Frontal Brain Activity Predicts Individual Performance in an Associative Memory Exclusion Test

Cyma Van Petten, Barbara J. Luka, Susan R. Rubin and John P. Ryan

Department of Psychology, University of Arizona, Tucson, AZ 85721, USA

Event-related brain potentials (ERPs) were recorded from 24 young adults during a recognition test including Old, New, and Recombined pairs composed of two words studied in different pairs. Recombined pairs called for a response of 'new'. Task difficulty was increased by repetition of some words during the study phase; a subject might study tower/pie, puppet/pie, drill/wreath and bee/wreath (pairs with a Common word), and at test, encounter the Common Recombined pair of puppet/wreath (in addition to Unique Recombined pairs composed of two words studied once). Individual accuracy in the Recombined conditions varied widely, but was unrelated to general memory ability as indexed by accuracy on the Old and New pairs. Posterior brain potentials showed graded amplitudes dependent on the oldness of both the individual words and their combinations (Old > Recombined > New, and Common > Unique), but were also unrelated to accuracy in the Recombined conditions. Amplitudes of ERPs recorded over prefrontal scalp accounted for a large proportion of the individual variability in differentiating studied combinations of words from recombinations of studied elements. The experimental design differentiates three possible roles of prefrontal cortex in source or associative memory tests: resolving a conflict between familiarity and a response of 'new', extended memory search, and evaluation of ambiguous memory signals.

Introduction

A close match between elements of a stimulus and a memory trace typically provides good evidence that the stimulus was indeed previously experienced. The exception to this rule occurs when new stimuli contain one or more familiar elements. When walking on a city street, one may be briefly misled into thinking a friend is ahead based on hair and clothing, only to see a complete stranger on closer view. Particularly compelling cases of false recognition may arise when a new stimulus consists only of familiar elements, but not in the combination previously experienced, as reported to occur in cases of false eyewitness testimony (Wells and Loftus, 1984; Lindsay and Johnson, 1989). In the laboratory, memory for combinations of attributes is the focus of *source memory* tests in which subjects are asked about the relationship between a stimulus and the context in which it was initially studied. A source recognition test may ask, for instance, whether a test sentence is spoken by the same voice in which it was initially heard or by a different voice [that also occurred during the study phase, but speaking other sentences (Glisky *et al.*, 1995)]. Source recognition tasks are typically contrasted with *item recognition* tests that simply ask whether a stimulus was studied or not, independent of any additional information.

Associative memory tests are conceptually similar to source memory tests, in that subjects are tested for their memory of the relationship between two elements of a studied stimulus, e.g. two words of a pair. The distinction between source memory and associative memory does not, to our knowledge, have a conventional definition. However, two major differences are apparent from even a cursory literature review. First, source memory tests always call for judgements about two qualitatively distinct attributes of a single stimulus, e.g. word identity (conveyed by the sequence of phonemes comprising the word) and the talker's voice (conveyed by nonphonemic frequency information in the acoustic signal), whereas the stimuli in associative memory paradigms are frequently composed of two elements drawn from the same domain, e.g. two words or two line drawings. The second general distinction between 'source' and 'associative' memory paradigms is the combinatorial mapping between the two elements of a stimulus. Source memory paradigms typically employ a many-to-one mapping between 'sources' and 'items'. For instance, if two voices are paired with a large number of different words, there may be interference when trying to remember a particular source/item combination. In contrast, associative memory paradigms typically do not reuse elements across stimuli, so that the mapping between elements is one-to-one. The present experiment uses a hybrid paradigm with aspects of both source and associative memory tasks. Like an associative memory task, the two elements to be remembered are drawn from the same domain (two words). To the extent that the results are similar to those observed in source memory paradigms, it will suggest that remembering cross-domain and within-domain conjunctions tap the same neural processes. But we also manipulate the mapping between elements; contrasts between the many-to-one and unique mapping conditions will indicate whether the re-pairing of especially familiar elements is a critical feature for the engagement of prefrontal cortex that has been the hallmark of source memory tests.

Several lines of evidence indicate that source memory performance is dependent on prefrontal cortex. Patients with damage to the frontal lobes are disproportionately impaired in source memory tasks relative to item memory tasks (Janowsky *et al.*, 1989; Johnson *et al.*, 1997b), although not amnesic in any traditional sense (Wheeler *et al.*, 1995; Swick and Knight, 1996). In healthy older adults and in children, source memory accuracy is correlated with neuropsychological measures thought to tap the integrity of prefrontal regions, but item recognition accuracy is not (Glisky *et al.*, 1995, 2001; Henkel *et al.*, 1998; Cycowicz *et al.*, 2001). Event-related brain potentials (ERPs) in healthy young adults show that studied items elicit a large positive potential at prefrontal scalp sites when subjects make source judgements, but not when they make item recognition judgements (Senkfor and Van Petten, 1998; Van Petten *et al.*, 2000; Senkfor, 2002). Hemodynamic imaging studies show greater blood flow in prefrontal regions during source than item memory tasks, although these studies have averaged across studied and unstudied items (Cabeza *et al.*, 1997; Rugg *et al.*, 1999; Raye *et al.*, 2000).

Although paired-associate learning is a traditional memory task, there have been very few contrasts between associative and single-item memory tests which can reveal whether associative

Cerebral Cortex Nov 2002;12:1180–1192; 1047–3211/02/$4.00

retrieval is like source retrieval showing a greater reliance on prefrontal cortex. Passingham has reviewed strong evidence that inferior prefrontal cortex is essential for associative learning, but stresses associations between stimuli and responses, rather than associations between stimulus elements (Passingham *et al.*, 2000). More relevant is Dimitrov's (1999) report of a deficit in cued-recall of word pairs in frontal patients, but there was no baseline of single-word recall (Dimitrov *et al.*, 1999). Most relevant is Donaldson and Rugg's examination of ERPs elicited when subjects judged word pairs as old, new or recombinations of studied words (exp. 1), or when they judged only the words comprising the pairs as old or new (exp. 2) (Donaldson and Rugg, 1998). In both experiments, old pairs elicited more positive prefrontal potentials than new or recombined pairs (which were very similar), but there was apparently little impact of the cross-experiment contrast between associative and non-associative response requirements. This last result may suggest that associative retrieval is very different from source retrieval, but the present experiment revisits this question.

Despite the general agreement that prefrontal cortex is both engaged by, and necessary for the accurate performance of source memory judgements, it is not yet clear why source recognition places greater demands on prefrontal cortex than old/new recognition judgements.

Many theorists have suggested that prefrontal cortex serves an important role in *memory control* processes, as distinct from retrieval *per se* (Luria, 1973; Moscovitch, 1994; Burgess and Shallice, 1996; Shimamura, 2000). In a long-term memory task, control processes might include setting response criteria contingent on current goals or instructions, evaluating the strength of retrieved information in light of the criteria, evaluating the logical consistency between bits of retrieved information, directing a more extended search of memory if necessary, revising response criteria as necessary, making a decision, and, finally, mapping the decision onto an overt response. This diverse set of cognitive processes is sometimes labeled 'memory monitoring', a label so broad that it has sometimes led to disagreement among researchers as to whether a given experimental manipulation should increase or decrease monitoring demands (Fletcher and Henson, 2001). Nonetheless, some empirical results have suggested a role for prefrontal cortex in specific control processes. High false alarm rates during old/new recognition tests (observed in patients with frontal damage and healthy older adults who perform poorly on standardized tests of 'frontal function') can be attributed to poor criterion setting (Schacter *et al.*, 1996; Rapscak *et al.*, 1998; Rubin *et al.*, 1999; Swick and Knight, 1999; Davidson and Glisky, 2002). Miller and colleagues observed increased blood flow in dorsolateral prefrontal cortex in a recognition test which required frequent shifts between liberal and strict criteria as compared to maintaining a constant criterion (Miller *et al.*, 2001). Henson and colleagues observed greater blood flow in a small bilateral region of dorsolateral prefrontal cortex accompanied by low- as compared to high-confidence recognition decisions, which they attributed to enhanced monitoring, 'engaged whenever the results of an initial retrieval attempt are ambiguous (i.e., close to the response critierion)' (Henson *et al.*, 2000, p. 918). However, as Henson *et al.* (2000) also note, it can be difficult to differentiate specific control processes from one another. In the case of initially ambiguous retrieval results, 'one might just use the information already obtained in an attempt to judge more precisely its relation to the response criterion' or '. . . one might make further retrieval attempts . . .' (Henson *et al.*, 2000, p. 918).

The supposition that the engagement of prefrontal cortex during source memory tests is due to a control process, as opposed to something intrinsic to source memory *per se*, is bolstered by a recent result of Hermann *et al.*, who observed increased blood flow in medial, inferior, and dorsolateral prefrontal cortex during a source memory task of discriminating words from two studied lists, but only when the old target and old distractor words were semantically related. These investigators thus concluded that 'monitoring and control of interference are specific components of the PFC contribution to human memory processing' (Hermann *et al.*, 2001, p. 99).

The present study attempts to differentiate the contributions of control processes that might be particularly critical during source or associative recognition tests. We consider three plausible alternatives. The first is instigating an extended search of memory to recover more information than was initially found. We have previously favored this interpretation of the prefrontal ERP effect observed during source memory tests, based on its temporal relationship to ERPs recorded over more posterior scalp (Senkfor and Van Petten, 1998; Van Petten *et al.*, 2000; Senkfor, 2002; Senkfor *et al.*, 2002). In these experiments, the earliest memory effect (beginning 200–400 ms poststimulus onset) is a spatially widespread positivity that is larger for recognized old items (hits) than for correct rejections, false alarms or misses (Van Petten and Senkfor, 1996; Rubin *et al.*, 1999). This early *old/new* effect varies little between item and source recognition tests, so that it can be regarded as an index of successful item retrieval. The early old/new effect is reduced in patients with medial temporal or diencephalic damage; this reduction is proportional to the density of their amnestic syndrome (Olichney *et al*, 2000). A prefrontal positivity elicited by studied items is of later onset and longer duration, and much larger in source than item memory tasks (Wilding *et al.*, 1995; Johnson *et al.*, 1997a; Trott *et al.*, 1997; Ranganath and Paller, 1999; Johansson *et al.*, 2002). In our studies, the prefrontal effect has been insensitive to whether or not the source information accompanying each item is successfully recovered. Instead, a late-onset positivity (~800 ms) over posterior scalp is elicited by items for which source retrieval is successful. These results thus suggest that successful retrieval *per se* is indexed by ERPs recorded over posterior brain regions, but some prefrontal control process is taxed by source memory tasks. Because the prefrontal old/new effect occurred after successful item retrieval, but slightly before the differentiation between successful and unsuccessful source retrieval, we have proposed that it indexes an *extended memory search* for source information in memory.

The second alternative is one suggested by Henson and co-workers: careful *evaluation* of the products of memory retrieval in order to make a decision (Henson *et al.*, 2000). Studied items in source recognition paradigms are likely to require close scrutiny, because they are composed of two studied elements, but recognition of the two elements in isolation is not sufficient to make a judgement about whether the combination of elements was studied or unstudied.

The third alternative is that source memory tests put high demands on a process of conflict resolution beween the initial results of memory retrieval and the need for categorical decision. When new combinations of studied attributes are presented at test, the combined memory signal may suggest that the entire package is old, but this suggestion must be resisted to obtain greater than chance accuracy in the source decision. A bias to respond 'old same' when both attributes have been studied — but not in their tested combination — is suggested by the consistent

observation of higher source accuracy for old–same than old–different trials across studies (Dodson and Shimamura, 2000). Numerous studies implicate prefrontal cortex in the selection of an appropriate response when an alternative response is particularly compelling or overlearned (MacDonald *et al.*, 2000). As applied to source memory tests, the *familiarity/response conflict hypothesis* is thus that prefrontal cortex is engaged when multiple attributes of a stimuli are both familiar and relevant to the memory decision, but the mapping between familiarity and behavioral response is not straightforward. The response conflict hypothesis is specific to source recognition tests, and is less relevant to source memory tests using a recall format in which studied source information does not recur at test. A handful of studies have suggested that source recall tests continue to rely on prefrontal regions to a greater extent than do item memory tests (Johnson *et al.*, 1997a,b; Rugg *et al.*, 1999) (also A.J. Senkfor, C. Van Petten and M. Kutas, submitted), so that a conflict between familiarity and response selection cannot be the sole cause of prefrontal activity in source memory tests. However, there have been no direct comparisons between prefrontal activity in source recall and recognition tests, so that response conflict may be a contributing factor in recognition tests. In the present study, the possibility of a conflict between familiarity and response selection is maximized by asking subjects to respond 'new' to word pairs composed of recombinations of studied elements.

The Present Study

Subjects study pairs of words, and at test are asked to respond 'old' to studied pairs, and 'new' to pairs including a new word, and to pairs composed of two studied words in a new combination (Recombined pairs). Accurate responses thus require associative memory for the original word pairings, and the familiar elements contained in both Old and Recombined pairs are likely to require high levels of evaluation before selecting differential responses of 'old' and 'new'. They may also stimulate extended searches of memory to recover the relevant differentiating information. Both the *search* and *evaluation* ideas predict large prefrontal potentials for Old and Recombined pairs as compared to New. However, the *familiarity/response conflict* hypothesis clearly predicts that the largest prefrontal potentials will be observed in response to the Recombined pairs.

A second stimulus manipulation was designed to help tease apart the three possible control processes, and make the results more comparable to those of source memory studies. Source memory paradigms typically include one stimulus attribute (e.g. a speaker's voice) which is presented many times during the study phase, paired with different items. To mimic this exceptionally high level of familiarity for one stimulus element, half of the pairs included a *Common* word that was paired with several different words during the study phase. The other half of the word pairs are *Unique pairs*, composed of words presented only once during the study phase (see Table 1 for stimulus examples). Although both Unique and Common pairs occurred in Old, New and Recombined versions at test, the excess familiarity of the Common words might make it particularly difficult to label pairs including these words as new. The familiarity/response conflict hypothesis thus predicts a strong impact of the Common versus Unique manipulation, specifically that both the Common New and Common Recombined pairs will elicit larger prefrontal ERPs than the analogous Unique conditions. In contrast, the presence of a Common word should generate no conflict between familiarity and response selection for studied pairs, so that the familiarity/response conflict idea predicts equivalent prefrontal ERPs in the Common Old and Unique Old conditions.

The evaluation hypothesis suggests that memory signals close to the criterion between 'old' and 'new' require close scrutiny. Because of the excess familiarity of Common words, pairs containing such words will yield a stronger memory signal, and be closer to the criterion threshold than pairs of Unique words. Even Common New pairs – those with one new word and one very familiar word – may thus require close evaluation before a response of 'new' can be selected. The evaluation hypothesis generally predicts that *all* pairs with at least one studied element (five of the six conditions) should elicit large prefrontal potentials. In contrast, the memory search hypothesis makes no obvious prediction about the impact of the Common/Unique manipulation; there seems little reason to suppose that especially familar elements would be more likely than less familiar elements to trigger an extended memory search.

Given the reliability of the early widespread old/new effect across studies [see reviews by (Rugg, 1995; Friedman and Johnson, 2000)], we also expected variation in this ERP component in the present experiment. Because both Old and Recombined pairs contain studied elements, they should elicit more positive ERPs than New pairs. Because all three varieties of Common pair (Old, New and Recombined) contain one word that was studied many times, each Common condition should elicit more positive ERPs than the parallel Unique condition. Our general predictions were thus that the objective familiarity of the test pairs would be mirrored by amplitudes of an early, spatially widespread positive potential, but that late prefrontal potentials would show a different pattern of results – the exact nature of which will depend on which memory control process is most critical for task performance.

Materials and Methods

Participants

Twenty-four young adults (13 men, 11 women) served as paid volunteers. Their mean age was 24.3 years (range 19–40); 23 were right-handed, one was left-handed; six of the right-handers reported a left-handed parent or sibling. All of the participants had normal visual acuity, and no history of learning disability, neurological or psychiatric disorder (by self-report). An additional seven participants began the experiment, but did not yield usable data: three had an unacceptably high rate of artifact rejection (>50%), two did not return for the second session of the experiment, and data from two subjects were lost due to experimenter error.

Stimuli

The stimulus set was composed of 786 concrete nouns; mean frequency of usage was 25 (SD = 64) when calculated as the sum of all regularly inflected forms in the Francis and Kucera count (Francis and Kucera, 1982); mean length was 5.8 (SD = 1.5) characters. The words were presented visually in pairs, one word above the other.

Each study block consisted of 48 word pairs. Twenty-four were *Unique pairs*, meaning that both the top and bottom words were presented only a single time during the study block, so that the conjunction of top and bottom word was unique to one study trial. Twenty-four were *Common pairs*, because the lower element (bottom word) of these pairs was repeated across the study block. In each study block, there were two such common bottom words; a participant might see 12 different top words each paired with the bottom word 'pie', and another 12 different top words each paired with the bottom word 'wreath', for instance. There were eight study blocks altogether, each followed by a recognition test for those studied words and their combinations. No words repeated across study blocks.

Each recognition block consisted of 64 word pairs. Sixteen were studied Unique pairs: *Unique Old*; 16 were studied Common pairs: *Common Old.* Eight pairs consisted of words not presented in the study block: *Unique New*. Eight pairs consisted of new top words, paired

Table 1
Example of stimulus arrangement for one study/test cycle

Study phase	Test phase		
Unique pairs ($n = 24$)	lantern motel tavern pancake	lantern motel tavern pancake	*Unique old* ($n = 16$)
	grape banner bracelet hail	grape hail bracelet banner	*Unique recombined* ($n = 8$)
		carpet sky pencil alligator	*Unique new* ($n = 8$)
Common pairs ($n = 12 + 12$)	tower pie drill wreath	tower pie drill wreath	*Common old* ($n = 8 + 8$)
	puppet pie bee wreath	puppet wreath bee pie	*Common recombined* ($n = 8$)
		cable pie piano wreath	*Common new* ($n = 8$)

with the bottom words that had repeated during the study block (four for each of the two studied bottom words): *Common New*. Finally, two recognition conditions included words which had been studied, but not in the combination presented at test. Eight *Unique Rearranged* pairs were formed by recombining the top and bottom elements of studied Unique pairs. Eight *Common Rearranged* pairs were formed by recombining the top and bottom elements of studied Common pairs, so that if a participant studied 'puppet pie' and 'bee wreath', the test stimuli would be 'puppet wreath' and 'bee pie'. For both of the Rearranged conditions, the spatial positions of the words presented at study were maintained during test. The relationships between stimuli presented during the study and recognition blocks are illustrated in Table 1.

Eight different stimulus lists were constructed so that, across subjects, individual words were equally likely to occur as elements of studied pairs, new pairs and rearranged pairs. Every 'top' word rotated across the six test conditions of Unique Old, Common Old, Unique New, Common New, Unique Rearranged and Common Rearranged. Bottom words occurring in the Unique Old, Unique New and Unique Rearranged conditions were similarly rotated. Every participant studied the same set of 16 bottom words in the Common conditions (two in each study/test cycle).

Procedure

The experiment was conducted in two sessions, each lasting ~2.5 h. Each session consisted of four study/test cycles; all procedures were identical between the two sessions. During the study phases, pairs were presented for 275 ms with a 4000 ms intertrial interval. During the test phases, pairs were presented for 275 ms with an 8000 ms intertrial interval.

For each study block, the assigned task on each trial was a relative size judgement. Participants were asked to form a mental image of each concrete noun, and to press a key with the right (or left) index finger if the top word was the larger object, or to press with left (or right) finger if the bottom word was the larger object. The mapping between response hand and 'top/bottom larger' was counterbalanced across subjects for the first session; in the second session the response mapping for each individual was reversed. Across all of the stimulus pairs, participants responded 'top larger' and 'bottom larger' with equal frequency. To ensure that this would be true for all conditions, the words used as common bottom words were in the middle of the size range of the total stimulus set.

In the recognition tests, participants were asked to respond 'old' only if a test pair corresponded exactly to one that they had studied. 'Old' was thus the correct response for Unique Old and Common Old pairs, whereas 'new' was the correct response for Unique New, Common New, Unique Rearranged and Common Rearranged pairs. The stimulus composition in each recognition test was such that 'old' was thus the correct response for 50% of the trials, and the correct response could not be produced merely by determining whether a bottom word was unique or common. Left and right keypresses for 'old' and 'new' were counterbalanced across subjects and sessions as described above for the study task.

The experimental study/test cycles were preceded by two practice cycles with 32 study and 48 test pairs representing all the conditions. Participants were forewarned that they would be tested for their memory of combinations of the studied words, and told about the presence of the rearranged pairs. They were encouraged to strive for accuracy during the recognition tests, but told that their reaction times would also be measured. Participants were given feedback about their performance in each condition at the end of the first study/test practice cycle, and encouraged to use the second practice cycle as an opportunity to improve. No feedback was offered once the experiment began.

Electrophysiological methods

Electroencephalograms (EEGs) were recorded from tin electrodes secured in an elastic cap (Electrocap International). Recording sites included midline sites FPz, Fz, FCz, Cz, CPz and Pz, together with lateral pairs FP1, FP2, F3, F4, FT7, FT8, T3, T4, P3, P4, T5, T6, O1 and O2, as defined by the 10–10 system (American Electroencephalographic Society, 1994). An additional active electrode was placed over the right mastoid. Vertical eye movements and blinks were monitored via an electrode placed below the right eye (Le). Interelectrode impedances were below 5 kΩ. The scalp sites and vertical EOG electrode were referenced to the left mastoid during recording, but re-referenced offline to an average of the left and right mastoids. Horizontal eye movements were monitored via a right-to-left bipolar montage at the external canthi of the two eyes. EEGs were amplified by a Grass Model 12 polygraph with half-amplitude cutoffs of 0.01 and 100 Hz, digitized online at a sampling rate of 250 Hz, and stored on compact disk along with stimulus codes for subsequent averaging. Trials with artifacts due to eye movements, blinks or amplifier saturation were rejected prior to averaging.

Results

All Participants

Accuracy

Mean accuracy is shown in Table 2, which indicates that labeling the recombined pairs as 'new' was more difficult than accepting the studied pairs as old, or rejecting the pairs with at least one new word. For the Unique pairs, Recombined accuracy was lower than accuracy for both Old [$F(1,23) = 8.40$, $P < 0.01$] and New [$F(1,23) = 83.4$, $P < 0.0001$]. For the Common pairs, Recombined accuracy was also lower than the Old and New conditions [Fs(1,23) > 76.8, Ps < 0.0001]. Repetition of the common 'bottom words' during the study (and test) phase increased the error rate on the recombined pairs by an average of 24% [Unique Recombined versus Common Recombined: $F(1,23) = 124.7$, $P < 0.0001$]. In an ANOVA with only the Old and New pairs, accuracy was generally higher for New than Old [$F(1,23) = 62.5$, $P < 0.0001$], and higher for Unique than Common pairs [$F(1,23) = 12.4$, $P < 0.02$]. A significant interaction between Old/New and Unique/Common reflected the fact Common pairs were more likely to receive 'old' responses, so that hit rates for Old pairs were higher, and correct rejection rates for New pairs lower when the pairs contained a common bottom word than when they did not [$F(1,23) = 9.87$, $P < 0.005$].

Reaction time

Analysis of the reaction times for all participants (shown in the right-hand column of Table 3) yielded main effects of Old/Recombined/New [$F(2,46) = 98.1$, $P < 0.0001$] and of Unique versus Common [$F(1,23) = 13.7$, $P < 0.005$], as well as an interaction [$F(2,46) = 4.97$, $P < 0.02$]. Pairwise comparisons showed that for both Unique and Common pairs, responses to the New pairs were fastest, followed by Old, with the slowest responses made to Recombined pairs [all Fs(1,23) > 9.82, all Ps < 0.005]. Comparisons between Unique and Common pairs showed that responses to Common pairs were slower than Unique when they were New [$F(1,23) = 6.73$, $P < 0.02$] or Recombined [$F(1,23) = 20.3$, $P < 0.001$], but not when they were

Table 2
Accuracy in percent correct

	Old	Recombined	New
Unique			
Mean (SE)	80.5 (2.1)	71.5 (3.1)	97.3 (0.7)
Range	62–93	50–95	85–100
Common			
Mean (SE)	82.3 (1.6)	47.2 (3.7)	91.3 (1.8)
Range	63–92	14–80	65–100

Table 3
Reaction times

	Poor performers	Good performers	All
Unique old	1544 (76)	1599 (80)	1572 (54)
Common old	1456 (67)	1798 (90)	1627 (66)
Unique recombined	1875 (103)	2019 (131)	1947 (83)
Common recombined	2027 (152)	2344 (114)	2185 (99)
Unique new	1258 (94)	1467 (100)	1363 (70)
Common new	1406 (80)	1590 (94)	1498 (64)

Mean and SE, correct responses only. Poor performers are the 12 participants with lower accuracy in the Recombined conditions; Good performers are the 12 participants with higher accuracy in those conditions; All are all 24 participants.

Old. However, these overall results are qualified by analyses which consider individual performance on the difficult Recombined pairs, presented below under 'Individual Variability'.

Event-related potentials

Figure 1 contrasts the ERPs elicited by Unique and Common pairs which elicited correct responses, separately for the Old, Recombined and New conditions. Over the posterior half of the head, all three columns show an enhanced positive potential for pairs with a Common word, maximal in the 300–600 ms poststimulus latency range. The Recombined and New conditions show an additional frontal difference between Common and Unique pairs that was absent in the Old condition. The frontal difference between Unique and Common also began ~300 ms, but persisted until ~1200 ms poststimulus. After 700 ms, differences between Unique and Common were largely restricted to the prefrontal and frontal sites. Topographic maps of the distribution of voltages in the 300–600 and 700–1000 ms latency windows are shown in Figure 2.

Mean amplitudes of ERPs from the midline electrode sites were measured in two latency windows, the 300–600 ms range, and a 700–1000 ms range, both relative to 200 ms of prestimulus activity. The impact of the Unique versus Common manipulation was analyzed by three ANOVAs for the Old, Recombined and New pairs, each taking scalp site as a repeated measure with six levels. An orthogonal trend analysis was included as an option in the ANOVAs, which forced the scalp sites to be taken as ordered (rather than simply different) levels of this anterior-to-posterior (AP) factor. In this type of analysis, graded changes in the Unique versus Common difference along the AP axis are indexed by interactions between Unique/Common and the linear component of the AP factor. In the 300–600 ms latency window, the main effect of Unique versus Common was significant for the Old [$F(1,23) = 8.65$, $P < 0.01$], Recombined [$F(1,23) = 7.97$, $P < 0.01$] and New pairs [$F(1,23) = 9.43$, $P < 0.005$]. For the Old pairs only, the main effect was accompanied by an interaction with the AP factor, reflecting the similarity of Unique and Common pairs at prefrontal and frontal sites as compared to more posterior sites [Unique/Common × AP_{linear}, $F(1,23) = 15.0$, $P < 0.005$, quadratic and higher order components nonsignificant]. The presence of both anterior and posterior differences for the Recombined and New pairs led to null interactions between Unique/Common and the AP factor (Fs < 1). Parallel analyses were performed on the seven lateral pairs of scalp sites, with the same pattern of results.

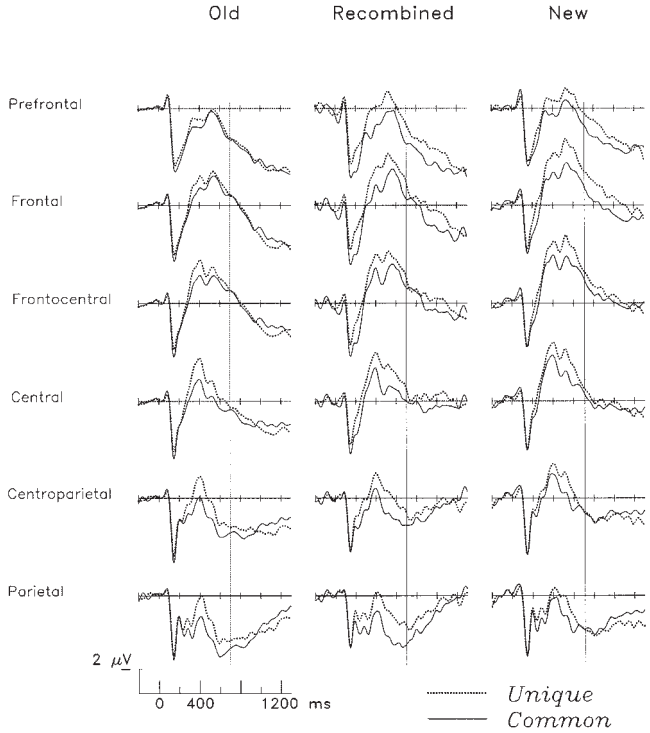

Figure 1. Grand average ERPs from all 24 subjects at the midline scalp sites. The vertical lines are drawn at 700 ms poststimulus onset.

The results from the 700–1000 ms window were different from those of the the earlier latency window. The main effect of Unique versus Common was not significant for any pair type. Because late differences between Unique and Common were restricted to prefrontal and frontal sites in the Recombined and New conditions, we expected to instead find interactions beween Unique/Common and the AP factor in these conditions only. This interaction was nonsignificant for the Old pairs ($F < 1$), and significant for the New pairs [Unique/Common × AP_{linear}, $F(1,23) = 13.9$, $P < 0.002$]. However, contrary to what Figure 1 suggests, the AP interaction was not significant for the Recombined pairs. This contrast between the figure and the statistical outcome is due to a large degree of intersubject variability, which is examined in the remainder of the paper.

Individual Variability

Accuracy

Across subjects, the probability of calling a Common Recombined pair 'old' or 'new' was close to 50%. However, examination of all the conditions indicates that this apparently chance level of performance cannot be attributed to simple forgetting of the words which made up recombined pairs, as participants were quite accurate in accepting pairs of words that occurred together in the study phase (Unique Old and Common Old), and extremely accurate in rejecting pairs containing two new words

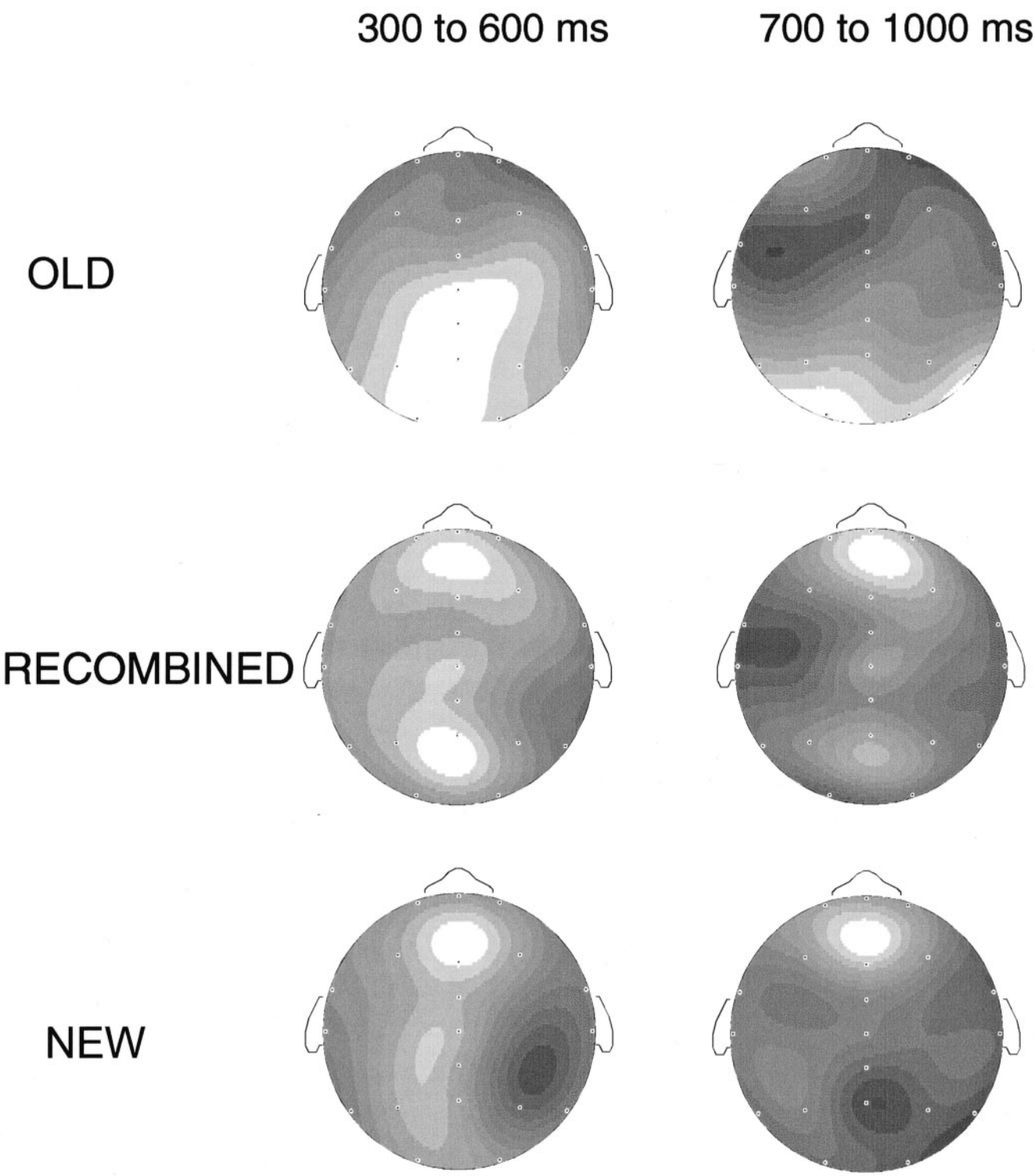

Figure 2. Voltage maps of the difference between Common and Unique pairs, for the three word pair types and two latency windows. Small white dots indicate electrode locations. Amplitude is normalized to a range of positive one (white) to negative one (black).

(Unique New) or one new word (Common New). Instead, the propensity to accept a recombined pair as old must be attributed to recognition of the individual words, but failure to correctly judge their pairing during the study phase. Table 2 shows the range of accuracy across participants for each condition. Although all conditions showed intersubject variability, the range of performance for the Recombined conditions was particularly striking. We subdivided the 24 participants into two groups of 12 based on their average accuracy in these two conditions (47% versus 71%). Considering only Unique Recombined, or only Common Recombined picked out the same subjects as *Good Performers* or *Poor Performers*. Figure 3

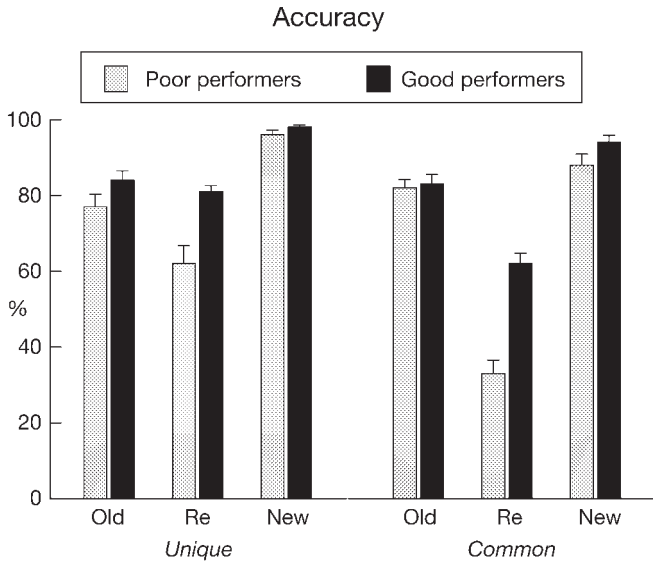

Figure 3. Percent correct in each of the six conditions. 'Re' refers to the Recombined pairs. Good and Poor performers are subgroups of 12 subjects each, defined by higher versus lower accuracy in labeling the Recombined pairs as 'new'.

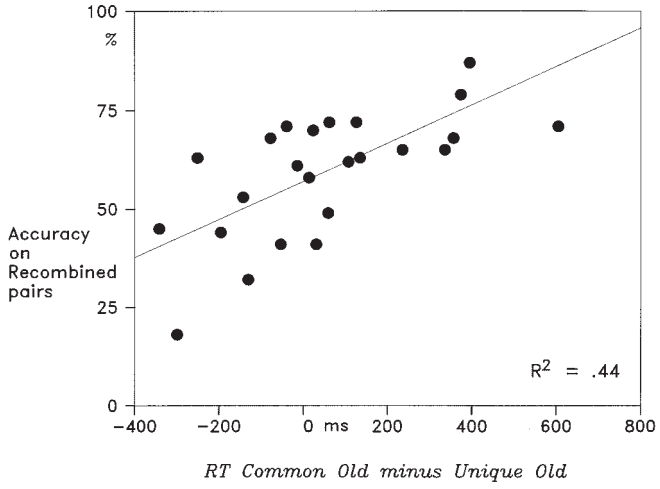

Figure 4. Each subject is represented by a dot. Accuracy on recombined pairs is an average of accuracy in the Unique Recombined and Common Recombined conditions. Reaction time (RT) is based on correct trials. The thin diagonal line and r^2 are from the regression equation relating Recombined accuracy to slowing on correct Old trials with a Common word as compared to correct Old trials with Unique words.

contrasts accuracy in all the conditions for these two subsets of participants, and suggests that the ability to reject the recombined pairs was not tied to overall memory ability. An ANOVA considering only accuracy for the Old and New pairs (with Unique versus Common as a second factor) showed no main effect or interactions involving the group factor. In contrast to the null group difference for Old and New pairs, accuracy on Unique Recombined [$F(1,22) = 13.7$, $P < 0.002$] and Common Recombined [$F(1,23) = 40.8$, $P < 0.0001$] were lower for the Poor than the Good Performers (not surprisingly, since the groups were formed on the basis of accuracies in these conditions).

Reaction Time

The first suggestion of a strategy difference among participants comes from the reaction times on correct trials, shown in Table 3: Good Performers were more likely than Poor Performers to slow down when encountering a pair with a Common (repeating) word. An ANOVA taking Group, Unique/Common and Old/Recombined/New as factors showed a nonsignificant trend ($P = 0.10$) for Good performers to be generally slower than than Poor performers. The three-way interaction of Group × Unique/Common × Old/Recombined/New [$F(2,44) = 4.23$, $P < 0.05$] was significant. This interaction was examined by separate ANOVAs on the Old, Recombined and New Pairs. For Old trials, the poor performers were faster when the pairs included a Common word, whereas the good performers were slower [Group × Unique/Common, $F(1,22) = 13.7$, $P < 0.002$]. Table 3 indicates that the Good performers also tended to slow down more than the Poor performers when encountering a Common word in a Recombined pair, but this group difference was not statistically significant, nor were there any group differences in reaction times for New pairs.

Correlations were used to evaluate the relationship between slowing on Common pairs and accuracy in the Recombined conditions across individual subjects. Three RT differences (Common minus Unique) were calculated for the Old, Recombined and New pairs. Figure 4 shows that reaction time slowing on the Old pairs was significantly correlated with accuracy ($r = 0.67$, $P < 0.0005$). As in the group analyses, slowing on Recombined or New pairs was not related to accuracy.

Event-related potentials at posterior sites

The preliminary ERP analyses reported above, as well as previous work from our laboratory, indicate a functional distinction between early posterior ERPs and late prefrontal ERPs in memory tasks. The individual difference analyses are thus based on maximally separable aspects of the ERP data: the 300–600 ms latency range at parietal, posterior temporal and occipital sites (P3,P4,T5,T6,O1, O2), and the 700–1000 ms latency range at prefrontal sites (Fpz, Fp1, Fp2). (Amplitudes of the prefrontal ERPs in the 300–600 ms latency window were also analyzed, but in contrast to the 700–1000 ms latency window did not correlate with behavioral performance. The functional significance of the early difference between more and less familiar items at prefrontal sites thus remains to be explained by future research.)

At posterior scalp sites, Figure 5 shows that Old pairs elicited more positive ERPs than Recombined pairs, which, in turn, elicited more positive ERPs than New pairs. Comparison of the right and left columns of Figure 5 also shows that pairs with a Common word elicited more positive ERPs than pairs of Unique words. The initial analysis considered ERP averages based on correct trials only, and yielded main effects of Old/Recombined/New [$F(2,46) = 24.7$, $P < 0.0001$] and Unique versus Common [$F(1,23) = 20.3$, $P < 0.0002$]. Both effects showed a small lateral asymmetry of being larger over the left than right scalp [$F(2,46) = 3.99$, $P = 0.05$ for Old/Recombined/New, and $F(1,23) = 5.87$, $P < 0.05$ for Unique versus Common]. There were no significant interactions between Old/Recombined/New and Unique/Common. Including trials with both correct and incorrect behavioral responses did not change the pattern of results (Fs > 22, Ps < 0.0001 for the two main effects, Fs > 4.92, Ps < 0.05 for the interactions with the Left/Right factor). [It is conventional in ERP memory research to analyze only trials with correct behavioral responses, with a corresponding decrease in signal-to-noise ratio (SNR) that comes from a reduced number of trials. In the present data, participants with low accuracy in the Recombined conditions will necessarily have lower SNRs than participants with higher accuracy in these conditions, if only correct trials

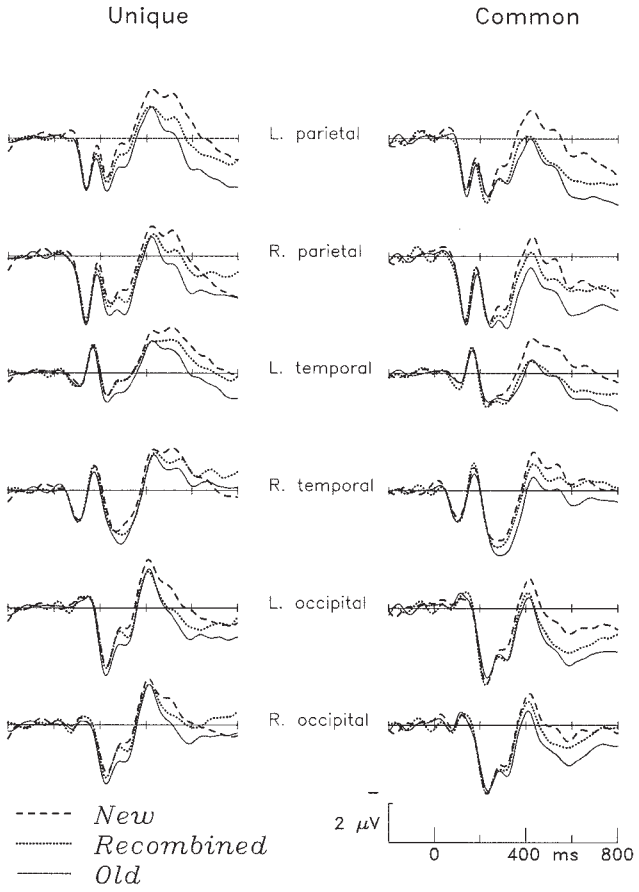

Figure 5. Grand average ERPs from all 24 subjects at the left and right parietal (P3, P4), posterior temporal (T5, T6) and occipital (O1, O2) scalp sites.

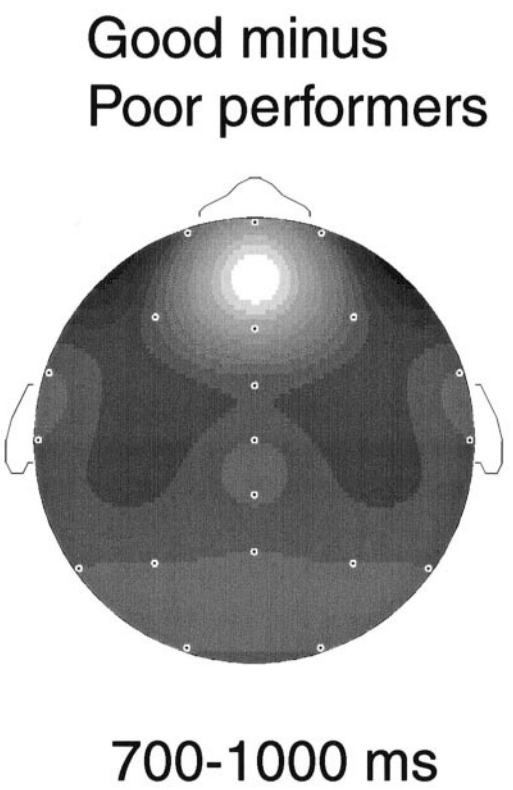

Figure 6. Voltage map of the difference between the ERPs of the Good and Poor performers, across all conditions. Small white dots indicate electrode locations. Amplitude is normalized to a range of positive one (white) to negative one (black).

are considered. We were concerned that differential SNRs might introduce spurious amplitude differences between those with higher and lower accuracy levels due to higher amplitude residual EEG activity (noise) in the poor performers. We thus conducted all the statistical analyses on two parallel sets of ERP averages, one including all artifact-free trials, and one including only artifact-free trials with correct behavioral responses. In all cases, the pattern of statistically significant and nonsignificant results was the same across the two sets of analyses. In the remainder of the paper, we thus report only those analyses based on correct trials, except for explicit comparisons between accurate and erroneous trials.]

Given that errors in the Recombined conditions consisted of labeling these pairs as 'old', we next examined whether the posterior ERPs could reliably discriminate between the truly Old pairs and the Recombined pairs, and whether the difference between Old and Recombined was larger in the Good than the Poor performers. An ANOVA taking Group, Old versus Recombined, and Unique versus Common as factors showed a main effect of Old versus Recombined [$F(1,22) = 8.83$, $P < 0.01$], but no significant interactions with the group factor [(Fs < 1.7]. The difference between the Recombined and New pairs was also reliable [$F(1,22) = 14.7$, $P < 0.002$], and similarly showed no significant interactions with Group. We also examined correlations between accuracy on the recombined pairs and the amplitudes of four ERP memory effects at the posterior scalp sites: Unique Old – Unique Recombined, Common Old – Common Recombined, Unique Old – Unique New, and Common Old – Common New. None were significant.

Overall, the posterior ERPs were sensitive to the 'oldness' of both the individual words making up the pairs, and to the combinations of words. The inclusion of a frequently studied element (the Common words) increased the posterior positivity as compared to pairs made up of two words studied only once (the Unique pairs). Truly old pairs elicited more positive ERPs than recombined pairs formed from two studied words, but the presence of two studied words — even in the wrong combination — yielded more positive ERPs than pairs that included at least one new element. However, the statistical analyses provided no suggestion that the posterior ERPs reflect the ability of some participants to successfully reject the recombined pairs, whereas other subjects were likely to erroneously accept the recombinations of studied elements as studied pairs.

Prefrontal ERPs

Figure 6 shows that the ERPs of Good versus Poor performers were distinguished by a focal prefrontal positivity in the late epoch, so that late prefrontal ERPs show promise in accounting for the wide range of performance variability on the Recombined pairs. The group difference appears to be laterally symmetric, like the bilateral effects in our previous source memory experiments (Senkfor and Van Petten, 1998; Van Petten *et al.*, 2000) (see also A.J. Senkfor *et al.*, submitted) and some reports from other laboratories (Wilding *et al.*, 1995), but unlike the rightward (Wilding and Rugg, 1996; Trott *et al.*, 1997) or leftward (Tendolkar and Rugg, 1998; Ranganath and Paller, 1999) asymmetries in late prefrontal memory effects in other reports.

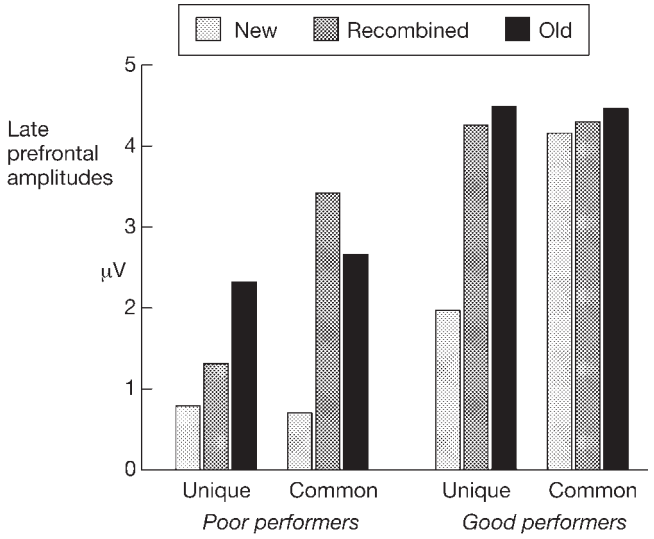

Figure 7. ERP amplitudes in the 700–1000 ms poststimulus latency window for each of the six conditions, averaged across the prefrontal scalp sites (Fpz, Fp1, Fp2). Good performers are the 12 subjects with higher accuracy in the Recombined conditions; Poor performers are the 12 subjects with lower accuracy in those conditions.

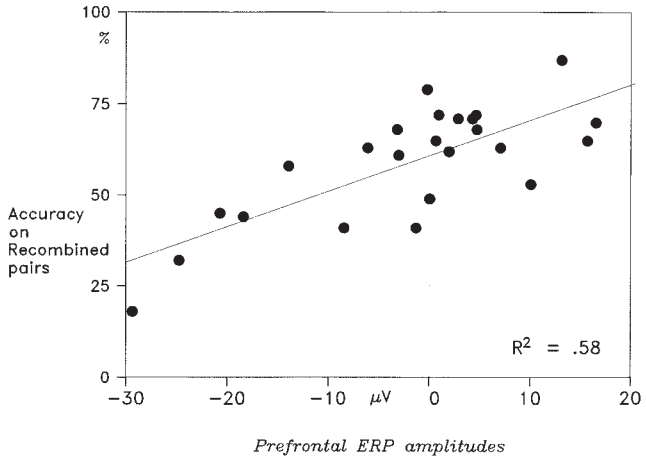

Figure 8. Each subject is represented by a dot. Accuracy on recombined pairs is an average of accuracy in the Unique Recombined and Common Recombined conditions. Prefrontal ERP amplitudes were measured in the 700–1000 ms postimulus latency window, averaged across scalp sites Fpz, Fp1 and Fp2 in the Common New and Common Recombined conditions. The thin diagonal line and r^2 are drawn from the regression equation: Recombined accuracy (in percent) = 61 + 4.78 (Common New ERP amplitude in microvolts) – 3.51 (Common Recombined ERP amplitude).

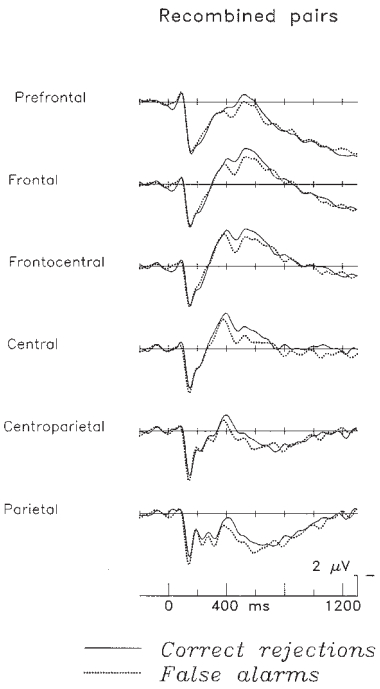

Figure 9. Grand average ERPs across all 24 subjects for Recombined trials which elicited a correct response (correct rejections) versus an incorrect response (false alarms). Each subject contributed equal weight to the grand average, regardless of accuracy. Statistical analyses showed no significant differences between correct rejections and false alarms at these scalp sites, or at the posterior (parietal, temporal, occipital) or prefrontal scalp sites.

Figure 7 shows prefrontal ERP amplitudes in the 700–1000 ms latency range broken down by condition for the Good and Poor performers. Good performers had equally large prefrontal positivities for all of the pairs including a Common word. In fact, the Good performers had large prefrontal amplitudes for all pairs including at least one studied word, including Unique Old and Unique Recombined. Poor performers had prefrontal positivities of more variable amplitude depending on pair type. The different patterns of amplitudes in the two groups were reflected in a three-way interaction of Group × Old/Recombined/New × Unique/Common [$F(2,44) = 3.61$, $P < 0.05$].

Correlations between Recombined accuracy and amplitudes of the prefrontal ERPs (averaged across Fpz, Fp1, Fp2) in each of the six conditions ranged from 0.12 to 0.50; the strongest correlations with accuracy were observed for prefrontal amplitudes in the Common New ($r = 0.50$, $P < 0.01$) and Unique Rearranged conditions ($r = 0.44$, $P < 0.02$). In a stepwise regression analysis, two of the ERP variables jointly accounted for 58% of the variance in performance [$F(2,21) = 14.6$, $P < 0.0005$]: prefrontal amplitudes in the Common New and Common Recombined conditions, as shown in Figure 8. [We averaged across the midline (Fpz), left (Fp1) and right (Fp2) prefrontal channels to provide the most stable measure, and because examination of the data suggested that the prefrontal differences between Good and Poor performers were bilateral in nature (see Fig. 5). Regressions on individual prefrontal sites picked out the same variables as predictors of Recombined accuracy, with $r^2 = 0.49$ for Fpz, 0.50 for Fp1, and 0.66 for Fp2. The predictive power of frontal potentials was slightly weaker, but still substantial: r^2s were 0.40, 0.51 and 0.52 for ERPs recorded at Fz, F3 and F4 respectively. A slightly more posterior site, Cz, yielded no significant relationships with Recombined accuracy, nor did the posterior cluster of P3, P4, T5, T6, O1 and O2.] A regression on accuracy in the Unique Recombined condition alone picked out the same predictor variables [$r^2 = 0.61$, $F(2,21) = 16.4$, $P < 0.0005$], as did a regression on accuracy in the Common Recombined condition [$r^2 = 0.47$, $F(2,21) = 9.4$, $P < 0.005$].

Prefrontal amplitudes in the six conditions were also entered as potential predictor variables for accuracy on the Old and New pairs, but these analyses yielded no significant outcomes. The prefrontal ERPs were thus linked solely with accuracy on the recombined trials, and not with general memory performance.

The analyses above indicate that prefrontal amplitudes are correlated with an individual's ability to judge the Recombined pairs as unstudied. A different way of examining the relationship

between accuracy and brain activity is to compare correct and incorrect trials within the same individuals. Figure 9 shows little difference between the ERPs elicited by Common Recombined pairs which received correct 'new' responses versus those that received erroneous 'old' responses, at either the posterior or the prefrontal sites. An ANOVA of posterior amplitudes in the 300–600 ms latency window, taking Good versus Poor performers, Unique versus Common, and correct/incorrect as factors yielded no significant main effect or interactions involving the correct/incorrect factor. Parallel analyses of posterior and prefrontal amplitudes in the 700–1000 ms latency window similarly yielded no significant differences between recombined trials with correct versus incorrect judgements.

Relationship between ERPs, reaction times and accuracy

Thus far, the analyses have pointed to two factors that are correlated with an individual's ability to reject the Recombined pairs: slowing down when encountering pairs with Common words as compared to Unique words (Fig. 4), and prefrontal ERP amplitudes (particularly those elicited by the Common New pairs). It is of some interest to evaluate whether these predictors are independent, or themselves related. The reaction time difference between Common Old and Unique Old was significantly correlated with prefrontal amplitudes in each of the six conditions (rs from 0.35 to 0.57, Ps < 0.05), with the strongest correlation for prefrontal amplitude in the Common New condition (r^2 = 0.33, P < 0.005). [We also examined correlations among prefrontal amplitudes and raw reaction times in the six conditions. These were low and nonsignificant. These null effects are important as they rule out the possibility that the prefrontal ERPs were strongly influenced by preparation to make a motor response (more typically associated with ERPs at central scalp sites), which might be expected to begin later on trials with slow reaction times.] We next returned to using accuracy on the Recombined pairs as a dependent variable, and asked whether prefrontal amplitudes and reaction time slowing could account for more of the variance in accuracy than the ERP or behavioral measure alone. This was not the case; reaction time slowing entered the regression equation first (r^2 = 0.44, P < 0.0005), and accounted for much the same variance as prefrontal ERP amplitudes, so that none of the ERP measures then met the minimum F-value of 4.0 for inclusion.

Discussion

Posterior ERPs and Raw Memory

The initial difference among conditions, in the 300–600 ms latency range, was much the same as the basic 'old/new effect' which has differentiated studied (and remembered) old items from unstudied (or forgotten) items in many previous experiments. At posterior scalp sites, the old/new effect reflected the *oldness* of test stimuli in a graded fashion, as shown by three contrasts: (i) pairs consisting of two studied elements elicited more positive ERPs than pairs with one or more new elements (Recombined > New, Fig. 5); (ii) pairs consisting of a studied combination of elements elicited more positive ERPs than new combinations of studied elements (Old > Recombined, Fig. 5); and (iii) pairs with a frequently studied element elicited more positive ERPs than pairs composed of elements studied only once (Common > Unique, Fig. 1). The early, posterior old/new effect thus seems to be a sensitive index of raw memory strength for both elements and combinations.

Some memory theories include a contrast between *familiarity*, roughly consistent with what we have called 'item memory', and *recollection*, defined as memory for an item and the context in which it was learned and thus like what we have called 'source memory' (Mandler, 1980; Yonelinas, 1999). It would not be appropriate to label the posterior old/new effect observed here as an index of familiarity in this sense, given that it discriminated Old from Recombined pairs. Instead, the graded amplitude of early positivity across Old, Recombined and New pairs is more consistent with the idea that familiarity and recollection are dependent on the same raw memory signal [a function attributed to the medial temporal lobe (Squire and Knowlton, 2000)].

However, the results showed that raw memory was insufficient for accurate judgements in the difficult recombined conditions. The Good and Poor performers showed the same posterior ERP gradient of Old > Recombined > New, the same accuracy on Old and New pairs, yet differed greatly in their Recombined accuracy rates (47% versus 71% on average). Comparing the individual trials which yielded correct and incorrect judgements of the Recombined pairs similarly revealed no difference over posterior scalp (Fig. 9), also suggesting that incorrect judgements cannot be attributed to faulty memory *per se*.

As predicted, ERPs recorded at prefrontal scalp sites proved more sensitive to the memory control processes which allowed participants to behaviorally distinguish the presence of familiar elements from studied pairs of elements. Before discussing the functional interpretation of the prefrontal ERPs, we turn to what the reaction time results reveal about the different strategies participants adopted to deal with the exclusion source memory task.

Familiarity-induced Slowing

In simple item recognition tasks, hits ('old' responses to studied items) are invariably faster than correct rejections ('new' responses to unstudied items). In source recognition tasks, the opposite result is seen: responses to studied items are substantially slower than those to new items. Reaction times to new items are much the same in item and source recognition tests. This pattern of results strongly suggests that source memory decisions about the context in which items were studied are postponed until after the items themselves have been recognized as old, and that new items largely escape the additional processing required for a source memory decision (Gronlund and Ratcliff, 1989; Johnson *et al.*, 1994; Senkfor and Van Petten, 1998).

In the present paradigm, the only stimuli that were unambiguously new were the Unique New pairs, comprised of two unstudied elements, and these received the fastest responses from all participants. The next fastest reaction time was observed for Common New pairs, in which the single unstudied word could signal that the pair was itself new. The remaining four conditions were comprised of two studied elements, so that an optimal strategy would mandate full analysis — with the attendant slowing — of the Unique Old, Common Old, Unique Recombined and Common Recombined pairs. When the full group of participants was divided into halves based on their accuracy for the difficult Recombined conditions, the Good performers showed evidence of this optimal strategy: their reaction times for both Old and Recombined pairs were slower than the corresponding reaction times for New pairs, in both the Unique and Common conditions [Fs(1,11) ranging from 5.00 to 319.8, Ps < 0.05]. We refer to this optimal strategy as *familiarity-induced slowing* — delaying a response after recognizing a studied word (or words) because old words may indicate an Old pair or a Recombined pair, and only additional analysis will

indicate which response should be made. The Poor performers showed a less consistent pattern of slower responses for pairs comprised of studied elements as compared to New pairs; their reaction times were slower for Unique Old, Unique Recombined and Common Recombined as compared to the corresponding New conditions [Fs(1,11) > 17.8, Ps < 0.002], but not for the Common Old pairs as compared to Common New (F < 1).

Additional evidence that familiarity-induced slowing was beneficial for accurate classification of the Recombined pairs comes from the contrast between pairs with familiar elements (Unique Old and Unique Recombined) and those with one especially familiar element that had been studied multiple times (the Common Old and Common Recombined pairs). The strategy of participants who performed poorly on the Recombined pairs might be characterized as leaping to conclusions — these participants tended to respond more quickly when a Common word occurred, and were likely to miscategorize the pairs with familiar words as old when they were in fact recombinations of studied elements. Slowing when encountering a Common word accounted for a substantial degree (44%, Fig. 4) of the individual variance in ability to reject the Recombined pairs.

Overall, the reaction time results suggest that good performance on the recombined conditions resulted from the engagement of some additional process after initial memory retrieval, but before emitting a response, and that this additional process was triggered by the presence of familiar items. The relationships between individual accuracy level and long-latency ERPs at prefrontal scalp sites strongly suggests that this additional process is based in prefrontal cortex.

Prefrontal Potentials and Memory Control Processes

We initially considered three possible accounts of the late prefrontal ERP effects apparent in source memory paradigms: resolution of response conflict when stimuli composed of familiar elements cannot simply be labeled as 'old', extended memory search and closer evaluation of the initial products of memory retrieval. A strong prediction of the response conflict hypothesis — that Recombined pairs should elicit the largest prefrontal positivity — was not confirmed for either the Good or the Poor performers (Fig. 7). The response conflict idea is less viable than either the search or evaluation ideas in accounting for the present results.

In our previous studies of source memory, the evidence in favor of the extended search hypothesis rested largely on the temporal relationships between the prefrontal effect and two memory effects observed over more posterior scalp. The prefrontal effect was maximal after the initial differentiation of old from new items, but before the late posterior effect indicating successful retrieval of the source information accompanying each item (Senkfor and Van Petten, 1998; Van Petten *et al.*, 2000; Senkfor, 2002; Senkfor *et al.*, 2002). The current results differ from these previous studies in showing no sign of a delayed posterior retrieval effect. We might have expected a late (post 700 ms) difference between correct Recombined trials as compared to incorrect Recombined trials, if indeed an extended search of memory revealed the information necessary to classify the recombined pairs.

The reliable appearance of a delayed source retrieval effect in other studies and its absence here suggests that retrieval of the word pair information occurred in a single step. The other piece of evidence suggesting a single retrieval phase for both the words and their combinations is, of course, the observation of a three-way difference between Old, Recombined and New pairs during the 300–600 ms epoch (Fig. 5), which also replicates the results of Donaldson and Rugg in a word pair recognition test (Donaldson and Rugg, 1998). There is thus one critical difference between source and associative memory tests. When a conjunction of cross-domain elements like a word plus a voice, or an object plus a location must be retrieved, the retrieval process appears to operate in two distinct stages. When the two elements are drawn from the same domain, as here, there is little evidence for staged retrieval or secondary memory search.

Given that participants retrieved all of the information that they ever would soon after stimulus onset, we are left with the third account of the prefrontal activity in recognition tests: close evaluation of retrieval products before making a categorical decision. Fortunately it is also consistent with the results. As compared to the New pairs with at least one unstudied element, both the Old and Recombined pairs were likely to generate a memory signal close to the criterion for responding 'old'. Both the Good and Poor performers generated larger prefrontal potentials for the Old and Recombined pairs as compared to the completely new pairs (Unique New, Fig. 7). All of the participants thus showed some sign of enhanced evaluation when confronted with the potentially problematic pairs. The difference between Good and Poor performers lay in the overall amplitudes of their prefrontal potentials, and, more critically, in how they treated the pairs composed of one old and one new word — the Common New pairs. It would appear that the Good performers adopted a lower threshold for what sort of stimuli required close scrutiny: all pairs with at least one old word elicited equivalently large prefrontal potentials. The Poor performers adopted an apparently less successful strategy of close evaluation only when a pair contained two studied words.

The present results provide a clear separation between memory retrieval *per se*, indexed by relatively early electrical brain activity recorded over posterior brain regions, and memory control processes which can lead to accurate judgements about familiar but new stimuli, indexed by activity over prefrontal regions. The results indicate that pausing when encountering familiar (but possibly new) stimuli, and enhanced prefrontal activity are both beneficial for judging familiar stimuli accurately. A fairly direct relationship between familiarity-induced slowing and the prefrontal potentials is suggested by the positive correlations between degree of slowing and prefrontal amplitudes, and by the regression analysis indicating that the RT and ERP variables accounted for the same individual variability in performance. The simplest account of the results is thus that the pause before responding is filled by prefrontal evaluative processes.

An intriguing aspect of the results is that although prefrontal activity predicted an individual's overall likelihood of correct response to the recombined pairs, there was no within-subject difference between prefrontal ERPs on correct and incorrect trials (Fig. 9). This dissociation indicates that prefrontal potentials reflected an individual's general strategy, but that even adoption of the optimal strategy (careful evaluation of all pairs with at least one studied word) did not guarantee a successful outcome on every trial. Indeed, a modified version of the lottery slogan could summarize the dissociation between subject-based accuracy and trial-based accuracy: 'You can't win if you don't play — but no matter how much you play, you can still lose.'

An Unresolved Question

An unresolved issue concerns the relationship between two memory control processes that we have suggested are critical for performance in source and associative recognition tests. The

present results are most consistent with the idea that close evaluation of memory traces near a response criterion drives accurate performance, and is prefrontally based. However, as described above, previous results in more canonical source recognition tests are much more consistent with a two-stage retrieval process, in which the second stage of extended memory search is instigated prefrontally (and perhaps followed by evaluation processes). Because these two control processes have not been securely identified within a single experiment (with the same participants), it is not presently possible to specify their anatomical similarity any more precisely than to say that both are dependent on the frontal lobe. In theory, a direct comparison between extended memory search and prolonged evaluation is possible, but this will be a difficult challenge as separating the two processes may require both spatially and temporally precise data.

Notes

We are grateful to John Gross, Roh Ohst, Jonathan Forster and David Sage for technical support. Financial support was provided by the National Institute of Aging (AG 14792).

Address correspondence to Cyma Van Petten, Department of Psychology, University of Arizona, Tucson, AZ 85721, USA. Email: vanpettc@u.arizona.edu.

References

American Electroencephalographic Society (1994) Guidelines for standard electrode position nomenclature. J Clin Neurophysiol 11:111–113.

Burgess PW, Shallice T (1996) Confabulation and the control of recollection. Memory 4:359–411.

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

Cycowicz YM, Friedman D, Snodgrass JG, Duff M (2001) Recognition and source memory for pictures in children and adults. Neuropsychologia 39:255–267.

Davidson PSR, Glisky EL (2002) Neuropsychological correlates of recollection and familarity in normal aging. Cognit Affect Behav Neurosci (in press).

Dimitrov M, Granetz J, Peterson M, Hollnagel C, Alexander G, Grafman J (1999). Associative learning impairments in patients with frontal lobe damage. Brain Cognit 41:213–230.

Dodson CS, Shimamura AP (2000) Differential effects of cue dependency on item and source memory. J Exp Psychol Learn Mem Cognit 26:1023–1044.

Donaldson DI, Rugg MD (1998). Recognition memory for new associations: electrophysiological evidence for the role of recollection. Neuropsychologia 36:377–395.

Fletcher PC, Henson RNA (2001) Frontal lobes and human memory. Insights from functional neuroimaging. Brain 124:849–881.

Francis WN, Kucera H (1982) Frequency analysis of English usage: lexicon and grammar. Boston, MA: Houghton Mifflin.

Friedman D, Johnson R, Jr (2000) Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microsc Res Techn 51:6–28.

Glisky EL, Polster MR, Routhieaux B (1995) Double dissociation between item and source memory. Neuropsychology 9:229–235.

Glisky EL, Rubin SR, Davidson PSR (2001) Source memory in older adults: an encoding or retrieval problem? J Exp Psychol Learn Mem Cognit 27: 1131–1146.

Gronlund SD, Ratcliff R (1989) Time course of item and associative information: Implications for global memory models. J Exp Psychol Learn Mem Cognit 15:846–858.

Henkel LA, Johnson MK, De Leonardis DM (1998) Aging and source monitoring: cognitive processes and neuropsychological correlates. J Exp Psychol Gen 127:251–268.

Henson RN, Rugg MD, Shallice T, Dolan RJ (2000) Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. J Cognit Neurosci 12:913–923.

Hermann M, Rotte M, Grubich C, Ebert AD, Schiltz K, Münte TF, Heinze HJ (2001) Control of semantic interference in episodic memory retrieval is associated with an anterior cingulate-prefrontal activation pattern. Hum Brain Map 13:94–103.

Jacoby LL (1991) A process dissociation framework: separating automatic from intentional uses of memory. J Mem Lang 30:513–541.

Janowsky JS, Shimamura AP, Squire LR (1989) Source memory impairment in patients with frontal lobe lesions. Neuropsychologia 27:1043–1056.

Johansson M, Stenberg G, Lindgren M, and Rosen I (2002) Memory for perceived and imagined pictures — an event-related potential study. Neuropsychologia 40:986–1002.

Johnson MK, Hashtroudi S, Lindsay DS (1993) Source monitoring. Psychol Bull 114:3–28.

Johnson MK, Kounios J, Reeder JA (1994) Time-course studies of reality monitoring and recognition. J Exp Psychol Learn Mem Cognit 20:1409–1419.

Johnson MK, Kounios J, Nolde SF (1997a) Electrophysiological brain activity and memory source monitoring. Neuroreport 8:1317–1320.

Johnson MK, O'Connor M, Cantor J (1997b) Confabulation, memory deficits, and frontal dysfunction. Brain Cognit 34:189–206.

Lindsay DS, Johnson MK (1989) The eyewitness suggestibility effect and memory for source. Mem Cognit 17: 349–358.

Luria AR (1973) The working brain: an introduction to neuropsychology. New York: Basic Books.

MacDonald AW, Jr, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835–1838.

Mandler G (1980) Recognizing: the judgement of previous occurrence. Psychol Rev 8: 252–271.

Miller MB, Handy TC, Cutler J, Inati S, Wolford GL (2001) Brain activations associated with shifts in response criterion on a recognition test. Can J Exp Psychol 55:162–173.

Moscovitch M (1994) Memory and working with memory: evaluation of a component process model and comparisons with other models. In: Memory Systems 1994 (Schacter DL, Tulving E, eds), pp. 269–310. Cambrige MA: MIT Press.

Olichney J, Van Petten C, Paller K, Salmon D, Iragui V, Kutas M (2000) Word repetition in amnesia: electrophysiological evidence of spared and impaired memory. Brain 123:1948–1963.

Passingham RE, Toni I, Rushworth MFS (2000) Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. Exp Brain Res 133:103–113.

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

Rapscak SZ, Kaszniak AW, Reminger SL, Glisky ML, Glisky EL, Comer JF (1998). Dissociation between verbal and autonomic measures of memory following frontal lobe damage. Neurology 50:1259–1265.

Raye CL, Johnson MK, Mitchell KJ, Nolde SF, D'Esposito M (2000) FMRI investigations of left and right PFC contributions to episodic remembering. Psychobiology 28:197–206.

Rubin SR, Van Petten C, Glisky EL, Newberg WM (1999) Memory conjunction errors in younger and older adults: event-related potential and neuropsychological evidence. Cognit Neuropsychol 16:459–488.

Rugg MD (1995). ERP studies of memory. In: Electrophysiology of mind: Event-related brain potentials and cognition, Vol. XV (Rugg MD, Coles MGH, eds), pp. 132–170. New York: Oxford University Press.

Rugg MD, Fletcher PC, Chua PML, Dolan RJ (1999) The role of the prefrontal cortex in recognition memory and memory for source: an fMRI study. Neuroimage 10:520–529.

Schacter DL, Curran T, Galluccio L, Milberg WP, Bates JF (1996). False recognition and the right frontal lobe: a case study. Neuropsychologia 8:793–808.

Senkfor AJ (2002) Episodic action memory: Characterization of time course and neural circuitry. Conscious Cognit (in press).

Senkfor AJ, Van Petten C (1998) Who said what? An event-related potential investigation of source and item memory. J Exp Psychol Learn Mem Cognit 24:1005–1025.

Senkfor AJ, Van Petten C, Kutas M (2002) Episodic action monitoring for real objects: an ERP investigation with Perform, Watch, and Imagine action encoding tasks versus a non-action encoding task. J Cognit Neurosci 14:402–419.

Shimamura AP (2000) The role of the prefrontal cortex in dynamic filtering. Psychobiol 28:207–218.

Squire LR, Knowlton B (2000) The medial temporal lobe, the hippo-

campus, and the memory systems of the brain. In: The new cognitive neurosciences (Gazzaniga MS, ed.). Cambridge, MA: MIT Press.
Swick D, Knight RT (1996) Is prefrontal cortex involved in cued recall? A neuropsychological test of PET findings. Neuropsychologia 34:1019–1028.
Swick D, Knight RT (1999) Contributions of prefrontal cortex to recognition memory. Electrophysiological and behavioral evidence. Neuropsychology 13:155–170.
Tendolkar I, Rugg MD (1998). Electrophysiological dissociation of recency and recognition memory. Neuropsychologia 36:477–490.
Trott C, Friedman D, Ritter W, Fabiani M (1997) Item and source memory: differential age effects revealed by event-related potentials. Neuroreport 8:3373–3378.
Van Petten C, Senkfor AJ (1996) Memory for words and novel visual patterns: repetition, recognition, and encoding effects in the event-related brain potential. Psychophysiology 33:491–506.
Van Petten C, Senkfor AJ, Newberg WM (2000) Memory for drawings in locations: spatial source memory and event-related potentials. Psychophysiology 37:551–564.
Wells GL, Loftus EF (eds) (1984) Eyewitness testimony: psychological perspectives. New York: Cambridge University Press.
Wheeler MA, Stuss DT, Tulving E (1995) Frontal lobe damage produces episodic memory impairment. J Int Neuropsychol Soc 1:525–536.
Wilding EL, Rugg MD (1996) An event-related potential study of recognition memory with and without retrieval of source. Brain 119:889–905.
Wilding EL, Doyle MC, Rugg MD (1995) Recognition memory with and without retrieval of context: an event-related potential study. Neuropsychologia 33:743–767.
Yonelinas AP (1999). The contribution of recollection and familiarity to recognition and source-memory judgements: A formal dual-process model and an analysis of receiver operating characteristics. J Exp Psychol Learn Mem Cognit 25:1415–1434.