Brain (1998), **121,** 1985–2002

Prefrontal cortex and recognition memory

Functional-MRI evidence for context-dependent retrieval processes

Anthony D. Wagner,[1] John E. Desmond,[1,2] Gary H. Glover[2] and John D. E. Gabrieli[1]

Departments of [1]Psychology and [2]Radiology, Stanford University, Palo Alto, USA

Correspondence to: Anthony Wagner, MGH-NMR Center, Bldg 149, 13th Street, Mail code 149 (2301), Charlestown, MA 02129–2060, USA.
E-mail: adwagner@nmr.mgh.harvard.edu

Summary

Functional neuroimaging studies of episodic recognition memory consistently demonstrate retrieval-associated activation in right prefrontal regions, including the right anterior and right dorsolateral prefrontal cortices. In theory, these activations could reflect processes associated with retrieval success, retrieval effort or retrieval attempt; each of these hypotheses has some support from previous studies. In Experiment 1, we examined these functional interpretations using functional MRI to measure prefrontal activation across multiple levels of recognition performance. Results revealed similar patterns of right prefrontal activation across varying levels of retrieval success and retrieval effort, suggesting that these activations reflect retrieval attempt. Retrieval attempt may include initiation of retrieval search or evaluation of the products of retrieval, such as scrutiny of specific attributes of the test item in an effort to determine whether it was encountered previously. In Experiment 2, we examined whether engagement of retrieval attempt is context-dependent by varying the context in which retrieval was performed; this was done by changing test instructions. Importantly, study and test stimuli were held constant, with only the test instructions varying across conditions. Results revealed that the pattern of right prefrontal activation varied across retrieval contexts. Collectively, these experiments suggest that right prefrontal regions mediate processes associated with retrieval attempt, with the probability of engaging these regions depending upon the retrieval context. Conflicting results across previous studies may be reconciled if the influence of retrieval context on the adopted retrieval strategy is considered. Finally, these results suggest that right prefrontal regions activated during recognition are not critical for successful performance as similar magnitudes of activation were present across multiple levels of performance. These findings reconcile imaging results with the selective effects of prefrontal lesions on retrieval-intensive episodic memory tests.

Keywords: episodic memory; explicit memory; prefrontal cortex; strategic processes; functional MRI

Abbreviations: AC = anterior commissure; ANOVA = analysis of variance; APC = anterior prefrontal cortex; BA = Brodmann area; DLPC = dorsolateral prefrontal cortex; ERP = event-related potential; fMRI = functional MRI; rCBF = regional cerebral blood flow

Introduction

Human memory consists of multiple forms of learning which differ in the component processes and neural networks that mediate their acquisition and retrieval (e.g. Cohen and Squire, 1980; Graf and Schacter, 1985; Roediger, 1990; Tulving and Schacter, 1990; Squire, 1992; Gabrieli, 1998). One important form of learning is episodic memory, which refers to memory for experiences that are associated with a specific spatial and temporal learning context (Tulving, 1972, 1983). Retrieval from episodic memory is thought to entail the conscious recollection of aspects of the past, and consists of multiple component processes including the representation of retrieval cues and products in working memory, the generation and initiation of a retrieval search, the recovery of information from episodic memory and the evaluation of the relevance of retrieved information to the current task goal (e.g. Raaijmakers and Shiffrin, 1981).

Retrieval from episodic memory is subserved by multiple brain regions, including the prefrontal cortex. Prefrontal lesions yield strategic processing deficits that result in select episodic retrieval impairments (Shallice, 1988; Milner *et al.*, 1991; Shimamura, 1994). Functional neuroimaging studies of episodic retrieval using PET or functional MRI (fMRI) have consistently revealed retrieval-associated activations in right anterior prefrontal cortex (APC), at or near Brodmann area (BA) 10, and right dorsolateral prefrontal cortex (DLPC), at or near BA 46 and 9. Both regions of activation have tended to be right lateralized, although there have been reports of left anterior activation (e.g. Schacter *et al.*, 1996*a*; Tulving *et al.*, 1996). Right APC and right DLPC activation generalizes across episodic retrieval tasks and stimulus forms, including word-stem cued recall, word and picture paired-associate recall, and recognition of faces, complex pictures, sentences and words (for reviews, see Buckner, 1996; Nyberg *et al.*, 1996*a*). Right APC and right DLPC activation has been reported during direct or intentional episodic retrieval but not during indirect or incidental retrieval (e.g. Squire *et al.*, 1992; Rugg *et al.*, 1997; Wagner *et al.*, 1997*b*). Activation in these regions does not appear restricted to episodic retrieval, however, as activation has also been noted during performance of working memory tasks (e.g. Petrides *et al.*, 1993; Fiez *et al.*, 1996; Gabrieli *et al.*, 1997).

Three functional interpretations of retrieval-associated right prefrontal activation have been proposed. The 'retrieval-attempt' hypothesis posits that right prefrontal activation reflects processes associated with attempts to retrieve the past (Kapur *et al.*, 1995; Nyberg *et al.*, 1995). Although not entirely specified, attempt processes may include initiation of a retrieval search or evaluation of the products of retrieval. Importantly, the probability of engaging these processes is thought to be independent of the success of such retrieval attempts. Alternatively, right prefrontal activation may reflect the extent of 'retrieval effort' associated with attempts to remember (Schacter *et al.*, 1996*a*), with such activation increasing with increasing retrieval effort. To the extent that poor memory for events necessitates greater effort to retrieve these memories, then this hypothesis predicts that right prefrontal activation should be inversely related to retrieval success. Finally, the 'retrieval-success' hypothesis posits that right prefrontal activation is associated with processes that accompany the successful recollection of a past experience (operationalized as a 'hit' in recognition of a correct retrieval in recall; Rugg *et al.*, 1996). Success processes may include integration of the retrieved item and contextual information or post-retrieval evaluation. Importantly, this account asserts that such processes cannot occur when retrieval has been unsuccessful.

Prior studies have failed to provide unambiguous support for any of these functional interpretations. Consistent with retrieval attempt, three PET studies of episodic recognition revealed similar right prefrontal activations across conditions varying in retrieval success (Kapur *et al.*, 1995; Nyberg *et al.*, 1995; Rugg *et al.*, 1997). However, apparently contradictory results come from three PET studies of recognition that demonstrated greater right prefrontal activation with greater retrieval success (Tulving *et al.*, 1994, 1996; Rugg *et al.*, 1996; see also Buckner *et al.*, 1998*b*). Finally, a PET study of word-stem cued recall provides support for the retrieval-effort hypothesis, with anterior prefrontal activation occurring when a low success (high effort) condition was compared with baseline, but not when a high success (low effort) condition was compared with baseline (Schacter *et al.*, 1996*a*).

Rugg *et al.* (1996) posit that these apparent inconsistencies may be reconcilable. First, they note that overall behavioural performance was poor in the Nyberg *et al.* (1995) study, raising the possibility that these authors may have been unable to detect success-related differences due to modest differences in success across conditions. Secondly, they posit that there is a non-linear relationship between the number of successful retrievals and prefrontal regional cerebral blood flow (rCBF) such that each successful retrieval elicits a nearly asymptotic rCBF response that persists for a period of time longer than the neural response due to the temporal blurring of the haemodynamic time-constant (Friston *et al.*, 1994). From this perspective, persisting rCBF responses in low success conditions reduce the rCBF differences between high and low conditions, even though the neural responses in these two conditions are markedly different.

Although Rugg *et al.* (1996) suggest that previous results may be reconciled within a single retrieval-success framework, the pattern of results across previous studies raises not only the possibility that right prefrontal activation may not be tightly coupled with success, but also that it may not be entirely independent of it either. The variable pattern of results obtained across previous studies may be due to between-study differences along two critical dimensions. First, as posited by Rugg *et al.* (1996), the magnitude of the behavioural differences between the various levels of success may be an important consideration when attempting to examine right prefrontal activation. Secondly, it is worth noting that the test instructions provided to participants varied across studies and even across conditions within studies (e.g. Tulving *et al.*, 1994, 1996). These differences in test instructions may result in differences in the adopted retrieval strategy. Both of these factors raise the possibility that a critical component of episodic retrieval that has not been extensively considered is the context in which retrieval is performed. Specifically, under certain testing situations participants may be more, or less, biased to engage in extensive retrieval search or post-retrieval evaluation. As the retrieval context varies, either due to differences in retrieval success or to task expectancies that arise from the test instructions, the adopted retrieval strategy may vary.

We conducted two fMRI studies aimed at further elucidating the functional significance of prefrontal activation during episodic retrieval. These studies were designed to test the attempt, effort and success hypotheses by considering the effects of performance levels and test instructions, on right prefrontal activation. In these studies, attempts were made (i) to ensure that the performance levels between conditions were considerably different, (ii) to implement conditions where the persisting-response interpretation could be functionally tested and (iii) to vary the retrieval context by varying the test instructions under which retrieval is performed. In Experiment 1, right prefrontal activation was examined during 'High', 'Low' and 'New' recognition conditions. In Experiment 2, test instructions were directly manipulated to consider the effects of retrieval context.

Experiment 1

In Experiment 1 prefrontal activation was examined as a function of success on a recognition memory test. Three levels of success (High, Low and New) were created by (i) varying the nature of operations engaged during encoding, using a levels-of-processing technique (Craik and Lockhart, 1972), and (ii) varying the density of studied items appearing on the recognition test. High performance was created by testing memory for words studied twice in a semantic manner, with target density at test being high (60 out of 66 test items were old). Low performance was created by testing memory for words studied once in a non-semantic manner, with target density also being high (60 out of 66 were old). The New condition was created by testing memory for words, the vast majority of which had not been studied, thus target density was very low (6 out of 66 were old). In three fMRI scans, activation during recognition (High, Low or New) was compared with that during a non-memory baseline condition consisting of silent reading of novel words (Read). In a fourth scan, blocks of High recognition success were directly compared with blocks of Low recognition success (Fig. 1). The logic of this experiment parallels that of the earlier PET studies (e.g. Kapur *et al.*, 1995; Nyberg *et al.*, 1995; Schacter *et al.*, 1996*a*). To the extent that retrieval success markedly differs across the three recognition conditions, each functional interpretation of right prefrontal activation predicts a unique pattern of results (Table 4). To verify that levels of retrieval success would markedly differ across the High, Low and New conditions, a separate behavioural study was conducted prior to scanning, to examine recognition performance and response latencies under these conditions.

Methods

Participants

For the behavioural study, 16 volunteers (aged 18–24 years) participated. Data from two participants were excluded and replaced because they failed to follow the task instructions. For the fMRI study, nine right-handed volunteers (seven female and two male, aged 19–34 years) participated. Data from two additional participants were collected but excluded, one due to a susceptibility artefact and the second due to failure to follow the task instructions. All participants gave informed, written consent and were from the Stanford community. The study was approved by the Institutional Review Board at Stanford University.

Materials

The stimuli were 528 abstract or concrete nouns (3–10 letters long); half were printed in upper case and half in lower case. Eight 60-item blocks were created such that each block contained: 15 upper-case abstract words (e.g. 'LOVE'), 15 lower-case abstract words (e.g. 'hope'), 15 upper-case concrete words (e.g. 'SHOE') and 15 lower-case concrete words (e.g. 'dog'). Blocks were matched for word length and word frequency. The remaining items were divided into eight 6-item blocks; each block contained three abstract and three concrete words. Three were in upper case and three were in lower case letters.

Two kinds of study lists were created: Low performance and High performance. Low performance lists consisted of 123 unique items, 120 items from two of the 60-item blocks plus three items from one of the 6-item blocks, randomly ordered once. High performance lists consisted of 123 unique items, 120 items from two other 60-item blocks plus the other three items from the one 6-item block; the 123 items appeared once in a random order and then a second time in a new random order.

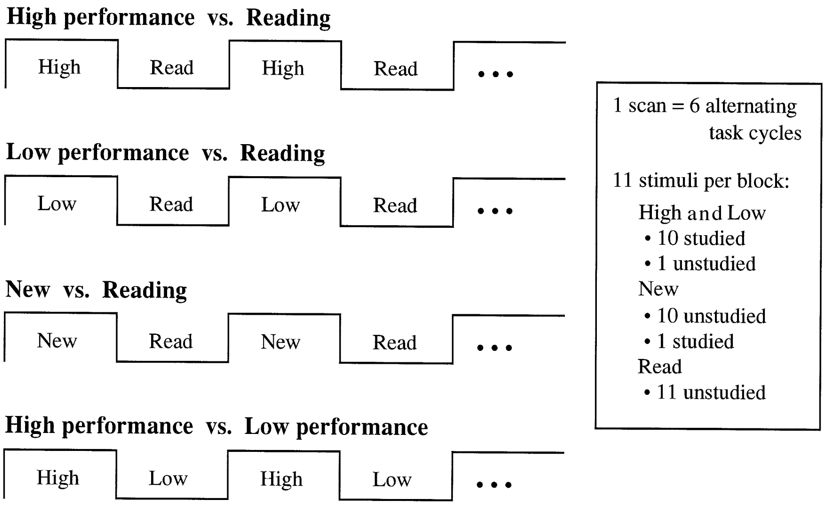

Fig. 1 Schematic diagram of the four recognition test scans in Experiment 1.

Four 132-item test lists were created; each contained 12 test blocks of 11 items each. Six of the test blocks consisted of a random sequence of 10 items from a given 60-item block plus one item from a given 6-item block. The other six test blocks consisted of items from a different 60-item block and a different 6-item block. Each test list alternated back and forth six times between the two kinds of test blocks. The four test lists were presented in the same fixed order. When combined with the study lists, there were four kinds of test conditions. The Low–Read sequence alternated between blocks of 10 studied items from the Low study list plus one unstudied item and blocks of 11 unstudied items. The High–Read sequence alternated between blocks of 10 studied items from the High study list plus one unstudied item and blocks of 11 unstudied items. The New–Read sequence alternated between blocks of 10 unstudied items plus one studied item (across the six blocks, three were Low and three were High) and blocks of 11 unstudied items. The High–Low sequence alternated between blocks of 10 studied items from the High study list plus one unstudied item and blocks of 10 studied items from the Low study list plus one unstudied item. The following counterbalancing was performed across participants: (i) each test condition appeared in each of the four possible test positions (first, second, third and fourth); (ii) the order of the two kinds of blocks in each test condition was counterbalanced (i.e. the Low and Read conditions were paired as Low–Read and as Read–Low); and (iii) items were counterbalanced across the High, Low, New and Read blocks.

Behavioural study procedure

The experiment consisted of two study phases, followed by four ‘yes–no’ recognition tests. A Low performance phase was immediately followed by a High performance phase. For both, individual words appeared centrally on a computer screen for 2000 ms, each followed by a 400-ms inter-stimulus interval. In the Low study phase, participants judged whether each word was printed in upper-case or lower-case letters. Half of the participants responded to words in upper case by pressing the slash key (‘/’) on the keyboard and did not respond otherwise, and half responded to words in lower case with the slash key. In the High study phase, participants judged whether the word represented an abstract or concrete concept; half were instructed to respond to abstract words and half responded to concrete words. Words on the High study list appeared twice in a spaced manner.

Following this study, participants advanced to four test conditions (Fig. 1). For all conditions, individual words again appeared centrally (2000 ms on; 400-ms inter-stimulus interval). The High–Read, Low–Read and New–Read test conditions consisted of six alternating-task cycles; each alternated between blocks of a recognition memory task and blocks of a reading baseline task. The fourth condition, High–Low, consisted of six alternations between blocks of High and Low recognition. Prior to High and Low blocks, an instruction cue was presented (‘Task—Old’). Participants were instructed to determine whether each presented test item had been previously studied, and to respond if they recognized the word as studied. Prior to New blocks, an instruction cue was presented (‘Task—New’). Again, participants were instructed to determine whether each test item had been previously studied (but, to control for the number of motor responses across conditions, participants were instructed to respond if they failed to recognize the item). Prior to Read blocks, an instruction card was presented (‘Task–Read’). Participants were instructed to read each word silently to themselves, and to respond upon completion. The computer recorded responses and response latencies.

fMRI study procedure

Imaging was performed with a 1.5-T whole-body MRI scanner (General Electric Medical Systems Signa, Rev. 5.5, Waukesha, USA). Two 5-inch-diameter local receive coils were bilaterally positioned over the prefrontal cortex to obtain the activation signal. Head movement was minimized by using a ‘bite bar’ formed with each participant’s dental impression. A T_2*-sensitive gradient echo spiral sequence (Noll *et al.*, 1995; Glover and Lai, 1998), which is relatively insensitive to pulsatility motion artefacts (Glover and Lee, 1995), was used for functional imaging [repetition time (TR) = 720 ms; echo time (TE) = 40 ms; flip angle = 65°]. For each scan, 120 images were acquired continuously over a 346-s session from each of eight coronal slices (6 mm thick; 1-mm inter-slice space; 2.35-mm in-plane resolution) situated 14–63 mm rostral to the anterior commissure (AC). Four interleaves were obtained for each image, with a total acquisition time of 2.88 s per image. T_1-weighted, flow compensated spin-warp anatomy images (TR = 500 ms; minimum TE) were acquired for all sections imaged during the functional scans.

Prior to functional data collection, participants were engaged in the Low and High study phases coincident with acquisition of T_1-weighted sagittal localizer and in-plane anatomy images. Following the study period, functional scans were conducted during performance of the four recognition test conditions. Participants responded by pressing a pneumatic bulb. Stimuli were generated from a computer and back-projected onto a screen located above the participant’s neck; visual images were viewed in a mirror mounted above the participant’s head.

fMRI data analysis

Image reconstruction was performed off-line by transferring the raw data to a Sun SparcStation (Sun Microsystems, Mountain View, Calif., USA). The data were resampled into a Cartesian matrix and then processed with a 2D fast Fourier transform. Once individual images were reconstructed, the time series from each pixel was correlated with a reference waveform. The reference waveform was calculated by convolving a square wave representing the time course of the alternating task conditions with a data-derived estimate of the haemodynamic response function (Friston *et al.*, 1994). For the present study, the task frequency was ~0.017 Hz (six cycles over 346 s). The resulting correlations were transformed into a Z-score map (SPM{Z}) (Friston *et al.*, 1994). Because this study tests particular hypotheses, pixels that satisfied the criterion of $Z > 1.96$ (representing significance at $P < 0.05$, two-tailed) were selected and overlaid on the corresponding T_1-weighted structural image. For display purposes, the SPMs were processed with a median filter with spatial width of 3 pixels to emphasize spatially coherent patterns of activation.

To analyse functional activation across participants, the SPM{Z} maps were averaged. Each section from each participant was transformed onto a corresponding standardized coronal section (Talairach and Tournoux, 1988) at the same distance rostral to the AC (see Desmond *et al.*, 1995). Following transformation, the average Z-score for each pixel in a section was computed across participants, and pixels that reached a statistical threshold of $P < 0.05$ (two-tailed) were displayed on each map. Stereotaxic coordinates for clusters of activation within these averaged SPMs (statistical parametric maps) were obtained using the coordinate system of the Talairach and Tournoux (1988) atlas. Activations exceeding a cluster size threshold of two pixels are described, with the reported coordinates corresponding to the centroid of the cluster. The maximum Z-score for each cluster is also reported.

To assess the pattern of activation across levels of recognition performance (High, Low and New), further analyses were conducted on 'regions of interest' identified as regions significantly associated with recognition at the group level: bilateral APC, right DLPC, bilateral frontal opercular cortex, medial prefrontal cortex and bilateral posterior inferior frontal gyrus. These anatomical regions were identified in each individual participant, and the regions of interest were specified as the subset of pixels within the region that demonstrated greater activation during any of the recognition conditions relative to the baseline for that individual. Once these regions of interest were specified, the mean Z-score across the region of interest was determined for each participant for each condition. For each region of interest, an analysis of variance (ANOVA) was performed to determine whether the magnitude of activation within the region varied systematically across the levels of recognition performance. Further, a t test was conducted on the mean Z-score over each region of interest from the High versus Low comparison to determine whether this value systematically differed from zero.

Results

Behavioural results

The behavioural and fMRI studies yielded performance results from five memory conditions: High (High–Read), Low (Low–Read) and New (New–Read) in the context of the reading (Read) baseline, and High (High–Low) and Low (Low–High) in the context of each other (Table 1). Unless otherwise indicated, an α-level of 0.05 was used for all statistical tests. The results revealed considerable differences in retrieval success (indexed by the number of hits) and retrieval effort (indexed by recognition accuracy and response latency) across the recognition conditions.

Numbers of hits and false alarms

Retrieval success declined across the conditions from High to Low to New. Hits were greater than false alarms [behavioural study $F(1,15) = 1257.68$; fMRI study $F(1,8) = 110.30$]. There was a significant memory-condition × item-type interaction [behavioural study $F(4,60) = 252.15$; fMRI study $F(4,32) = 135.19$]; hits were greater in the High–Read than the Low–Read conditions [behavioural study $F(1,15) = 1149.68$; fMRI study $F(1,8) = 229.34$] which were, in turn, greater than those in the New–Read condition [behavioural study $F(1,15) = 51.58$; fMRI study $F(1,8) = 28.64$], whereas false alarms were greater in the New–Read than in the High–Read conditions [behavioural study $F(1,15) = 9.42$; fMRI study $F(1,8) = 34.65$] which, in turn, did not differ from those in the Low–Read condition (both $F < 1.0$). Neither hits nor false alarms differed between the High–Read and High–Low conditions (both $F < 1.04$), nor between the Low–Read and Low–High conditions (both $F < 1.0$).

Accuracy

Accuracy ($P_{\text{hits}} - P_{\text{false alarms}}$) differed across the memory conditions [behavioural study $F(4,60) = 62.41$; fMRI study $F(4,32) = 24.05$]; accuracy was superior in the High–Read condition than in the New–Read condition [behavioural study $F(1,15) = 16.02$; fMRI study $F(1,8) = 9.36$] which was, in turn, superior to that in the Low–Read condition [behavioural study $F(1,15) = 36.42$; fMRI study $F(1,8) = 10.37$]. Accuracy was higher in the High–Low than in the High–Read condition in the behavioural study [$F(1,15) = 4.71$], whereas it was comparable in these conditions in the fMRI study ($F < 1.0$). In both studies, accuracy was similar across the Low–High and the Low–Read conditions (both $F < 1.0$).

Response latencies

Response latencies were only collected during the behavioural study. They differed across the memory conditions [$F(4,60) = 8.12$]; mean median reaction times were shorter in the High–Read than in the New–Read condition [$F(1,15) = 6.39$] which, in turn, were shorter than those in the Low–Read condition [$F(1,15) = 4.25$]. Response latencies were similar in the High–Read and High–Low conditions ($F < 1.0$), and in the Low–Read and Low–High conditions [$F(1,15) = 1.42$, $P > 0.20$]. Thus, both accuracy and response latency measures

Table 1 *Experiment 1: recognition performance from behavioural and fMRI studies*

	Test condition				
	High–Read	Low–Read	New–Read	High–Low	Low–High
Behavioural study					
Hits	0.92 (55.2)	0.25 (15.2)	0.59 (3.6)	0.91 (54.8)	0.27 (16.2)
False alarms	0.22 (1.3)	0.08 (0.5)	0.10 (6.3)	0.10 (0.6)	0.10 (0.6)
Accuracy	0.70	0.17	0.49	0.82	0.18
Reaction time (ms)	798.7	958.8	886.8	804.2	917.2
fMRI study					
Hits	0.77 (46.3)	0.25 (14.9)	0.63 (3.8)	0.81 (48.4)	0.24 (14.3)
False alarms	0.11 (0.7)	0.09 (0.6)	0.22 (12.9)	0.07 (0.4)	0.11 (0.7)
Accuracy	0.66	0.16	0.42	0.73	0.13

Numbers in parentheses correspond to the raw number of hits and false alarms for each condition.

Table 2 *Experiment 1: regions of increased activation during episodic recognition*

Region	Talairach coordinates			$Z_{\max}$	BA
	x	*y*	*z*		
High > Read					
Medial prefrontal/anterior cingulate	–2	21	41	6.36	6, 8, 32
R anterior prefrontal	32	56	2	5.31	10, 11
R dorsolateral prefrontal	46	35	19	5.10	46, 9
R posterior inferior prefrontal	41	14	35	4.50	9, 8, 44
R frontal operculum	35	21	–3	4.32	47
L posterior inferior prefrontal	–44	14	31	3.66	9, 8, 44
L frontal operculum	–31	21	4	3.60	47
L anterior prefrontal	–43	49	1	3.36	10
Low > Read					
R frontal operculum	36	21	–1	5.70	47
R anterior prefrontal	29	56	–7	5.46	10, 11
L frontal operculum	–30	21	3	5.13	47
Medial prefrontal/anterior cingulate	–2	21	41	4.98	6, 8, 32
L posterior inferior prefrontal	–43	14	28	4.62	9, 8, 44
R dorsolateral prefrontal	45	35	18	4.35	46, 9
R posterior inferior prefrontal	40	14	32	3.93	9, 8
New > Read					
R anterior prefrontal	26	56	–6	5.85	10, 11
R frontal operculum	34	21	–3	4.95	47
Medial prefrontal/anterior cingulate	2	21	37	4.74	6, 8, 32
L posterior inferior prefrontal	–44	14	27	3.69	9, 8, 44
R posterior inferior prefrontal	39	14	35	3.42	9, 8
R dorsolateral prefrontal	44	35	18	2.79	46, 9

R = right; L = left.

indicate that retrieval effort differed across the High, Low and New conditions.

The fMRI results

Analysis of the functional data revealed greater activation in a number of prefrontal cortical regions during recognition relative to baseline, including the right APC and right DLPC (Table 2). Region-of-interest analyses revealed that the magnitudes of activation across the three levels of performance (High, Low and New) relative to baseline were comparable (Table 3). These results are inconsistent with predictions from the retrieval-success and retrieval-effort hypotheses, and are generally consistent with those from the retrieval-attempt perspective (Table 4). Importantly, the demonstration of activation in the New condition cannot be accounted for by the success hypothesis.

Right APC

In all three scans in which recognition was compared with the baseline condition, recognition resulted in greater right APC activation (Fig. 2). Right APC activation extended approximately from 49–63 mm rostral to the AC, and was situated in the right middle frontal gyrus, frontomarginal sulcus and lateral orbital sulcus (at or near BA 10 and 11;

Table 3 *Experiment 1: mean* Z*-score for each region of interest and level of recognition performance*

	Mean Z-score across region of interest			
	New–Read	Low–Read	High–Read	High–Low
R anterior prefrontal	1.41	1.30	1.91	0.19
L anterior prefrontal	0.95	1.18	1.56	0.13
R dorsolateral prefrontal	1.53	1.56	1.74	0.07
R frontal operculum	1.46	1.70	1.65	–0.03
L frontal operculum	1.23	1.48	1.45	–0.01
Medial prefrontal/anterior cingulate	1.47	1.63	1.76	–0.05
R posterior inferior prefrontal	1.37	1.52	1.83	0.22
L posterior inferior prefrontal	1.22	1.59	1.63	0.09

R = right; L = left.

Table 4 *Experiment 1: predicted and resulting pattern of right APC and right DLPC activation*

Hypothesis	Predicted pattern from behavioural results
Retrieval success	
Number of hits	High > Low > New = Read
Retrieval effort	
Accuracy (inversely related)	Low > New > High > Read
Response latency	Low > New > High > Read
Retrieval attempt	
Independent of performance	High = Low = New > Read
	Actual functional results
Right APC activation	High = Low = New > Read
Right DLPC activation	High = Low = New > Read

Table 2). At the group level, this frontopolar activation appeared to be greater in the High–Read comparison than in the Low–Read and New–Read comparisons. To assess whether this difference was systematic across subjects, the mean Z-score for the right APC region was computed for the High–Read, Low–Read and New–Read comparisons for each participant. The region-of-interest analysis indicated that the memory condition did not systematically affect the magnitude of right APC activation [$F(2,16) = 1.43$, $P > 0.25$; Table 3]. Consistent with this finding, the High–Low scan failed to reveal a significant difference in right APC activation; the region-of-interest analysis demonstrated no difference either [$t(8) = 1.25$, $P = 0.25$].

At the group level, a significant cluster of retrieval-associated activation was present in the left APC only in the High–Read comparison. However, as with the right APC, there was considerable individual variability in the pattern of activation across memory conditions, with the region-of-interest analysis failing to reveal an effect of memory condition [$F(2,16) = 1.15$, $P > 0.30$; Table 3]. Further, the High–Low scan failed to reveal a significant difference in left APC activation, and the region-of-interest analysis demonstrated no difference either [$t(8) = 0.86$, $P > 0.40$].

Right DLPC

All levels of recognition resulted in greater right DLPC activation relative to baseline (Fig. 2), with this activation situated in right middle frontal gyrus and inferior frontal sulcus (at or near BA 46, 9; Table 2) extending approximately from 28–35 mm rostral to the AC. At the group level, the magnitude of this activation appeared to decline across the High–Read, Low–Read and New–Read comparisons. However, there was considerable individual variability in the pattern of activation across the three memory conditions. The region-of-interest analysis indicated no effect of memory condition ($F < 1.0$; Table 3). The High–Low scan failed to reveal a significant difference in right DLPC activation, and the region-of-interest analysis demonstrated no difference either [$t(8) = 0.84$, $P > 0.40$].

Other regions demonstrating retrieval-associated activation

Recognition resulted in greater activation in bilateral frontal opercular cortex, medial prefrontal cortex (near the anterior cingulate) and bilateral posterior inferior frontal gyrus relative to baseline (Fig. 2). For all regions, the region-of-interest analysis indicated that the memory condition did not affect the magnitude of activation (all $F < 1.05$, $P > 0.35$; Table 3). Further, the High–Low scan failed to reveal a significant difference in any region, with the region-of-interest analysis also demonstrating no differences (all P values > 0.10).

Frontal opercular activation was present in the anterior insular and inferior frontal cortices (at or near BA 47) situated approximately 21–28 mm rostral to the AC (Table 2). Although previous studies of episodic retrieval have demonstrated activation in this region (e.g. Buckner *et al.*, 1995, 1996), frontal opercular activation is not specific to episodic retrieval as activation has been noted during performance of other tasks, including: (i) semantic classification of words (e.g. Wagner *et al.*, 1997*a*); (ii) target detection of auditorily presented verbal and non-verbal stimuli (Fiez *et al.*, 1995); and (iii) maintenance of verbal material in working memory (e.g. Desmond *et al.*, 1996; Fiez *et al.*,

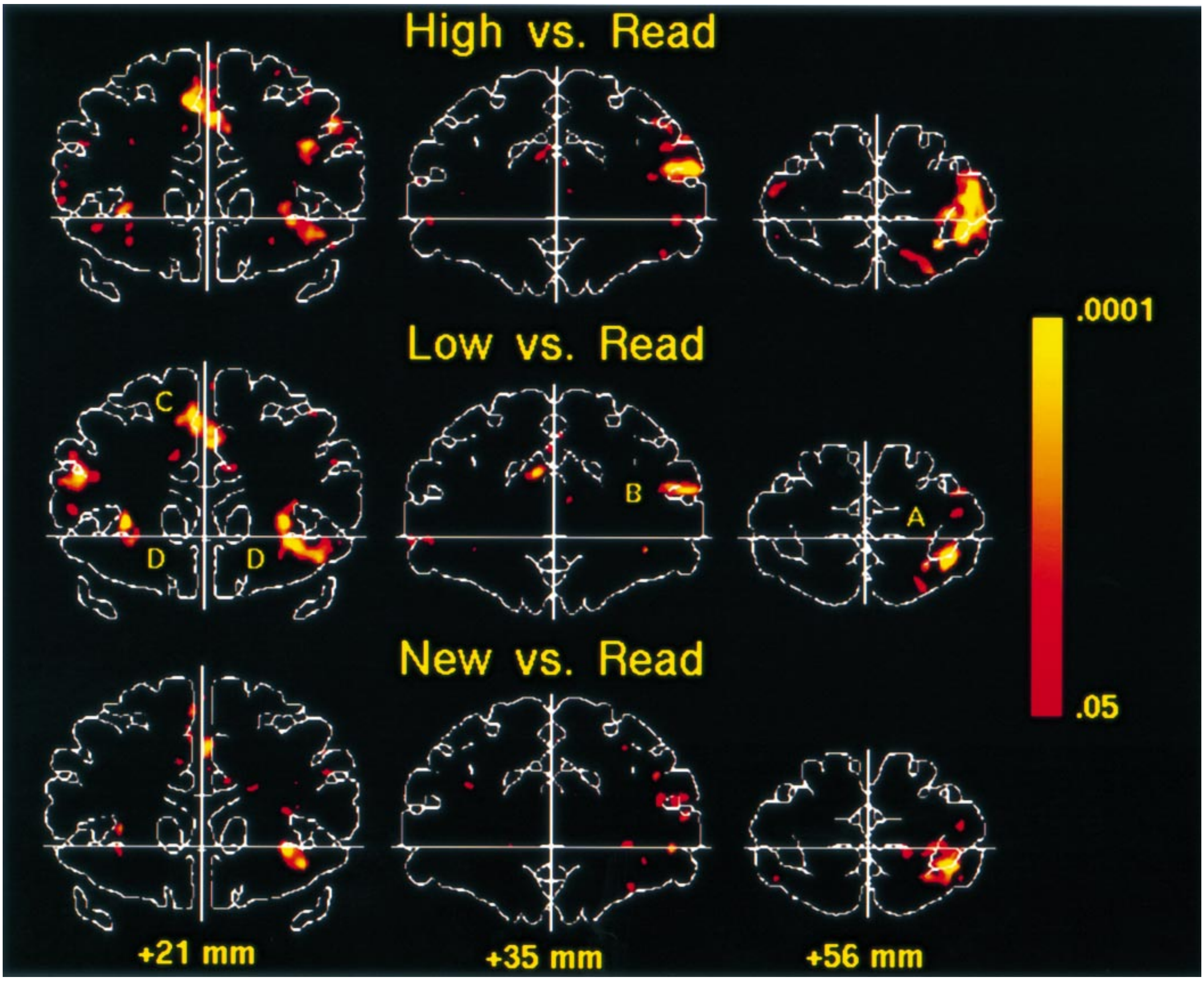

Fig. 2 Composite coronal images of activation in three sections corresponding to +21, +35 and +56 mm rostral to the AC. Each column displays the composite from one section (indicated at the bottom), and each row displays data from one condition (indicated above). All composites reflect *Z*-scores averaged across all participants. Regions that show greater activation during the memory condition relative to the reading baseline are overlaid on the corresponding coronal section from the stereotaxic atlas of Talairach and Tournoux (1988). In all conditions, greater activation was present in right APC (A), right DLPC (B), medial prefrontal cortex near the anterior cingulate (C) and bilateral frontal opercular cortex (D). For all figures, the left side of the image corresponds to the left side of the brain.

1996). The functional significance of frontal opercular activations has yet to be resolved; one suggestion is that this region mediates a high-level articulatory or phonological coding (Fiez *et al.*, 1995), which may differ across the recognition and reading of words.

Posterior inferior frontal activation was present bilaterally in inferior frontal gyrus, middle frontal gyrus and inferior frontal sulcus (at or near BA 9, 8 and 44) situated approximately 14–21 mm rostral to the AC. The location of this activation in the left hemisphere falls near the region frequently visualized during phonological working memory tasks (e.g. Awh *et al.*, 1996) and has been previously visualized during episodic retrieval of verbal stimuli (e.g. Buckner *et al.*, 1996; Wagner *et al.*, 1997*d*). This activation is consistent with the possibility that the phonological and articulatory demands were greater during recognition relative to baseline.

Medial superior frontal cortex activation bordered the anterior cingulate (at or near BA 6, 8 and 32) and was situated approximately 14–21 mm rostral to the AC. As with frontal opercular activation, previous PET studies of episodic and semantic retrieval have visualized activation in a similar region (e.g. Raichle *et al.*, 1994). This medial prefrontal region may correspond to the anterior extent of the supplementary motor area, or pre-supplementary motor area (Picard and Strick, 1996), and may mediate the internal generation of motor programmes required for responding (e.g. Buckner *et al.*, 1996). The functional significance of this activation within the context of retrieval is not yet known.

Regions of decreased activation

Two regions of decreased activation were present across all three memory conditions relative to baseline. Posteriorly, there was a decrease in the left middle frontal gyrus (at or near BA 8, 9 and 46) situated approximately 14–28 mm rostral to the AC (with High–Read, Z_{max} = –4.29 for cluster x, y, z = –44, 28, 31, and Z_{max} = –3.48 for –38, 21, 43; with Low–Read, Z_{max} = –4.23 for –42, 28, 30, and Z_{max} = –3.60 for –40, 21, 41; with New–Read, Z_{max} = –4.08 for –42, 28, 29, and Z_{max} = –4.41 for –34, 21, 50). Anteriorly, there was a decrease in medial orbital frontal cortex (at or near BA 10, 32), situated approximately 42–49 mm rostral to the AC (with High–Read, Z_{max} = –5.28 for –5, 49, –3; with Low–Read, Z_{max} = –5.01 for –4, 49, –5; with New–Read, Z_{max} = –4.05 for –5, 49, –4).

Experiment 2

The results of Experiment 1 are consistent with studies supporting right prefrontal activation as a reflection of retrieval attempt, rather than retrieval success or effort. Resolution of between-study inconsistencies is critical to understanding the functional contributions of right prefrontal activation. We propose that the retrieval context plays a pivotal role in influencing the nature of processing engaged during retrieval. Depending on the retrieval context, participants may be more or less biased to engage in extensive retrieval search or post-retrieval evaluation. Aspects of the retrieval context that may influence the probability of engaging in these processes include on-line monitoring of retrieval performance and expectations that arise from test instructions. For example, in two previous studies supporting the retrieval-success interpretation, participants were provided with different instructions during the high and low success conditions (Tulving *et al.*, 1994, 1996). An alternative interpretation of the results from these studies is that different retrieval strategies were adopted in the old and new test contexts because of the differences in test instructions (see below).

In Experiment 2 we tested the hypothesis that right prefrontal activation reflects the strategic search or post-retrieval evaluative processes postulated by attempt theories, and that the probability of engaging in these processes varies across retrieval contexts. In two scans, High and New recognition were directly compared (Fig. 3). In the High–New (Standard 1) scan, recognition took place in the context of 'standard' test instructions (i.e. by determining whether the test item was old, as in Experiment 1). The '1' denotes the number of new and old items in the High and New blocks, respectively. In the High–New (Biasing 1) scan, recognition took place in the context of 'biasing' test instructions. These instructions, modelled after those of Tulving *et al.* (1994, 1996), informed participants of the relative probabilities of old and new items in each condition. In the High condition, participants were instructed to monitor for the few new items; in the New condition, participants were instructed to monitor for the few old items. Importantly, the study conditions and test materials were identical in these two scans, with the critical difference being the test instructions provided.

In Experiment 2 the success hypothesis (Rugg *et al.*, 1996) was further examined. Demonstration of right prefrontal activation in the New condition in Experiment 1 is inconsistent with this hypothesis. In Experiment 2, two standard-instruction scans [High–New (Standard 1) and High–New (Standard 0)] were used to compare High and New conditions directly. The New condition in the High–New (Standard 0) scan was a pure condition (each block consisted of 11 new items). To the extent that right prefrontal activation reflects retrieval success, both scans should reveal greater right prefrontal activation in the High than the New condition.

Methods

Participants

Six right-handed volunteers (two females and four males) from the Stanford community (aged 19–30 years) participated. All participants gave informed, written consent.

Materials and procedure

The stimuli were a subset of those used in Experiment 1. Two study lists were created; each consisted of 192 unique items, 180 items from three of the 60-item blocks plus 12 items from two of the 6-item blocks. Three 132-item test lists were created in a manner identical to that in Experiment 1. There were three test conditions. The High–New (Standard 1) and High–New (Biasing 1) sequences alternated between blocks of 10 studied items plus one unstudied item and blocks of 10 unstudied items plus one studied item, and the High–New (Standard 0) sequence alternated between blocks of 10 studied items plus one unstudied item and blocks of 11 unstudied items. Across participants, the test position of the High–New (Standard 1) and High–New (Standard 0) conditions was first and second equally often.

The imaging procedure was identical to that in Experiment 1, with the exceptions that (i) a prototype whole-head receive coil was used for five participants and local receive coils were used for one participant, and (ii) for two participants, functional images were corrected for motion-related artefacts using the automated image registration algorithm for the in-plane dimensions (AIR 2.0; Woods *et al.*, 1992). The study phase consisted of 'abstract versus concrete' decisions for words presented twice in a spaced manner. Three recognition scans were conducted (Fig. 3). In the High–New (Standard 1)

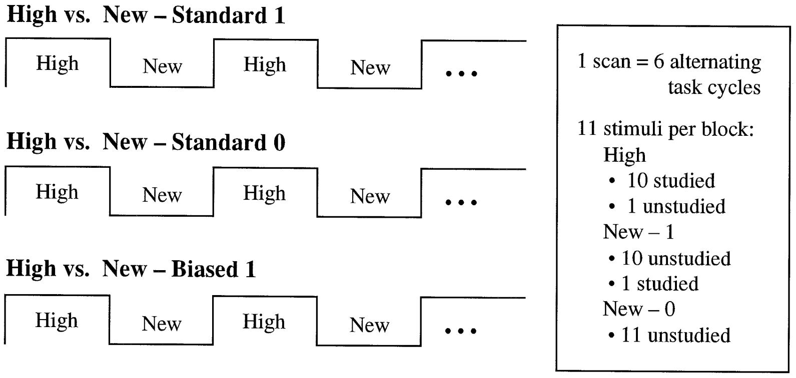

Fig. 3 Schematic diagram of the three recognition test scans in Experiment 2.

Table 5 *Experiment 2: recognition performance*

	Performance condition					
	High (S1)	New (S1)	High (B1)	New (B1)	High (S0)	New (S0)
Hits	0.90 (54.2)	0.89 (5.3)	0.86 (51.4)	0.81 (4.8)	0.85 (50.8)	–
False alarms	0.11 (0.7)	0.10 (6.2)	0.28 (1.7)	0.05 (3.2)	0.06 (0.3)	0.09 (5.8)
Accuracy	0.79	0.79	0.58	0.75	0.79	–

High (S1) and New (S1) are from the High–New (Standard 1) condition; High (B1) and New (B1) are from the High–New (Biasing 1) condition; High (S0) and New (S0) are from the High–New (Standard 0) condition. Numbers in parentheses correspond to the raw number of hits and false alarms for each condition.

and High–New (Standard 0) scans, the instructions for High and New blocks were identical to those in Experiment 1. In the High–New (Biasing 1) scan, participants were informed of the nature of the test blocks. Prior to High blocks, an instruction cue was presented (‘Task–New’) which informed the participant that the majority of the test words would be studied, but that there would be a minority of new, unstudied items. Participants were instructed to note the appearance of these new items by making a response. Prior to New blocks, an instruction cue was presented (‘Task–Old’) which informed the participant that the majority of the test words would be new, but that there would be a minority of old, studied items. Participants were instructed to note the appearance of these studied items by making a response. Due to concerns that the biasing instructions might have an impact on subsequent performance in the standard conditions, the High–New (Biasing 1) scan was always performed last.

Results

Behavioural results

The three scans yielded performance measures for six memory conditions: High (Standard 1), New (Standard 1), High (Biasing 1), New (Biasing 1), High (Standard 0) and New (Standard 0). The results revealed considerably more hits in the High than in the New conditions, and lower accuracy in the High (Biasing 1) condition (Table 5).

Numbers of hits and false alarms

A two × two × two repeated-measures ANOVA [instructions (Standard 1 and Biasing 1) × memory condition × item type] on the data from the High–New (Standard 1) and High–New (Biasing 1) conditions revealed that hits were greater than false alarms [$F(1,5) = 232.10$]. The memory-condition × item-type interaction was significant [$F(1,5) = 226.61$]; hits were greater in the High than in the New conditions [$F(1,5) = 367.75$], whereas false alarms were marginally greater in the New than in the High conditions [$F(1,5) = 4.46$, $P < 0.10$].

Accuracy

A two × two repeated-measures ANOVA (instructions × memory-condition) on the data from the High–New (Standard 1) and High–New (Biasing 1) conditions revealed that accuracy did not differ across instructions [$F(1,5) = 3.09$, $P > 0.10$] or across memory condition [$F(1,5) = 1.03$, $P > 0.35$]. However, the instructions × memory-condition interaction was reliable [$F(1,5) = 8.10$]; accuracy was similar in the High (Standard 1) and New (Standard 1) conditions ($F < 1.0$), but lower in the High (Biasing 1) than in the New (Biasing 1) condition [$F(1,5) = 15.22$]. An analysis of hits and false alarms revealed that this accuracy difference reflects a significantly greater false alarm rate in the High (Biasing 1) condition [$F(1,5) = 50.36$]. A separate analysis of recognition accuracy in the High (Standard 1) and High (Standard 0) conditions revealed no difference between these two conditions ($P > 0.50$).

The fMRI results

The pattern of right prefrontal activation varied across retrieval contexts; right prefrontal activation did not differ across High and New recognition under standard instructions, whereas under biasing instructions it was greater during the High condition. These results suggest that: (i) right prefrontal activation does not reflect retrieval success and that failures to reveal activation when comparing conditions of differing frequencies of success do not reflect a persisting-response mechanism; (ii) retrieval attempt is an incomplete account of retrieval-associated right prefrontal activation; and (iii) right prefrontal activation is sensitive to retrieval context, with recruitment of retrieval-attempt processes varying across contexts.

Right APC

Right APC activation did not differ across conditions when standard instructions were given [High–New (Standard 1) and High–New (Standard 0)], but was greater during the High condition when biasing instructions were given [High–New (Biasing 1); Figs 4 and 5]. Five participants demonstrated activation in the anterior extent of the right middle frontal gyrus, in the frontomarginal sulcus and in the lateral orbital gyrus (at or near BA 10 and 11), situated approximately 56–63 mm rostral to the AC (Table 6). The sixth participant demonstrated no difference in any prefrontal region in any scan comparison. To examine the pattern of right APC activation across the three scan comparisons, a repeated-measures ANOVA was conducted on the mean *Z*-scores for the right APC region of interest [scan comparisons: High–New (Standard 1), High–New (Standard 0) and High–New (Biasing 1)]. This region-of-interest analysis revealed a marginally significant effect of scan comparison [$F(2,5) = 2.89$, $P = 0.10$; Table 6]; there was a trend towards a greater right APC activation in the High–New (Biasing 1) contrast compared with the High–New (Standard 1) contrast [$F(1,5) = 3.53$, $P < 0.10$], and a greater activation in the High–New (Biasing 1) contrast compared with the High–New (Standard 0) contrast [$F(1,5) = 5.01$]. The High–New (Standard 1) and High–New (Standard 0) conditions did not differ ($F < 1.0$). When the region-of-interest analysis was re-computed excluding the participant who failed to demonstrate activation in any scan, the main effect was significant [$F(2,4) = 6.71$]: there was a greater right APC activation in the High–New (Biasing 1) contrast compared with the High–New (Standard 1) contrast [$F(1,4) = 11.44$], and compared with the High–New (Standard 0) contrast [$F(1,4) = 8.47$]. Again, the High–New (Standard 1) and High–New (Standard 0) conditions did not differ from each other ($F < 1.0$).

Right DLPC

In the standard-instruction scans, right DLPC activation did not differ across conditions (Fig. 4). In contrast, there was greater activation in a focal region of right DLPC during the High condition when recognition was performed under biasing instructions. This activation in the right middle frontal gyrus (at or near BA 9 and 46) was situated approximately 28–35 mm rostral to the AC (Table 6). A repeated measures ANOVA was conducted on the mean *Z*-scores for the right DLPC region of interest (Table 6). This region-of-interest analysis revealed a significant effect of scan comparison [$F(2,5) = 6.52$]; there was a greater activation in the High–New (Biasing 1) contrast compared with the High–New (Standard 1) contrast [$F(1,5) = 7.06$], and compared with the High–New (Standard 0) contrast [$F(1,5) = 11.87$]. The High–New (Standard 1) and High–New (Standard 0) conditions did not differ from each other ($F < 1.0$). Exclusion of the participant who failed to demonstrate activation in any scan did not change this pattern. No other prefrontal region demonstrated a significant difference in activation in any of the three scans.

General discussion

In these two fMRI studies we examined right prefrontal activation across recognition conditions that varied in retrieval success, retrieval effort and retrieval context. The results of both experiments are inconsistent with the retrieval-success and retrieval-effort hypotheses, at least as initially postulated, as neither success nor effort systematically affected the pattern of activation in the right APC and right DLPC. The results are generally consistent with the retrieval-attempt hypothesis; those from Experiment 2 indicate that retrieval attempt is not constant across all episodic retrieval situations. Rather, the probability of engaging attempt processes is context-dependent. These results contribute to a growing body of seemingly contradictory findings; between-study differences in memory performance and persisting success-induced haemodynamic responses have been posited to account for these inconsistencies (Rugg *et al.*, 1996). However, the present experiments suggest that neither is a likely candidate for resolving the variability of right prefrontal activation across studies. Consideration of the influence of retrieval context on the adopted retrieval strategy may serve to reconcile these inconsistencies.

Levels of recognition performance

The present studies are inconsistent with the retrieval-success hypothesis for the following reasons: (i) comparison of New blocks with baseline revealed right prefrontal activation even though there were only 3.8 successful retrievals in the New condition (Experiment 1); (ii) two different comparisons of New and High conditions failed to demonstrate differential right prefrontal activation even though there were over 50

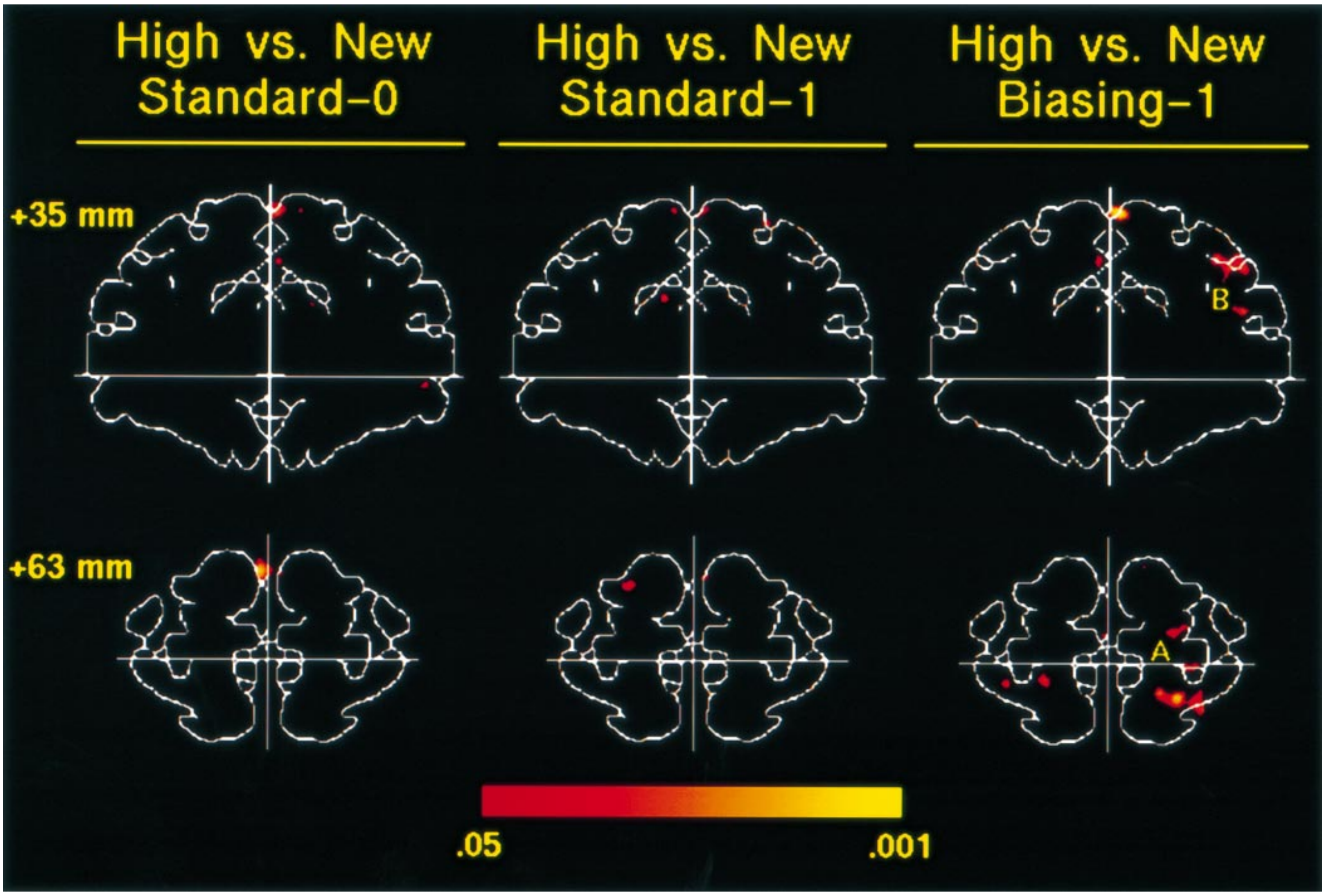

Fig. 4 Composite coronal images of activation in two sections corresponding to +35 and +63 mm rostral to the AC. Each row displays the composite from one section (indicated at left), and each column displays data from one condition (indicated at top). All composites reflect *Z*-scores averaged across all participants. Regions that show greater activation during the High condition relative to the New condition are overlaid on the corresponding coronal section. Greater right APC (A) and right DLPC (B) activation was present in the High condition only during the Biasing instructions scan.

successful retrievals in the High conditions and only 5.3 or zero successful retrievals in the New conditions (Experiment 2, standard instructions); (iii) High, Low and New conditions produced a similar pattern of right prefrontal activation when compared with baseline, even though retrieval success markedly differed (Experiment 1); and (iv) comparison of High and Low conditions failed to demonstrate differential right prefrontal activation even though there were 34 more successful retrievals in the High condition (Experiment 1). These results indicate that retrieval success is not a necessary condition for right prefrontal activation, and that the magnitude of activation in these regions is not directly tied to the frequency of retrieval success.

Previous reports of similar right prefrontal activation across differing levels of retrieval success may be partially attributable to overall poor levels of memory performance (Rugg *et al.*, 1996). However, the present experiments rule out this interpretation as similar right prefrontal activation was present across conditions where memory behaviour differed markedly. In Experiment 1, there was a substantial decline in the number of successful retrievals from the High (46.3) to Low (14.9) to New (3.8) conditions. In Experiment 2, there were significantly more successful retrievals in the High (Standard 1) and High (Standard 0) conditions (54.2 and 50.8, respectively) relative to the New (Standard 1) and New (Standard 0) conditions (5.3 and zero, respectively). In neither case did different levels of retrieval success result in reliable differences in right prefrontal activation.

The persisting-response interpretation and retrieval success

Rugg *et al.* (1996) also posit that previous demonstrations of similar activation in comparisons of high and low success conditions are not due to an insensitivity of the right prefrontal cortices to retrieval success, but to temporal blurring of success-induced haemodynamic responses. The present results are inconsistent with this interpretation: (i) it seems unlikely that the right prefrontal activation observed when comparing New with baseline conditions is the product of persisting responses to the 3.8 retrieved items in the New

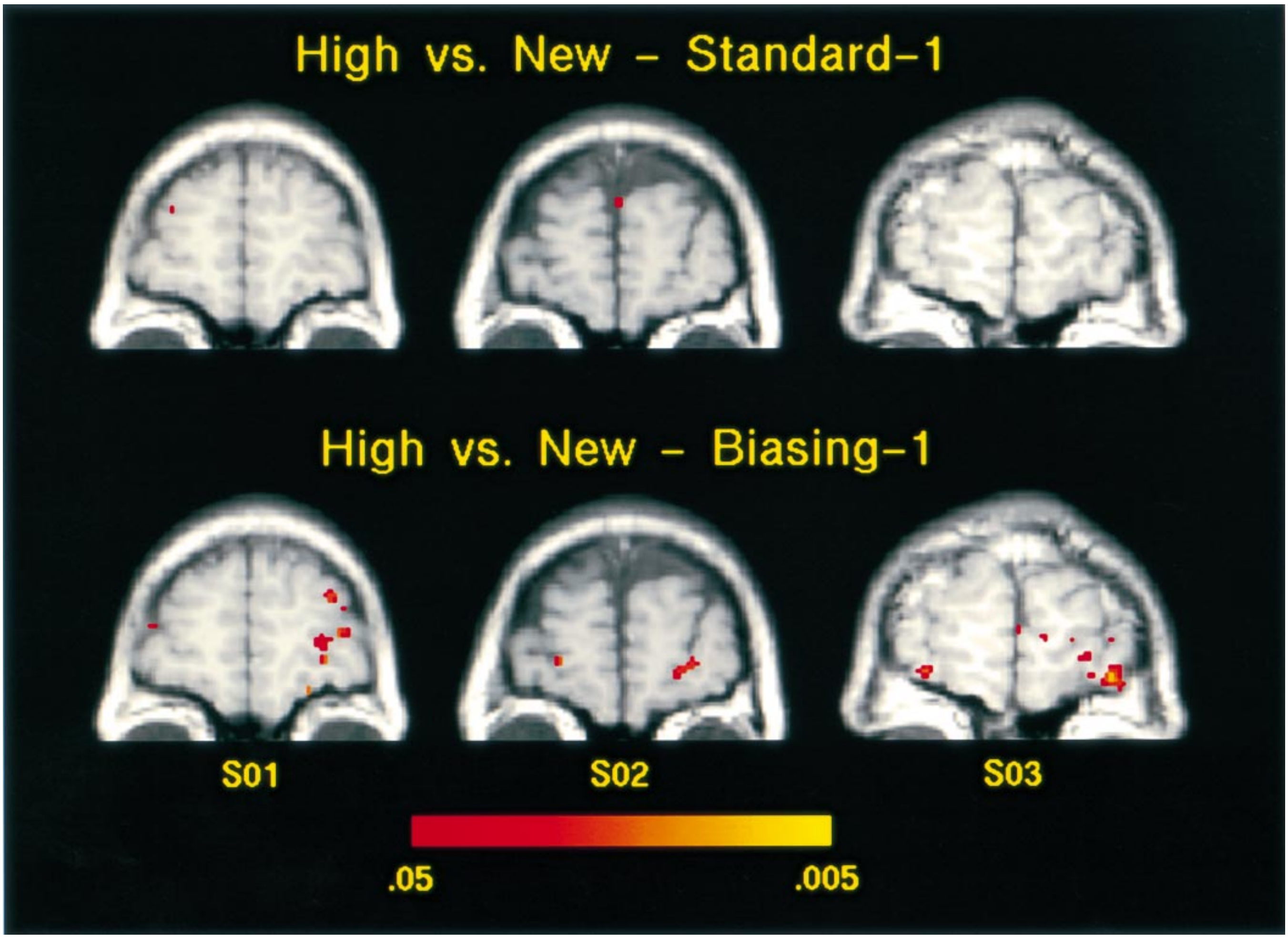

Fig. 5 Functional activation maps for one anterior section from each of three participants (S01, S02 and S03). Each column displays data from one section from one participant (indicated below each column), and each row displays data from one condition (indicated above each row). None of the subjects showed significant differences in right APC activation in the High versus Low condition comparison under Standard instructions (top row), but all three showed greater activation during the High condition under Biasing instructions (bottom row).

Table 6 *Experiment 2: regions of increased activation during High (relative to New) recognition*

Region	Talairach coordinates			Z_{max}	BA	Mean *Z*-score across region of interest
	x	*y*	*z*			
High > New (Biasing 1)						
Right APC	20	63	–9	3.28	10, 11	1.70
Right DLPC	37	35	35	2.77	9, 46	1.53
High > New (Standard 1)						
Right APC	No differential activation					0.94
Right DLPC	No differential activation					0.85
High > New (Standard 0)						
Right APC	No differential activation					0.80
Right DLPC	No differential activation					0.65

condition (Experiment 1); (ii) it seems unlikely that persistence of the responses to 5.3 successful retrievals in the New (Standard 1) condition could produce activation that did not differ from that produced by 54.2 successful retrievals in the High (Standard 1) condition (Experiment 2); and (iii) failure to visualize activation when comparing High (Standard 0) and New (Standard 0) conditions cannot be attributed to persisting activation as there were no successful retrievals in the New (Standard 0) condition (Experiment 2).

The persisting-response interpretation also is inconsistent with a linear haemodynamic response model (Boynton *et al.*, 1996; Dale and Buckner, 1997). The persisting-response interpretation posits that, although the neural response of right prefrontal cortices is linearly related to retrieval success, the haemodynamic response is not a linear function of the neural response. This suggestion is not supported by demonstrations that the haemodynamic response, at least as measured by fMRI blood oxygen level dependent signal, operates approximately as a shift-invariant linear system such that the observed signal elicited by a number of events is closely approximated by the sum of the signals elicited by each event (Boynton *et al.*, 1996; Dale and Buckner, 1997). Thus, temporal blurring of success-induced responses does not appear to be the source of results that are consistent with the retrieval-attempt hypothesis.

Additional evidence that right prefrontal activation is not dependent on retrieval success comes from an event-related fMRI study that examined right prefrontal haemodynamic responses associated with recognition judgements for new and old words presented in an intermixed manner at test (Buckner *et al.*, 1998*a*). In that study, new items correctly identified as new, and old items correctly identified as old, elicited similar haemodynamic responses in the right APC. This finding complements the present demonstrations that right prefrontal activation is greater when comparing New recognition with the baseline, and does not differ when comparing New and High recognition conditions under standard test instructions. Collectively, these data provide strong evidence indicating that retrieval success is not necessary for retrieval-associated right prefrontal activation.

Retrieval context affects right prefrontal activation

Consideration of the context in which retrieval is conducted may provide a framework for reconciling the apparently contradictory results across studies. Experiment 2 revealed two different patterns of right prefrontal activation, one consistent with retrieval success and the other consistent with retrieval attempt, in a single group of participants under conditions where the encoding and the test stimuli were held constant and only the retrieval context varied. Differences in the pattern of activation across studies may reflect across-study differences in retrieval context.

The hypothesis that the processes engaged during retrieval vary across retrieval contexts builds upon the attempt hypothesis. The original formulation of the retrieval-attempt hypothesis implies that these processes will be similarly engaged across all conditions requiring episodic retrieval (Kapur *et al.*, 1995; Nyberg *et al.*, 1995). This position is most strongly asserted during characterization of attempt as retrieval mode, which is thought to be a state change associated with thinking back in time; retrieval mode ‘sets the stage’ for the recovery of specific episodes from long-term memory (e.g. Wheeler *et al.*, 1997). The present results are inconsistent with this formulation of retrieval attempt as a context-invariant state change. Rather, these results extend the attempt interpretation by demonstrating that engagement of attempt processes varies across retrieval contexts.

Why did the standard and biasing retrieval contexts produce different patterns of activation? One possibility is that the biasing instructions induced different retrieval strategies in the High and New conditions, whereas the standard instructions induced the same strategy in the two conditions. By informing participants in the New (Biasing 1) condition that most of the test items would be new and that their task is to identify the few old items, participants may have adopted a retrieval strategy of only assessing the relative familiarity of each test probe (e.g. Jacoby, 1991). The old items were probably considerably more familiar than the new items due to their having been conceptually processed during learning (Wagner *et al.*, 1997*c*) and they may therefore have been more easily discriminable through familiarity alone. Such a strategy would render an organized retrieval search, or careful evaluation of the products of retrieval, unnecessary. In contrast, under conditions where participants are aware that most stimuli will be old [High (Biasing 1)], or where they have no expectations about the nature of the test lists (the standard instructions conditions), participants may engage in search and evaluative retrieval processes during each recognition attempt. This interpretation assumes that it is easier to discriminate a familiar item from a group of unfamiliar items, than it is to discriminate an unfamiliar item from a group of familiar items. The behavioural data provide some support for this interpretation, as participants were significantly less accurate in the High (Biasing 1) condition.

Alternatively, the biasing instructions may have increased the probability of retrieval attempts in the High (Biasing 1) condition. Participants were significantly less accurate in this condition, frequently failing to identify the new items. The biasing instructions, however, informed participants that they should be able to identify a few of the test probes as new. Such performance expectations may have encouraged participants to persist in their retrieval attempts even when those attempts met with failure. In contrast, participants probably did not have expectations about the nature of the test lists under the standard instructions. For each test probe they may have initiated a few attempts to retrieve. If these attempts resulted in failure, the attempts are aborted. This interpretation is related to the retrieval-effort hypothesis, as it suggests that in some situations additional retrieval attempts may be necessary in order to perform the task (Schacter *et al.*, 1996*a*). However, this interpretation differs from the retrieval-effort hypothesis as it suggests that differences in memory performance need not produce differences in right prefrontal activation. Differences in activation may arise only to the extent that the retrieval context results in a bias to continue to engage in retrieval attempts. In this way, retrieval context plays an influential role by guiding participants towards a particular retrieval strategy.

Experiment 2 reconciles the conflicting results between studies supporting the attempt hypothesis and some of those supporting the success hypothesis (Tulving *et al.*, 1994, 1996). However, the results from this experiment cannot directly account for demonstrations that retrieval success affects right prefrontal activation (Rugg *et al.*, 1996). One possibility is that the results of Rugg *et al.* (1996) reveal yet another factor that determines the retrieval context: expectations that arise from participants' on-line monitoring of their retrieval performance. In the study of Rugg *et al.* (1996), new and high recognition conditions were used during entirely separate scanning epochs. It is possible that in the new epochs, participants noticed their low level of retrieval success during the run of 20 consecutive unstudied trials and shifted from a strategy consisting of retrieval search, or post-retrieval evaluation, to a more minimalist strategy based on familiarity. Such a strategy shift may be less likely, however, when new blocks regularly alternate with high blocks, as in the present study, or when new and old items are intermixed, as in the event-related fMRI study of Buckner *et al.* (1998*a*). From this perspective, although right prefrontal activation is not dependent on retrieval success, retrieval success may play a role in establishing the retrieval context in which participants attempt to retrieve.

The notion of context-dependent retrieval processes may also assist in integrating the current results with those of Schacter *et al.* (1996*a*). Although there are numerous methodological differences between the present study and that of Schacter *et al.* (1996*a*), perhaps the most significant difference is the nature of the retrieval task (recognition or cued recall). In the context of cued recall, participants may have the expectation that they should be able to retrieve a previous experience when provided with the test probe, as the instructions inform them that each probe can be used to recall something from the past. Such expectations may encourage participants to persist with their retrieval attempts when faced with failure. In contrast, in recognition participants probably do not have the expectation that they should recognize every test probe. Thus, they may be more likely to abort retrieval attempts earlier in recognition than in cued recall. It may be that the pattern of right prefrontal activation across different levels of retrieval success partially depends upon the retrieval task. Tasks that encourage repeated retrieval attempts under conditions of poor performance may yield a different pattern of activation from tasks that allow for termination of such attempts after a few failures.

Further evidence that retrieval attempt is context dependent comes from neuroimaging studies of veridical and illusory recognition (Schacter *et al.*, 1996*c*; Johnson *et al.*, 1997; Schacter *et al.*, 1997). In those studies, participants made recognition decisions for studied words (veridical recognition) and for novel words that were conceptually related to the studied words (illusory recognition). The recognition test was either blocked by illusory and veridical conditions (Schacter *et al.*, 1996c, 1997) or had intermixed illusory and veridical trials (Johnson *et al.*, 1997; Schacter *et al.*, 1997). Behaviourally, participants appear to struggle more with illusory than with veridical recognition judgements within a blocked context (response latencies are longer during illusory recognition), but not within an intermixed context (Johnson *et al.*, 1997). Schacter *et al.* (1997) suggest that participants 'might engage in more careful scrutiny of specific attributes of the test item' in the illusory versus the veridical condition when in the blocked retrieval context but not when in the intermixed context. Consistent with the context-dependent retrieval-process hypothesis, there was a trend towards greater right prefrontal activation when comparing illusory with veridical recognition in the blocked context (Schacter *et al.*, 1996*c*; see also Schacter *et al.*, 1997), but not in the intermixed context (Johnson *et al.*, 1997; Schacter *et al.*, 1997).

Electrophysiological evidence

The present blocked-design fMRI experiments, together with the event-related fMRI study of Buckner *et al.* (1998*a*) and previous PET studies (Kapur *et al.*, 1995; Nyberg *et al.*, 1995), suggest that retrieval-associated right prefrontal activation does not depend on retrieval success. However, a number of investigations measuring event-related potentials (ERPs) have demonstrated differential activity at electrode sites overlying prefrontal cortex across conditions varying in retrieval success (for a review, see Allan *et al.*, 1998). For example, a positive-going modulation situated over right frontal electrode sites was greater when recognition judgements were accompanied by a correct source judgement than when accompanied by an incorrect one (Wilding and Rugg, 1996). Similarly, a positive-going modulation situated over bilateral frontal sites was greater when recognition was associated with conscious recollection (as indexed by subjective judgements) than when not associated with conscious recollection (Duzel *et al.*, 1997). These results suggest that right prefrontal activation is greater under conditions where recognition is accompanied by memory for specific aspects of a previous event. It is difficult to reconcile these ERP results with fMRI and PET demonstrations that retrieval success is not necessary for, and does not directly modulate, right prefrontal activation. One possibility is that the neural source of the ERP effects is different from the right APC and right DLPC regions frequently visualized during fMRI and PET studies and is undetected by these methods. Alternatively, these ERP effects appear to be demonstrated under recognition conditions that require an additional judgement above and beyond recognition, such as judgements of source or conscious recollection. Differences across fMRI, PET and ERP studies may partially reflect these differences in retrieval-task demands. Future studies that use either PET or fMRI in conjunction with ERP may serve to clarify the source of these inconsistencies.

When are right prefrontal cortices relevant to recognition performance?

Repeated demonstrations of prefrontal activation in PET and fMRI studies of recognition contrast with reports of modest or no deficits in standard yes–no recognition tests following prefrontal lesion (Shimamura, 1994; Wheeler *et al.*, 1995). Such dissociations suggest that regions visualized during memory performance may be correlated with, but are not necessary for, the memory behaviour being measured (Blaxton *et al.*, 1996). The present finding that the magnitude of right prefrontal activation does not systematically vary with recognition performance is consistent with the possibility that these regions are not critical for recognition performance under standard testing conditions. This pattern can be contrasted with the greater medial temporal activation associated with greater retrieval success in recognition (Nyberg *et al.*, 1996*b*; Rugg *et al.*, 1997) and cued recall (Schacter *et al.*, 1996*a*). This medial temporal activation possibly reflects conscious recollection of the past that can support a recognition judgement.

Although prefrontal lesions typically result in minimal deficits on standard recognition tasks, marked impairments can occur under select recognition conditions where accurate recognition performance requires extensive retrieval search or post-retrieval evaluation. For example, Schacter *et al.* (1996*b*) report a patient (B.G.) who demonstrates impaired recognition memory due to a right frontal lesion. B.G.'s recognition deficits manifest as increased illusory recognition. Under conditions where foils are similar to targets, participants may need to initiate multiple retrieval searches or extensively evaluate the products of memory in order to recollect specific episodic information that would allow discrimination between foils and targets. Prefrontal cortices may be critical for recognition performance in such a condition. Along these same lines, studies of the effects of Parkinson's disease on recognition performance revealed that although recognition is usually unaffected when performed shortly after study (Brown and Marsden, 1990), recognition is impaired when performed after long delays (Stebbins *et al.*, 1998). Under long retention intervals, participants may not be able to discriminate targets from foils based on differential familiarity. Rather, participants may need to recollect specific aspects of the study event. These requirements would place greater demands on the prefrontal regions thought to be impaired in Parkinson's disease.

Conclusions

The present fMRI studies indicate that right APC and right DLPC regions mediate processes that are associated with attempting to recollect the past, and that the probability of engaging these processes varies across retrieval contexts. Multiple factors appear to affect the retrieval context, including test instructions, on-line monitoring of retrieval success and task designs. Thus, variable right prefrontal activation may be expected across studies. The present results complement neuropsychological demonstrations that prefrontal cortices are not necessary for standard recognition performance, as right prefrontal activation did not systematically vary with performance. However, these prefrontal regions may be critical for recognition performance under contexts where discrimination between novel and studied stimuli requires retrieval of specific aspects of the past, such as when foils are closely associated with studied items or when the study–test delay is lengthened.

Acknowledgements

Supported by grants from the National Institutes of Health (NIA AG12995 and AG11121). We thank G. H. Bower, R. L. Buckner, L. Nyberg, R. A. Poldrack, M. D. Rugg, D. E. Rumelhart, D. L. Schacter, E. Tulving and B. Tversky for insightful comments on an earlier version of this manuscript.

References

Allan K, Wilding EL, Rugg MD. Electrophysiological evidence for dissociable processes contributing to recollection. [Review]. Acta Psychol 1998; 98: 231–52.

Awh E, Jonides J, Smith EE, Schumacher EH, Koeppe RA, Katz S. Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. Psychol Sci 1996; 7: 25–31.

Blaxton TA, Zeffiro TA, Gabrieli JD, Bookheimer SY, Carrillo MC, Theodore WH, et al. Functional mapping of human learning: a positron emission tomography activation study of eyeblink conditioning. J Neurosci 1996; 16: 4032–40.

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 1996; 16: 4207–21.

Brown RG, Marsden CD. Cognitive function in Parkinson's disease: from description to theory. [Review]. Trends Neurosci 1990; 13: 21–9.

Buckner RL. Beyond HERA: contributions of specific prefrontal brain areas to long-term memory retrieval. [Review]. Psychonom Bull Rev 1996; 3: 149–58.

Buckner RL, Petersen SE, Ojemann JG, Miezin FM, Squire LR, Raichle ME. Functional anatomical studies of explicit and implicit memory retrieval tasks. J Neurosci 1995; 15: 12–29.

Buckner RL, Raichle ME, Miezin FM, Petersen SE. Functional anatomic studies of memory retrieval for auditory words and visual pictures. J Neurosci 1996; 16: 6219–35.

Buckner RL, Koutstaal W, Schacter DL, Dale AM, Rotte M, Rosen BR. Functional-anatomic study of episodic retrieval: II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. Neuroimage 1998a; 7: 163–75.

Buckner RL, Koutstaal W, Schacter DL, Wagner AD, Rosen BR. Functional-anatomic study of episodic retrieval using fMRI: I. Retrieval effort versus retrieval success. Neuroimage 1998b; 7: 151–62.

Cohen NJ, Squire LR. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociations of knowing how and knowing what. Science 1980; 210: 207–10.

Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. J Verbal Learn Verbal Behav 1972; 11: 671–84.

Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 1997; 5: 329–40.

Desmond JE, Sum JM, Wagner AD, Demb JB, Shear PK, Glover GH, et al. Functional MRI measurement of language lateralization in Wada-tested patients. Brain 1995; 118: 1411–9.

Desmond JE, Gabrieli JDE, Sobel N, Rabin LA, Wagner AD, Seger CA, et al. An fMRI study of frontal cortex and cerebellum during semantic and working memory tasks [abstract]. Soc Neurosci Abstr 1996; 22: 1111.

Duzel E, Yonelinas AP, Mangun GR, Heinze H-J, Tulving E. Event-related brain potential correlates of two states of conscious awareness in memory. Proc Natl Acad Sci USA 1997; 94: 5973–8.

Fiez JA, Raichle ME, Miezin FM, Petersen SE, Tallal P, Katz WF. PET studies of auditory and phonological processing: effects of stimulus characteristics and task demands. J Cognit Neurosci 1995; 7: 357–75.

Fiez JA, Raife EA, Balota DA, Schwarz JP, Raichle ME, Petersen SE. A positron emission tomography study of the short-term maintenance of verbal information. J Neurosci 1996; 16: 808–22.

Friston KJ, Jezzard P, Turner R. Analysis of functional MRI time-series. Hum Brain Mapp 1994; 1: 153–71.

Gabrieli JDE. Cognitive neuroscience of human memory. [Review]. Annu Rev Psychol 1998; 49: 87–115.

Gabrieli JDE, Prabhakaran V, Rypma B, Wagner AD, Kang E, Desmond JE. Attention to thought: functional magnetic resonance imaging (fMRI) evidence [abstract]. Abstr Psychonom Soc 1997; 2: 28.

Glover GH, Lai S. Self-navigated spiral fMRI: interleaved versus single-shot. Magn Reson Med. In press 1998.

Glover GH, Lee AT. Motion artifacts in fMRI: comparison of 2DFT with PR and spiral scan methods. Magn Reson Med 1995; 33: 624–35.

Graf P, Schacter DL. Implicit and explicit memory for new associations in normal and amnesic subjects. J Exp Psychol Learn Mem Cogn 1985; 11: 501–18.

Jacoby LL. A process dissociation framework: separating automatic from intentional uses of memory. J Mem Lang 1991; 30: 513–41.

Johnson MK, Nolde SF, Mather M, Kounios J, Schacter DL, Curran T. The similarity of brain activity associated with true and false recognition memory depends on test format. Psychol Sci 1997; 8: 250–7.

Kapur S, Craik FI, Jones C, Brown GM, Houle S, Tulving E. Functional role of the prefrontal cortex in retrieval of memories: a PET study. Neuroreport 1995; 6: 1880–4.

Milner B, Corsi P, Leonard G. Frontal-lobe contribution to recency judgements. Neuropsychologia 1991; 29: 601–18.

Noll DC, Cohen JD, Meyer CH, Schneider W. Spiral K-space MR imaging of cortical activation. J Magn Reson Imaging 1995; 5: 49–56.

Nyberg L, Tulving E, Habib R, Nilsson L-G, Kapur S, Houle S, et al. Functional brain maps of retrieval mode and recovery of episodic information [see comments]. Neuroreport 1995; 7: 249–52. Comment in: NeuroReport 1995; 7: 9–10.

Nyberg L, Cabeza R, Tulving E. PET studies of encoding and retrieval: the HERA model. [Review]. Psychonom Bull Rev 1996a; 3: 135–48.

Nyberg L, McIntosh AR, Houle S, Nilsson L-G, Tulving E. Activation of medial temporal structures during episodic memory retrieval [see comments]. Nature 1996b; 380: 715–7. Comment in: Nature 1996; 380: 669–70.

Petrides M, Alivisatos B, Meyer E, Evans AC. Functional activation of the human frontal cortex during the performance of verbal working memory tasks. Proc Natl Acad Sci USA 1993; 90: 878–82.

Picard N, Strick PL. Motor areas of the medial wall: a review of their location and functional activation. [Review]. Cereb Cortex 1996; 6: 342–53.

Raaijmakers JGW, Shiffrin RM. Search of associative memory. Psychol Rev 1981; 88: 93–134.

Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, et al. Practice-related changes in human brain functional anatomy during nonmotor learning. Cereb Cortex 1994; 4: 8–26.

Roediger HL 3d. Implicit memory: retention without remembering. [Review]. Am Psychol 1990; 45: 1043–56.

Rugg MD, Fletcher PC, Frith CD, Frackowiak RS, Dolan RJ. Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. Brain 1996; 119: 2073–83.

Rugg MD, Fletcher PC, Frith CD, Frackowiak RS, Dolan RJ. Brain regions supporting intentional and incidental memory: a PET study. Neuroreport 1997; 8: 1283–7.

Schacter DL, Alpert NM, Savage CR, Rauch SL, Albert MS. Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. Proc Natl Acad Sci USA 1996a; 93: 321–5.

Schacter DL, Curran T, Galluccio L, Milberg WP, Bates JF. False recognition and the right frontal lobe: a case study. Neuropsychologia 1996b; 34: 793–808.

Schacter DL, Reiman E, Curran T, Yun LS, Bandy D, McDermott KB, et al. Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography [see comments]. Neuron 1996c; 17: 267–74. Comment in: Neuron 1996; 17: 191–4.

Schacter DL, Buckner RL, Koutstaal W, Dale AM, Rosen BR. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. Neuroimage 1997; 6: 259–69.

Shallice T. From neuropsychology to mental structure. Cambridge: Cambridge University Press; 1988.

Shimamura AP. Memory and frontal lobe function. In: Gazzaniga MS, editor. The cognitive neurosciences. Cambridge (MA): MIT Press; 1994. p. 803–13.

Squire LR. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans [published erratum appears in Psychol Rev 1992; 99: 582]. [Review]. Psychol Rev 1992; 99: 195–231.

Squire LR, Ojemann JG, Miezin FM, Petersen SE, Videen TO, Raichle ME. Activation of the hippocampus in normal humans: a functional anatomical study of memory. Proc Natl Acad Sci USA 1992; 89: 1837–41.

Stebbins GT, Gabrieli JDE, Masciari F, Monti L, Goetz CG. Delayed recognition memory in Parkinson's disease: a role for working memory? Neuropsychologica. In press, 1998.

Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme; 1988.

Tulving E. Episodic and semantic memory. In: Tulving E, Donaldson W, editors. Organization of memory. New York: Academic Press; 1972. p. 381–403.

Tulving E. Elements of episodic memory. Oxford: Clarendon Press, 1983.

Tulving E, Schacter DL. Priming and human memory systems. Science 1990; 247: 301–6.

Tulving E, Kapur S, Markowitsch HJ, Craik FI, Habib R, Houle S. Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition [see comments]. Proc Natl Acad Sci USA 1994; 91: 2012–5. Comment in: Proc Natl Acad Sci USA 1994; 91: 1989–91.

Tulving E, Markowitsch HJ, Craik FI, Habib R, Houle S. Novelty and familiarity activations in PET studies of memory encoding and retrieval. Cereb Cortex 1996; 6: 71–9.

Wagner AD, Buckner RL, Koutstaal WK, Schacter DL, Gabrieli JDE, Rosen BR. An fMRI study of within- and across-task item repetition during semantic classification [abstract]. Abstr Cognit Neurosci Soc 1997a; 4: 68.

Wagner AD, Desmond JE, Demb JB, Glover GH, Gabrieli JDE. Semantic repetition priming for verbal and pictorial knowledge: a functional MRI study of left inferior prefrontal cortex. J Cognit Neurosci 1997b; 9: 714–26.

Wagner AD, Gabrieli JD, Verfaellie M. Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. J Exp Psychol Learn Mem Cogn 1997c; 23: 305–23.

Wagner AD, Poldrack RA, Desmond JE, Glover GH, Gabrieli JDE. An fMRI study of prefrontal activation during retrieval of verbal and non-verbal episodic memory [abstract]. Soc Neurosci Abstr 1997d; 23: 492.

Wheeler MA, Stuss DT, Tulving E. Frontal lobe damage produces episodic memory impairment. J Int Neuropsychol Soc 1995; 1: 525–36.

Wheeler MA, Stuss DT, Tulving E. Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. [Review]. Psychol Bull 1997; 121: 331–54.

Wilding EL, Rugg MD. An event-related potential study of recognition memory with and without retrieval of source [published erratum appears in Brain 1996; 119: 1416]. Brain 1996; 119: 889–905.

Woods RP, Cherry SR, Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET images. J Comput Assist Tomogr 1992; 16: 620–33.

Received February 10, 1998. Revised April 25, 1998.
Accepted May 21, 1998