Electrophysiologically Dissociating Episodic Preretrieval Processing

Emma K. Bridger and Axel Mecklinger

Abstract

■ Contrasts between ERPs elicited by new items from tests with distinct episodic retrieval requirements index preretrieval processing. Preretrieval operations are thought to facilitate the recovery of task-relevant information because they have been shown to correlate with response accuracy in tasks in which prioritizing the retrieval of this information could be a useful strategy. This claim was tested here by contrasting new item ERPs from two retrieval tasks, each designed to explicitly require the recovery of a different kind of mnemonic information. New item ERPs differed from 400 msec poststimulus, but the distribution of these effects varied markedly, depending upon participants' response accuracy: A protracted posteriorly located effect was present for higher performing participants, whereas an anteriorly distributed effect occurred for lower performing participants. The magnitude of the posterior effect from 400 to 800 msec correlated with response accuracy, supporting the claim that preretrieval processes facilitate the recovery of task-relevant information. Additional contrasts between ERPs from these tasks and an old/new recognition task operating as a relative baseline revealed task-specific effects with nonoverlapping scalp topographies, in line with the assumption that these new item ERP effects reflect qualitatively distinct retrieval operations. Similarities in these effects were also used to reason about preretrieval processes related to the general requirement to recover contextual details. These insights, alongside the distinct pattern of effects for the two accuracy groups, reveal the multifarious nature of preretrieval processing while indicating that only some of these classes of operation are systematically related to response accuracy in recognition memory tasks. ■

INTRODUCTION

Several models of episodic memory retrieval are characterized by the view that there exist multiple stages at which control can determine successful retrieval and subsequent behavior (e.g., Moscovitch & Winocur, 2002; Schacter, Norman, & Koutstaal, 1998). In one model, for example, Burgess and Shallice (1996) outline the separate but complementary contributions made by processes that occur before and after retrieval has occurred. Subject to current retrieval demands, the former specify details that are goal-relevant whereas the latter determine whether memory outputs meet such specifications. Of principal interest in the current report is the influence of those processes that occur before retrieval. The role of these mechanisms is highlighted most often in cases in which they are presumed to be impaired, most commonly in confabulating patients who pathologically retrieve irrelevant or erroneous memory traces (see Metcalf, Langdon, & Coltheart, 2007, for a review). Equally compelling support for the influential role of these processes, however, would come from demonstrating a relationship between the engagement of these processes and episodic memory performance in adults with intact control mechanisms.

Saarland University

One way of measuring preretrieval processes is to measure the neural activity elicited by new items employed in recognition memory tests. It is typically assumed that, because these items have not been presented in a previous study phase, they provide an uncontaminated index of processing engaged before, but not during, successful retrieval (Hornberger, Morcom, & Rugg, 2004; Rugg & Wilding, 2000). Furthermore, if such measures of brain activity are compared across recognition tasks with distinct requirements, they should index the way in which preretrieval processing is specified to the task in hand. This logic has been applied to studies in which ERPs have been employed as indices of brain activity and in which reliable differences between ERPs elicited by new items have been reported (for comparable fMRI contrasts, see Hornberger, Rugg, & Henson, 2006; Woodruff, Uncapher, & Rugg, 2006). Retrieval requirements have been manipulated in various ways, including by contrasting test phases in which study and test formats either do or do not match (Hornberger et al., 2004; Herron & Rugg, 2003), by changing response requirements depending upon the information associated with each item (Dzulkifli & Wilding, 2005; Wilding, 1999), or by manipulating the degree to which contextual details need to be retrieved (Werkle-Bergner, Mecklinger, Kray, Meyer, & Düzel, 2005; Ranganath & Paller, 1999, 2000; Johnson, Kounios, & Nolde, 1997). In line with the variety of manipulations that have been employed and the assumption that these contrasts index processes specific to each retrieval task, there is a considerable degree of heterogeneity in the timing and distribution of these effects across studies. This variety highlights the diverse manner with which retrieval cues can be processed in light of the particular task demands.

In one recent report, Bridger, Herron, Elward, and Wilding (2009) developed the argument that the specificity of these effects across experiments comes about because they reflect operations that relate to the recovery of information relevant to the goals at hand. This assertion was supported by the outcomes of analyses between ERPs elicited by new items from two recognition memory tasks, for which the task relevance of different kinds of mnemonic information was manipulated by a simple change in response requirements. Specifically, responses at retrieval were based on those employed in Jacoby's (1991) exclusion task. In each test phase, new items were intermixed with old items that were associated with one of two study contexts which, in the paradigm reported by Bridger et al. (2009), related to the judgment required for that item at encoding. Participants were only ever required to endorse one class of old items as old (hereafter termed as "targets") depending upon the associated study context and to reject the remainder of old items (nontargets) by responding to them on the same response key as new items. Task requirements were thus manipulated by changing the class of information that required a target response. Reliable differences between new item ERPs from the two tasks were observed over occipital sites from 500 to 1400 msec poststimulus but only for a subgroup of participants whose ability to discriminate between targets and nontargets was above the group median. Notably, across all participants, there was a positive relationship between this behavioral index of response accuracy and the extent to which processing of new items differed between the two recognition tasks (operationalized simply as the magnitude of the difference between new item ERPs at occipital electrodes during the critical time windows). Taken together, these data comprise the first demonstration of a relationship between the engagement of preretrieval processes and increases in behavioral accuracy on a memory task. This is a necessary demonstration if the assumption that processes engaged before retrieval can influence successful retrieval is shown to be true.

These data points were also used to reason about the efficacy of strategies that participants might use to complete these tasks. One approach, for example, is to rely equally on the recovery of both target and nontarget information in each test phase to correctly endorse or reject each class of old item (Jacoby, 1991). An alternative strategy is to judge whether each item is associated with target information and to use this information, alongside the failure to recover such details, to respond appropriately (Herron & Rugg, 2003). Bridger et al. reasoned that, because only the second of these options assumes that retrieval processing should be affected by changing target designation, a systematic relationship between response accuracy and the extent to which retrieval processing differed with target designation implied that the second strategy was more likely to benefit correct responding in that task. The embedded—and critical—assumption here is that this is beneficial because the processes reflected in the ERP contrast increase the likelihood that task-relevant information is retrieved by biasing the retrieval process toward target details in each test phase. In line with this is the observation that, at the behavioral level, it was participants' ability to accurately discriminate between the two classes of information rather than the speed or efficiency with which they responded to items that was related to the engagement of preretrieval processing. Moreover, this pattern suggests that, in tasks requiring binary discriminations of this kind, the benefits associated with prioritizing task-relevant information outweigh those associated with relying on the recovery of more than one type of information (Bridger et al., 2009). The differences between new item ERPs were therefore taken to reflect the engagement of qualitatively distinct retrieval mechanisms associated with maximizing the recovery of different memory contents (in this case, target information) in each task.

It is also possible, however, to accommodate these data within accounts that claim that new item ERP effects reflect a quantitative change in the engagement of the same process in the two tasks rather than the presence of distinct retrieval operations (Bridger et al., 2009). Recourse to one account of this kind often occurs when one task appears to be more difficult than another on the basis of the associated behavioral performance, allowing for the prospect that any changes in retrieval processing reflect a simple increase in the effort with which the same retrieval resources are engaged (Rosburg, Mecklinger, & Johansson, 2011; Rugg & Wilding, 2000). This particular possibility cannot be entirely excluded when interpreting the data reported by Bridger and colleagues because of a consistent trend toward higher accuracy in one of the two exclusion tests in that study. Moreover, in the case that there were no behavioral indicators of changes in task difficulty, the possibility that any ERP differences might simply reflect the engagement of controlled preretrieval processing in one task but not the other would nonetheless remain. Accounts of this kind persist so long as contrasts are limited to simple pairwise comparisons because these cannot in and of themselves adjudicate between quantitative and qualitative accounts of retrieval cue processing. With this in mind, the experiment described here was designed to provide an explicit test of the claim that pre-retrieval operations—as indexed by new item ERP effects—index distinct retrieval operations which facilitate the recovery of task-relevant information rather than reflect the differential engagement of the same process.

The current study contained a number of measures designed to enable this. First, recognition memory tasks were employed which (i) extensive piloting ensured were comparably difficult (to circumvent potential retrieval effort explanations) and (ii) were explicitly designed to encourage the recovery of one class of information in each test phase. At study, participants were required to make one of two judgments on items (in this instance, judging the pleasantness of the item or how easy it would be to draw). For a third class of items, they made both these judgments sequentially. In subsequent test phases, all old items were intermixed with new items, and participants made a yes/no response for each item. In one test phase, participants were to respond yes if they had previously performed a pleasantness judgment on the item at study and no if not. In the remaining test phase, a comparable yes/no judgment was required, addressing whether items had previously been encountered in the drawing judgment task. The critical modification from the exclusion task, therefore, was the addition of items associated with both study contexts. These items were included to reduce the degree to which it is useful to recover nontargeted information in each test phase because the recovery of this information alone is no longer sufficiently indicative of a "no" or reject response. Correct responding nonetheless remains contingent upon the recollection of target information in each test phase, making prioritizing the recovery of target details in each test phase an especially valid strategy for completing this task. This emphasis on one class of diagnostic information in each test is reflected in previous reports of this kind of task (see Gallo, Weiss, & Schacter, 2004, for the earliest such report), which refer to this as the criterial recollection task. If new item ERPs reflect preretrieval processes that relate to the recovery of particular classes of information then differences between new item ERPs from the two criterial recollection tasks employed here should replicate the pattern reported by Bridger et al. and show a relationship between the engagement of these processes and the likelihood of responding correctly in these tasks.

A second feature of the current design was included to address the inability of pairwise contrasts to arbitrate between accounts that state that these effects reflect the engagement of qualitatively distinct processes and those that state that they reflect a change in the engagement of the same process. Participants completed a third recognition task in which the same types of studied information were present but for which only simple old/new judgments were required. This task thus provided a class of new items for which processing was assumed to be relatively neutral, without an emphasis on the recovery of a particular class of information, and to which ERPs from the criterial tasks could be separately contrasted. These separate contrasts are assumed to provide indices of genuine task-specific retrieval effects, because they should reflect processes engaged exclusively toward the recovery of information associated with pleasantness or drawing judgments in each case. Evidence that these task-specific effects differ significantly in their scalp topographies would provide convergent support for the claim that participants engage in distinct retrieval processes in each test phase rather than change the degree to which they engage in the same preretrieval process. These contrasts were also presumed to provide insight into a complementary aspect of preretrieval processing by highlighting the general manner with which processing in the criterial tasks differed from the old/new recognition task. In line with the assumption that the requirement to prioritize the recovery of particular contextual details is common to the criterial but not to the old/new recognition task, commonalities in the time course and scalp distribution of the differences between ERPs from these tasks should also shed light on the nature of general preretrieval control processes engaged in these tasks.

METHODS

Participants

Thirty-two native German speakers (13 men) were recruited from the student population of Saarland University. Participants were compensated €8/hr or in course credit for their time. All were right-handed and reported no diagnosis of dyslexia or neurological problems. The mean age of these participants was 23 years (range = 18–28 years). A further seven students (five men) also participated but were excluded from the final analyses, because they either failed to follow instructions in one of the test phases (2) or were unable to successfully discriminate between target and nontarget items as indicated by a discrimination value (p[target hit] − p[nontarget false alarm]) at or below zero.

Materials and Design

Four hundred eighty German nouns with a frequency of 1–60 occurrences per million and a word length of four to nine letters were taken from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). All words were independently verified as concrete by 39 native German speakers. Raters made two further judgments about each word: Using a scale of 1–4, they were required to determine (i) the ease with which the item depicted by the word could be drawn and (ii) the pleasantness of the item. The modes derived from these ratings along with frequency per million were used to allocate words to eight sets of 60 words matched for ease of drawing, pleasantness, and frequency.

The experiment was split into two separate study–test cycles. Each cycle began with a study phase in which participants were presented with 150 words. For 55 of these words, participants were required to rate how difficult it would be to draw the object denoted by each word and for a further 55 words, participants were required to determine how pleasant the object denoted by the word was. The remaining 40 words were presented twice (in succession), and participants were required to complete both tasks for these words. After study, participants performed two criterial recollection tasks and an old/new recognition task. In each criterial recollection task, 20 pleasant items, 20 drawing items, and 20 items from both tasks were randomly intermixed with 30 new items, and participants were required to make a yes/no response to each item. In the pleasant criterial task, this decision depended upon whether the item had been encoded in the pleasantness task at study, whereas in the Drawing criterial task, the decision referred to whether each item had been in the drawing judgment task at study. In the old/new recognition task, 15 pleasant items and 15 drawing items were intermixed with 30 new items, and participants' yes/no judgments were made on the basis of whether each item was old or new.

The unequal number of old and new items across the three tasks precluded a complete counterbalance of items across all experimental conditions. Instead, for each study–test cycle, four sets of 60 words (matched for frequency, ease of drawing, and pleasantness ratings) were allocated to the following groups: (a) 60 old items in the pleasant criterial task, (b) 60 old items in the Drawing criterial task, (c) 60 new items for both criterial recollection tasks, and (d) all 60 (old and new) items in the old/new recognition task. Table 1 depicts the experiment conditions that comprise these groups. Rotating word sets across these four groups as well as three times within each set created 12 counterbalanced lists that ensured all items served at least once in all the possible study conditions for the two criterial recollection tests and that each set of 60 was encountered in each of the three retrieval tasks.

Procedure

Participants were initially fitted with an EEG cap (see recording parameters below), a process which took approximately 1 hr. The experiment began once participants had completed a short practice phase in which they were familiarized with the task requirements in both the study and test phases. All instructions were provided both verbally and on the computer screen. Participants were asked to repeat each set of instructions for the three memory tests back to the experimenter to ensure they understood the separate task requirements.

In each study phase, cues immediately preceding each word specified which judgment to make on the subsequent word ("ANGENEHM?" for the pleasantness judgment and "ZEICHNEN?" for the judgment task). Responses were made on a computer keyboard with the following four-point scales; "very pleasant, pleasant, unpleasant, very unpleasant" and "very easy to draw, easy to draw, difficult to draw, very difficult to draw." Task cues were presented for 600 msec, followed by a blank screen for 500 msec and then the word for 300 msec. The screen remained blank until the participant made their response and the next task cue appeared 1250 msec later. The order in which these tasks were encountered was pseudorandomized, such that the same task was never encountered more than three times in succession. For 50% of those items for which participants performed both tasks, the drawing judgment was made first. Participants received a self-terminated break midway through each study phase.

Test trials began with a fixation crosshair for 500 msec followed by a blank screen for 500 msec. Words were presented for 300 msec before a blank screen was presented until the participant responded. The next trial began 1000 msec after a response was made. Participants made yes/no responses depending upon the particular task instructions (see Table 1). In keeping with the nomenclature employed in standard exclusion tasks, items which had only been presented once at study are hereafter referred to as targets if participants were required to respond "yes" to them and nontargets if they were to respond "no." In each study–test cycle participants always performed the two criterial recollection tasks before the old/new recognition task, but the order in which the pleasant and drawing criterial tasks were encountered was balanced across study–test cycles. Hand-to-response mappings remained constant within each experiment run and half of the participants responded "yes" with the left index finger and "no" with the right index finger.

Table 1. Experimental Conditions Generated by the Three Retrieval Tasks (Pleasant, Drawing, and Old/New) and Four Study Conditions (Pleasant, Drawing, Both, and New)

	Retrieval Task		
	Pleasant	*Drawing*	*Old/New*
Study	**Pleasant**	Pleasant	**Pleasant**
	Drawing	**Drawing**	**Drawing**
	Both	**Both**	–
	New	New	New

Items to which participants were required to respond "yes" in each retrieval task are highlighted in **bold**.

Electrophysiological Recording Parameters and Data Processing

Continuous EEG was recorded from 58 scalp locations based on the extended international 10–20 system (Jasper, 1958). EEG was acquired referenced to the left mastoid and rereferenced off-line to the average of the mastoid signals. EEG signals were band-pass filtered from DC-250 Hz and digitized at a sampling rate of 500 Hz. EOG activity was assessed using signals recorded from four additional electrodes above and below the right eye (vertical EOG) and on the outer canthi (horizontal EOG). Electrode impedances were kept below 5 kΩ. Off-line, a digital bandpass filter (0.03–500 Hz) was applied, and epochs were created beginning 200 msec before and ending 1600 msec after the onset of stimulus presentation. Waveforms were corrected relative to a 200-msec prestimulus baseline period. EOG blink and movement artifacts were corrected using the modified linear regression algorithm (Gratton, Coles, & Donchin, 1983) embedded in the EEProbe Software package. ERP analyses were conducted only on items to which a correct response was given. For illustrative purposes only, a 12-Hz low-pass filter has been applied to all ERP waveforms depicted in the figures.

RESULTS

Behavior

Study

Mean RTs for responding at study were 1214 msec (SD = 223 msec) for pleasantness judgments and 1226 msec (SD = 230 msec) for drawing judgments. Response times did not differ for the two encoding tasks ($p > .29$).

Test

Table 2 displays the likelihoods of a correct response in each criterial recollection task and each word category along with corresponding mean RTs. These data followed the pattern typically reported in this task: Accuracy was greatest for correctly rejected new items, which were more likely than correct responses to both items ($t(1, 31) = 10.53$, $p < .001$), which were more likely than target hits ($t(1, 31) = 6.82$, $p < .001$). These, in turn, were more likely than correctly rejected nontargets ($t(1, 31) = 2.10$, $p < .05$). As a necessary step in determining how ERP new item effects relate to accuracy and in line with the approach reported by Bridger et al. (2009), participants were allocated to a high- or low-accuracy group according to their performance in the two criterial tests.

Table 2. Mean Proportions of Correct Responses along with Associated Mean RTs for Each Word Category in the Two Criterial Recollection Tasks ($n = 32$)

	Word Status			
Retrieval Task	*Pleasant*	*Drawing*	*Both*	*New*
Pleasant				
p(correct)	.71 (.10)	.66 (.15)	.79 (.09)	.94 (.06)
RT (msec)	976 (194)	1016 (213)	926 (182)	829 (145)
Drawing				
p(correct)	.65 (.15)	.70 (.12)	.81 (.10)	.95 (.06)
RT (msec)	1022 (274)	964 (201)	937 (187)	816 (136)

*SD*s are in parentheses.

This was operationalized using an index of each participant's ability to discriminate between targets and nontargets (p[target hit] − p[nontarget false alarm]) collapsed across the two criterial tasks. Individuals were then split into two groups, depending upon whether their discrimination score fell above or below the median (.38). The mean target/nontarget discrimination scores were .52 and .20 for the subsequent high and low groups, respectively. When submitted to an ANOVA with factors of retrieval task (pleasant, drawing) and accuracy group (high, low), these data revealed a main effect of accuracy group ($F(1, 30) = 78.62$, $p < .001$) but no effect of retrieval task ($F(1, 30) = 0.03$, $p = .87$) or interaction ($F(1, 30) = 1.26$, $p = .27$).

Table 3 shows the probabilities of correct responses along with associated RTs for both accuracy groups and all three retrieval tasks. A global ANOVA with factors of retrieval task (pleasant, drawing), word category (target, nontarget, both, and new) and accuracy group (high, low) revealed no main effects or interactions with the retrieval task factor (all relevant p values > .37). Main effects of word category ($F(3, 90) = 90.45$, $p < .001$) and accuracy group ($F(1, 30) = 69.44$, $p < .001$) were moderated by an interaction ($F(3, 90) = 4.67$, $p < .05$). Bonferroni-corrected t tests (adjusted α level: 0.05/12 = 0.004) indicated that, in both groups, correct responses to old items were less likely than to new items (all ps < .004). In the high-accuracy group, accuracy for nontargets did not differ significantly from that for targets or "both" items (both $p > .004$) but was higher for "both" items than targets ($p < .004$). The pattern was comparable in the low group, except that correct responses were also significantly more likely to a "both" item than to a nontarget ($p < .001$). In line with the critical comparisons between new item ERPs across retrieval tasks and accuracy group, mean probabilities of a correct response to a new item in the three retrieval tasks were submitted to an ANOVA with factors of retrieval task (pleasant, drawing, old/new) and accuracy group (high, low). Correct rejections were more likely in the high group ($F(1, 30) = 7.98$, $p < .01$). There was also a main effect of retrieval task ($F(2, 60) = 42.37$, $p < .001$) because the likelihood of correctly rejecting a new item was greater in the criterial than the old/new recognition task (pleasant, $t(31) = 7.27$, $p < .001$; drawing, $t(31) = 6.89$, $p < .001$) but did not differ for the two criterial tasks ($t < 1$).[1] An ANOVA for old responses in the old/new recognition task with factors of accuracy group and word category (pleasant, drawing) revealed that the likelihood of making a correct old response was greater in the high-accuracy group ($F(1, 30) = 14.60$, $p < .01$).

Mean response times for the criterial tasks were submitted to the same global ANOVA employed for response probabilities and revealed a main effect of word category ($F(3, 90) = 52.17$, $p < .001$), which was moderated by an interaction with accuracy group ($F(3, 90) = 6.09$, $p < .01$). Bonferroni-corrected pairwise t tests (α level = 0.004) revealed that correct rejections were faster than

Table 3. Mean Proportions of Correct Responses along with Associated Mean RTs for Each Word Category in the Criterial and Old/New Recognition Tasks

			Word Category			
Retrieval Task	*Accuracy Group*		*Target*	*Nontarget*	*Both*	*New*
Pleasant	High	*p*(correct)	.77 (.08)	.74 (.12)	.82 (.08)	.97 (.02)
		RT (msec)	1007 (198)	1081 (216)	961 (189)	836 (160)
	Low	*p*(correct)	.65 (.09)	.58 (.13)	.75 (.10)	.91 (.07)
		RT (msec)	944 (191)	952 (195)	891 (173)	822 (133)
Drawing	High	*p*(correct)	.78 (.10)	.75 (0.10)	.85 (.07)	.97 (.05)
		RT (msec)	1000 (195)	1120 (307)	978 (219)	835 (160)
	Low	*p*(correct)	.63 (.08)	.55 (.11)	.76 (.10)	.93 (.07)
		RT (msec)	929 (207)	924 (200)	895 (145)	796 (110)
			Pleasant	*Drawing*		*New*
Old/New	High	*p*(correct)	.90 (.08)	.86 (.08)		.91 (.05)
		RT (msec)	789 (142)	808 (139)		850 (170)
	Low	*p*(correct)	.77 (.09)	.77 (.13)		.83 (.12)
		RT (msec)	766 (122)	773 (122)		816 (128)

The data are separated for the two accuracy groups. *SD*s are shown in parentheses.

all other responses (all *p*s < .004) in the low-accuracy group but that response times did not differ for the classes of old items. In the high-accuracy group, correct rejections were also faster than responses to all other items (all *p*s < .001), but responses to nontargets were reliably longer than responses to targets ($p < .001$) and "both" items ($p < .001$). Speed of responding did not differ between targets and "both" items in the high-accuracy group. A planned ANOVA with factors of retrieval task (pleasant, drawing, old/new) and accuracy group (high, low) revealed no differences in response times for correct rejections from the three retrieval tasks.

In summary, participants in the high-accuracy group were reliably more likely to make a correct response than those in the low-accuracy group. More accurate responders were also slower to respond to nontargets. Critically, across all participants, there was no evidence of changes in difficulty for the two criterial tasks. Participants were, however, less likely to correctly reject new items in the old/new recognition task.

ERP Analyses

All analyses of ERP data are restricted to contrasts between ERPs elicited by correctly responding to new items from the three retrieval tasks. To select time windows and electrode sites that encapsulate real differences between correctly rejected new item ERPs, the following procedure was followed. The mean amplitudes of ERPs for the three retrieval tasks were quantified for 100 msec bins from 200 to 1600 msec after stimulus presentation and were used to conduct separate paired *t* tests (pleasant vs. drawing, pleasant vs. old/new and drawing vs. old/new) in each time window at each electrode. *p* Values of < 05 but > .01 were regarded as significant only if neighboring electrodes and/or neighboring time windows showed a significant condition effect at $p < .02$ (see Rosburg et al., 2011, for a comparable approach). The outcomes of these analyses were used to guide the selection of time windows and electrode grids that best encompassed the majority of ERP effects (from all three contrasts) within the following constraints: Time windows were to comprise epochs of equal duration, and electrode grids were to sample equally from left, midline, and right hemisphere sites and anterior, central, and posterior sites. Mean amplitude data from three time windows (400–800, 800–1200, and 1200–1600 msec) and a 3 × 3 electrode grid comprising three frontal (F5, Fz, F6), three central (C5, Cz, C6), and three parietal (P5, Pz, P6) electrodes were chosen on the basis of the outcomes of these contrasts.

In line with the primary hypothesis, analyses began with paired comparisons between new item ERPs from the two criterial tasks. To determine whether effects were influenced by accuracy group, initial ANOVAs including factors of accuracy group (high, low), retrieval task (pleasant, drawing), anterior/posterior (frontal, central, parietal), and laterality (left, midline, right) were conducted in each epoch. The Geisser–Greenhouse correction for violations of sphericity (Greenhouse & Geisser, 1959) was employed where necessary. Only significant interactions including the retrieval task factor are of concern and these are shown for each time window in Table 4. Where accuracy group interacted with retrieval task, subsequent analyses were conducted separately for the two groups, and to anticipate, these revealed distinct new item ERP effects in each group. To determine whether either of these effects was systematically related to response accuracy, correlational analyses between the mean amplitudes of these effects and response accuracy were conducted at the electrode locations and time points where between-group differences were present.

The next set of analyses comprised comparisons between ERPs from the old/new recognition task and those from each criterial task. In a first step, these were averaged across the two criterial tasks to highlight the overall pattern with which ERPs from the old/new recognition and criterial tasks differ from each other. ANOVAs for these analyses employed the same factors as the preceding global ANOVA, except that the retrieval task factor was updated (criterial, old/new) and the outcomes of these contrasts (shown in Table 4) were deconstructed using the same approach as for the criterial pairwise contrasts. Table 4 also shows the outcomes of the comparisons between new item ERPs from the separate criterial tasks and the old/new recognition task, which reveal changes in the pattern of interactions with location factors for the two contrasts. In a next step, the topographies of these task-specific pairwise comparisons were directly contrasted against one another in each time window to determine whether the variations in retrieval task × location factors for the two contrasts reflect robust differences in scalp topography. Demonstrating changes in the scalp distribution of the two criterial task-specific contrasts would support the claim that not entirely the same cognitive processes were engaged in the two cri-

Table 4. Outcomes of Global ANOVAs for Pairwise Comparisons in Each Time Window

	Epoch		
	400–800 msec	*800–1200 msec*	*1200–1600 msec*
Pleasant vs. Drawing			
Task × Group	^{1,30}7.64*		
Task × Anterior/Posterior	$^{1.3,39.1}$15.17**	$^{1.4,43.1}$13.59**	$^{1.4,43.2}$4.59*
Task × Laterality × Group	$^{1.78,53.4}$4.48*	$^{1.9,57.0}$3.42*	$^{1.8,53.1}$4.20*
Criterial vs. Old/New			
Task			^{1,30}6.87*
Task × Anterior/Posterior	$^{1.4,40.7}$10.46**	$^{1.5,44.7}$6.80**	
Task × Laterality	$^{2,58.8}$7.16**		
Task × Anterior/Posterior × Laterality	$^{3.5,104.7}$3.25*		
Pleasant vs. Old/New Recognition			
Task			^{1,30}5.48*
Task × Anterior/Posterior	$^{1.3,38.0}$16.20**	$^{1.4,42.5}$11.32**	
Task × Laterality	$^{2,58.9}$5.05*		
Task × Anterior/Posterior × Laterality	$^{3.5,105.7}$3.16*		
Drawing vs. Old/New Recognition			
Task			^{1,30}5.61*
Task × Anterior/Posterior	$^{1.5,46.3}$5.51*	$^{1.6,49.8}$5.02*	$^{1.6,47.2}$4.49*
Task × Laterality	$^{2,61.8}$6.98**		

Geisser–Greenhouse corrected degrees of freedom for deviations from sphericity are shown in superscript.

*$p < .05$.

**$p < .01$.

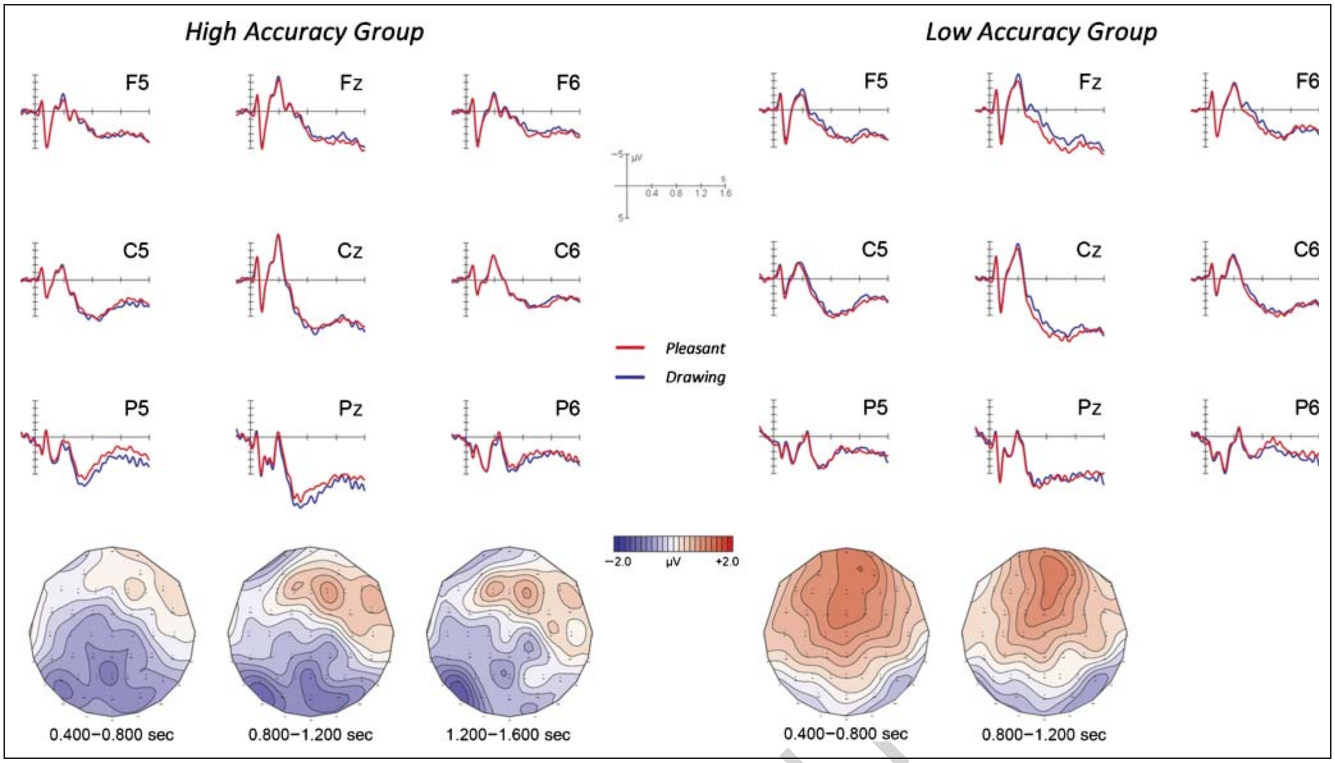

Figure 1. Top: Grand-averaged ERP waveforms elicited by new items in the two criterial recollection tasks for the high-accuracy (left) and low-accuracy (right) groups. Data are shown for the nine electrode locations over frontal (F5, Fz, F6), central (C5, Cz, C6), and posterior (P5, Pz, P6) scalp sites used in all statistical analyses. Bottom: Topographic maps showing the scalp distributions of the differences between neural activity elicited by new items for the high-accuracy (left) and low-accuracy (right) groups. Topographic maps are computed on the basis of the difference scores obtained by subtracting mean amplitudes for ERPs elicited by new items in the drawing criterial task from those in the pleasant criterial task and are shown for those time windows in which there was a reliable interaction between retrieval task and factors of scalp site in each group.

terial tasks (Wilding, 2006). Task-specific subtractions were thus extracted (pleasant minus old/new and drawing minus old/new) for each epoch and subjected to ANOVAs with factors of retrieval task subtraction (pleasant, drawing), anterior/posterior (frontal, central, parietal), and laterality (left, midline, right).

The additive nature of ANOVAs means that it is possible to obtain reliable interactions between condition and location factors simply because the engagement of activity is significantly greater in one condition rather than because there is a change in the way in which activity is distributed over the scalp (Wilding, 2006; McCarthy & Wood, 1985). The recommended approach to overcoming this problem and ensuring such outcomes reflect topographical differences is to determine whether reliable interactions between factors of condition and scalp location remain when amplitude differences are removed (Picton et al., 2000). Thus, where task-specific contrasts interacted with scalp location, follow-up analyses were conducted on data rescaled using the vector length method and submitted to ANOVAs with factors of retrieval task subtraction (pleasant, drawing) and electrode site (9) in each time window. Reliable interactions of this kind would thus confirm the presence of distinct scalp topographies in each time window. Comparable analyses were also used to determine whether the scalp distribution of differences between new item ERPs (pleasant vs. drawing; criterial vs. old/new; pleasant subtraction vs. drawing subtraction) change over time. These time course analyses comprised pairwise comparisons between rescaled subtraction data from adjacent epochs and employed factors of epoch (2) and electrode site (9). The mean trial numbers (range in brackets) forming individual subjects' average ERPs for the high-accuracy group were 51 (37–59), 50 (28–60), and 47 (30–57) for the pleasant, drawing, and old/new recognition tasks, respectively. The means for the same conditions in the low-accuracy group were 48 (35–57), 50 (42–58), and 44 (31–58).

Criterial Task New Item ERPs

Figure 1 shows the grand-averaged waveforms for ERPs elicited by new items in the pleasant and drawing criterial recollection tasks for the high- (left) and low-accuracy (right) groups. The figure also contains topographic maps depicting the scalp distribution of the differences between correct rejections from the two test phases for each of the two groups. The left panel shows these data for the high-accuracy group, in which pleasant new item ERPs are more negative than drawing new item ERPs over left posterior sites, from approximately 500 msec poststimulus until the end of the recording epoch. For the low-accuracy group, however, depicted in the right panel, pleasant new item ERPs are relatively more positive than drawing new item ERPs from 500 to 1300 msec at anterior and central sites.

Table 4 shows reliable interactions, including the factors of accuracy group and retrieval task in each time window, prompting follow-up contrasts within each group. In the 400–800 msec time window, retrieval task interacted with anterior/posterior for both the high- ($F(1.29, 19.33) = 7.68$, $p < .01$) and low-accuracy group ($F(1.31, 19.61) = 7.88$, $p < .01$). Analyses at each level of anterior/posterior for the high-accuracy group revealed a main effect of retrieval task at posterior sites only ($F(1, 15) = 11.35$, $p < .01$). These analyses in the low-accuracy group revealed a main effect of retrieval task at anterior ($F(1, 15) = 6.10$, $p < .05$) and central sites ($F(1, 15) = 6.12$, $p < .05$). Although there were no interactions with laterality in this time window for either group, analyses at each level of this factor were included to endeavor to break down the initial interaction between accuracy group, retrieval task, and laterality presented in Table 4. In the high-accuracy group, this revealed interactions between retrieval task and anterior/posterior at all three levels of laterality (all Fs(<1.61, <24.11) > 4.76, $p < .05$). In the low-accuracy group, there were main effects of retrieval task at left and midline sites (all Fs$(1, 15) > 4.72$, $p < .05$) and interactions between retrieval task and anterior/posterior at all three levels of laterality ($F(<1.9, <28.48) > 4.66$, $p < .05$). The initial interaction with accuracy group, retrieval task, and laterality is likely to reflect the reversal in polarity of the two effects, a difference which was most robust over left hemisphere and midline sites.

From 800 to 1200 msec, retrieval task interacted with anterior/posterior in the high- ($F(1.29, 18.43) = 4.42$, $p < .05$) and low-accuracy group ($F(1.59, 23.84) = 9.62$, $p < .01$). The anterior/posterior interaction breakdown in the high-accuracy group revealed a main effect of retrieval task at posterior sites ($F(1, 15) = 7.63$, $p < .05$) as well as interactions between retrieval task and laterality at anterior ($F(1.72, 25.83) = 3.93$, $p < .05$) and central sites ($F(1.78, 26.66) = 3.58$, $p < .05$). These latter interactions reflect the relatively greater difference between new item ERPs over left hemisphere sites for this group. For the low-accuracy group, there was a main effect of retrieval task at anterior sites only ($F(1, 15) = 4.91$, $p < .05$). In the final 1200–1600 msec time window, retrieval task interacted with laterality for the high group ($F(1.7, 25.55) = 4.97$, $p < .05$), whereas there were no significant effects for the low group. Analysis at each level of laterality for the high group revealed a main effect of retrieval task over the left hemisphere sites only ($F(1, 15) = 7.36$, $p < .05$). Combined, the outcomes of these analyses confirm the impression provided by Figure 1 that new item ERPs show a temporally protracted left posterior difference in the high-accuracy group beginning around 400 msec but an anteriorly distributed effect with a reversed polarity in the low-accuracy group.

The preceding analyses reveal between-group differences in the way in which new item ERPs were processed according to the specific retrieval requirements. If these changes in retrieval processing are systematically related to the accuracy of retrieval judgments, either positively or negatively, then it should also be possible to observe a relationship between response accuracy and the engagement of these processes at the individual subject level. To investigate this, subtraction data were computed from the nine electrode sites (F5, Fz, F6, C5, Cz, C6, P5, Pz, and P6) in the three time windows used in all preceding analyses. Criterial recollection subtraction amplitudes were calculated by subtracting the mean amplitude of ERPs from the pleasant task from the mean amplitude of drawing ERPs. For all participants, these values were plotted against discrimination scores (p[target hit] − p[nontarget false alarm]) averaged across the two criterial recollection tasks. Across all participants, significant positive correlations were obtained in the 400–800 msec time window at electrodes P5 ($n = 32$, $r = 0.36$, $p < .05$), Pz ($n = 32$, $r = 0.52$, $p < .01$), and Cz ($n = 32$, $r = 0.49$, $p < .05$). Figure 2 depicts scatterplots of these relationships at electrode Pz, where the effect was largest. There were no significant correlations at anterior sites from 400 to 800 msec or at any location in the latter two epochs. The significant correlations in the early time window indicate that response accuracy increased with the extent to which the new item ERPs over posterior sites differed from one another from 400 to 800 msec.

Criterial versus Old/New Recognition Task New Item ERPs

Figure 2 shows, for all participants, the grand-averaged waveforms for ERPs elicited by new items in the old/new recognition task, plotted separately against ERPs from the two criterial tasks. These contrasts are assumed to reflect processes that relate to the recovery of contextual details associated with the pleasantness task in one case and drawing-related details in the other. From approximately 500 msec, items from both criterial recollection tasks are more negative than old/new recognition new item ERPs. Although the distribution of both these effects is maximal over central sites, this effect also extends over posterior and right hemisphere sites for pleasant ERPs. From 800 msec onward, criterial task ERPs become more positive relative to old/new recognition task ERPs. From 1200 msec, for both contrasts, this relative positivity occurs over both frontal and posterior sites. In the preceding 800–1200 time window, however, only ERPs from the pleasant task are relatively more positive than old/new recognition ERPs, and this effect is most pronounced at frontal sites.

Table 4 shows the outcomes of the contrasts between criterial task and old/new recognition ERPs both when collapsed across and when separated for the two criterial tasks. There were no interactions with accuracy group

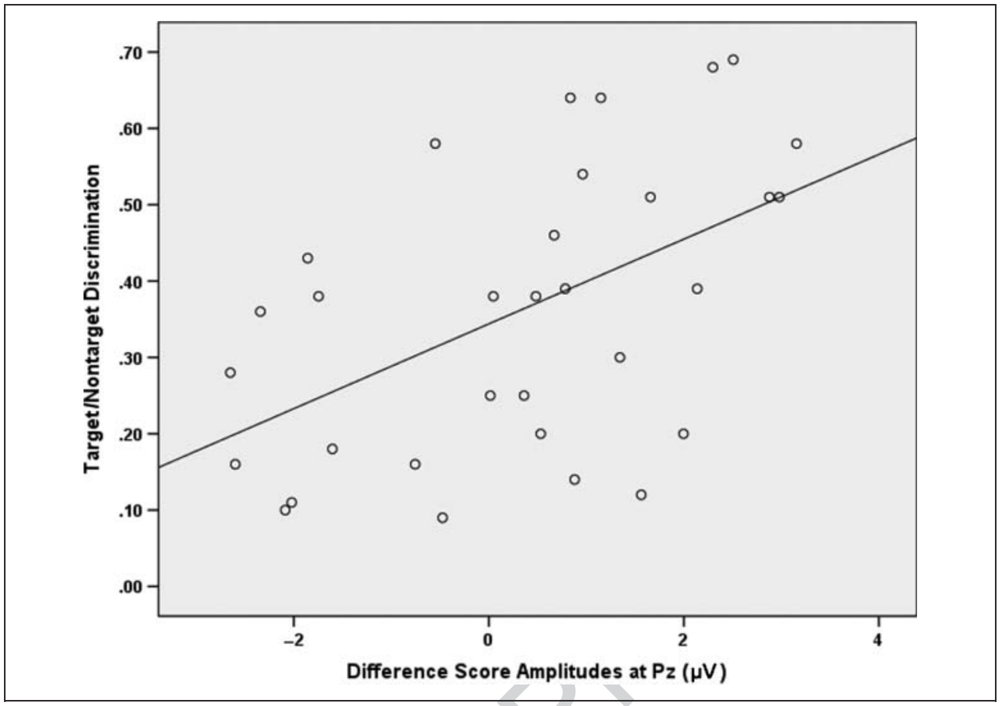

Figure 2. Scatterplot showing the relationship between target/nontarget discrimination and difference score amplitudes (drawing ERPs minus pleasant ERPs) at electrode Pz in the 400–800 msec window for all 32 participants.

for any of these contrasts and subsequent subtraction and rescaled analyses are thus collapsed across this factor. To characterize the general pattern with which criterial task ERPs (collapsed across the two tasks) differed from old/new recognition ERPs, follow-up analyses were conducted at each level of anterior/posterior in the 400–800 and 800–1200 msec epochs. In the 400–800 msec window, these revealed main effects of retrieval task at central ($F(1, 31) = 5.83$, $p < .05$) and posterior ($F(1, 31) = 4.56$, $p < .05$) sites and retrieval task by laterality interactions at frontal ($F(1.9, 59.1) = 4.00$, $p < .05$), central ($F(1.9, 60.4) = 9.06$, $p < .001$), and posterior ($F(1.7, 54) = 4.79$, $p < .05$) sites. These interactions with laterality came about because effects were present predominantly over the right hemisphere and the midline for the central and posterior analyses, whereas the effect was greatest over the left hemisphere at anterior sites. Comparable follow-up contrasts in the 800–1200 msec revealed a main effect of retrieval task at anterior sites only ($F(1, 31) = 5.30$, $p < .05$). Combined, these analyses indicate a change in the pattern of differences between criterial and old/new recognition task ERPs over time; whereas from 400-800 msec, criterial task ERPs elicit a relative negativity over central sites, they demonstrate an anterior positivity from 800 msec onward, which becomes more diffusely distributed across the scalp from 1200 msec.

The preceding analyses reveal the overall manner with which ERPs in the criterial task diverge from those in the old/new recognition task. Table 4 also reports the outcomes for these contrasts separated for the two criterial tasks. To determine whether the topographies of these effects differ from one another, the distributions of the two effects were directly compared with one another via a series of subtraction analyses. In the 400–800 msec window, retrieval task subtraction interacted with anterior/posterior ($F(1.31, 40.74) = 15.45$, $p < .05$) in line with the shift in distribution of the relative negativity of the two effects in this time window; whereas this was greatest at central sites for the drawing contrast, it was most negative at posterior sites for the pleasant contrast for which it was also relatively positive at anterior sites. Retrieval task also interacted with anterior/posterior in the subsequent 800–1200 msec time window ($F(1.44, 44.71) = 13.74$, $p < .001$). Whereas both effects demonstrated a relative positivity in this time window, the distribution was maximal at anterior sites for the pleasant task-specific contrast. The same interaction term was also significant in the final 1200–1600 msec time window ($F(1.47, 45.5) = 4.67$, $p < .05$), because of the anterior maximum of the pleasant task-specific contrast compared with the posterior drawing task-specific contrast (see Figure 3).

Topographical Analyses

Rescaled task-specific subtraction comparisons revealed interactions between retrieval task and electrode site, confirming that the effects were associated with reliably distinct scalp topographies in all three time windows (400–800, $F(3.9, 120.82) = 5.29$, $p < .01$; 800–1200, $F(4.47, 138.55) = 4.70$, $p < .01$; 1200–1600, $F(4.2, 130.15) = 2.69$, $p < .05$). To determine whether the distributions of each of the two task-specific subtraction contrasts changed with time, the same rescaled data were submitted to pairwise comparisons between adjacent time windows. Reliable Epoch × Electrode location interactions were revealed for the pleasant contrast when the 400–800 and 800–1200 window ($F(3.83, 118.62) = 4.12$, $p < .01$) and the 800–1200 and 1200–1600 epochs ($F(3.7, 114.54) = 5.34$, $p < .01$) were compared. These same interaction terms were marginal for the drawing task-specific contrasts (400–800 vs. 800–1200, $F(4.04, 125.23) = 2.32$, $p = .06$; 800–1200 vs. 1200–1600, $F(4.02, 124.68) = 2.38$, $p = .06$), and significant when collapsed across criterial task (criterial vs. old/new; 400–800 vs. 800–1200, $F(3.79, 117.58) = 3.27$, $p < .05$; 800–1200 vs. 1200–1600, $F(3.67, 113.61) = 3.92$, $p < .05$).

To determine whether or how the scalp distributions of the differences between criterial task ERPs changed over time, analyses were conducted on rescaled subtraction data (drawing minus pleasant), calculated separately for the two groups. For the high-accuracy group, there were no interactions between retrieval task and electrode site for any of the comparisons between adjacent time windows (all ps $> .16$). Similarly, there was no evidence of reliable changes in scalp distribution over the two time windows in which criterial task ERPs differed from one another for the low-accuracy group (400–800 vs. 800–1200 msec time windows; $p = .57$). Thus, whereas there was no evidence that the scalp distribution of differences between criterial task new item ERPs changed over time for either accuracy group, this was not the case for the differences between criterial and old/new recognition new item ERPs, the pattern of which changed over successive epochs.

Within-group Task-specific Analyses

In a final analysis, task-specific subtraction (criterial minus old/new) contrasts were conducted separately for the high- and low-accuracy groups for the 400–800 msec time window to provide support for the critical assumption that higher performing participants engaged distinct retrieval processing in each criterial test in this epoch. Figure 4 shows the contrasts in this time window for the high- and low-accuracy group. Whereas the broad negativity of the effects in the high group extends over left posterior sites for the pleasant effect it demonstrates a right-sided and more anterior distribution for the drawing effect. The differences in distribution are more marked for the low performing group; whereas the central negativity is again evident for the drawing effect, the pleasant effect was characterized by a prefrontally distributed positivity for this group. Demonstrating that topographic differences remain for the high-accuracy group when low performers are removed is necessary in line with the interpretation that it is the engagement of distinct preretrieval operations in this group, which is related to an increase in response accuracy. In line with this critical assumption, retrieval task

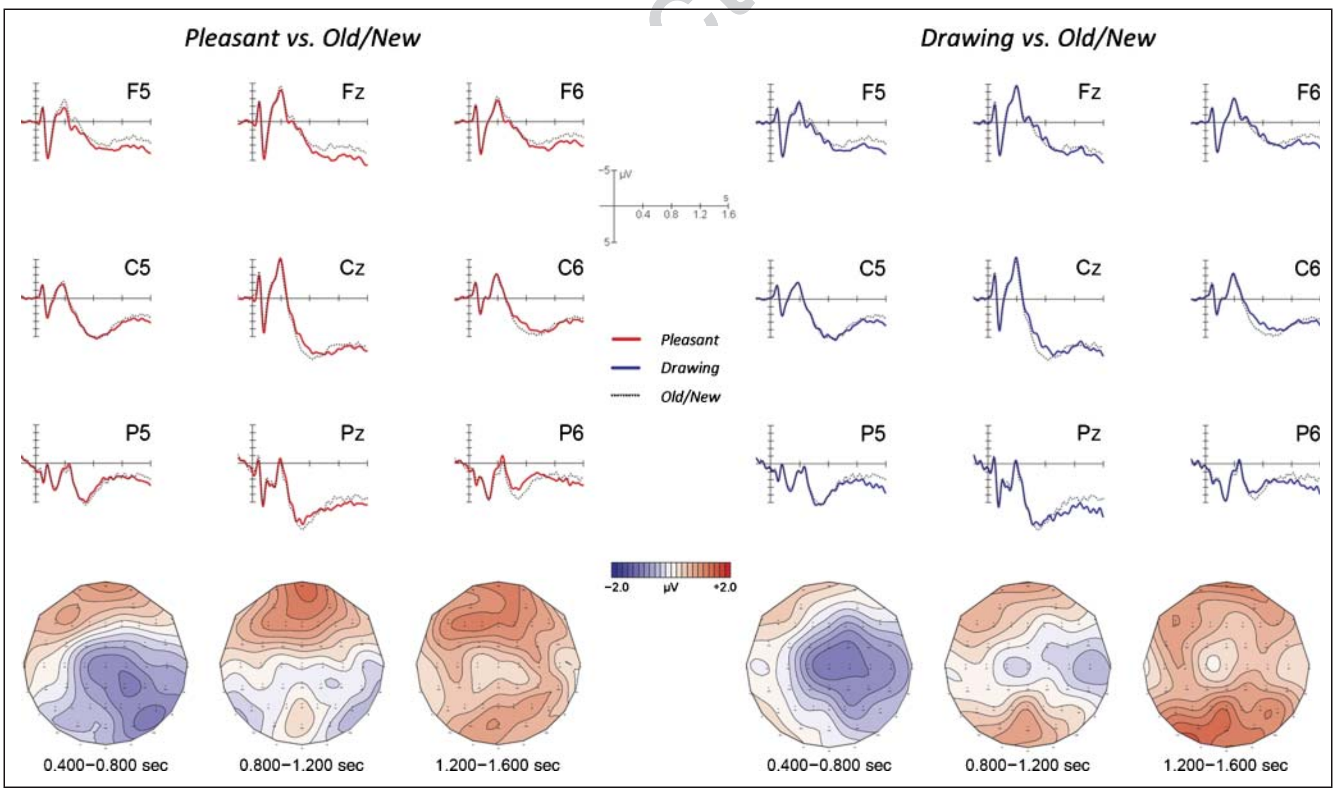

Figure 3. Top: Grand-averaged ERP waveforms elicited by new items in the old/new recognition task contrasted against new items from the pleasant criterial (left) and drawing criterial task (right) for all participants. Bottom: Topographic maps showing the scalp distributions of these two effects computed by subtracting mean amplitudes for ERPs elicited by new items in the old/new recognition task from the two criterial tasks.

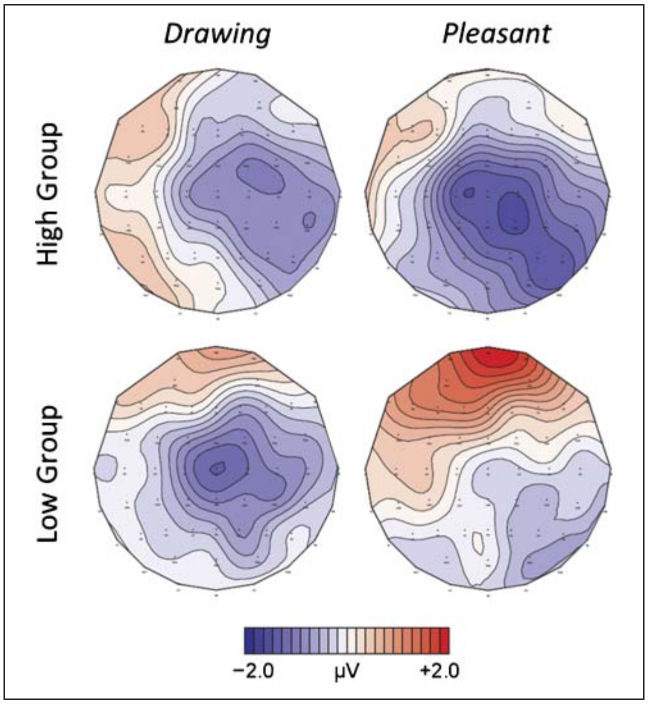

Figure 4. Topographic maps showing the scalp distributions of the two task-specific contrasts (criterial recollection minus old/new recognition) for the high- and low-accuracy groups in the 400–800 msec time window.

interacted with anterior/posterior in the high- ($F(1.31, 19.6) = 7.35, p < .01$) and low-accuracy group ($F(1.31, 19.61) = 7.88, p < .01$) and follow-up ANOVAs on rescaled data confirmed the presence of significant interactions between retrieval task and electrode site for both the high- ($F(3.91, 58.62) = 2.86, p < .05$) and the low-accuracy group ($F(3.3, 49.43) = 5.16, p < .01$). Evidence that participants who made more correct responses engaged distinct preretrieval operations in each criterial task supports the claim that the correlation between the magnitude of criterial task ERP differences and response accuracy reflects changes in the engagement of qualitatively distinct retrieval operations.

DISCUSSION

New Item ERP Effects: Criterial Recollection Tasks

If new item ERP contrasts index preretrieval processes that increase the likelihood that task-relevant information is recovered (Bridger et al., 2009) the extent to which these processes are engaged should correlate with accuracy in memory tasks in which facilitating the recovery of this information favors correct responding. This hypothesis was tested by contrasting ERPs elicited in two criterial recollection tasks designed to encourage the explicit recovery of target information in each test phase. ERPs elicited by new items from the two tasks differed from 400 msec onward, but the way in which ERPs differed depended upon the likelihood that participants made a correct response. In the low-accuracy group, ERPs from the pleasant criterial task demonstrated an anterior positivity from 400 to 1200 msec, whereas ERPs elicited by these same items in a high-accuracy group were more negative over left posterior sites from 400 to 1600 msec. This pattern comprises evidence of reliable changes in new item ERP processing across all participants, in line with the assumption that the parameters of the criterial test encouraged the engagement of distinct retrieval processes in each test phase.

Although there were reliable ERP new item effects in both the high- and low-accuracy groups, the engagement of only one of these effects was systematically related to increases in response accuracy; across all participants, the magnitude of the posterior criterial ERP effect from 400 to 800 msec correlated with each individual's ability to respond correctly to targets and nontargets. This pattern joins those reports in which retrieval cue processes have been shown to relate to incremental variations in accuracy in recognition tasks (Bridger et al., 2009; see also Rosburg et al., 2011). Nonetheless the current data differ from previous findings in a notable way. Bridger et al. reported that the magnitude of retrieval cue processing ERP differences correlated with accuracy across all participants, such that the effect was reliable only for a higher performing subset of participants whilst there was no evidence of differential retrieval cue processing for participants who responded comparatively poorly. The current data replicate the initial correlation but also provide evidence that low performers engaged distinct retrieval cue processing operations but that the amplitude of this effect did not correlate with variations in accuracy.

The current findings provide important support for the claim that the processes reflected by new item ERP contrasts can positively influence the recovery of particular kinds of mnemonic information. This claim is supported first by the absence of behavioral evidence that the criterial tasks varied in difficulty, effectively removing the possibility that any observed neural differences reflect changes in retrieval effort. A second reason to take this view follows directly from the logic underpinning criterial recollection tasks. Correct responding in these tasks is contingent upon the successful recovery of criterial information, and a strategy in which one prioritizes the recovery of criterial information in each test phase, therefore, should be associated with increases in correct responding. In line with this, we observed a correlation between the degree to which participants engaged distinct retrieval processes in each test and their ability to respond correctly to targets and nontargets. The differences between new item ERPs over posterior sites from 400 to 800 msec are likely, therefore, to reflect retrieval operations concerned with recovering imagery-related information in the drawing criterial test and valence-based semantic details in the pleasant criterial test.

In previous reports of this kind, the stability of new item ERP effects over time has led to the suggestion that these effects represent the maintenance of internal retrieval cue processing that is biased in line with sought-for information (Bridger et al., 2009; Hornberger et al., 2004). The left posterior new item ERP effect here is in keeping with such an account, in terms of both the outcomes of the time course analyses, which indicated that it remained stable from 400 to 1600 msec and the duration over which its amplitude correlated with response accuracy, which comprised a 400-msec epoch until 800-msec poststimulus. This time window is particularly important because it maps onto what is known about the time course of recollection. Evidence from both electrophysiological indices of retrieval success (for a review, see Rugg & Curran, 2007) and data from response-deadline paradigms (Mecklinger, Brunnemann, & Kipp, 2011; Yonelinas & Jacoby, 1994) converge to show that recollection occurs by 800 msec poststimulus. In the current study, therefore, ERP new item effects in a task that required criterial information to be recollected were no longer directly linked to response accuracy beyond the point at which recollection should have occurred. This correspondence strongly implicates that these processes are tied to the requirement to recollect contextual details and that they are not sustained once recollection fails. Moreover, this pattern supports the claim that not only can ERPs to new items reflect the engagement of operations that influence the degree to which retrieval is successful but that these influential operations can occur before retrieval.

This point is also significant because it highlights an inconsistency between the current data and that reported by Bridger et al. (2009), where the relationship between new item ERP effects and accuracy was most robust in the epochs occurring after the point by which retrieval should occur. This discrepancy raises the possibility that there may be instances in which maintaining retrieval cue processing beyond the time by which retrieval should have occurred is useful, perhaps because it might influence the likelihood of retrieval on the subsequent trial. One of the few task differences which might influence whether this occurs is the relative proportion of new items in the two tasks. Whereas new items represented a relative minority in the exclusion task reported by Bridger et al. (2009), they comprised half of all the items presented in the current criterial tasks. It may be that the extent to which it is efficient to sustain retrieval cue processing beyond the time point at which retrieval should have occurred is greater when the likelihood that the subsequent retrieval trial contains an old item is relatively high as was the case in the exclusion task in Bridger et al. (2009). Although reasoning along these lines remains somewhat speculative, it raises the possibility that the characterization of retrieval cue processes may be sensitive to manipulations as subtle as the ratio of old: new items.

The principal difference between the high- and low-accuracy group in the current study was the ability to discriminate between targets and nontargets, and the observation that this measure correlated with the magnitude of differences between ERPs to new items supports Bridger et al.'s (2009) assertion that prioritizing the recovery of one class of information confers benefits when binary discriminations of this kind are required. An additional behavioral difference between the two accuracy groups is also worth noting: participants in the high-accuracy group showed a significant slowing when correctly responding to nontargets. Such slowing is likely to reflect a greater sensitivity to the intrinsic structure of the criterial recollection task. Remember that, for nontargets, participants are required to a make a "new" judgment to an old item. Delaying responding therefore could reflect mechanisms that ensure these items are not also associated with as-yet-unrecovered criterial information (a possibility that arises because of the presence of items associated with both kinds of information) and/or to overcome response conflict arising from the requirement to respond "new" to an old item. The latter point might be especially salient here because nontargets constitute a minority of old items (33%) rather than the 50% they typically comprise in standard exclusion task paradigms. Regardless of which account best explains the data, slowed responding for nontargets corresponds with the assumption that higher performers were more sensitive to the role of criterial information in this task and could adjust their responding accordingly.

New Item ERPs: Criterial versus Old/New

The current design also allowed additional contrasts to be made between new item ERPs from each of the criterial tasks with those from a "baseline" old/new recognition test. These contrasts provide separate indices of the retrieval processes specific to each of the two criterial tasks and, for the first time, provide additional insight into the nature of pre-retrieval processing in two ways. First, a direct comparison of the scalp distribution of ERP effects from the pleasant and drawing tasks (criterial ERPs minus recognition ERPs) provided evidence that the topographies of the two effects were at least partially nonoverlapping, as would be expected if participants were engaging distinct retrieval processes in the two criterial tasks rather than simple changes in the degree to which the same process was engaged. Evidence for nonoverlapping topographies in the 400–800 msec time window was also observed when contrasts were limited to the high-accuracy group. This is an important finding, because it supports the functional interpretation of the correlation between the magnitude of the criterial ERP effects and response accuracy. According to this account, the correlation reflects the fact that higher performing participants were more likely to engage in retrieval cue processes that were qualitatively distinct from each other rather than simple changes in the engagement of the same process. Demonstrating that criterial task-specific effects were associated with distinct scalp distributions in the group of participants who were more likely to make a correct response provides important convergent support for this account. Evidence for distinct criterial task-specific effects in the low-accuracy group is also important but before discussing the functional significance of those effects, we turn our attention to the second way in which these contrasts can complement current understanding of preretrieval processing, by providing insight into processing that might be common to the two criterial tests.

In principle, criterial retrieval requirements are likely to require a greater degree of specificity than old/new recognition task demands. Whereas participants can make correct judgments in the recognition task without retrieving specific contextual details, the parameters of both criterial tasks explicitly require participants to recollect information for at least a proportion of old items. Contrasts between ERPs from these types of task, therefore, might also reflect processes related to the general requirement to retrieve associated information. These contrasts should be interpreted cautiously, however, because they also correspond to a decrease in response accuracy in the old/new recognition task, which may come about because of the relaxed retrieval demands and/or a more liberal response criterion for this task. Although the possibility that commonalities reflected in the criterial versus recognition ERP effects correspond primarily to changes in response accuracy cannot be entirely discounted, one aspect of these differences supports the position that at least a proportion of these effects relate to the heightened retrieval requirements in the criterial tasks.

From 400 to 800 msec, new item ERPs from both criterial tasks showed a negativity at central and posterior sites relative to recognition task ERPs. New item ERP effects with a centrally-maximal but broad distribution around 400–800 msec have been previously reported in studies in which contrasts were made between retrieval phases that prompted different levels of retrieval specificity. These comprise reports in which contrasts were made between tasks for which the test format (e.g., word or picture) of items either did or did not match the studied format (Hornberger et al., 2004; Robb & Rugg, 2002). Across these studies, new item ERPs were more negative when the study format of to-be-recovered information did not match the test format. The sensitivity of these effects to the match between study and test format led to the proposal that this negativity represents the extent to which retrieval processes are focused upon the conceptual characteristics of retrieval cues, because it is only at this level that perceptually nonmatching cues and targets overlap with one another (Hornberger et al., 2004). This reasoning can be extended to the current data in so far as the recovery of conceptual information is required in the two criterial tasks but not in the recognition task where responding can, in principle, be made on the basis of perceptual matching only. We tentatively propose, therefore, that the relative negativity for criterial ERPs from 400 to 800 msec reflects the engagement of operations that emphasize conceptual details associated with each item, in line with the assumption that stressing the conceptual aspects of retrieval cues might enable the recovery of semantic information necessary in both criterial tasks.

The idea that emphasizing processing associated with to-be-recovered information during test benefits episodic retrieval is drawn from the principle of transfer appropriate processing, which states that the likelihood that retrieval occurs is a function of the extent to which retrieval cues are processed in line with encoding operations (Morris, Bransford, & Franks, 1977). With respect to the current processes, therefore, engaging and maintaining internal representations that correspond to the class of sought-for information should increase the likelihood that this information is recovered. This would hold both for the recapitulation of generic semantic operations as well as the re-engagement of processes biased toward more particular characteristics, such as judging valence or attention to visual or physical characteristics of items. In addition, if the central negativity does in part capture aspects of preretrieval processing that are common to the two criterial tasks, the fact that this could occur during the same time window in which criterial task-specific processes are present, introduces the possibility that participants are able to engage in more than one class of preretrieval process at the same time during a recognition task. A challenge for future research, therefore, would be to investigate this possibility using comparable retrieval manipulations that are not confounded by behavioral differences.

New Item ERP Effects Associated with Low Task Performance

We have argued that the left posterior criterial ERP effect observed in the early window for the high-accuracy group reflects retrieval operations associated with recovering different kinds of information because the magnitude of this effect correlates with response accuracy in tasks that require the recovery of specific kinds of information. What might the processes reflected by the anterior criterial ERP effect in the low-accuracy group, which do not correlate with accuracy, represent? In a number of previous studies, frontally-based positive waveforms have been associated with increases in the engagement of retrieval effort (Rosburg et al., 2011; Dzulkifli, Sharpe, & Wilding, 2004; Ranganath & Paller, 2000) in line with the position that prefrontal control mechanisms should be engaged to a greater degree as retrieval becomes more difficult or requires a greater degree of specification (Mecklinger, 2010). It is reasonable to hypothesize here that lower performers, aware of their poor performance, might have simply increased the engagement of control mechanisms. This is unlikely to reflect the changes in activity over anterior sites for this group, however, because there was no behavioral indication that one of the tasks was more difficult than the other. We favor an alternative account, which rests on the fact that the outcomes of the task-specific contrasts strongly suggest that this anterior positivity primarily reflects processes engaged in the pleasant criterial task.

The task-specific pleasant criterial contrast from 400 to 800 msec depicts the onset of a longer-lasting positivity over frontal sites (Figure 3, bottom left). The outcomes of the within-group contrasts in this time window, moreover, indicate that this anterior positivity arises predominantly in the low-accuracy group and is absent in the high-accuracy group (see Figure 4). Together, these observations indicate that the anterior difference between criterial new item ERPs in the low group reflects processes intrinsic to the pleasantness task. Whereas both study tasks require the recovery of item semantics, a unique component of the pleasantness judgment is to attend to the valence of each item. It is possible that the response requirements of the pleasantness criterial task may have caused some individuals to attend to the valence of new items. In line with this interpretation is the similarity between the frontal distribution of the current effect and the results of explicit investigations of the neural correlates of valence processing. Dolcos and Cabeza (2002), for example, report that ERPs elicited by items varying in valence differentiate over anterior sites from 500 to 800 msec poststimulus. König and Mecklinger (2008) also report temporarily extended frontal positive slow waves associated with the processing of positively valenced images from 250 to 1000 msec. Moreover, mechanisms associated with valence processing have been shown to anatomically dissociate from the automated processing of arousing stimuli. Kensinger and Corkin (2004) report the hemodynamic correlates of processes that predicted whether neutral, negative, and nonarousing or negative and arousing words were later remembered. Whereas the encoding of arousing items initiated an amygdalar–hippocampal network, this dissociated from a pFC–hippocampal network that responded to valenced information only, highlighting the presence of frontally mediated processes associated with the processing and encoding of valenced information.

Although those studies linking frontally based modulations and valence processing have typically employed paradigms that differ markedly from the current task, we argue that similarities in polarity and anterior distribution between those reports and the current effect are in keeping with the notion that the anterior positivity reported here is a reflection of top–down item valence processing. Although in the current study this effect was, to some extent, present for all participants from 400 msec onward, evidence that it was engaged to a greater extent by the low-accuracy group is in line with the notion that these participants tended to emphasize the valence of new items in the pleasant criterial task. Processing of this kind might detract from the ability to engage more efficient retrieval cue strategies, such as those presumed to be engaged by the higher performing participants. Although we favor an account of this kind, plausible alternatives remain, such as the likelihood that the effect represents enhanced valence processing in light of the paucity of other semantic characteristics available to these participants. Whereas the current data are unable to distinguish between possibilities of this kind, we stress that they nonetheless provide, for the first time, a characterization of the retrieval cue processing associated with the requirement to recover one particular class of information.

Concluding Remarks

The present findings replicate and extend those reported by Bridger et al. (2009), and support the proposal that preretrieval processes facilitate the retrieval of task-relevant details because they relate to accuracy in tasks in which prioritizing the recovery of this information is especially valuable for correct responding. Contrasts unique to the current design allowed further insight into these operations by providing indices of processes unique to each criterial task. These contrasts revealed task-specific effects with distinct scalp distributions for all participants. That this was also the case when contrasts were limited to high performing participants supported the assumption that participants who made more accurate responses engaged qualitatively rather than quantitatively distinct retrieval processes in each test phase. Distinct scalp distributions were also found for the criterial task-specific effects in the low-accuracy group, but these differences are likely to reflect a frontal effect that was specific to the processing engaged in the pleasantness task. Finally these contrasts also provide insight into operations that may be common to the requirement to explicitly recover contextual information. Together, these findings emphasize the multifaceted manner with which retrieval cues can be processed in line with task demands and support those models of controlled retrieval which state that specification processes engaged before retrieval can influence correct responding in memory tasks.

Acknowledgments

We thank Anna-Lena Scheuplein, Marie Schwartz, and Sandra Schappert for invaluable assistance with stimuli preparation, translation, and data collection. This research was supported by the German Research Foundation under Grant DFG-IRTG-1457 and was conducted in the International Research Training Group "Adaptive Minds" hosted by Saarland University, Saarbrücken, Germany.

Reprint requests should be sent to Dr. Emma K. Bridger, Department of Psychology, Saarland University, 66123 Saarbrücken, Germany, or via e-mail: e.bridger@mx.uni-saarland.de.

Note

1. Readers will note that Table 3 also shows a higher overall "hit" rate in the old/new recognition than the criterial task and that corresponding increases in hits and false alarms are in line with the adoption of a more liberal criterion in the old/new recognition task. It is not possible to explicitly test this, however, because of the absence of response bias measures that are not

contingent upon assumptions about the models that underpin responding in recognition tasks (Snodgrass & Corwin, 1988). This problem is especially compounded here because the mapping of old/new items onto yes/no responses is confused both within and across recognition tasks.

REFERENCES

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database*. University of Pennsylvania, Linguistic Data Consortium.

Bridger, E. K., Herron, J. E., Elward, R. L., & Wilding, E. L. (2009). Neural correlates of individual differences in strategic retrieval processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1175–1186.

Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory, 4,* 359–411.

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective & Behavioral Neuroscience, 2,* 252–263.

Dzulkifli, M. A., Sharpe, H. L., & Wilding, E. L. (2004). Separating item-related electrophysiological indices of retrieval effort and retrieval orientation. *Brain and Cognition, 55,* 433–443.

Dzulkifli, M. A., & Wilding, E. L. (2005). Electrophysiological indices of strategic episodic retrieval processing. *Neuropsychologia, 43,* 1152–1162.

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory and Language, 51,* 473–493.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Greenhouse, G. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 49,* 95–112.

Herron, J. E., & Rugg, M. D. (2003). Retrieval orientation and the control of recollection. *Journal of Cognitive Neuroscience, 15,* 843–854.

Hornberger, M., Morcom, A. M., & Rugg, M. D. (2004). Neural correlates of retrieval orientation: Effects of study–test similarity. *Journal of Cognitive Neuroscience, 16,* 1196–1210.

Hornberger, M., Rugg, M. D., & Henson, R. N. (2006). fMRI correlates of retrieval orientation. *Neuropsychologia, 44,* 1425–1436.

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

Jasper, H. A. (1958). The Ten-Twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

Johnson, M. K., Kounios, J., & Nolde, S. F. (1997). Electrophysiological brain activity and memory source monitoring. *NeuroReport, 8,* 1317–1320.

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, U.S.A., 101,* 3310–3315.

König, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion, 8,* 162–173.

McCarthy, G., & Wood, C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203–208.

Mecklinger, A. (2010). The control of long-term memory: Brain systems and cognitive processes. *Neuroscience and Biobehavioral Reviews, 34,* 1055–1065.

Mecklinger, A., Brunnemann, N., & Kipp, K. (2011). Two processes for recognition memory in children of early school age: An event-related potential study. *Journal of Cognitive Neuroscience, 23,* 435–446.

Metcalf, K., Langdon, R., & Coltheart, M. (2007). Models of confabulation: A critical review and a new framework. *Cognitive Neuropsychology, 24,* 23–47.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behaviour, 16,* 519–533.

Moscovitch, M., & Winocur, G. (2002). The frontal cortex and working with memory. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 188–209). London: Oxford University Press.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37,* 127–152.

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

Ranganath, C., & Paller, K. A. (2000). Neural correlates of memory retrieval and evaluation. *Cognitive Brain Research, 9,* 209–222.

Robb, W. G., & Rugg, M. D. (2002). Electrophysiological dissociation of retrieval orientation and retrieval effort. *Psychonomic Bulletin & Review, 9,* 583–589.

Rosburg, T., Mecklinger, A., & Johansson, M. (2011). Electrophysiological correlates of retrieval orientation in reality monitoring. *Neuroimage, 54,* 3076–3084.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11,* 251–257.

Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences, 4,* 108–115.

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology, 49,* 289–318.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Werkle-Bergner, M., Mecklinger, A., Kray, J., Meyer, P., & Düzel, E. (2005). The control of memory retrieval: Insights from event-related potentials. *Cognitive Brain Research, 24,* 599–614.

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia, 37,* 441–454.

Wilding, E. L. (2006). The practice of rescaling scalp-recorded event-related potentials. *Biological Psychology, 72,* 325–332.

Woodruff, C. C., Uncapher, M. R., & Rugg, M. D. (2006). Neural correlates of differential retrieval orientation: Sustained and item-related components. *Neuropsychologia, 44,* 3000–3010.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology, 48,* 516–535.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

During the preparation of your manuscript, the questions listed below arose. Kindly supply the necessary information.

1. Please provide publisher name and location of Baayen et al., 1993.
2. Please check if subsections of Results were captured properly.

END OF ALL QUERIES