2000; Henson et al., 1999a; Henson et al., 1999b). Based on the behavioral data shown in Table 1, we would expect the effect to be the largest for old/different items, and this is in fact what we observed (Figure 2). Other studies have also reported activation in anterior ventrolateral PFC associated with post-retrieval monitoring processes (Henson et al., 1999b; Rugg et al., 2003). However, as noted above, the activation in anterior ventrolateral PFC could also be explained in terms of the engagement of other control processes that are more generally recruited during difficult tasks. Alternatively, it is conceivable that anterior ventrolateral PFC activity directly reflects the successful retrieval of contextual (voice) information. This alternative account, however, runs counter to the widely held view that the PFC implements control processes that influence the reactivation of representations in more posterior neocortical regions (Fuster, 1997; Miller, 2000; Ranganath, 2006; Ranganath and Blumenfeld, in press; Ranganath and Knight, 2003).

Another hypothesis that has been advanced is that anterior ventrolateral PFC is involved in the controlled retrieval of semantic information that contributes to “cue specification” (Dobbins et al., 2002). In that study, anterior ventrolateral PFC showed increased activation during semantic encoding tasks and during a forced-choice source memory task relative to a fixation baseline. However, activation during old-new recognition did not differ from fixation (Dobbins et al., 2002). The authors interpreted this activation pattern to suggest a role for anterior ventrolateral PFC in controlled semantic retrieval, but this interpretation relies on the tenuous assumption that participants consistently engaged in controlled semantic retrieval on source memory trials and never on old-new recognition memory trials. In the present study, changes in retrieval cue processing on specific test trials presumably reflected the need to orient towards the perceptual, rather than the semantic attributes of retrieval cues. Accordingly, our data run
counter to what one would expect if anterior ventrolateral PFC were solely engaged in semantic retrieval processes that support cue specification.

In contrast to BA 47/12, we found that activation in BA10 and BA 46 differed between the two test conditions for new and old items. As noted earlier, accuracy and RTs to new items were nearly identical between the two test conditions (Table 1), so between-test activation differences seen in these two regions cannot be attributed to nonspecific effects related to time-on-task or difficulty. Given that the specific test primarily required additional processing for old items, it seems unlikely that, on these trials, participants would devote additional resources to new items late in the retrieval process. Furthermore, participants were highly accurate at rejecting new items (see performance on general test), suggesting that there was little need to dedicate additional processing to items judged new. These observations argue for the view that the differential activation across tasks (but comparable across old and new items within tasks) reflects a change in processes that are engaged prior to retrieval, and perhaps in pursuit of task-appropriate information. For example, activation in right anterior and dorsolateral PFC may be correlated with the adoption of a retrieval orientation.

As noted earlier, activation in right dorsolateral PFC (BA 46) is often attributed to retrieval monitoring (e.g., Henson et al., 2000; Henson et al., 1999a; Henson et al., 1999b; Rugg et al., 1999; Rugg et al., 2003). Perhaps the strongest evidence in favor of the monitoring hypothesis came from a study in which activation elicited by old and new items was compared between tests that required either item recognition or source memory decisions (Rugg et al., 2003). In that study, right dorsolateral prefrontal activation was significantly enhanced during processing of old items in the source test as compared with the item test. Although not directly tested, activity in this region on new trials did not seem to differ across test conditions, prompting the authors to conclude that, “this region supports post-retrieval monitoring of retrieved information.” The location of activation in BA 46 reported by Rugg et al. (x=48, y=42, z=24 mm) is very close to the region identified in the present study.

Although the present study showed that right dorsolateral prefrontal activation was enhanced for old items in the source memory test, a similar increase was also observed for new items. This finding contradicts the post-retrieval monitoring hypothesis and is at odds with the findings of Rugg and colleagues (Rugg et al., 2003). One way to resolve these discrepancies, however, stems from consideration of the different retrieval demands in the two studies, and the observation that people flexibly engage retrieval processes depending upon the particular testing conditions. In Rugg et al. (2003), study items were presented in red or green, and all test items – to which either source or recognition memory judgments were required – were presented in black. In the present study, by contrast, 50% of old test items were “copy cues” (i.e., physically identical to the items that were studied). Thus, only in the present study did retrieval cues carry information germane to the source judgment (same/different voice) that was required. It is therefore possible that the presence of copy cues in the present study encouraged participants to specify the processing of all test items in order to make source judgments. In contrast, the absence of task-relevant contextual (color) information in the Rugg et al. study might have encouraged the adoption of a strategy whereby cue specification processes relevant to source (in this case color) judgments were engaged only after an item was identified as studied.

This explanation reconciles the disparate findings described above, providing a common functional interpretation, but locating it at different retrieval processing stages across the different experiments. The key assumption is that, if retrieval cues sometimes contain information that is diagnostic for a specific decision, this will encourage participants to engage in cue specification processes for all test stimuli. When cues are never diagnostic for source decisions, a somewhat more restrictive approach will be employed. If this assumption is correct,
then one should see right dorsolateral and anterior PFC activation increase during processing of new items in a source memory test with copy cues (e.g., a source recognition test), relative to a test in which there are no copy cues (e.g., a source recall test). This prediction could be tested in a future experiment.

Activation in right anterior PFC (BA 10) has been attributed to “retrieval mode”—a cognitive set that ensures subsequent stimuli will be treated as cues for episodic retrieval (Duzel et al., 1999; Lepage et al., 2000). This claim is open to dispute, however, as it has been drawn primarily on the basis of findings in PET studies of episodic retrieval. In PET studies, it is difficult to distinguish unequivocally between set-related activity (e.g. retrieval mode) that is maintained throughout a task, and activity that is initiated in response to items presented during the task (Rugg & Wilding, 2000). Velanova and colleagues, however, demonstrated that right frontopolar activation was sustained across blocks of trials during recognition memory tasks (Velanova et al., 2003). The design they employed is not susceptible to the criticisms directed at the PET studies described above. This finding suggests that right anterior PFC is involved in sustained (“set-related”) processing during episodic retrieval tasks.

Because the current findings were obtained in an event-related fMRI study, they converge with others (Dobbins et al., 2002; Dobbins et al., 2003; Nolde et al., 1998a; Rugg et al., 2003; Simons et al., 2006; Simons et al., 2005a; Simons et al., 2005b) in suggesting that right anterior PFC can also be engaged in response to stimuli presented during retrieval tasks, and that the processes supported by the region are engaged flexibly according to specific retrieval demands. Ascribing both set-related and item-related processing operations to anterior PFC is not incompatible, but the functional relationship between these two classes of retrieval processing remains to be specified (Herron & Wilding, 2005).

We note that results from recent studies (Gilbert et al., 2006; Simons et al., 2006; Simons et al., 2005a; Simons et al., 2005b) suggest that there may be important functional differences between lateral and medial regions of anterior PFC (BA 10). According to one proposal, medial regions of anterior PFC may specifically contribute to control processes relevant to making attributions regarding internally-generated information about a previous event, whereas lateral regions may be more generally involved in retrieval cue processing (Simons et al., 2006; Simons et al., 2005a; Simons et al., 2005b). Our study could not directly test this idea, because it only required attributions about external stimuli. Nonetheless, the fact that between-test activation differences were observed in lateral and not medial anterior PFC is consistent with this distinction.

Other recent results have suggested that there may be functional differences between memory control operations implemented by prefrontal regions in the left and right hemispheres. For example, in many previous studies (e.g., Dobbins et al., 2002; Lundstrom et al., 2003; Mitchell et al., 2004; Nolde et al., 1998a; Ranganath et al., 2000), contrasts of activity between source memory and item memory tests have yielded relatively left-lateralized activation (see Nolde et al., 1998b; Ranganath, 2004, for review). However, this pattern has not been observed in all such comparisons (Cabeza et al., 2003), as some kinds of source memory tests, such as recency judgments, tend to elicit right-lateralized activation (Cabeza et al., 2000; Cabeza et al., 1997; Dobbins et al., 2003). In the present study, we did not observe evidence for qualitative differences between left and right PFC recruitment—although anterior and dorsolateral prefrontal activations were observed in the right hemisphere, qualitatively similar, but subthreshold effects were seen in the left hemisphere. However, our study was not designed to test hypotheses about laterality in the PFC (Cabeza et al., 2003; Dobbins et al., 2004; Habib et al., 2003; Mitchell et al., 2004; Nolde et al., 1998b), and more focused research will be necessary to investigate the relationship between laterality of prefrontal activation and memory control operations.
Prefrontal contributions to mnemonic and nonmnemonic processing

Although right frontopolar and dorsolateral PFC probably implement control processes that contribute to accurate memory retrieval, it is likely that these regions contribute to other domains as well (Cabeza and Nyberg, 2000; Fletcher and Henson, 2001; Ranganath et al., 2003; Ranganath and Knight, 2003). For example, in one recent study, right anterior and dorsolateral PFC activation was associated with decision-making processes in both an episodic memory and a lexical decision task (Dobbins and Han, 2005). In the episodic memory task, participants studied words and were tested with pairs of words. On “forced-choice” trials, participants were instructed to determine which of the items was studied, whereas on “same-different” trials, participants were instructed to determine whether the two items were of the same class (i.e., both old or both new). Even when the content of the test trials was equated, right frontopolar and dorsolateral prefrontal activation was increased during the same-different judgment as compared with the forced-choice judgment. A very similar result was observed during a lexical decision task that did not have an episodic retrieval component, prompting the authors to suggest that these regions implement domain-general control processes. Based on the assumptions that underlie signal detection theory, the authors argued that the same-different judgment required participants to classify the two items under separate decision criteria and then use the outcomes of each classification to make a final decision. Consequently, they reasoned that frontopolar and dorsolateral PFC activation during episodic retrieval might more generally reflect the need to assess the outcomes of multiple cognitive operations (Braver and Bongiolatti, 2002; Ramnani and Owen, 2004).

Although their hypothesis is not quite consistent with the present results (participants made a similar yes-no decision in both test conditions), the results of Dobbins and Han, along with those of other imaging studies (Braver et al., 2001; Cabeza et al., 2002; Ranganath et al., 2003), suggest the need to frame theories of prefrontal contributions to memory retrieval in the context of a more domain-general framework (e.g., Fletcher and Henson, 2001). For example, one recent model proposes a unified framework to explain lateral prefrontal activation during episodic memory encoding, retrieval, working memory, and other non-memory tasks (Blumenfeld and Ranganath, in press; Ranganath and Blumenfeld, in press). Building on the widely held premise that the PFC plays a role in accentuating or inhibiting the activation of posterior cortical representations (Cohen and Servan-Schreiber, 1992; Curtis and D’Esposito, 2003; Fuster, 1997; Goldman-Rakic, 1987; Kimberg et al., 1997; Miller, 2000; Thompson-Schill et al., 1997; Wagner, 1999), this framework suggests that, within lateral PFC, there is a rostro-caudal hierarchy (Christoff and Gabrieli, 2000; Fuster, 2004; Koechlin et al., 2003).

More caudal prefrontal regions (e.g., ventrolateral PFC) may be critical for resolving conflict by activating or inhibiting representations of competing stimuli or responses (Jonides and Nee, 2006; Thompson-Schill et al., 1997; Wagner et al., 2001). More rostral PFC subregions (e.g., dorsolateral and anterior PFC) may activate or inhibit representations of cognitive sets or abstract rules that specify the stimulus-response mappings that are currently relevant (Bunge et al., 2003; Sakai and Passingham, 2003). According to this framework, ventrolateral PFC subregions should contribute to conflict resolution when information retrieved from memory does not strongly specify a particular response. Anterior and possibly dorsolateral prefrontal regions should activate relevant cognitive sets in order to determine how a retrieval cue should be specified and how to weight different kinds of evidence in order to make accurate decisions. Thus, different prefrontal regions may implement complementary control processes that support goal-directed memory retrieval processing.

Acknowledgments

This research was supported by NIH grants R01 MH68721 and PO1 NS40813, and by grants from the UC Davis Academic Senate and Alzheimer’s Disease Center. Dr. Edward Wilding was supported by the Wellcome Trust and a
travel award from Cardiff University. We thank Dr. Karen Mitchell and an anonymous reviewer for their helpful suggestions and Ms. Silvia Rancilio for inspiring this study.

References


Blumenfeld RS, Ranganath C. Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. The Neuroscientist, in press


Neuroimage. Author manuscript; available in PMC 2009 July 7.


Rajkowska G, Goldman-Rakic PS. Cytoarchitectonic definition of prefrontal areas in the normal human cortex: II. Variability in locations of areas 9 and 46 and relationship to the Talairach coordinate system. Cerebral Cortex 1995;5:323–337. [PubMed: 7580125]


Figure 1.
Regions showing increased activation during specific test trials, relative to general test trials, during processing of (A) new items and (B) old/different items. Boxes highlight the locations of regions that were interrogated in ROI analyses (right BA 10, right BA 46, and bilateral BA 47/12).
Figure 2.
Patterns of activation in right anterior, right dorsolateral, and bilateral ventrolateral PFC. Mean parameter estimates from each ROI during the specific (gray bars) and general (open bars) tests are separately plotted for old/same, old/different, and new items. Error bars depict the standard error of the mean.
### Table 1

#### Behavioral results.

<table>
<thead>
<tr>
<th></th>
<th>Old/Same</th>
<th>Old/Different</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Proportion Correct</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Test</td>
<td>0.85 (0.07)</td>
<td>0.63 (0.18)</td>
<td>0.90 (0.10)</td>
</tr>
<tr>
<td>General Test</td>
<td>0.92 (0.04)</td>
<td>0.87 (0.07)</td>
<td>0.89 (0.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Old/Same</th>
<th>Old/Different</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B. Reaction Time (ms)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Test</td>
<td>1741.52 (156.99)</td>
<td>2056.97 (269.21)</td>
<td>1634.41 (195.41)</td>
</tr>
<tr>
<td>General Test</td>
<td>1410.6 (158.73)</td>
<td>1479.1 (166.49)</td>
<td>1628.14 (187.45)</td>
</tr>
</tbody>
</table>

Note: Standard deviation is shown in parentheses.
Table 2

Local maxima of activated regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>t(13)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New: specific - general</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. Middle Frontal Gyrus</td>
<td>6</td>
<td>5.71</td>
<td>32</td>
<td>7</td>
<td>66</td>
</tr>
<tr>
<td>R. Superior Frontal Gyrus</td>
<td>10</td>
<td>5.16</td>
<td>7</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>R. Frontomarginal Gyrus</td>
<td>11</td>
<td>4.81</td>
<td>18</td>
<td>63</td>
<td>−14</td>
</tr>
<tr>
<td>R. Lateral Orbital Gyrus</td>
<td>11</td>
<td>4.39</td>
<td>24</td>
<td>56</td>
<td>−14</td>
</tr>
<tr>
<td><strong>Old different: specific - general</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate Sulcus</td>
<td>32</td>
<td>9.75</td>
<td>7</td>
<td>14</td>
<td>46</td>
</tr>
<tr>
<td>L. Superior Frontal Gyrus</td>
<td>6</td>
<td>5.04</td>
<td>−10</td>
<td>7</td>
<td>52</td>
</tr>
<tr>
<td>L. Superior Frontal Gyrus</td>
<td>6</td>
<td>4.89</td>
<td>−10</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>Precuneus</td>
<td>7</td>
<td>8.15</td>
<td>−7</td>
<td>−70</td>
<td>38</td>
</tr>
<tr>
<td>Precuneus</td>
<td>18</td>
<td>6.53</td>
<td>−7</td>
<td>−70</td>
<td>28</td>
</tr>
<tr>
<td>Precuneus</td>
<td>18</td>
<td>4.33</td>
<td>7</td>
<td>−70</td>
<td>28</td>
</tr>
<tr>
<td>R. Middle Frontal Gyrus</td>
<td>46</td>
<td>6.11</td>
<td>46</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>R. Middle Temporal Gyrus</td>
<td>21</td>
<td>5.76</td>
<td>49</td>
<td>−46</td>
<td>0</td>
</tr>
<tr>
<td>R. Anterior Inferior Frontal Gyrus</td>
<td>47/12</td>
<td>5.48</td>
<td>32</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>L. Anterior Inferior Frontal Gyrus</td>
<td>47/12</td>
<td>5.4</td>
<td>−35</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>L. Superior Frontal Sulcus</td>
<td>6</td>
<td>4.83</td>
<td>−32</td>
<td>−7</td>
<td>63</td>
</tr>
<tr>
<td>L. Superior Frontal Sulcus</td>
<td>6</td>
<td>4.15</td>
<td>−32</td>
<td>−18</td>
<td>60</td>
</tr>
<tr>
<td>R. Middle Frontal Gyrus</td>
<td>9</td>
<td>4.71</td>
<td>28</td>
<td>10</td>
<td>38</td>
</tr>
</tbody>
</table>

Note: R=Right; L=Left.