Behavioral/Systems/Cognitive

Functional–Neuroanatomic Correlates of Recollection: Implications for Models of Recognition Memory

Itamar Kahn,1,2 Lila Davachi,3 and Anthony D. Wagner1
1Department of Psychology and Neurosciences Program, Stanford University, Stanford, California 94305-2130, 2Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and 3Department of Psychology, New York University, New York, New York 10003

Recognition decisions can be based on familiarity, the sense that an item was encountered previously (item memory), and on recollection, the conscious recovery of contextual information surrounding a previous encounter with the item (e.g., source memory). Recognition with recollection is thought to depend on multiple mechanisms, including prefrontal “control” processes that guide retrieval and recapitulation mechanisms that reactivate posterior neocortical representations that were present at encoding. However, uncertainty remains regarding the precise nature of prefrontal contributions to recollection and the selectivity of recapitulation to veridical recollection. The present event-related functional magnetic resonance imaging study sought to examine whether regions showing “old–new” effects support processes sensitive to recollection success or recollection attempt and whether recapitulation of neocortical representations emerge during veridical recollection as well as during false recognition (i.e., false alarms) or whether false recognition resembles familiarity-based responding. Results revealed that multiple left prefrontal cortical regions were engaged during attempts to recollect previous contextual (source) details, regardless of the nature of the to-be-recollected details and of source recollection outcome (successful vs unsuccessful). Recapitulation effects were observed in regions sensitive to the encoding task, suggesting that veridical recollection entails the reactivation of processes or representations present during encoding. Importantly, in contrast to leading models of recognition memory, false alarms also appeared to be based partially on recollection, as revealed through false recapitulation effects. Implications for neural and cognitive models of recognition are considered.

Key words: episodic memory; explicit memory; declarative memory; prefrontal cortex; recapitulation; neuroimaging

Introduction

Recognition memory can be based on familiarity (the sense that an item was encountered previously) and/or recollection (the conscious recovery of contextual information surrounding a previous encounter with an item) (Yonelinas, 2002). Recognition with recollection depends on multiple mechanisms, including prefrontal cortical (PFC) “control” processes that guide retrieval attempts, and recapitulation mechanisms that reactivate neocortical representations present at encoding (Rugg and Wilding, 2000; Buckner and Wheeler, 2001; Wagner, 2002).

In humans, leverage on the neural mechanisms supporting recognition comes partially from neuroimaging comparisons between correctly recognized old items (Hits) and correctly rejected new items (CRs). Event-related functional magnetic resonance imaging (fMRI) has revealed “old–new” effects in left PFC and parietal cortices (Nolde et al., 1998; Konishi et al., 2000; McDermott et al., 2000; Maril et al., 2003), suggesting that these regions are sensitive to or signal successful retrieval (Henson et al., 1999; Konishi et al., 2000; Donaldson et al., 2001; Wheeler and Buckner, 2003), and event-related potentials have revealed a parietal-situated old–new effect that tracks recollection (Smith, 1993; Wilding et al., 1995; Duzel et al., 1997). Other evidence, however, suggests that PFC and parietal old–new effects reflect differential engagement of control processes brought to bear in attempts to recollect, regardless of the success of recollection (Ranganath et al., 2000; Dobbins et al., 2002). Resolving this apparent conflict between “retrieval success” and “recollection attempt” hypotheses will be a key step toward understanding PFC and parietal contributions to recognition.

Recollection is hypothesized to entail the reactivation of representations that were present at encoding (Naya et al., 2001). Consistent with this hypothesis, neuroimaging studies have demonstrated that domain-specific (e.g., visual vs auditory) perceptual cortices that were engaged during encoding are re-engaged during the retrieval of domain-specific event details (Nyberg et al., 2000; Wheeler et al., 2000; Vaidya et al., 2002). However, these previous studies used multiple study events per item, raising the possibility that such effects do not reflect episodic recapitulation. Rather, such effects may reflect a top-down attentional modulation of domain-specific representational layers during attempts to recollect caused by having acquired semantic knowledge about the encoding context of an item. A key question is whether reca-
pituation effects emerge during the recollection of single episodes.

Neural recapitulation can also provide a critical test of models of recognition. A central assumption of leading dual-process models is that false recognition [i.e., false alarms (FAs)] reflects above-criterion familiarity in the absence of recollection (Jacoby, 1991; Yonelinas et al., 1996); Hits are based on recollection and/or above-criterion familiarity, whereas FAs are thought to be recollection-free. To the extent that this assumption is valid and that neural recapitulation responses reflect recollection, FAs should not be accompanied by such responses. Alternatively, given behavioral evidence of false recollection (Roediger and McDermott, 1995), it remains possible that FAs are accompanied by illusory recapitulation responses.

The present fMRI experiment examined whether regions showing old–new effects support processes sensitive to recollection success or attempt and whether neural correlates of recapitulation emerge during veridical recollection of single episodes, and whether such responses are selective to veridical recognition or generalize to false recognition.

Materials and Methods

Subjects. Eight female and nine male right-handed, native-English speakers (18–33 years of age) were paid $65 for their participation. Data were acquired but discarded from three additional participants, two because they failed to respond to 15% or more of the trials and one because of performance levels that resulted in multiple conditions with fewer than 15 events (thus resulting in unstable measurements). Informed consent was obtained in a manner approved by the institutional review boards at the Massachusetts Institute of Technology and Massachusetts General Hospital.

Behavioral procedures. Subjects performed two interleaved incidental encoding tasks identical to those explored previously by Davachi et al. (2003); they were subsequently scanned while making item-recognition decisions combined with a source-recollection judgment. Across eight (nonscanned) study lists, 200 visually presented adjectives were encoded via an orienting task requiring mental imagery ("image" task) and 200 via an orienting task requiring orthographic-to-phonological transformation ("Read" task) (Fig. 1A). On each trial, a 500 msec cue (place/read) signaled the encoding task to be performed on an adjective that was then presented for 500 msec. During Image trials, subjects generated a mental image of a spatial scene (i.e., a place) described by the adjective (e.g., for "dirty," the subject might imagine a garbage dump). During Read trials, subjects covertly pronounced the word backward (e.g., "happy" might be pronounced "ip-pæ"). After a 4000 msec fixation period, during which subjects performed the indicated task, the fixation cross changed color, signaling subjects to indicate their level of task success by pressing one of four buttons: 1, unsuccessful; 2, partially successful; 3, succeeded with effort; 4, succeeded with ease. To ensure that the fMRI retrieval effects do not reflect differential task success at encoding, analyses were restricted to trials on which the encoding task was performed successfully (i.e., received a response of 3 or 4) (Davachi et al., 2003). Across subjects, the assignment of items to conditions was counterbalanced.

Approximately 20 hr after encoding, subjects returned for an event-related fMRI scanning session. Subjects were initially scanned while engaging in a parahippocampal place area (PPA) localizer task designed to identify parahippocampal voxels that respond differentially to the perception of visual scenes (Epstein and Kanwisher, 1998). Subsequently, eight critical memory-retrieval scans were conducted, during which memory for the encoded words was tested (Fig. 1B). Specifically, recognition of studied items and recollection of the source (Imaged or Read) associated with each item were indexed by a one-step memory test. During this test, subjects were presented all 400 studied words (Old items) as well as 400 unstudied lures (New items). On each trial, a test word was presented for 750 msec, followed by a fixation cross for 1750 msec. During this combined 2500 msec window, subjects indicated whether they recognized the word as having been studied and which encoding task was performed with the item when studied. Specifically, the subject made one of three responses: (1) "Old–Imaged" or (2) "Old–Read" indicated that the subject recognized the item as having been studied and recollected while encoding task was performed with the item, whereas (3) "New" indicated that the subject did not recognize the item as studied. Thus, measures of item recognition (recognized vs forgotten) and source recollection (source correct vs source incorrect) were obtained for each Old word, and measures of CRs and false recognition (FA–Imaged and FA–Read) were obtained for New words (Fig. 1C).

The order of conditions at test (Imaged, Read, and New) was determined using a sequencing program designed to maximize the efficiency of the event-related design (Dale, 1999). Trials were "jittered" using variable duration (2–18 sec) fixation null events.

fMRI procedures. Functional data were acquired using a gradient-echo echo-planar sequence (1.5 T; repetition time = 2 sec; echo time = 40 msec; 21 axial slices, 3.125 × 3.125 × 5 mm; 1 mm skip; 210 volumes; n = Siemens Sonata; Siemens AG, Erlangen, Germany). High-resolution T1-weighted (magnetization prepared-rapid acquisition gradient echo) structural images were collected for anatomical visualization. A bite-bar minimized head motion. Visual stimuli were projected onto a screen viewed through a mirror; responses were collected using a magnet-compatible response pad.

Data were analyzed using SPM99 (Wellcome Department of Cognitive Neurology, London, UK), with standard preprocessing procedures that included slice-timing and motion correction (Davachi and Wagner, 2002). Structural and functional images were normalized to templates based on the Montreal Neurological Institute (MNI) stereotaxic space. Images were resampled into 3 mm cubic voxels and spatially smoothed with an 8 mm full-width half-maximum isotropic Gaussian kernel. Statistical analyses were performed using the general linear model. Trials were modeled using a canonical hemodynamic response function and its first-order temporal derivative. Effects were estimated using a subject-specific fixed-effects model, with session-specific effects and low-frequency signal components treated as confounds. Linear contrasts yielded subject-specific estimates that were entered into second-level random-effects analyses. As detailed below, targeted voxel-based contrasts permitted the identification of neural signals associated with performance of the recognition memory task, as well as signals that differed across retrieval outcome (i.e., successful and unsuccessful trials). Regions consisting of at least five contiguous voxels that exceeded an uncorrected threshold of $p < 0.001$ were considered reliable.

Targeted region-of-interest (ROI) analyses were performed to further characterize the nature of the retrieval patterns in PFC and parietal regions that, a priori, were expected to show old–new effects. Unless oth-
erwise stated, the ROIs were defined based on the group-level voxel-wise contrast of all retrieval trials > fixation; note that this approach ensures that the ROIs were unbiased with respect to the retrieval conditions, thus permitting ANOVA tests of the activation patterns across conditions. For ROI definition, each ROI included all significant voxels within 6 mm of each maximum. For each subject, signal from the ROI was calculated by selectively averaging data with respect to peristimulus time per condition. The time point corresponding to peak activation in the ROI was identified by averaging across conditions and subjects; the time point associated with the maximum value was identified as the peak. Subsequently, peak activation values in the ROI were then determined for each condition for each subject. The memory conditions consisted of (1) Old items that were recognized and accompanied by correct source recollection (Item plus Source), (2) Old items that were recognized without correct source recollection (Item Only), (3) Old items that were forgotten (Miss), (4) New items that were correctly rejected as novel lures (Correct Rejections; CRs), and (5) New items that were falsely embraced as studied (False alarms; FAs) (Fig. 1C). Moreover, all Hits (Item plus Source and Item Only) and FAs were further subdivided based on the nature of the source response (Imaged or Read) indicated by the subject. The resulting ROI data were subjected to mixed-effect ANOVAs, treating Encoding task (Imaged and Read) and Memory condition as repeated measures and subjects as a random effect. Interaction analyses further assessed between-region differences in the patterns of activation across conditions.

Previous imaging data have revealed parahippocampal cortical activation during scene imagery (O’Craven and Kanwisher, 2000; Davachi et al., 2003), and left posterior ventrolateral prefrontal–premotor cortical activation during tasks requiring the generation of a novel phonological representation based on orthographic inputs (Clark and Wagner, 2003). Imaging investigations of episodic retrieval suggest that regions engaged during encoding are likely to be re-engaged when subjects successfully recollect specific details from the study episode (Nyberg et al., 2000; Wheeler et al., 2000). Thus, we expected a priori that parahippocampal and left posterior ventrolateral PFC–premotor regions would show task-sensitive recollection success effects for the Imaged and Read items, respectively. To complement the voxel-based analyses of the effects of recollection success (i.e., Item plus Source > Item Only, per task), additional targeted ROI analyses were performed. For these analyses, ROIs in parahippocampal and ventrolateral–premotor cortices were identified from the task-specific recollection contrasts, and the nature of the activation patterns in these ROIs across the retrieval conditions was subsequently explored with ANOVA.

Results

Recognition performance

Behavioral effects were considered reliable at an α-level of 0.05. Recognition response probabilities differed across Memory condition (Item plus Source, Item Only, and FAs) for Imaged ($F_{(2,32)} = 33.65$) and Read trials ($F_{(2,32)} = 10.56$) (Fig. 2A). Item recognition with source recollection (Item plus Source) was greater than recognition without recollection (Item Only) and the corresponding FA rate (Imaged, $F_{(1,16)} > 34.14$; Read, $F_{(1,16)} > 12.96$). Item Only recognition levels for Imaged (0.5) and Read trials (0.41) were higher than the corresponding FA rate ($F_{(1,16)} > 20.1$), when correcting for the opportunity to make such a response (Yonelinas and Jacoby 1995; Davachi et al., 2003). Finally, corrected recognition (collapsed across Item plus Source and Item Only trials) was superior after Imaged (0.55) than after Read encoding (0.38) ($F_{(1,16)} = 17.21$), and recognition with recollection (Item plus Source) was higher after Imaged than after Read encoding ($F_{(1,16)} = 7.4$) (Fig. 2A).

Reaction times (RTs) (Fig. 2B), analyzed using a two-way ANOVA of Encoding task (Imaged, Read, New) × Response type (Imaged, Read, New), revealed (1) an effect of the Encoding task ($F_{(2,32)} = 3.29$) such that RTs were faster for Imaged compared with Read items ($F_{(1,16)} = 6.5$), (2) an effect of Response type ($F_{(2,32)} = 5.73$) such that RTs were faster during Imaged and New compared with Read responses ($F_{(1,16)} > 4.75$), and (3) an interaction ($F_{(4,64)} = 4.52$). Post hoc contrasts revealed a nonsignificant RT difference between CRs relative to Hits (1687 msec) ($F_{(1,16)} = 3.58$; $p = 0.068$) and reliably faster RTs for CRs relative to FAs (1737 msec) ($F_{(1,16)} = 7.78$). Per task, RTs differed across Memory condition (Item plus Source, Item Only, Miss, and FA) for Imaged ($F_{(3,48)} = 6.48$) but not for Read trials ($F < 1$). For Imaged words, RT was reliably faster for Item plus Source compared with Item Only, Miss, and FA trials accompanied by the erroneous claim that the novel item had been imaged (FA–Imaged; $F_{(1,16)} > 5.92$). RTs did not differ reliably across Item Only, Miss, and FA–Read trials ($F_{(1,16)} < 3.7$; $p > 0.065$). The fact that RTs were faster on Item plus Source trials relative to all other trial types indicates that any observed fMRI retrieval responses associated with successful recollection cannot reflect a longer duty cycle or retrieval effort. Moreover, any similarities between the neural patterns of Item plus Source and FA trials cannot reflect duty cycle effects, because RTs differed across these conditions during Imaged trials and were comparable during Read trials.

Neural old–new effects

Initial fMRI analyses identified neural responses that differed between studied items that were correctly recognized as old (i.e., Hits) and unstudied items correctly recognized as new (i.e., CRs). Contrasting Hits (collapsed across correct and incorrect source recollection) with CRs revealed greater activation in left inferior parietal, left precuneus, and posterior cingulate cortices, as well as in the left frontopolar, ventrolateral, and posterior dorsolateral PFC (Fig. 3A; supplemental Table 1, available at www.jneurosci.org). These findings converge with previous demonstrations of left-lateralized old–new effects during episodic retrieval (Nolde et al., 1998; Konishi et al., 2000; McDermott et al., 2000; Kensing et al., 2003; Maril et al., 2003; Wheeler and Buckner, 2003) and with observations of differential activation in similar structures during source recollection versus temporal recency decisions (Dobbins et al., 2003) and source recollection versus
novelty detection (Dobbins et al., 2002; Dobbins and Wagner, 2003). In the reverse contrast, greater activation during CRs relative to Hits was observed in a number of regions, including right inferior parietal, bilateral lateral temporal, and right dor-

solation and frontopolar PFC (supplemental Table 1, available at www.jneurosci.org). Similar right PFC regions have been observed during comparisons of familiarity-based versus recollection-based recognition (Henson et al., 1999; Eldridge et al., 2000), temporal recency versus source recollection decisions (Dobbins et al., 2003), and Hits versus CRs when the target density of old to new test probes is <50:50, and CRs versus Hits when the target density of old to new test probes is >50:50 (Herron et al., 2004).

Importantly, in the currently observed regions showing old–new effects (i.e., Hits > CRs), subsequent voxel-based comparisons of item recognition with recollection (Item plus Source) to recognition without recollection (Item Only) failed to reveal activation differences that tracked recollection outcome. This was the case even when lowering the threshold to a lenient level (p < 0.01), suggesting that these left lateralized regions are sensitive to “perceived oldness” (Wheeler and Buckner, 2003) or “perceived familiarity” (i.e., Hits > CRs) but are insensitive to recollection success (i.e., Item plus Source ≈ Item Only) (cf. Cansino et al., 2002).

**Retrieval success versus recollection attempt**

To further explore the effect of Memory condition (Item plus Source, Item Only, and CRs trials), ROI analyses were conducted to examine the responses of a priori predicted frontal and parietal structures observed previously to show old–new effects (Fig. 3B–E). As described in Materials and Methods, ROIs were defined from an unbiased contrast comparing all retrieval trials to the fixation baseline. For each ROI, a one-way ANOVA of Memory condition (Item plus Source, Item Only, and CR) was performed separately for Imaged and Read trials, and a two-way ANOVA was performed with factors of Recollection success (Item plus Source vs Item Only) and Encoding task (Imaged vs Read). Two classes of regions were revealed: those showing an old–new effect and (1) insensitivity to Recollection success for both Imaged and Read items or (2) a Recollection success effect that depended on the Encoding task.

Left frontopolar–anterior ventrolateral PFC [approximately Brodmann’s area (BA) 10/47] (Fig. 3B), left mid-ventrolateral–dorsolateral PFC (approximately BA 9/45) (Fig. 3C), and anterior cingulate cortex (ACC; approximately BA 32; MNI coordinates of −6, 21, and 45) showed a reliable effect of Memory condition for both Imaged (F(2,32) values > 5.02; p < 0.02) and Read trials (F(2,30) valuesa > 4.74; p < 0.02). In each region, activation (1) was greater during Hits relative to CRs, (2) did not differ according to Encoding task (F(1,15) values < 1.05), (3) did not differ according to Recollection success (F values < 1), and (4) with the exception of frontopolar cortex, did not demonstrate a Recollection success × Encoding task interaction (F(1,15) values < 1.75; p > 0.21) (Fig. 3B,C). Although the interaction was reliable in frontopolar cortex (F(1,15) = 4.68; p < 0.05), post hoc comparisons indicated that activation did not differ according to Recollection success for either the Imaged (F(1,15) = 2.18; p > 0.15) or Read trials (F(1,15) = 2.51; p > 0.13). Collectively, the ROI analyses (and the above-described voxel-based contrasts) revealed that these regions showed old–new effects, but were insensitive to recollection outcome or encoding task. This pattern suggests that these regions are sensitive to perceived familiarity, being engaged during recollection attempts only for items eliciting above-criterion familiarity.

*The degrees of freedom are lower for the Read contrasts because one subject had to be excluded from this subset of the analyses as a result of having an insufficient number of trials in the Read–Item Only condition.
As with the preceding regions, left posterior ventrolateral–dorsolateral PFC (approximately BA 44/6/8) (Fig. 3D), left inferior parietal cortex (approximately BA 40/7) (Fig. 3E), and precuneus (approximately BA 7; coordinates of −9, −69, and 51) showed a reliable effect of Memory condition for both Imaged and Read trials ($F(1,15)$ values $> 7.66; p < 0.005$). However, each of these structures also showed greater activation for old items embraced as Read (Imaged–Item Only trials and Read–Item plus Source trials) relative to old items embraced as Imagined or CRs ($F$ values $> 8.12; p < 0.01$). Confirming this pattern, these regions demonstrated a reliable Recollection outcome $\times$ Encoding task interaction ($F_{(1,15)}$ values $> 23.31; p < 0.0005$), but no reliable effects of Recollection success ($F_{(1,15)}$ values $< 4.14; p > 0.06$) nor of Encoding task ($F$ values $< 1$). Thus, for Read-encoded items, activation was greater during Item plus Source trials relative to Item Only trials ($F_{(1,15)}$ values $> 12.33; p < 0.005$), whereas for Image encoded items, the opposite activation pattern was observed ($F_{(1,15)}$ values $> 6.31; p < 0.05$). Collectively, the ROI analyses complemented the voxel-wise analyses, indicating that these regions showed old–new effects and were insensitive to recollection outcome across encoding tasks. Thus, these regions were sensitive to perceived familiarity, being engaged during recollection attempts only for items eliciting above-criterion familiarity. However, unlike left frontotopical and mid-ventrolateral–dorsolateral PFC, these structures demonstrated the additional characteristic that activation differed according to recollection outcome depending on the encoding task: when subjects claimed to have previously read the item, activation during Hits was greater compared with when subjects claimed to have previously imaged the item, regardless of the actual task at encoding. We return to the effect of the encoding task when we consider recapitulation responses below.

Responses to Misses and FAs in regions showing old–new effects

Previous fMRI investigations of neural old–new effects have tended to lack sufficient power to permit comparison of Hits and CRs to Misses and FAs (but see Wheeler and Buckner, 2003; Weis et al., 2004). Such comparisons are critical for determining whether old–new effects are restricted to trials performed accurately (i.e., Hits > CRs) or whether they generalize to the contrast between perceived familiar items (i.e., Hits and FAs) relative to perceived novel items (i.e., Misses and CRs), regardless of the true memory status of the test probes, or to the contrast between studied items (i.e., Hits and Misses) compared with unstudied items (i.e., CRs and FAs), regardless of memory accuracy.

The present recognition performance levels were sufficient to permit a test of this question. Voxel-based comparisons revealed greater activation for Hits relative to Misses in similar left-lateralized structures as detected by the Hits versus CRs contrast, including inferior parietal, frontopolar, and posterior ventrolateral PFC, as well as left anterior insular and bilateral fronto-opercular cortex. ROI analyses confirmed these patterns, with activation during Misses being comparable with that during CRs ($F$ values $< 2.43; p > 0.12$) (Fig. 3B–E), indicating that the responses in regions showing old–new effects did not simply track the studied–unstudied dimension.

Intriguingly, given theoretical accounts of the mnemonic bases of FAs, the voxel-based contrast of Hits to FAs failed to reveal differential cortical responses between these two trial types in regions showing old–new effects. This outcome suggests that regions showing old–new effects demonstrate a generalized response to recognized compared with unrecognized items regardless of the true study status of the test probe. Because the pattern of activation during FAs is a central question of interest, especially in regions thought to be associated with recollection (or the recapitulation of encoding-based representations), we expand on FA-related activations below.

Neural recapitulation effects

Item recognition accompanied by recollection is hypothesized to re-engage regions that were engaged during encoding of the subsequently recollected event attributes. Accordingly, we predicted that in regions demonstrating recapitulation, the effect of Recollection success at retrieval will differ depending on the task performed at encoding; the encoding task presumably influences the nature of the episodic features bound to the item during learning and thus reinstated at retrieval (Johnson et al., 1997; Senkfor and Van Petten, 1998; Gonsalves and Paller, 2000b; Nyberg et al., 2000; Wheeler et al., 2000; Cycowicz et al., 2001). Consistent with this prediction, voxel-based comparisons, performed separately on Imaged trials and Read trials, revealed regions that were differentially engaged during recognition with recollection (Item plus Source) compared with without recollection (Item Only) (supplemental Table 2, available at www.jneurosci.org). These regions included left parahippocampal cortex (PHc; approximately BA 36) in the Imaged condition, and left premotor–posterior ventrolateral PFC (approximately BA 6/44) in the Read condition (Fig. 4A). Right PHc (approximately BA 36; coordinates of 30, −36, and −24) was also observed in the Imaged condition, at a slightly more lenient threshold ($p < 0.005$).

Targeted ROI analyses were subsequently performed on the PHc and left premotor–posterior ventrolateral PFC regions observed in the voxel-based recollection success contrasts. These analyses aimed to further assess the broader pattern of activation in these structures using a two-way ANOVA with factors of Recollection success (Item plus Source, Item Only, and Miss) and Encoding task (Imaged vs Read), and a one-way ANOVA, performed separately for Imaged and Read trials, with Memory condition (Item plus Source, Item Only, Miss, and CR) as the factor.

Left PHc showed an effect of Recollection success ($F_{(2,30)} = 8.51; p < 0.005$) and a Recollection success $\times$ Encoding task interaction ($F_{(2,30)} = 6.65; p < 0.005$) (Fig. 4A). For Image-encoded items, left PHc demonstrated an effect of Memory condition ($F_{(3,48)} = 10.26; p < 0.0001$), with post hoc contrasts revealing greater activation during Item plus Source compared with Item Only, Miss, and CR trials ($F_{(1,16)}$ values $> 13.67; p < 0.001$). These latter three conditions did not reliably differ ($F_{(1,16)}$ values $< 1.48; p > 0.23$), indicating that during Imaged trials this response was selective to recognition with recollection. Moreover, for Read-encoded items, activation did not differ according to Memory outcome ($F_{(3,45)} = 1.17; p > 0.32$). A qualitatively similar pattern was observed in right PHc, which showed a Recollection success $\times$ Encoding task interaction ($F_{(2,30)} = 7.57; p < 0.005$): activation was greatest during Item plus Source trials for Image-encoded items ($p < 0.05$), but did not differ across the Memory conditions for Read-encoded items ($p > 0.35$). Thus, Recollection success effects in left and right PHc were restricted to Imaged trials.

Left premotor–posterior ventrolateral PFC also showed a Recollection success $\times$ Encoding task interaction ($F_{(2,30)} = 10.38; p < 0.0005$), but, in contrast with PHc, the effect of Memory condition ($F_{(3,45)} = 7.07; p < 0.001$) was observed for Read-encoded items, with post hoc contrasts revealing greater activation for Item plus Source compared with Item Only, Miss, and CR trials ($F_{(1,15)}$ values $> 10.82; p < 0.005$) (Fig. 4A). No other
conditions differed reliably for Read items \( F(1,15) \) values < 1), indicating that during Read trials this response was selective to recognition with recollection. Moreover, for Image-encoded items, the effect of Memory condition was not reliable \( F(3,48) = 2.45; p < 0.08 \); if anything, there was greater activation during Item Only versus Item plus Source trials and comparable activation between Item Only and Miss trials. Thus, the Recollection success effect in left premotor–ventrolateral PFC was restricted to Read trials.

A Region (left PHc vs premotor–ventrolateral PFC) × Task (Imaged vs Read) × Recollection success (Item plus Source vs Item Only) interaction \( F(1,15) = 13.97; p < 0.005 \) confirmed that the functional patterns in the left PHC and premotor–ventrolateral PFC were distinct. This outcome indicates that the Recollection success effects in left PHC and premotor–ventrolateral PFC were tied to the task performed at encoding, as would be expected if these regions were engaged during the recapitulation of visuo-perceptual and phonological representations that were present, respectively, during encoding.

Finally, although multiple regions were observed in the voxel-based Recollection success contrasts (i.e., Item plus Source > Item Only, per task) (supplemental Table 2, available at www.jneurosci.org), the above analyses focus on PHC and left premotor–ventrolateral PFC because of a priori expectations that these regions would show task-selective effects. The posterior PHC is known to be engaged during scene imagery (O’Craven and Kanwisher, 2000), and greater PHC encoding activation is predictive of subsequent recollection of having performed the Image task during study (Davachi et al., 2003). In contrast, left posterior ventrolateral PFC–premotor cortex is thought to subserve the assembly of novel phonological representations (Clark and Wagner, 2003), processes that are differentially required during the Read task. Confirming that the currently observed regions were differentially engaged by the two tasks at encoding, we took advantage of the fact that we previously collected fMRI measures of neural activation during performance of the Image and Read tasks in an independent sample of subjects (Davachi et al., 2003). Importantly, the present Recollection success effects overlapped with the bilateral PHC regions demonstrating an Image > Read pattern of encoding activation and the left prefrontal–PFC region demonstrating a Read > Image pattern of encoding activation (Fig. 4B). Moreover, Recollection success effects in bilateral PHC overlapped with the parahippocampal regions detected by the "PPA localizer" scan.

**Neural responses during FAs**

A voxel-based contrast of FAs to CRs revealed differential activation during FAs in the same regions that were differentially engaged when comparing Hits with CRs. This finding suggests that many of the neural computations supporting veridical recognition are qualitatively similar to those supporting false recognition. That is, to the extent that a test probe elicits above-criterion familiarity, the computations subserved by these structures appear to be recruited to guide recollection attempts.

Leading models of recognition memory assume that FAs are based exclusively on recognition without recollection (i.e., above-criterion familiarity), rather than on familiarity, recollection, or both (Jacoby, 1991; Yonelinas et al., 1996). We tested this assumption by comparing trials on which subjects falsely embraced new items as having been studied in either the Image or the Read task. To the extent that FAs are recollection-free, regions showing task-selective recapitulation effects for studied items should fail to differentiate according to the false source (Imaged or Read) being misattributed to new items. In contrast, to the extent that false recollection partially contributes to FAs, the pattern of task selectivity (Imaged vs Read) seen for Item plus Source trials in the PHC and left premotor–ventrolateral PFC (Fig. 4A) should also be seen for FAs mistakenly attributed to having been encountered in the Imaged versus Read tasks, respectively. The sensitivity of these ROIs to false recollection was assessed using a two-way ANOVA, with factors of the Encoding task (Imaged and Read) and Memory condition (Item plus Source, Item Only, and FA).

The left PHC showed an effect of Encoding task \( F(1,15) = 6.61; p < 0.05 \), of Memory condition \( F(2,30) = 5.76; p < 0.01 \), and a reliable interaction \( F(2,30) = 9.41; p < 0.001 \). Importantly, activation during FA–Image trials was reliably greater than that during Imaged–Item Only and FA–Read trials \( F(1,15) \) values > 7.28; \( p < 0.05 \), suggesting that false recollection of imagery-based information accompanied FA–Imaged trials. (Note that the pattern of neural activation did not track the pattern of reaction times, arguing against the interpretation that differential neural activation was caused by differential retrieval effort.) In
contrast, FA–Read trials did not differ from Read–Item plus Source, Read–Item Only, and Imaged–Item Only trials ($F_{(1,15)}$ values < 2.95; $p > 0.1$), suggesting that this false recollection effect was selective to FAs to which subjects falsely indicated that they had imagined a spatial referent of the item. Interestingly, activation during FA–Imaged trials was weaker than that during Imaged–Item plus Source trials ($F_{(1,15)} = 4.46$; $p < 0.05$) (Fig. 5A), suggesting that, although false recollection may have accompanied some FA–Imaged decisions, the probability of this occurring was lower than that during veridical recollection. A qualitatively similar pattern was observed in right PHc (Encoding task × Memory condition; $p < 0.01$), although the false recollection effect was less robust (FA–Imaged vs Imaged–Item Only; FA–Imaged vs FA–Read; $p < 0.1$).

A complementary pattern was observed in left premotor–posterior ventrolateral PFC, which demonstrated an effect of Encoding task ($F_{(1,15)} = 5.31$; $p < 0.05$) and a Memory × Encoding task interaction ($F_{(2,30)} = 11.85$; $p < 0.0005$). Activation during FA–Read trials was reliably greater than that during Read–Item Only and FA–Imaged trials ($F_{(1,15)}$ values > 7.69; $p < 0.01$), but did not differ from that during Read–Item plus Source trials ($F_{(1,15)} = 1.09$; $p > 0.27$) (Fig. 5B). In addition, activation during FA–Read trials was reliably greater than that during Imaged–Item plus Source ($F_{(1,15)} = 9.74$; $p < 0.01$), whereas activation during FA–Imaged trials did not differ reliably from that during Imaged–Item plus Source, Imaged–Item Only, and Read–Item Only trials ($F_{(1,15)}$ values < 3.11; $p > 0.1$). Collectively, this pattern of activation is consistent with the conclusion that FA–Read trials were accompanied by false recollection.

**Discussion**

Recollecting the past depends on multiple mechanisms, including cognitive control processes that guide retrieval attempts and recovery processes that recapitulate episodic details (Rugg and Wilding, 2000; Buckner and Wheeler, 2001; Dobbins et al., 2002). The present results advance the understanding of recollection in three important ways. First, the findings offer a resolution to the debate regarding whether left PFC regions (known to show old–new effects) are sensitive to retrieval success (Konishi et al., 2000) or recollection attempt (Ranganath et al., 2000; Dobbins et al., 2003). Our data indicate that both hypotheses have merit, because left PFC subregions were sensitive to familiarity-based retrieval success but were insensitive to recollection-based success. These structures appear to support control processes that guide recollection attempts when familiarity levels fall above an internal decision criterion. Second, successful recollection elicited representation-specific activation in premotor–posterior prefrontal and parahippocampal cortices that paralleled that seen at encoding. This outcome was observed for items encountered once, suggesting a recapitulation rather than attentional modulation interpretation. Finally, false recognition was accompanied by recaptulation responses, suggesting that FAs can be based partially on illusory recollection.

**Familiarity-gated recollection attempt**

Consistent with previous studies (Nolde et al., 1998; Konishi et al., 2000; Maril et al., 2003), our data revealed left-lateralized old–new effects in left frontopolar, ventrolateral, and posterior dorsolateral PFC, inferior parietal cortex, and precuneus (Fig. 3A; supplemental Table 1, available at www.jneurosci.org). Such effects could reflect processes that are sensitive to successful retrieval (Henson et al., 1999; Habib and Lepage, 2000; Konishi et al., 2000; Donaldson et al., 2001) or processes that are brought to bear during attempts to recollect, regardless of recollection outcome (Ranganath et al., 2000; Dobbins et al., 2002, 2003). Herefore, adjudicating between these competing hypotheses has been complicated because success accounts emerged primarily from studies of yes–no recognition (Nolde et al., 1998; Henson et al., 1999; Konishi et al., 2000; McDermott et al., 2000), whereas attempt accounts emerged primarily using forced-choice recognition (Dobbins et al., 2002, 2003; cf. Ranganath et al., 2000).

The present data (from a one-step yes–no recognition, plus source recollection paradigm) offer a resolution that accommodates both hypotheses: specific left PFC and, to a lesser extent, parietal subregions mediate control processes that guide recollection attempts, with these processes being gated/disengaged when the recognition probe is perceived to be of low familiarity. Two observations motivate this conclusion. First, the left PFC and parietal regions showing old–new effects (1) also showed greater activation during FAs compared with CRs, but did not show differential activation during (2) Hits versus FAs and (3) Misses versus CRs. Thus, these regions were insensitive to the true memory status of the probe, because they were engaged to the extent that above-criterion familiarity (perceived familiarity) was elicited. Recently, Wheeler and Buckner (2003) reported greater activation during FAs compared with CRs in left ventrolateral PFC and parietal cortices; the present findings extend such perceived oldness/familiarity effects to left frontopolar and dorsolateral PFC. Second, regardless of the encoding task, these regions were insensitive to recollection outcome (Item plus Source vs Item Only), although parietal cortex was modulated by response type (Imaged vs Read).

It is possible that this null effect of recollection outcome in left
PFC and parietal regions emerged because the recognition test indexed source memory in a forced-choice manner. Accordingly, because subjects could not respond “old” without designating a source, some source responses reflect guessing. The behavioral results, however, revealed Item plus Source rates that were well above guessing, suggesting that recollection was present. Moreover, and more compellingly, recollection success effects were detected in other brain structures (bilateral PHc and left premotor–posterior ventrolateral PFC), indicating that the experimental design was sensitive to such effects when present. Finally, although it remains possible that other regions may have shown recollection success effects were guesses removed, a previous study that demonstrated recollection success effects in medial PFC and right parietal regions (Cansino et al., 2002) also failed to observe such effects in the currently noted left PFC and parietal regions showing perceived familiarity effects.

Collectively, our data indicate that left-lateralized PFC and, to a lesser extent, parietal cortices mediate processes that guide attempts to recollect, such as maintaining and elaborating on retrieval cues and monitoring the products of recollection attempts (Rugg and Wilding, 2000; Dobbins et al., 2002). Engagement of these processes depends on perceived familiarity, suggesting that these regions have early access to familiarity signals and are gated (in an automatic or controlled manner), depending on the expected utility of effortful recollection attempt.

Neural recapitulation supports recollection

Separate analyses on Imaged and Read trials revealed regions that were differentially engaged during recognition with recollection compared with recognition without recollection. Bilateral PHc (Fig. 4A) was differentially engaged during accurate recollection of having engaged in scene imagery at encoding, whereas the left premotor–posterior ventrolateral PFC (Fig. 4A) was differentially engaged during recollection of having performed the Read task. These structures were also differentially active during the performance of the Image and Read tasks at encoding (Fig. 4B), suggesting that task-sensitive recollection effects (Fig. 4A) mark the recapitulation of neocortical representations that were present at encoding.

An alternative interpretation of such task-sensitive recollection effects is that they reflect a consequence of top-down attentional orienting to domain-specific representational layers (i.e., cortical structures that differentially represent particular kinds of features), rather than the recapitulation of episodic details. From this perspective, retrieved knowledge about the general context associated with an item (e.g., visuoperceptual vs verbal) permits recruitment of attentional mechanisms that bias specific representational layers in attempts to recollect experiential details (regardless of the outcomes of such attempts). Attentional orienting may be particularly prevalent when items are encoded multiple times (Nyberg et al., 2000; Wheeler et al., 2000; Vaidya et al., 2002), because semantic knowledge about the context in which an item appeared may be abstracted across the multiple encoding events, allowing for recovery of this general knowledge even in the absence of experience-specific recollection. To diminish this possibility, in the present experiment, items were encoded once, eliminating the possibility of acquiring semantic knowledge about the encoding context of an item and demanding recollection of event-specific/trial-unique details. By definition, at retrieval, any knowledge of the task context must reflect episodic recollection because there was only a single episode in which the item appeared in the task context. Nevertheless, questions may still remain as to whether the neural effects reflect episodic recollection of the context per se or recollection of the representations (i.e., visuospatial imagery or phonological codes) elicited by this context.

False recognition and recollection

The recapitulation responses observed in PHc and left premotor–ventrolateral PFC during veridical recollection were also observed during false recognition accompanied by an erroneous Imaged or Read judgment, respectively. This parallel between the activation patterns during veridical recollection and false recognition (Fig. 5) provides important new evidence that FAs may be partially based on false recollection. Moreover, weaker recapitulation responses were observed in bilateral PHc during false recognition (FA–Imaged) relative to veridical recollection (Imaged–Item plus Source), suggesting that false recognition is not always accompanied by recollection (or is accompanied by recollection of fewer details).

Behavioral data indicate that false recognition can be accompanied by illusory recollection when novel recognition probes are conceptually related to studied items (i.e., “related lures”) (Roe-diger and McDermott, 1995; Schacter et al., 1998). Previous neuroimaging data indicate that false recognition of related lures can be accompanied by medial temporal lobe activation that resembles that seen during veridical recognition, whereas regions that represent domain-specific experiential details have been selectively or differentially engaged during veridical recognition (Schacter et al., 1996; Cabeza et al., 2001). However, the present findings suggest that the same neocortical recapitulation effects that occur during veridical recollection can also occur during false recognition, but to a lesser extent (Gonsalves and Paller, 2000a), lending support to the conclusion that false recollection is partially based on erroneously triggered recollection. Intriguingly, this false recollection effect emerged within the context of a paradigm that did not purposefully manipulate study/lure similarity, although the sheer number of experimental stimuli raises the possibility that this effect nevertheless stems from similarity between experienced and novel stimuli.

The current observation of false recollection during FAs, as indexed by illusory recapitulation, has important implications for models of recognition. A critical assumption of leading dual-process models is that FAs are based on above-criterion familiarity in the absence of recollection (Jacoby, 1991; Yonelinas et al., 1996). To the extent that neural recapitulation effects mark recollection, then, at a minimum, the present findings indicate that there are instances in which this assumption is violated. One possibility is that such effects emerge when subjects are forced to make source decisions during recognition, as in the present experiment. It remains an open question as to whether illusory recollection also emerges during the performance of simple recognition tasks, during which old and new items can be discriminated based solely on familiarity. A critical goal for future research is to determine the conditions in which FAs are partially based on recollection, because theoretically important estimates of recollection and familiarity in healthy and clinical populations rest on this assumption (Yonelinas, 1997; Yonelinas et al., 2002).

References

