Prefrontal Cortex Activity Associated with Source Monitoring in a Working Memory Task

**Karen J. Mitchell, Marcia K. Johnson, Carol L. Raye,
and Erich J. Greene**

Abstract

■ Using functional magnetic resonance imaging (fMRI), we investigated prefrontal cortex (PFC) activity during remembering specific source information (format, location judgments) versus remembering that could be based on undifferentiated information, such as familiarity (old/new recognition [ON], recency judgments). A working memory (WM) paradigm with an immediate test yielded greater activation in the lateral PFC for format and location source memory (SM) tasks than ON recognition; this SM-related activity was left lateralized. The same regions of PFC were recruited in Experiment 2 when information was tested immediately and after a filled delay. Substituting recency for location judgments (Experiment 3) resulted in an overall shift in task context that produced greater right PFC activity associated with ON and recency tasks compared to the format task, in addition to left SM-related activity. These data extend to WM previous findings from long-term memory (LTM) indicating that the left and right PFC may be differentially involved in memory attributions depending on the specificity of information evaluated. The findings also provide evidence for the continuity of evaluative processes recruited in WM and LTM. ■

INTRODUCTION

The term *source memory* (SM) refers to memory for aspects of an event that help differentiate it from other events. According to the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), during remembering, activated source information contributes to an attribution that a mental experience is an episodic memory, as opposed to, for example, knowledge or a belief. These remembered and inferred details can include perceptual information, such as color or size, spatiotemporal details, how we felt or what we thought at the time, and so on (i.e., SM subsumes memory for context). Source attributions can also be based on undifferentiated information, such as familiarity or recency. Both specific and undifferentiated features are experienced on continua. Specific information varies in vividness, a concept that typically encompasses both how intense or clear and how rich or embellished mental experiences are (e.g., ranging from just a dim recollection of a car that was some shade of red to the experience of ''seeing again'' a bright fire-engine red convertible, or from remembering that the speaker was a man to remembering that it was Tony; Dodson, Holland, & Shimamura, 1998). Undifferentiated mental experiences, such as familiarity or recency, vary in strength. The processes involved in making attributions about the origins of mental experiences (source moni-

toring) capitalize on average differences in specific and less differentiated information characteristic of different types of events. *Monitoring* encompasses processes involved in the revival and evaluation of information. Both revival processes (e.g., reactivation and retrieval) and evaluation processes (e.g., is this information characteristic of Source A?) contribute to performance in the typical study of long-term memory (LTM).

Specific and undifferentiated information are sometimes referred to as *recollection* and *familiarity*, respectively (Yonelinas, 1999; Jacoby, 1991). These terms are descriptively useful as shorthand. However, the underlying assumption in some recollection/familiarity models is that recollection is all-or-none and familiarity varies in degree (Yonelinas, 1999) or that both can be modeled with all-or-none assumptions (discrimination and detection in multinomial models, e.g., Batchelder & Riefer, 1990). Such all-or-none assumptions are contrary to the assumption in the source monitoring framework (Mitchell & Johnson, 2000; Johnson et al., 1993) that recollection, like familiarity, occurs in varying degrees (e.g., Qin, Raye, Johnson, & Mitchell, 2001). A formal two-dimensional signal detection model (STREAK) that makes assumptions consistent with the source monitoring framework has recently been described by Rotello, Macmillan, and Reeder (in press). It should be noted that according to the source monitoring framework, undifferentiated information is not necessarily just weak specific information, for example, one can have a strong

Yale University

feeling of familiarity accompanied by no recollection or only a faint recollection of some specific detail(s). Although the relation between familiarity and more specific information is an important focus of interest in the memory literature (see also, e.g., Banks, 2000), it is beyond the scope of this article.

Empirically, SM is often contrasted with item recognition memory (e.g., "Where did you see this?" vs. "Did you see this?"). Again, this is a useful dichotomy, although, strictly speaking, item recognition as explored in laboratory tasks typically requires some source information (e.g., say "yes" to words from the list just presented; e.g., Johnson et al., 1993; Anderson & Bower, 1974). More generally, item recognition often involves some mix of more specific and less differentiated information (e.g., Jacoby, 1991; Mandler, 1980). Still, contrasting SM and old/new (ON) recognition tasks provides a way of manipulating the amount of specific (episodic) information required.

SM has been investigated over several decades using a variety of cognitive behavioral paradigms, and there is a rich set of theoretical ideas and empirical findings from this domain clarifying the nature of episodic memory (see, e.g., Hintzman, 2000; Kelley & Jacoby, 2000; Mitchell & Johnson, 2000; Johnson et al., 1993, for reviews). For example, most theories of episodic memory agree that remembering involves both heuristic and systematic (or more automatic and more controlled) processes. Heuristic and systematic are relative terms that apply to cognitive processes in general. With respect to remembering, for example, retrieval (which may involve self-cueing) is more systematic than reactivation (a relatively automatic response to cues, e.g., Johnson, 1992; Moscovitch, 1989). Undifferentiated information, such as familiarity or recency, although it could be processed systematically, is typically more likely to be processed heuristically, and specific detail often requires relatively more systematic processing. Studies exploring the mechanisms underlying true and false memories (Mitchell & Johnson, 2000; Johnson & Raye, 1981) have demonstrated that shifting the information people monitor during remembering (e.g., from less differentiated to more specific) can influence the likelihood of source misattributions (e.g., Marsh & Hicks, 1998; Multhaup, 1995; Jacoby, Kelley, Brown, & Jasechko, 1989; Lindsay & Johnson, 1989).

More recently, studies using neuroimaging have begun to explore the neural correlates of SM (e.g., Slotnick, Moo, Segal, & Hart, 2003; Cansino, Maquet, Dolan, & Rugg, 2002; Dobbins, Foley, Schacter, & Wagner, 2002; Dobbins, Rice, Wagner, & Schacter, 2003; Konishi et al., 2002; Suzuki et al., 2002; Ranganath, Johnson, & D'Esposito, 2000; Raye, Johnson, Mitchell, Nolde, & D'Esposito, 2000; Henson, Shallice, & Dolan, 1999; Rugg, Fletcher, Chua, & Dolan, 1999; Nolde, Johnson, & D'Esposito, 1998).[1] These studies have tended to focus on the prefrontal cortex (PFC) as a region of particular

interest, in part because of neuropsychological evidence that damage to the PFC disrupts SM more than ON recognition memory (e.g., Shimamura, Janowsky, & Squire, 1990).
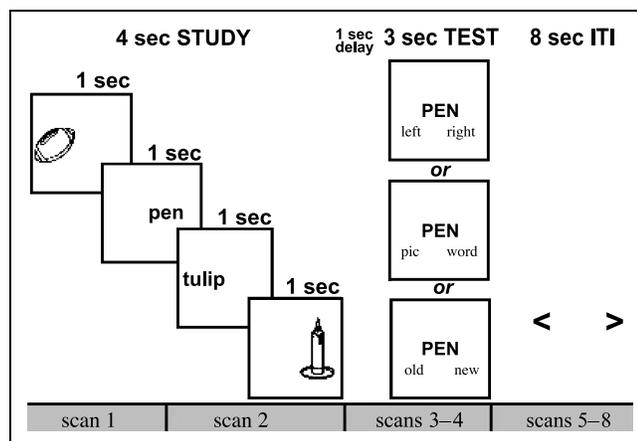
Based on the source monitoring framework (e.g., Johnson et al., 1993) and on a review of early neuroimaging studies (Nolde, Johnson, & Raye, 1998), we suggested the working hypothesis that the right PFC may support primarily heuristic processing and the left PFC (or left and right together) may support more systematic processing (Johnson & Raye, 1998, 2000; Nolde, Johnson, & D'Esposito, 1998). Consistent with this latter prediction, several LTM studies examining various types of SM (e.g., location, size, cognitive operation performed) have found SM judgments are associated with increased activity in the left PFC (and sometimes right as well), including the superior, middle, and inferior frontal gyri (GFs, GFm, and GFi) primarily in Brodmann's areas (BA) 9, 10, 44, 46, and 47 (e.g., Slotnick et al., 2003; Dobbins et al., 2002, 2003; Ranganath et al., 2000; Raye et al., 2000; Henson et al., 1999; Rugg et al., 1999; Nolde, Johnson, & D'Esposito, 1998). Left PFC activity associated with long-term SM has been found with auditory and visual words, nameable pictures, and abstract shapes, suggesting a role for the left PFC that is not specific to type of material.

The present experiments further clarify the role of PFC in source monitoring. As noted above, remembering in LTM tasks involves both reactivation/retrieval and evaluation processes. Here we minimized the reactivation/retrieval aspects of the tasks by using a working memory (WM) paradigm where source judgments about information that is still active in WM should reflect primarily evaluation.

The present studies also address the issue of whether the same brain areas are involved in WM and LTM tasks. Several recent studies have investigated continuity of processes across WM and LTM for item memory (e.g., Nyberg et al., 2003; Ranganath, Johnson, & D'Esposito, 2003; Braver et al., 2001), and evidence is mounting that under a range of circumstances, many of the same brain areas are recruited (see, e.g., Cabeza & Nyberg, 2000; Duncan & Owen, 2000, for reviews). With respect to SM, Cabeza, Dolcos, Graham, and Nyberg (2002) compared, within the same subjects, LTM for words using a remember/know procedure and WM for words and their location (i.e., whether the word had been presented in the left column, the right column, or was new); thus, both tasks required monitoring for specific memorial detail. Several left PFC areas were identified in both tasks, relative to a fixation baseline, including an area of left BA 9 similar to left dorsolateral PFC regions identified in LTM SM studies. However, there were differences between the tasks other than whether the information was currently active in WM. The present research is the first to ask, using exactly the same tasks, whether the same left PFC regions are involved in SM

when information is still active in WM as when information must be revived/reactivated.

Our goal in Experiment 1 was to assess whether source attributions based on specific information currently active in WM show left PFC activity as is found in LTM SM tasks. To maximize the chance that the information was, indeed, likely to be in WM, each trial presented only four items (two words, two line drawings; one of each on the left of the screen, one of each on the right), and participants were probed immediately for ON or SM for a single item from that set (see Figure 1). For generality, we included two source tasks, format (picture, word; PW) and location (left, right; LR). The three test types were randomly intermixed. This design allowed us to compare ON recognition and source judgments about information in WM while equating encoding conditions and general memory factors, such as the set to remember (e.g., "retrieval mode," Lepage, Ghaffar, Nyberg, & Tulving, 2000; Nyberg et al., 1995), type and amount of to-be-remembered information, and number of response options. After finding the expected pattern of greater left PFC activation for source than ON tests in Experiment 1, in Experiment 2 we included an immediate test condition, as in Experiment 1, and a delay condition. On delayed tests, the retention interval was filled by one or two intervening trials (up to 36 sec) so that the information tested was unlikely to be currently active in WM (e.g., Cowan, 1999). Thus, as in typical LTM tasks, ac-

curate responding in the delayed condition required the reactivation, as well as evaluation, of memorial information about the target event (Johnson, 1992; Johnson, Raye, Mitchell, Greene, & Anderson, 2003, Experiment 2). Assuming evaluation of source information in WM and LTM are similar and subserved by the left PFC, we predicted, for both immediate and delayed trials, SM-related activity in areas of the left GFm/GFi. In Experiment 3, we tested the hypothesis that the right PFC is recruited for memory judgments requiring undifferentiated information (e.g., Dobbins et al., 2003; Kensinger, Clarke, & Corkin, 2003; Nolde, Johnson, & Raye, 1998; Johnson et al., 1993; Johnson & Raye, 1998, 2000) by shifting the task context so that accurate performance on the majority of trials could be based on less differentiated qualities (familiarity and/or recency).

RESULTS

Experiment 1

Behavioral Results

Table 1 shows the proportions of hits (H) and false alarms (FA), along with d', which we used as our primary measure of accuracy. As expected, accuracy was high, with a marginal effect of condition, $F(2,14) = 3.33$, $MSE = 1.00$, $p = .07$; ON was greater than LR, no other pairwise comparisons were significant. Analyses of response times (RTs; measured in milliseconds) on correct trials showed an effect of condition, $F(2,14) = 10.19$, $MSE = 23,099$, $p < .01$. As shown in Table 1, participants took significantly longer on the PW test than both the LR and ON tests, which did not differ from each other. Finding longer RTs in PW than ON is consistent with previous data indicating that specific source information often takes time to revive and evaluate (e.g., McElree, Dolan, & Jacoby, 1999; Hintzman, Caulton, & Levitin, 1998; Gronlund, Edwards, & Ohrt, 1997; Johnson, Kounios, & Reeder, 1994).[2]

Functional Magnetic Resonance Imaging Results

As shown in Figure 2, there was activity in large bilateral regions of the GFm and GFi that included BA 9 extending into BA 46, BA 45 on the left (Figure 2A; PW > LR > ON) and primarily BA 9 on the right (Figure 2B; PW > LR > ON [LR vs. ON, $p = .08$]). Not shown in the figure is a more inferior, but adjacent, left region of the GFi, GFm [BA 10, 46; $x = -38$, $y = 45$, $z = 3$; 7 voxels; max $F(14,98) = 3.95$] that showed the pattern PW = LR > ON (LR vs. ON $p = .08$). These areas are generally consistent with those found in long-term studies of SM (e.g., Dobbins et al., 2002, 2003; Ranganath et al., 2000; Henson et al., 1999; Rugg et al., 1999; Nolde, Johnson, & D'Esposito, 1998). The only other PFC region identified in this analysis (not shown) was located in the left GFd, GFs [BA 8, 6; $x = -6$, $y = 20$, $z = 51$;

Figure 1. Procedure and trial timeline (see text for full details). The test types, which were explained in advance to participants, were the following: LR, the probe always represented an old item and participants were to indicate whether the item represented by the word originally appeared to the left (left-hand button press) or to the right (right-hand button press); PW, the probe always represented an old item and participants were to indicate whether the item originally appeared as a picture (left-hand button press) or a word (right-hand button press); or ON, the probe represented an item previously presented (i.e., an old item), or a new item and participants indicated old (left-hand button press) or new (right-hand button press). In Experiment 3, an EL judgment was used in place of LR and participants responded with a left-hand button press for early if the item was one of the first two presented and with a right-hand button press for late if it was one of the last two presented.

Table 1. Behavioral Data: Hits, False Alarms, *d'*, and Correct RT

Experiment 1	Hits		False Alarms		*d'*		RT	
Condition								
LR	0.83 (0.04)		0.11 (0.02)		2.46 (0.40)		1422 (79)	
PW	0.91 (0.03)		0.10 (0.03)		3.07 (0.34)		1644 (71)	
ON	0.92 (0.02)		0.04 (0.02)		3.75 (0.28)		1306 (63)	

Experiment 2	Hits		False Alarms		*d'*		RT	
Condition	Immediate	Delay	Immediate	Delay	Immediate	Delay	Immediate	Delay
LR	0.87 (0.04)	0.75 (0.04)	0.10 (0.03)	0.43 (0.05)	2.47 (0.20)	0.90 (0.17)	1442 (88)	1720 (141)
PW	0.81 (0.02)	0.89 (0.03)	0.08 (0.03)	0.29 (0.05)	2.21 (0.16)	1.81 (0.14)	1776 (79)	2028 (131)
ON	0.88 (0.09)	0.70 (0.04)	0.06 (0.03)	0.06 (0.03)	2.69 (0.28)	1.99 (0.17)	1574 (97)	1612 (96)

Experiment 3	Hits		False Alarms		*d'*		RT	
Condition								
EL	0.92 (0.02)		0.11 (0.02)		2.70 (0.16)		1631 (113)	
PW	0.83 (0.06)		0.06 (0.01)		2.67 (0.23)		1790 (79)	
ON	0.91 (0.04)		0.03 (0.01)		3.35 (0.17)		1498 (79)	

Numbers in parentheses are the standard error. RTs are in milliseconds.

85 voxels; max $F(14,98) = 7.68$]. It also showed the pattern PW > LR > ON.

As in many LTM studies, the SM-related PFC activity in the present WM experiment appeared to be left-lateralized (see Figure 2). To further assess laterality, we used as regions-of-interest (ROI), the functionally defined left and right PFC areas shown in Figure 2A and B and applied them as masks for individual subject's *t* maps that resulted from contrasts of each source task versus ON (i.e., PW-ON and LR-ON). We then counted, for each participant in each ROI, the number of suprathreshold voxels showing SM > ON. There were significantly more suprathreshold voxels on the left ($M = 66.25$) than the right ($M = 14.00$), $F(1,7) = 11.75$, $MSE = 929.11$, $p = .01$, for the PW-ON contrast, with all eight participants showing this pattern. The same was true for the LR-ON contrast ($M = 53.00$ and 8.63, for left and right, respectively), $F(1,7) = 6.16$, $MSE = 1277.71$, $p < .05$, with six of the eight subjects showing this pattern.

Experiment 2

As expected, in Experiment 1, the left PFC showed greater activity for both PW and LR SM tests than ON recognition. We also saw increased SM-related activity in a smaller right homologue region, as seen in many long-term SM studies. Having obtained evidence for left-lateralized PFC activity associated with source judgments made about information in WM, Experiment 2 com-

pared, within subjects, an immediate test (where the relevant information is presumably currently active) to a delayed test (where the relevant information is unlikely to still be active). On half of the trials, the test probe was delayed by one or two intervening trials (up to 36 sec). Although this is shorter than the typical LTM retention interval, relevant information is unlikely to still be active after the filled delay. Thus, delay trials should be more likely than immediate trials to require reactivation (Johnson et al., 2003; Cowan, 1999), as would be required in long-term episodic memory tasks.

Behavioral Results

Due to equipment malfunction, behavioral data from one participant were missing ($n = 8$). A 3 (condition) × 2 (retention interval) analysis of variance (ANOVA) conducted on *d'* scores (see Table 1) showed main effects of condition, $F(2,14) = 4.80$, $MSE = 0.36$, $p < .05$, and interval, $F(1,7) = 99.40$, $MSE = 0.10$, $p < .0001$, which were superceded by a significant interaction, $F(2,14) = 4.28$, $MSE = 0.34$, $p < .05$. Whereas accuracy did not differ between the tests immediately ($p > .10$), accuracy on the LR test was significantly worse than either the PW or ON test when delayed by one or two intervening trials (both $ps < .01$). Analysis of correct RTs (see Table 1) revealed main effects of condition, $F(2,14) = 18.47$, $MSE = 28,761$, $p < .001$, and interval, $F(1,7) = 8.34$, $MSE = 51,533$, $p < .05$, and a significant interaction, $F(2,14) = 6.28$, $MSE = 11,005$, $p < .05$. For

Figure 2. PFC areas showing a significant Condition × Time interaction in Experiment 1 and associated within-trial time courses. In all figures, for the time courses, the *x*-axis represents time within a trial (sec), the *y*-axis represents mean percent signal change from T1. Abbreviations of brain areas follow Talairach and Tournoux (1988). BA areas are listed in order of approximate size (with approximately equal areas of activation indicated by a slash). Talairach coordinates are reported for the local maximum in each area, but the slice ($z = 33$) was chosen to show representative activations. (A) Left GFm, GFi, BA 9, 46/45, 6/8 [$x = -48, y = 18, z = 31$; 206 voxels; max $F(14,98) = 8.98$]; (B) right GFm, BA 9, 46 [$x = 45, y = 22, z = 32$; 39 voxels; max $F(14,98) = 6.84$].
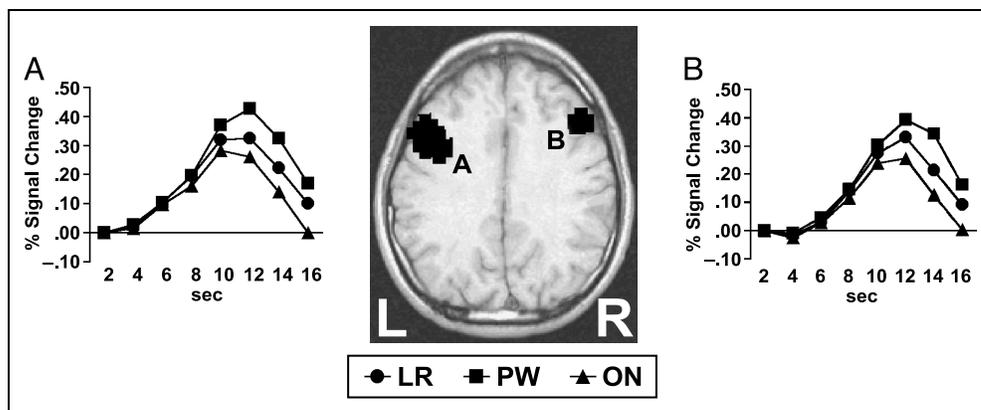
both immediate and delay conditions, participants were slower on the PW test than either the LR or ON test (all *ps* < .05). LR and ON did not differ at either delay interval (*ps* > .05).

Functional Magnetic Resonance Imaging Results

Figure 3 shows the GFm, GFi areas that demonstrated a significant Condition × Time interaction and their associated timelines. In all cases, the PW condition showed greater activation than the ON condition but

LR did not differ significantly from ON (although in the area shown in Figure 3C, the difference between LR and ON at a delay was $p = .07$).[3] Subsequent analyses of the percent signal change during immediate and delay trials separately suggest that there was little difference in the general pattern of activity in these areas as a function of retention interval. Furthermore, these areas were not identified as showing a 3-way interaction (Condition × Time × Retention Interval).[4] Thus, these results suggest that similar processes are engaged in evaluation whether the information is still
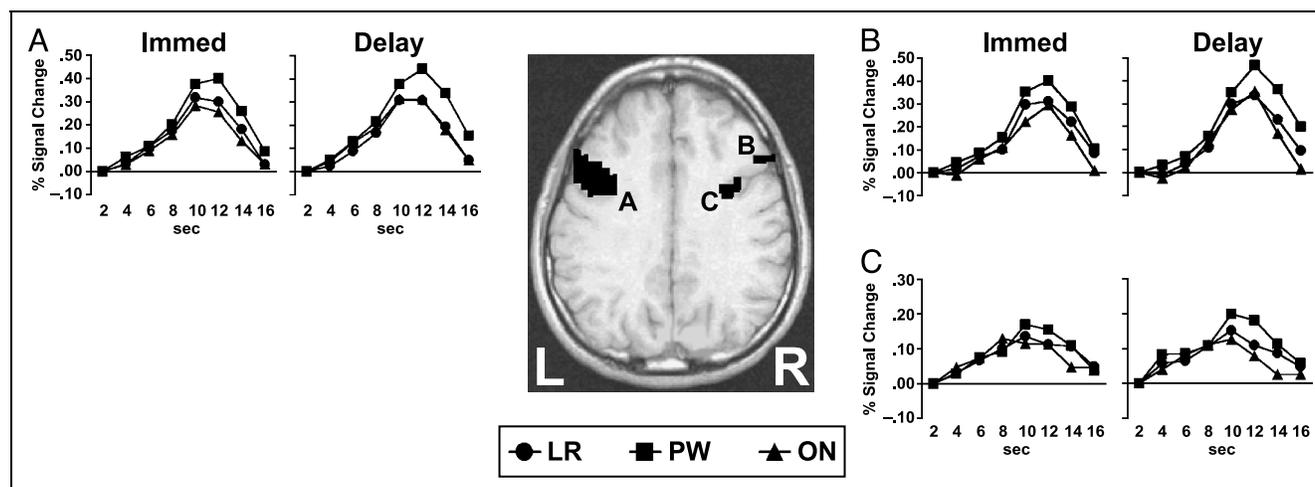
Figure 3. PFC areas showing a significant Condition × Time interaction in Experiment 2 and associated within-trial time courses. Talairach coordinates are reported for the local maximum in each area, but the slice ($z = 32$) was chosen to show representative activations. (A) Left GFm, GFi, GPrC, BA 9, 45, 46, 13 [$x = -41, y = 6, z = 35$; 114 voxels; max $F(14,112) = 6.59$]; (B) right GFm, GFi, BA 9, 46, 45 [$x = 52, y = 22, z = 31$; 13 voxels; max $F(14,112) = 3.46$]; (C) right GFm, GFi, GPrC, BA 9, 6 [$x = 33, y = 5, z = 35$; 17 voxels; max $F(14,112) = 4.83$].

active in WM or must be reactivated after a delay. To be clear, these data say nothing about WM processes that maintain information (e.g., refreshing or rehearsing) or about processes that revive information (e.g., reactivation or retrieval);[5] rather, these findings suggest that whether information has never left WM or has been revived, the evaluation processes involved in making source attributions about it are similar.

Source-related PFC activation again appeared to be left lateralized. We used the ROI masks from Experiment 1 (see Figure 2) and again conducted an analysis of number of suprathreshold voxels found in individual subject's t tests of each SM task versus ON. Because no PFC area in the main analysis showed an interaction between condition and delay interval, we collapsed across delay. As in Experiment 1, for the PW-ON contrast, there were more suprathreshold voxels on the left ($M = 24.22$) than the right ($M = 6.22$), $F(1,8) = 7.70$, $MSE = 189.38$, $p < .05$, with eight of the nine participants showing this pattern. There was no significant difference in number of suprathreshold voxels between the left and right ROIs for the LR-ON comparison. As one might expect from the timelines shown in Figure 3, very few voxels showed an LR > ON difference in the t tests, although the laterality trend was in the expected direction with numerically more such voxels on the left ($M = 8.56$) than on the right ($M = 0.89$), with five of the nine participants showing this pattern.

Two more medial PFC regions not shown in Figure 3 demonstrated a weaker SM effect: left GFs, GFd [BA 9, 8; $x = -15$, $y = 43$, $z = 34$; 10 voxels; max $F(14,112) = 4.49$] and GFd, ACC, GFs [BA 8, 32, 6; $x = 2$, $y = 20$, $z = 47$; 16 voxels; max $F(14,112) = 3.61$]. A superior area of right GFm [BA 6; $x = 25$, $y = 9$, $z = 51$; 7 voxels; max $F(14,112) = 3.92$] showed activity that was greater in LR than either PW or ON.

Experiment 3

In Experiments 1 and 2, as predicted, we found left-lateralized GFm/GFi activity that was greater for SM than for ON recognition. However, we did not find, in either experiment, greater activity in the right PFC associated with ON recognition compared to SM, as might be expected if the right PFC subserves judgments based on undifferentiated information, such as familiarity or recency (e.g., Kensinger et al., 2003; Johnson & Raye, 1998; Nolde, Johnson, & Raye, 1998). However, greater activity in ON than SM tasks has been found in the right PFC in some long-term SM studies. Raye et al. (2000, Experiment 1C) reported greater activity for ON than SM (format judgments) in an area of the right PFC, in addition to the main finding of greater activity in the left PFC for SM than ON recognition. Dobbins et al. (2003) compared recency judgments and source judgments (cognitive operation performed at encoding: pleasant/unpleasant vs. concrete/abstract) and found greater

right PFC activity for recency judgments and greater left PFC activity for the cognitive operation judgment. Such findings provide support for the idea that the right and left PFC differ in the specificity of information they retrieve and/or monitor (see also Kensinger et al., 2003; Suzuki et al., 2002).

Perhaps we found SM > ON activity in the left PFC, but not ON > SM activity in the right PFC, in Experiments 1 and 2, because in WM tasks, specific detail is highly available and familiarity may not be a salient factor. Alternatively, the extent to which less differentiated information versus more specific information is used in attributions about information in WM may depend, as it does in LTM, on the overall task context (e.g., Johnson et al., 1997). In Experiments 1 and 2, two-thirds of the trials were source trials (PW and LR) and one-third were ON trials. Thus, participants may have been inclined to monitor for specific detail to some extent on all trial types. Changing the overall task context so that undifferentiated information more often would be adequate to perform the task may show more right PFC activity associated with ON judgments. We tested this hypothesis in Experiment 3 by replacing the LR trials in the Experiment 1 design with trials on which participants made an early/late (EL) judgment (on EL trials, the first two study items were defined as early and the second two as late). Because relatively undifferentiated information (familiarity, fluency, recency) could be used on two-thirds of the trials (ON and EL), we predicted that there would be an area(s) of right lateral PFC in which ON would show greater activity than SM. If so, this would provide additional evidence for differential hemispheric engagement in monitoring less versus more specific information, extending previous findings from LTM (e.g., Dobbins et al., 2003; Kensinger et al., 2003; Raye et al., 2000, Experiment 1C) to WM.

Behavioral Results

Data were missing from one participant due to equipment malfunction ($n = 8$). Analysis of d' scores (see Table 1) showed an effect of condition, $F(2,14) = 6.48$, $MSE = 0.18$, $p = .01$; accuracy on ON trials was higher than EL and PW, which did not differ. Analysis of correct RTs also showed an effect of condition, $F(2,14) = 14.92$, $MSE = 11,412$, $p < .001$; participants took significantly longer on the PW test than both the EL and ON tests and longer on EL than on ON (all ps < .05).

Functional Magnetic Resonance Imaging Results

We predicted that shifting the task context to one in which judgments could be made based on undifferentiated information on the majority of the trials would result in ON recognition-related activity in right lateral PFC. That is what we found. Figure 4A shows an area of right GFm (BA 6, 8) that showed greater activity on ON

Figure 4. PFC areas showing a significant Condition × Time interaction in Experiment 3 and associated within-trial time courses. Talairach coordinates are reported for the local maximum in each area, but the slice ($z = 44$) was chosen to show representative activations. (A) Right GFm, BA 6, 8 ($x = 42$, $y = 14$, $z = 45$; 17 voxels; max F (14, 112) = 3.63); (B) right GFm, GFs, BA 6 ($x = 20$, $y = 7$, $z = 44$; 16 voxels; max F = 4.14); (C) left GFs, BA 6 ($x = -22$, $y = 10$, $z = 45$; 6 voxels; max F (14, 112) = 4.01).
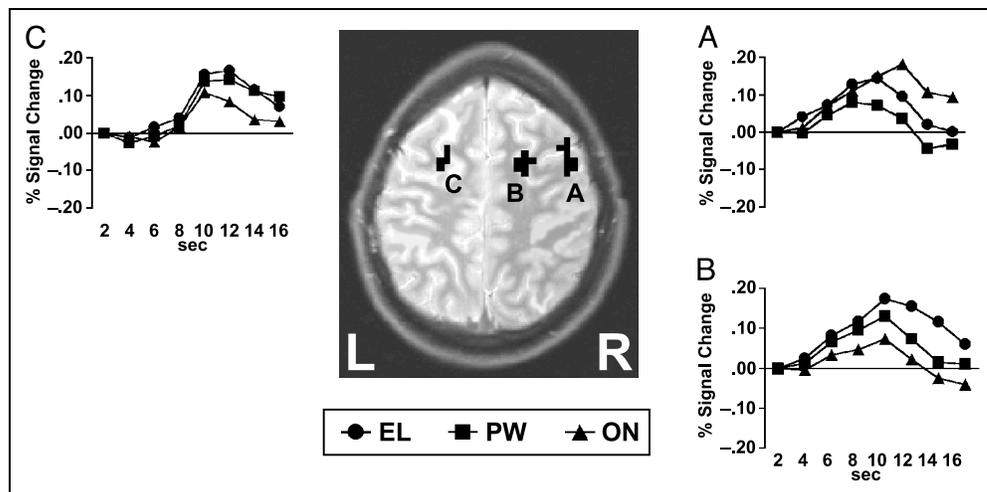

and EL trials than on PW trials. In addition, Figure 4B shows an area of more medial right GFm, GFs (BA 6) that demonstrated greater activity in the EL than either PW or ON conditions. Other studies comparing recency judgments to ON recognition also have found recency-related activity in this general area (e.g., Cabeza et al., 1997, 18, −6, 44; Suzuki et al., 2002, 24, 6, 60). There were no left PFC areas that showed activity that was greater on ON trials than in the other conditions.

As in Experiments 1 and 2, PW-related activity was found in the left PFC. Figure 4C shows an area of left GFs (BA 6) in which there was greater activity on both the PW and EL than ON trials. Figure 5 shows a more lateral area, left GFm (BA 8, 9), where activity was greater for PW than either ON or EL. There was also an area of left GFs, BA 8 [not shown; $x = -15$, $y = 36$, $z = 46$; 36 voxels; max $F(14,112) = 4.89$], with a similar time course. Note that the left and right PFC areas identified in this experiment were toward the more superior extent of those identified in Experiments 1

and 2. This suggests a shift in the particular areas within the left and right PFC engaged in evaluation as a function of overall task context; replication is needed to confirm this possibility.

We directly tested our prediction about the relative roles of left and right lateral PFC in monitoring specific and undifferentiated information, respectively, by conducting 3 (condition: PW, EL, ON) × 2 (hemisphere) ANOVAs on the mean percent change (from Time 1) at times 5, 6, and 7. We first compared the region seen in Figure 4C (left) to the one seen in Figure 4A (right) and found, as expected, a significant Condition × Hemisphere interaction, $F(2,16) = 15.93$, $MSE = 0.002$, $p < .001$. Post hoc LSD tests for condition differences within each hemisphere showed that, as predicted, mean percent signal change in PW (0.13) and EL (0.14) were greater than ON (0.07) on the left, and that the change in ON (0.15) was greater than both EL (0.09) and PW (0.02) on the right, and EL was greater than PW (all $ps < .05$). Likewise, a comparison of the left region in Figure 5

Figure 5. Left GFm area [BA 8, 9; $x = -41$, $y = 14$, $z = 37$; 36 voxels; max $F(14,112) = 4.94$] showing a significant Condition × Time interaction in Experiment 3 and associated within-trial time courses.
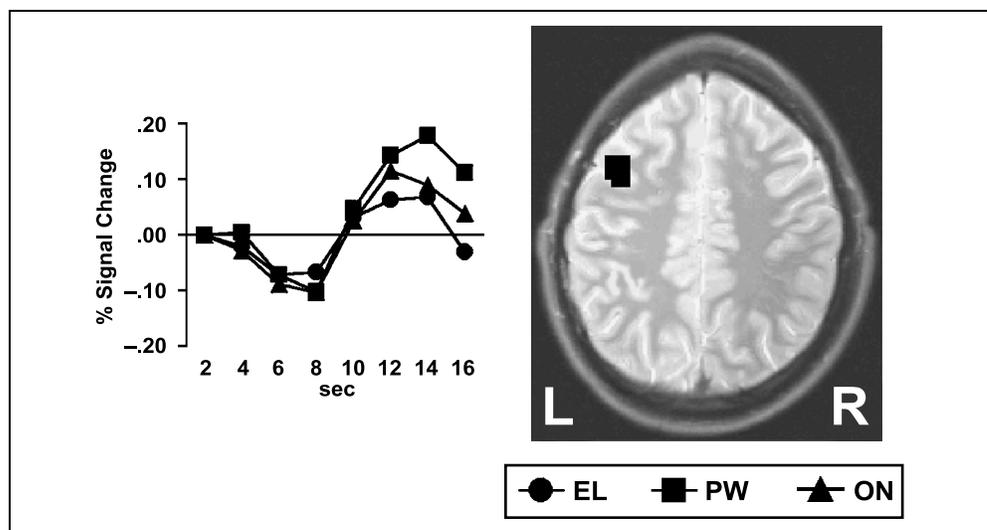
to the right lateral region in Figure 4A showed a significant Condition × Hemisphere interaction, $F(2,16) = 18.21$, $MSE = 0.002$, $p < .0001$. Post hoc LSD tests showed that on the left, the percent signal change in PW (0.13) was greater than EL (0.05), which did not differ from ON (0.08), and on the right, ON (0.15) was greater than both EL (0.09) and PW (0.02) ($ps < .05$).

There was one small area of the right PFC (not shown) [GFi, BA 45, $x = 44$, $y = 22$, $z = 9$; 8 voxels; max $F(14,112) = 3.40$], in which activity was greater in PW than either ON or EL.

The difference between the time courses in the two right PFC areas shown in Figure 4 suggests that ON recognition is not simply an easier discrimination than recency along the same dimension. In addition, in the left PFC, activity associated with the early–late task sometimes looked more like ON (Figure 5) and sometimes more like PW (Figure 4C). Together, these patterns are consistent with findings indicating that there are multiple bases for temporal judgments (e.g., Marshuetz, submitted; Suzuki et al., 2002).

Experiments 1–3: Comparisons between Source Tasks

Although we assume that neural activity will differ depending on what is being monitored (e.g., format or location), the source monitoring framework (e.g., Johnson et al., 1993) is agnostic about whether such differences are reflected in PFC activity. That is, different PFC areas might be associated with monitoring different types of information (presumably in circuits involving other areas) or a specific region of PFC might monitor various features with differences between features reflected primarily in the activity of more posterior regions. The former possibility is consistent with models that associate different PFC regions with discrete types of information (e.g., Goldman-Rakic, 1987), whereas the latter possibility is consistent with models that propose functional specificity of PFC based on process (e.g., Petrides, 2000; D'Esposito et al., 1998). Alternatively, PFC might be organized by both process and type of material (e.g., Johnson et al., 2003), perhaps opportunistically (e.g., Duncan & Owen, 2000; Miller, 2000). Furthermore, PFC may have a more hybrid organization in which some processes are organized by types of information and some are not. For example, different types of information may be represented and/or maintained in different PFC areas but evaluated with respect to task agendas by the same PFC area for a given task.

When we conducted analyses that compared PW directly to LR and to EL, areas of the left PFC, primarily in GFi, showing greater activity in PW were identified in each experiment (Experiment 1: $x = -53$, $y = 21$, $z = 14$; Experiment 2: $x = -45$, $y = 18$, $z = 14$; Experiment 3: $x = -41$, $y = 28$, $z = -13$). In addition, posterior regions, primarily in the GOm, GTi, GF, also showed

PW > LR/EL (Experiment 1: $x = -50$, $y = -62$, $z = -8$; Experiment 2: $x = -50$, $y = -58$, $z = -8$; Experiment 3: $x = -46$, $y = -50$, $z = -3$). In contrast, there was a region of the right GFs/GFm ($x = 19$, $y = 6$, $z = 52$) that showed LR>PW for Experiment 1. This is near the region of GFm reported earlier that showed LR > PW and ON in Experiment 2 and similar to one found in other WM tasks with a spatial component (e.g., Cabeza et al., 2002). Although suggestive, our findings do not clearly differentiate among the various possible organizations of PFC with respect to types of information monitored. More systematic studies are required to explore the relative sensitivity of various PFC and more posterior areas as the representation and monitoring of different types of information are independently varied.

DISCUSSION

The findings from these three experiments are consistent with the hypothesis associating memory judgments based on specific information with the left PFC and memory judgments based on undifferentiated information with the right PFC (Dobbins et al., 2003; Kensinger et al., 2003; Johnson & Raye, 1998, 2000; Nolde, Johnson, & D'Esposito, 1998). Moreover, the findings are consistent with the hypothesis that such PFC activity reflects evaluation processes common to both WM and LTM tasks.

In Experiment 1, we found GFm, GFi activity that was greater for two WM source tasks (remembering whether an item was presented as a picture or a word or on the left or right) than for ON recognition. Consistent with several long-term SM studies, the extent of this activation was greater in the left than in the right PFC. In Experiment 2, this pattern was evident on both immediate and delay trials. Although the delay was brief (maximum 36 sec), it was filled with other trials, making it unlikely that the information needed for the judgment was still active. Rather, it was likely that information needed to be reactivated in the delay condition (Johnson et al., 2003; Cowan, 1999). Thus, our delay condition provides an intermediate data point between our WM task and the LTM tasks used in previous SM studies (e.g., Slotnick et al., 2003; Dobbins et al., 2002, 2003; Ranganath et al., 2000; Raye et al., 2000; Henson et al., 1999; Rugg et al., 1999; Nolde, Johnson, & D'Esposito, 1998). Together with results from these previous LTM studies, the present findings support the conclusion that lateral PFC, particularly left, is involved in memory attributions that require more specific memorial detail (i.e., source information) both when that information is still active in WM and when it must be reactivated.

In Experiment 3, substituting an early–late judgment for the left–right judgment in our WM procedure changed the overall task context and thereby the pattern of brain activity (see also Johnson et al., 1997). When the majority of judgments (ON and EL) could be made based on undifferentiated information (familiarity or recency),

we identified a region of the right PFC associated with making such judgments. These findings support the idea that the right PFC is sensitive to undifferentiated information (e.g., familiarity, recency, fluency) when items are presumably still active in WM, as it is in tasks requiring reactivation of information from LTM (Dobbins et al., 2003; Kensinger et al., 2003; Suzuki et al., 2002; Raye et al., 2000). In contrast to the proposal that the right PFC subserves monitoring in general (Cabeza, Locantore, & Anderson, 2003; Rugg, Henson, & Robb, 2003), these results add to evidence suggesting that the pattern of activity observed in the left and right PFC depends, at least in part, on the qualities of mental experiences that are being monitored (see also Dobbins, Simons, & Schacter, 2004).

Our strategy in these experiments was to equate encoding across conditions (e.g., by randomly intermixing test conditions) and to simplify the processing involved (e.g., by minimizing reactivation/retrieval requirements compared to typical LTM tasks) to examine brain activity as a function of whether undifferentiated or specific information was evaluated at test. We cannot assume that we have completely disentangled type of information (undifferentiated and specific) from type of processing (heuristic and systematic) because they tend to be correlated. This is a general problem that is typically ignored; for example, across studies in the literature, familiarity and recollection are variously treated as processes or as types of information. Additional studies designed to assess contributions of specific processes and types of information to PFC activity in SM tasks would be useful.

In summary, the present results provide evidence that the left PFC is particularly likely to be recruited when the task requires monitoring of specific features, such as modality or location, not only when reactivation is required, but also when the information is highly likely to be currently active in WM. The right PFC is more likely to be recruited when judgments assess undifferentiated information, such as familiarity or recency. These results are consistent with a role for both the left and right PFC in the evaluation of activated information. Furthermore, these findings, along with previous studies investigating long-term SM, point to the commonality in neural substrates of processes engaged flexibly across WM and LTM tasks.

METHODS

Participants

For all experiments reported here, participants were healthy, right-handed college-aged students (Experiment 1: $n = 8$, 6 women, $M = 23.8$ years; Experiment 2: $n = 9$, 7 women, $M = 21.5$ years; Experiment 3: $n = 9$, 6 women, $M = 22.8$ years). Data from one additional participant were excluded in Experiments 1 and 2 because of excessive head movement. All participants reported being in good health, with normal (or cor-rected-to-normal) vision and no history of psychiatric diagnosis or primary degenerative neurological disorder; none were taking psychotropic medications. All participants were paid. The Human Investigation Committee of Yale University Medical School approved the protocol; informed consent was obtained from all participants.

Task and Design

In all experiments, during scanning, stimuli were projected onto a screen at the foot of the scanner, which participants viewed through a mirror mounted on the head coil. Each trial was 16 sec (see Figure 1). Participants saw four items presented sequentially for 1 sec each. They were told to study each item for an upcoming test. Two of the items were concrete nouns printed in lower case (i.e., words), two were black-and-white line drawings of common objects (i.e., pictures). One of each item type was presented on the left of the screen and one of each on the right. Format and location were pseudorandomly assigned, so that, across runs, each format/location appeared nearly equally often in each ordinal position in each testing condition. After 1 sec of unfilled time, a probe word appeared in capital letters. The test probe was accompanied by cues, presented in smaller font to the left and right below the probe word, that indicated to the participant which aspect of their memory was being tested (see Figure 1). In Experiment 1 probes always tested items from the current trial. In Experiment 2, for half of the trials, probe items were from the current trial (immediate test) and for half of the trials probe items were from one or two trials back (i.e., the target could have been presented up to 36 sec prior to the test probe, delay test). Participants were aware of the delay manipulation. Experiment 3 included only immediate trials and the procedure was exactly like Experiment 1 except that participants made EL judgments instead of LR. For the EL task, participants were instructed to press the left button for *early* if the item had been one of the first two items presented on that trial and the right button for *late* if the item had been one of the last two items presented on that trial.

In all experiments, the test probe was followed by an 8-sec intertrial interval that included two arrows presented sequentially for 1400 msec each. Participants pressed one of two buttons to indicate which way the arrow pointed. The arrows provided a task common to all conditions to allow time for the hemodynamic response associated with encoding and remembering and to decrease variability among participants from uncontrolled mental activity between trials (e.g., Raye, Johnson, Mitchell, Reeder, & Greene, 2002).

Stimuli were common objects whose names were one to three syllables in length (e.g., toad, toothbrush, microscope). Pictures were chosen from Snodgrass and Vanderwart's (1980) corpus and similar drawings provided by Friedman (see Cycowicz, Friedman, Roth-

stein, & Snodgrass, 1997). Words were chosen from among the remaining picture labels and additional concrete nouns representing objects that could have been presented as line drawings. For the test probes, names of the pictures were either the original labels or single-word names reached by agreement among the experimenters. Items in all classes—studied, tested, pictures, words, new items—were equated between conditions on published norms for that class (Snodgrass & Vanderwart, 1980; Cycowicz et al., 1997), including number of syllables ($M = 1.81$) and Kucera–Francis frequency ($M = 23.95$; Kucera & Francis, 1967) for words and picture labels, familiarity ($M = 3.31$) for pictures and words, and name agreement ($M = 0.53$, Snodgrass & Vanderwart's, 1980 H statistic]), image agreement ($M = 4.12$), and complexity ($M = 3.06$) for pictures. Across participants, each item appeared equally often for each type of test. Test probes were taken equally often from each ordinal study position, format (PW), and location (LR).

Experiments 1 and 3 were simple single-factor designs in which condition had three levels (Experiment 1: LR, PW, and ON; Experiment 3: EL, PW, and ON) manipulated within subjects. There were five runs of 24 trials each (eight per condition for a total of 40 trials per condition per participant); in each run, LR/EL, PW, and ON trials were pseudorandomly intermixed, with a different order for each participant. Experiment 2 was a 3 (condition: LR, PW, ON) × 2 (retention interval: Immediate, Delay) design, with both factors varied within subjects. Trials were presented in a pseudorandom fashion, with a different trial order for each participant (each of 5 runs had 8 trials PW, LR [4 immediate, 4 delay], 9 trials ON [3 immediate, 3 delay, 3 row]).

Imaging Details

The imaging details were the same for Experiments 1 and 2. T1-weighted anatomical images were acquired for each participant using a 1.5-T GE Signa scanner at the Yale University School of Medicine. Functional scans were acquired with a single-shot echo-planar gradient-echo pulse sequence (TR = 2000 msec, TE = 35 msec, flip angle = $80°$, field of view = 24). Experiment 3 was run on a 3-T Siemens TRIO scanner at the new Magnetic Resonance Research Center at Yale University with the same scanning parameters, except that the anatomical images were T2-weighted and the TE for the functional scans was 25 msec. In all experiments, the 24 axial slices (slice thickness 3.8 mm, resolution 3.75 × 3.75 mm in plane) were aligned with the AC–PC line. Each run began with 12 blank seconds to allow tissue to reach steady state magnetization and was followed by a 1-min rest interval. One volume was collected every 2 sec, or eight full brain scans for each trial; thus there were, for each person, a total of 320 images collected in each condition in Experiments 1 and 3 and 160 in each

Condition × Retention interval cell in Experiment 2 (120 for ON).

Analyses of Behavioral Data

Responses that required a left-handed response (old, picture, left, early) were coded as targets and those requiring a right-handed response (new, word, right, late) as lures. Thus, hits and false alarms could be calculated and converted to d' scores, which we used as our measure of accuracy. Before calculating d', scores were adjusted as follows: $p(H) = 1$ was recalculated as $1 - 1/(2N)$; $p(FA) = 0$ was recalculated as $1/(2N)$, where $N =$ the maximum number of hits or false alarms possible (Macmillan & Creelman, 1991). An ANOVA with condition (and retention interval, Experiment 2) as within-subjects factor(s) was followed by an LSD post hoc test to determine which conditions differed. Unless otherwise noted, $p < .05$.

Functional Magnetic Resonance Imaging Analyses

Data were motion-corrected using a six-parameter automated algorithm (AIR; Woods, Cherry, & Mazziotta, 1992). A 12-parameter AIR algorithm was used to coregister participants' images to a common reference brain. Data were mean-normalized across time and participants and spatially smoothed (3-D, 8-mm FWHM Gaussian kernel).

The arrow task was included as a "filler" task to equate participants' cognitive activity during the intertrial interval (e.g., Raye et al., 2002). Regions were not identified by using the intertrial interval as a baseline, rather the analyses compared experimental conditions directly. The functional magnetic resonance imaging (fMRI) data were analyzed using ANOVA with participant as a random factor (NIS software, Laboratory for Clinical Cognitive Neuroscience, University of Pittsburgh, and the Neuroscience of Cognitive Control Laboratory, Princeton University). Run (1–5), condition (LR or EL, PW, ON), retention interval (immediate, delay; Experiment 2 only),[6] and time within trial (volumes 1–8) were fixed factors. The hemodynamic response that indexes brain activity in fMRI responds slowly, rising to a peak 4–6 sec after the critical event. Because encoding conditions were held constant, differences between the conditions should be seen as differences in the fMRI signal peaking at seconds 10–14 (volumes 5–7), corresponding to the test period in each trial. Brain regions in the Condition × Time interaction were identified that had a minimum of six spatially contiguous voxels, each significant at $p < .0001$ (Experiment 1) or $p < .001$ (Experiments 2 and 3) (Forman et al., 1995). Less stringent alpha levels were used in Experiments 2 and 3 because they had the built-in opportunity for replication, there were fewer trials per condition (Experiment 2) and a specific interaction between condition (ON, PW) and hemisphere (left,

right) was predicted (Experiment 3). Once areas of activation were identified, pairwise post hoc analyses were conducted on the mean percent change within trials (from Time 1) at Times 5, 6, and 7 to determine which conditions significantly differed at test within each area ($p < .05$, unless otherwise noted).

The resultant F maps were transformed to Talairach space using AFNI (Cox, 1996), and areas of activation were localized using Talairach Daemon software (Lancaster, Summerlin, Rainey, Freitas, & Fox, 1997), as well as manually checked with the atlases of Duvernoy (1999), Mai, Assheuer, and Paxinos (1997), and/or Talairach and Tournoux (1988). Because our focus concerns differential participation of regions of PFC in SM tests versus ON recognition, we report only PFC regions; a list of all brain regions identified can be obtained from the authors.

To subsequently quantify and test laterality differences in SM-related activity in PFC in Experiments 1 and 2, we used the left and right ROIs found in Experiment 1 (Figure 2A and B) as masks and counted the number of suprathreshold voxels in each participant's t maps that compared each SM task to ON (i.e., PW-ON and LR-ON). Cutoffs were $t \geq 1.89$ and $t \geq 1.86$, the group critical t at $p = .05$, $df = 7$ and $df = 8$ in Experiments 1 and 2, respectively. We then submitted these voxel counts for each participant to an ANOVA with hemisphere as the within-subjects factor. We tested for the specific predicted interaction between condition (ON, PW) and hemisphere (left, right) in Experiment 3 by submitting individuals' mean percent change during the test period (Times 5–7) for functionally defined ROIs on the left and on the right to ANOVAs with condition and hemisphere as within-subjects factors.

For the direct comparisons of PW to LR/EL discussed at the end of Experiment 3, the identified brain areas were obtained from ANOVAs, which included run (1–5), condition (LR or EL, PW), and time within trial (volumes 1–8) as fixed factors; for all three experiments, the regions noted had a minimum of six spatially contiguous voxels, each significant at $p < .001$.

Acknowledgments

Notes

1. There are also several event-related potential studies examining the neural correlates of long-term source memory (e.g., Ranganath & Paller, 1999; Allan & Rugg, 1998; Senkfor & Van Petten, 1998; Johnson, Kounios, & Nolde, 1996). In general, the findings are consistent with findings from fMRI and PET studies implicating PFC in source memory.

2. The PW task is particularly interesting because it represents a common source monitoring situation in which people must discriminate things they saw from things they learned about through language (e.g., read or heard about). However, it does raise the issue of the match between study and test format. The test probes were always words. Although there was a change from lower case to upper case between study and test, the probe was a better physical match for items originally presented as words than items originally presented as pictures. It might be that participants based their format decisions on the fluency created by a "perceptual match" between the target and probe. To evaluate this possibility, we examined correct RTs as a function of whether the target item had originally been presented as a picture or a word. There was an interaction of format of original presentation (P, W) and test condition (PW, LR, ON), $F(2,14) = 5.02$, $MSE = 18477$, $p < .05$, because responses were faster to P than to W items in both the PW condition ($M= 1606$ vs. 1798, for pictures and words, respectively) and the LR condition ($M = 1342$ vs. 1518, for pictures and words, respectively), but they did not differ in the ON condition ($M = 1464$ vs. 1385, for pictures and words, respectively, $p > .10$). This pattern is consistent with the idea that participants were evaluating test probes on a different basis on SM and ON tests, with amount of pictorial information apparently being relatively more salient in the source tasks than ON. However, similar analyses conducted for data from Experiments 2 and 3 showed no interaction of format and condition, and the main effect of format was not significant ($ps > .10$). Thus, it does not appear that judgments were based on the extent to which the test probe physically matched the study item in any of the experiments.

3. The timelines shown in Figures 2 and 3 suggest that the SM–ON difference tended to be somewhat greater for the PW than the LR task, especially when delay trials were included. The behavioral data showed poorer performance on LR than ON or PW after a delay in Experiment 2. Together, this pattern suggests that format is more salient than location and that location information becomes unavailable faster than format information in this task.

4. The only PFC area that showed a Condition × Time × Retention Interval interaction, after dropping the threshold to $p < .01$, was an inferior part of ACC ($x = -8$, $y = 36$, $z = 20$). It showed greater deactivation for ON than the source tests (which did not differ from each other) immediately but not after a delay. A very superior region of GFs (BA 6/8) showed uninterpretable time courses for both immediate and delay trials.

5. Two interesting regions potentially related to revival demonstrated a Retention Interval × Time within-trial interaction such that there was greater activity on delay than on immediate trials at test: precuneus ($x = 13$, $y = -67$, $z = 34$), a region often associated with episodic retrieval (e.g., Cabeza, Dolcos, Prince, et al., 2003; see Cabeza & Nyberg, 2000, for a review) and an area of the posterior cingulate gyrus ($x = 7$, $y = -33$, $z = 39$), which is sometimes found to be associated with retrieval when contextually specific cues are used as probes (e.g., temporal cues, Fujii et al., 2002, or one word of a learned pair, Krause et al., 1999). In contrast, as might be expected, an area of left parietal cortex (LPs/LPi, $x = -42$, $y = -55$, $z = 51$) showed greater test-related activity on immediate than delay trials (e.g., see Cabeza & Nyberg, 2000, for a review). These findings suggest that we were successful in manipulating the revival requirements between the immediate and delay conditions in Experiment 2.

6. For all three experiments, both correct and incorrect trials were included in the fMRI analysis because we were interested in the processes engaged in making ON and SM judgments. In Experiment 2, new test items on the ON test were neither immediate nor delay. Thus, the images corresponding to these trials were deleted from the analysis of fMRI data reported in Experiment 2 so that retention interval could be entered as a factor. Other ANOVAs conducted on the fMRI data with those trials included showed the same patterns of activity.

REFERENCES

Allan, K., & Rugg, M. D. (1998). Neural correlates of cued recall with and without retrieval of source memory. *NeuroReport, 9,* 3463–3466.

Anderson, J. R., & Bower, G. H. (1974). A propositional theory of recognition memory. *Memory and Cognition, 2,* 406–412.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11,* 267–273.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548–564.

Braver, T. S., Barch, D. M., Kelley, W. M., Buckner, R. L., Cohen, N. J., Miezin, F. M., Snyder, A. Z., Ollinger, J. M., Akbudak, E., Conturo, T. E., & Petersen, S. E. (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *Neuroimage, 14,* 48–59.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage, 16,* 317–330.

Cabeza, R., Dolcos, F., Prince, S. E., Rice, H. J., Weissman, D. H., & Nyberg, L. (2003). Attention-related activity during episodic memory retrieval: A cross-function fMRI study. *Neuropsychologia, 41,* 390–399.

Cabeza, R., Locantore, J. K., & Anderson, N. D. (2003). Lateralization of prefrontal activity during episodic memory retrieval: Evidence for the production-monitoring hypothesis. *Journal of Cognitive Neuroscience, 15,* 249–259.

Cabeza, R., Mangels, J., Nyberg, L., Habib, R., Houle, S., McIntosh, A. R., & Tulving, E. (1997). Brain regions differentially involved in remembering what and when: A PET study. *Neuron, 19,* 863–870.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience, 12,* 1–47.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex, 12,* 1048–1056.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29,* 162–173.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65,* 171–237.

D'Esposito, M., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., and Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Brain Research: Cognitive Brain Research, 7,* 1–13.

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron, 35,* 989–996.

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia, 41,* 318–333.

Dobbins, I. G., Simons, J. S., & Schacter, D. L. (2004). fMRI evidence for separable and lateralized prefrontal memory monitoring processes. *Journal of Cognitive Neuroscience, 16,* 908–920.

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1121–1136.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.

Duvernoy, H. M. (1999). *The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply* (2nd ed.). New York: Springer-Verlag.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647.

Fujii, T., Okuda, J., Tsukiura, T., Ohtake, H., Miura, R., Fukatsu, R., Suzuki, K., Kawashima, R., Itoh, M., Fukuda, H., & Yamadori, A. (2002). The role of basal forebrain in episodic memory retrieval: A positron emission tomography study. *Neuroimage, 15,* 501–508.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology, Section 1: The nervous system* (vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.

Gronlund, S. D., Edwards, M. B., & Ohrt, D. D. (1997). Comparison of the retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1261–1274.

Henson, R. N. A., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain, 122,* 1367–1381.

Hintzman, D. L. (2000). Memory judgments. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 165–177). New York: Oxford University Press.

Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory and Cognition, 26,* 449–462.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326–338.

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience, 4,* 268–280.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Johnson, M. K., Kounios, J., & Nolde, S. F. (1996). Electrophysiological brain activity and memory source monitoring. *NeuroReport, 7,* 2929–2932.

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1409–1419.

Johnson, M. K., Nolde, S. F., Mather, M., Kounios, J., Schacter, D. L., & Curran, T. (1997). The similarity of brain activity associated with true and false recognition memory depends on test format. *Psychological Science, 8,* 250–257.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67–85.

Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Sciences, 2,* 137–145.

Johnson, M. K., & Raye, C. L. (2000). Cognitive and brain mechanisms of false memories and beliefs. In D. L. Schacter & E. Scarry (Eds.), *Memory, brain, and belief* (pp. 35–86). Cambridge: Harvard University Press.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., & Anderson, A. W. (2003). fMRI evidence for an organization of prefrontal cortex by both type of process and type of information. *Cerebral Cortex, 13,* 265–273.

Kelley, C. M., & Jacoby, L. L. (2000). Recollection and familiarity: Process-dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 215–228). New York: Oxford University Press.

Kensinger, E. A., Clarke, R. J., & Corkin, S. (2003). What neural correlates underlie successful encoding and retrieval? A functional magnetic resonance imaging study using a divided attention paradigm. *Journal of Neuroscience, 23,* 2407–2415.

Konishi, S., Uchida, I., Okuaki, T., Machida, T., Shirouzu, I., & Miyashita, Y. (2002). Neural correlates of recency judgment. *Journal of Neuroscience, 22,* 9549–9555.

Krause, B. J., Horwitz, B., Taylor, J. G., Schmidt, D., Mottaghy, F. M., Herzog, H., Halsband, U., & Müller-Gärtner, H.-W. (1999). Network analysis in episodic encoding and retrieval of word-pair associates: A PET study. *European Journal of Neuroscience, 11,* 3293–3301.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lancaster, J. L., Summerlin, J. L., Rainey, L., Freitas, C. S., & Fox, P. T. (1997). The Talairach daemon, a database server for Talairach atlas labels. *Neuroimage, 5,* S633.

Lepage, M., Ghaffar, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 506–511.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition, 17,* 349–358.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, UK: Cambridge University Press.

Mai, J. K., Assheuer, J., & Paxinos, G. (1997). *Atlas of the human brain.* San Diego, CA: Academic Press.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1137–1151.

Marshuetz, C. (submitted). *Order information in working memory: An integrative review of evidence from brain and behavior.*

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 563–582.

Miller, E. K. (2000). The prefrontal cortex: No simple matter. *Neuroimage, 11,* 447–450.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.

Moscovitch, M. (1989). Confabulation and the frontal systems: Strategic versus associative retrieval in neuropsychological theories of memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 133–160). Hillsdale, NJ: Erlbaum.

Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology and Aging, 10,* 492–497.

Nolde, S. F., Johnson, M. K., & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: An event-related fMRI study. *NeuroReport, 9,* 3509–3514.

Nolde, S. F., Johnson, M. K., & Raye, C. L. (1998). The role of prefrontal cortex during tests of episodic memory. *Trends in Cognitive Sciences, 2,* 399–406.

Nyberg, L., Marklund, P., Persson, J., Cabeza, R., Forkstam, C., Petersson, K. M., & Ingvar, M. (2003). Common prefrontal activations during working memory, episodic memory, and semantic memory. *Neuropsychologia, 41,* 371–377.

Nyberg, L., Tulving, E., Habib, R., Nilsson, L. G., Kapur, S., Houle, S., Cabeza, R., & McIntosh, A. R. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport, 7,* 249–252.

Petrides, M. (2000). Frontal lobes and memory. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (2nd ed., vol. 2, pp. 67–84). Amsterdam: Elsevier.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: Comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1110–1115.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience, 20:RC108,* 1–5.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41,* 378–389.

Ranganath, C., & Paller, K. A. (1999). Frontal brain potentials during recognition are modulated by requirements to retrieve perceptual detail. *Neuron, 22,* 605–613.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Nolde, S. F., & D'Esposito, M. (2000). fMRI investigations of left and right PFC contributions to episodic remembering. *Psychobiology, 28,* 197–206.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: Dorsolateral PFC activity associated with refreshing just-activated information. *Neuroimage, 15,* 447–453.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (in press). Sum–difference theory of remembering and knowing: A two-dimensional signal detection model. *Psychological Review.*

Rugg, M. D., Fletcher, P. C., Chua, P. M.-L., & Dolan, R. J. (1999). The role of the prefrontal cortex in recognition memory and memory for source: An fMRI study. *Neuroimage, 10,* 520–529.

Rugg, M. D., Henson, R. N. A., & Robb, W. G. K. (2003). Neural correlates of retrieval processing in the prefrontal cortex

during recognition and exclusion tasks. *Neuropsychologia, 41,* 40–52.

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1005–1025.

Shimamura, A. P., Janowsky, J. S., & Squire, L. R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia, 28,* 803–813.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J., Jr. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research, 17,* 75–82.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Suzuki, M., Fujii, T., Tsukiura, T., Okuda, J., Umetsu, A., Nagasaka, T., Mugikura, S., Yanagawa, I., Takahashi, S., & Yamadori, A. (2002). Neural basis of temporal context memory: A functional MRI study. *Neuroimage, 17,* 1790–1796.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging.* New York: Thieme.

Woods, R. P., Cherry, S. R., & Mazziotta, J. C. (1992). Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography, 16,* 620–633.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1415–1434.