The Influence of Criterion Shifts on Electrophysiological Correlates of Recognition Memory

N. Azimian-Faridani and E. L. Wilding

Abstract

The claim that event-related potentials (ERPs) index familiarity was assessed by acquiring ERPs during a recognition memory task in which participants were instructed to adopt different decision criteria in separate retrieval phases. In one, the instructions were to respond “old” only when confident that this was the correct response, and to respond “new” otherwise (the conservative condition). In the other, the instructions were to respond new only when confident that this was the correct response (the liberal condition). The rationale for this approach was that the level of familiarity licensing an old response would be higher in the conservative than in the liberal condition, and if ERPs index familiarity, this would be reflected in changes to the putative ERP index. This index comprises relatively more positive-going neural activity for correct judgments to old than to new items, which is evident from 300 to 500 msec poststimulus at mid-frontal scalp locations. In keeping with task instructions, participants made more old responses in the liberal than in the conservative condition. There were reliable mid-frontal ERP old/new effects in both conditions, and the ERPs evoked by correct judgments to words in the conservative condition were relatively more positive-going than those in the liberal condition. This finding is consistent with the view that the mid-frontal ERP old/new effect indexes familiarity, and in combination with other ERP findings, provides strong support for dual-process accounts of recognition memory.

INTRODUCTION

According to dual-process models of recognition memory, the ability to discriminate between items encountered at study and items presented at test only is supported by two processes—recollection and familiarity (Yonelinas, 2002; Jacoby, 1991; Mandler, 1980, 1991; Jacoby & Dallas, 1981). Recollection is a process that entails recovery of qualitative information about a prior event, whereas familiarity provides relatively undifferentiated (quantitative) information about a previous occurrence: sufficient to indicate whether an event occurred previously but insufficient to support judgments about qualitative details (Yonelinas, 2002). Familiarity is commonly considered to vary in strength in a continuous fashion and can be modeled within a signal detection framework (Yonelinas, 2002; Yonelinas & Jacoby, 1994). According to perhaps the most influential current dual-process accounts (Jacoby, 1998; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998), recollection and familiarity operate independently at the time of retrieval.

Dual-process accounts of recognition memory receive support from a number of different sources. In addition to findings from a body of behavioral studies (for a comprehensive review, see Yonelinas, 2002), there are also findings from animal and patient studies that are consistent with the view that two processes contributing to recognition memory are supported by different regions of the medial-temporal lobe (e.g., Aggleton et al., 2005; Fortin, Wright, & Eichenbaum, 2004; Holdstock et al., 2002; Yonelinas, Kroll, Quamme, et al., 2002; Aggleton & Brown, 1999; Vargha-Khadem et al., 1997), although there are also contradictory findings and interpretations (Wixted & Squire, 2004; Yonelinas, Quamme, et al., 2004; Manns & Squire, 2000; Zola et al., 2000). There are also findings from functional magnetic resonance imaging and computational modeling studies of memory retrieval that are consistent with the claim that familiarity and recollection rely on different neural substrates (Davachi, Mitchell, & Wagner, 2003; Henson, Cansino, Herron, Robb, & Rugg, 2003; Norman & O’Reilly, 2003; Wheeler & Buckner, 2003; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000).

The findings in some event-related potential (ERP) studies of memory retrieval are also relevant to the question of the processes that support recognition memory judgments. In studies where participants have been asked to make remember/know (R/K) or forced-choice source judgments, there is only limited support for dual-process accounts (Gurran, 2004; Duarte, Ranganath, Winward, Hayward, & Knight, 2004; Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Wilding &
The strongest evidence concerns the existence of a functional dissociation between two ERP old/new effects that are held to index recollection and familiarity, respectively. ERP old/new effects are differences between the ERPs associated with correct responses to old and new test items (Rugg, 1994). The two effects in question comprise relatively more positive-going ERPs for correct judgments to old than to new test items. The effect with the longer history has been linked to recollection, is largest at left parietal scalp locations, and is evident from approximately 500 to 800 msec poststimulus. The second effect is commonly more positive-going ERPs for correct judgments to old and new test items (Rugg, 1996; Wilding, Doyle, & Rugg, 1995; Smith, 1993). The absence of changes in the size of the mid-frontal effect across depth of processing is problematic for a familiarity account. Rugg, Mark, et al. (1998) demonstrated that only the magnitude of the mid-frontal old/new effect was influenced by a depth of processing manipulation: It was larger for words subjected to the deeper of the two encoding conditions. These data support a dual-process account of recognition memory in so far as the left parietal and mid-frontal ERP old/new effects behaved differently according to the experimental manipulation.

The absence of changes in the size of the mid-frontal effect across depth of processing is problematic for a familiarity account, however, in light of behavioral data indicating that the manipulation increases the likelihood of responses made on the basis of recollection and familiarity (for review, see Yonelinas, 2002). Similar concerns apply to the fact that the mid-frontal effect is insensitive to full/divided attention encoding manipulations as well as to the confidence that is associated with recognition memory judgments at test (Curran, 2004). There are behavioral data in both cases which suggest that familiarity should vary according to these manipulations, being greater in the case of the full attention and high confidence conditions (e.g., Yonelinas & Jacoby, 1994; Gardiner & Java, 1993; Jacoby & Kelley, 1992).

One interpretation of these findings is that the mid-frontal old/new effect indexes a process that occurs downstream of the familiarity signal (see Tsivilis, Otten, & Rugg, 2001) and may behave in an all-or-none fashion. An alternative explanation, however, is that the null results described above reflect the fact that manipulations such as depth of processing typically influence recollection to a greater degree than familiarity (Yonelinas, 2002), and that the performance differences across the conditions of interest (e.g., deep vs. shallow encoding; Rugg, Mark et al., 1998) have not been of sufficient magnitude to elicit changes in the amplitude of the mid-frontal ERP old/new effect. It may also be important...
that the mid-frontal ERP old/new effect is, on average, smaller than the left parietal old/new effect, thus it is less likely to index differences between critical conditions (for related comments concerning statistical power, see Curran, 2004).

In summary, the points raised in the preceding two paragraphs emphasize that some of the apparent inconsistencies in the findings in respect of the mid-frontal ERP old/new effect may be a consequence of the kinds of manipulations that have been employed. It follows from this observation that perhaps a more rigorous test of the link between the mid-frontal effect and familiarity than has been achieved to date would stem from the use of a manipulation that is known to affect responses made on the basis of familiarity to a greater extent than those made on the basis of recollection.

The experiment described here was designed with this goal in mind, and this was achieved by varying test phase instructions across study–test blocks so as to encourage participants to adopt different response criteria. Manipulations that have been introduced to influence the response criterion that participants adopt affect responses made on the basis of familiarity to a markedly greater degree than recollection (Yonelinas, 2001; Postma, 1999; Hirshman & Henzler, 1998; Gardiner, Richardson-Klavehn, & Ramponi, 1997; for review, see Yonelinas, 2002), making this approach a good candidate for a test of the familiarity account of the mid-frontal ERP old/new effect. The central assumption is that encouraging changes in the placement of response criteria will result in variations in the levels of familiarity that participants will accept as a basis for an old or a new judgment. If the mid-frontal effect is an index of familiarity, then the amplitude of this effect should vary systematically according to the decision criterion that participants adopt, and this can be understood by referring to the upper panel of Figure 4: If a relatively liberal decision criterion is adopted, then, on average, the level of familiarity sufficient to justify an old judgment will be lower than the level of familiarity when a relatively conservative criterion is adopted. The same logic applies to correct rejections: The mean level of familiarity associated with new judgments should increase as increasingly conservative criteria are adopted. This approach thus provides a rigorous assessment of the link between the mid-frontal ERP old/new effect and familiarity, and by extension, an assessment of the extent to which ERPs acquired during recognition memory tasks provide support for some dual-process accounts of recognition memory.

METHODS
Participants

Thirty undergraduates (15 women, all native English speakers) from Cardiff University were paid £7.50/hr for their participation. All were right-handed and reported normal or corrected-to-normal vision. The average age was 22 (age range 19–30). Data from six participants (three women) were discarded prior to analysis. One participant failed to follow the task instructions and another was unable to perform the task above a preset criterion (see below). Four participants did not contribute sufficient trials to the conditions of interest following artifact rejection (for criteria, see below). All participants gave informed consent.

Stimuli

Stimuli were 640 low-frequency words taken from the MRC psycholinguistic database (4–9 letters, frequency 1–7 per million, www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm). One complete task list comprised all 640 words. The words were divided randomly into four equal subgroups of 160 words, each of which contained the words for one study–test cycle. Each subgroup was separated into two lists, one comprising 80 words to be presented at study, the other comprising all 160 words to be presented at study and at test. Thus, within each list, there were an equal number of old and new words presented in each test phase. Three further complete task lists were created by rotating the words within each subgroup and the order in which the subgroups were presented, such that across task lists all words were encountered as old and new stimuli at test, and in one of the first two as well as the last two study–test cycles. The order of presentation of items in each study and test cycle was determined randomly for each participant. All stimuli were presented in upper case letters in central vision on a PC monitor located 1 m away from participants. All stimuli were presented in white against a black background. They subtended maximum visual angles of 0.6° (vertical) and 2.2° (horizontal).

Procedure

Participants were fitted with an electrode cap before the experiment (see below). They were seated in a sound-attenuated booth facing a monitor with their thumbs resting on response keys. The participants first read through an instruction sheet and the instructions were also explained verbally prior to the start of the experiment.

Study phase trials began with a hash mark (#), which was displayed for 500 msec and followed by a blank screen (250 msec), after which the study word was displayed for 300 msec. Participants indicated via keypress whether the study word referred to an animate or inanimate entity. The next trial began 1000 msec after the participant made a response. Participants were reminded of which key on the response box corresponded to which response before commencing each study phase.

Test phase trials began with a hash mark (#), which was displayed for 1000 msec and followed by a blank screen (100 msec). The test word was then displayed for
300 msec, to which the participant was required to make an old/new response via keypress with the right/left hand. The next trial began 1200 msec after the participant responded. In two consecutive test phases, participants were instructed to make an old response only when confident that the test word was old (the “conservative” condition). In the other test phases, they were instructed only to make a new response when confident that the test word was new (the “liberal” condition). Participants were reminded of which key on the response box corresponded to which response immediately before each test phase. The hand used for old responses was crossed across participants, as was the order in which the liberal and conservative conditions were completed. There was a short practice session before each condition and a short break after each study–test block.

**Electrophysiological Recording Procedure**

Electroencephalogram (EEG) readings were recorded from 25 silver/silver chloride electrodes housed in an elastic cap. The sites were located at midline (Fz, Cz, Pz), as well as left and right hemisphere locations (FP1/FP2, F7/F8, F5/F6, F3/F4, T3/T4, C5/C6, C3/C4, T5/T6, P5/P6, P3/P4, O1/O2; Jasper, 1958). Additional electrodes were placed on the mastoid processes. Electrooculogram (EOG) readings were recorded from above and below the right eye (VEOG) and from the outer canthi (HEOG). Trials containing large EOG artifact were rejected, as were trials containing A/D saturation or baseline drift (difference between first and last data point) exceeding ±80 μV. Other EOG blink artifacts were corrected using a linear regression estimate (Semlitsch, Anderer, Schuster, & Presslich, 1986). EEG was recorded continuously at 166 Hz (6 msec per point) with Fz as the reference electrode, and was re-referenced computationally off-line to a linked mastoid reference into baseline-corrected epochs of 1002 msec (167 data points), each including a 102 msec prestimulus baseline and a 900 msec post-stimulus period. The data from Fz were reclaimed. EEG and EOG were recorded with a bandwidth of 0.03–40 Hz (±80 dB). Participants were excluded if they showed little evidence of the ability to discriminate between old and new test words \( p(\text{hit}) - p(\text{false alarm}) < .1 \) in either the liberal or the conservative response conditions, or if they contributed less than 12 trials to any response category entering into an analysis. The averaged ERPs were subjected to a 7-point binomially weighted smoothing filter prior to analysis.

**RESULTS**

**Behavior**

Table 1 shows the probabilities of an old response to old and to new test words in the liberal and conservative conditions. The data were subjected to analysis of variance (ANOVA) with factors of condition (conservative/liberal) and word status (old/new). The analysis revealed that although the probability of an old response is greater in the liberal than in the conservative condition \( F(1,23) = 42.00, p < .001 \), and greater for old than for new test words \( F(1,23) = 891.59, p < .001 \), across conditions there is a greater difference in the probability of an old response for new than for old words [Condition × Status interaction: \( F(1,23) = 7.97, p < .025 \)]. The measures of discrimination \( d_a \) and criterion \( c_a \) were computed for each participant, using an estimated slope of 0.8. Analyses of these measures revealed that discrimination was statistically equivalent for the two groups (2.30 and 2.10 in the conservative and liberal conditions, respectively) and that, as intended, participants were more likely to make an old response in the liberal condition than in the conservative condition \( [c_a(\text{conservative}) = 0.34, c_a(\text{liberal}) = -0.50] \).

Table 2 shows the reaction time (RT) data for correct responses to old and new test words, separated according to condition (conservative/liberal). ANOVA with factors of condition and category revealed reliable main effects [category: \( F(1,23) = 24.62, p < .001 \); condition: \( F(1,23) = 47.95, p < .001 \)] as well as a reliable interaction \( [F(1,23) = 6.03, p < .025 \). These outcomes reflect the fact that although correct responses to old words are faster than those to new words, the slower responses in the liberal condition are more marked for new than for old words.

**Event-related Potentials**

Event-related potentials could, in principle, be obtained for eight response categories: correctly and incorrectly identified old and new words separated according to condition (conservative/liberal). Too few participants made sufficient false alarms in the conservative condition (six participants) and misses in the liberal condition (five participants) to permit analysis of the ERP data for these response categories. As a consequence, the data from the conservative and liberal conditions were first analyzed separately for different numbers of participants for the 300–500 msec time window. For the conservative condition, the data were analyzed for the 19 participants.

**Table 1. Probabilities of “Old” Responses to Old and New Test Words in the Conservative and Liberal Conditions**

<table>
<thead>
<tr>
<th>Status</th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative</td>
<td>.78 (.13)</td>
<td>.09 (.08)</td>
</tr>
<tr>
<td>Liberal</td>
<td>.91 (.06)</td>
<td>.30 (.13)</td>
</tr>
</tbody>
</table>

SDs are in brackets.
who contributed sufficient artifact-free trials to the hit, correct rejection, and miss response categories. Mean trial numbers for these response categories were 95, 118, and 32, respectively. For the liberal condition, the data were analyzed for all 24 participants, each of whom contributed sufficient trials to the hit, correct rejection, and false alarm categories. Mean trial numbers for these response categories were 120, 91, and 38, respectively. The ERP data were also analyzed over the 500–700 and 700–900 msec epochs, where the analyses were restricted to the ERPs evoked by hits and by correct rejections in the two criterion conditions.

The ERP data were analyzed over these three time windows (500–500, 500–700, and 700–900 msec) because they span the poststimulus epoch, overlap with those employed in previous studies (e.g., Curran, 1999), and capture the principal divergences between the ERPs associated with correct and incorrect responses separated according to condition, as Figures 1–3 show. For the three time windows, the initial analyses were restricted to six electrode sites at anterior and posterior midline, left and right hemisphere locations (F3, Fz, F4, P3, Pz, P4; for a similar strategy, see Rugg, Mark, et al., 1998). These sites cover anterior and posterior locations at which functionally distinct memory processes have been identified in previous studies (Rugg, Mark, et al., 1998). In each epoch, the initial analysis therefore included the factors of category, location in the anterior/posterior plane (AP), and location in the left–right plane (LR: left hemisphere, midline, right hemisphere). Significant effects involving category were followed up by all possible paired contrasts between the three response categories, with all other factors as above. Subsequent interactions involving category and the AP dimension were followed up by separate analyses at posterior and anterior locations for all analyses reported below. In all cases, only reliable effects involving factors of condition and/or category are reported.

For the 300–500 msec time window, focused analyses of old/new effects (separated according to condition) were also restricted to mid-frontal electrode locations (F3, Fz, F4), because these locations cover the scalp region where the putative electrophysiological index of familiarity is located (Curran, 1999, 2000). For this direct analysis (including the data from all 24 participants) and for the analyses over the 500–700 and 700–900 msec time windows that were restricted to hits and correct rejections, mean trial numbers for hits were 102 (conservative) and 120 (liberal), whereas for correct rejections the numbers were 117 (conservative) and 91 (liberal).

Table 2. Reaction Times for Correct Responses to Old and to New Test Words in the Conservative and Liberal Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Hit</th>
<th>Correct Rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative</td>
<td>899 (130)</td>
<td>959 (184)</td>
</tr>
<tr>
<td>Liberal</td>
<td>1000 (202)</td>
<td>1168 (237)</td>
</tr>
</tbody>
</table>

SDs are in brackets.

Azimian-Faridani and Wilding 1079
the fact that the relative positivity associated with hits tends to be smallest at right hemisphere sites.

The critical directed analyses within the 300–500 msec window were conducted in order to assess how the midfrontal old/new effect was influenced by the criterion manipulation. The analysis included the factors of category (hit/correct rejection), condition (conservative/liberal), and site (F3, Fz, F4). There were three reliable effects of interest. First, a main effect of condition \( F(1,23) = 8.48, p < .01 \), reflecting the fact that the ERPs elicited by words in the conservative condition are more positive-going than those elicited in the liberal condition. Second, a main effect of response category \( F(1,23) = 23.12, p < .001 \), moderated by an interaction between this factor and site \( F(1.9,42.9) = 8.65, p < .01 \). These terms, in part, reflect the fact that the relatively greater positivity for hits than for correct rejections is smallest at F4 (mean hit – correct rejection amplitudes are 1.0, 1.1, and 0.8 \( \mu \text{V} \) at F3, Fz, and F4, respectively).

**500–700 msec**

The initial analysis for this epoch included the factors of condition (conservative/liberal), category (hit/correct rejection), AP, and LR. The highest order interactions revealed by this initial ANOVA were between condition, category, and LR \( F(1,23) = 5.29, p < .05 \) as well as condition, AP, and LR \( F(1.4,32.9) = 3.93, p < .05 \). The separate analysis of the ERP old/new effects in the conservative condition revealed a three-way interaction between category, AP, and LR \( F(1.4,32.6) = 5.47, p < .05 \), reflecting the fact that although ERPs elicited by hits are more positive-going than those elicited by correct rejections at anterior \( F(1,23) = 9.79, p < .01 \) and posterior sites \( F(1,23) = 50.76, p < .001 \), at posterior sites only is there statistical evidence for left-lateralized old/new effects [interaction between category and LR: \( F(1.6,36.7) = 8.44, p < .01 \)]. The mean amplitudes of the old/new effects at P3, Pz, and P4 in the conservative condition are 3.6, 3.3, and 2.6 \( \mu \text{V} \), respectively. The separate analyses at anterior locations revealed no reliable old/new effects, whereas at posterior locations hits were more positive-going than correct rejections \( F(1,23) = 5.55, p < .05 \), and these differences were largest at P3 [interaction between category
The performance data showed that participants in both conditions were able to discriminate old from new test words at levels above chance. Participants were also more likely to make old responses to test words in the liberal than in the conservative test conditions. The performance data in this experiment do not permit a direct assessment of the degree to which this shift in criterion influenced responses based upon recollection and familiarity, respectively, but Yonelinas (2002) has shown that the results from a number of previous studies are broadly consistent with the view that manipulations of criterion influence responding based upon familiarity to a considerably greater degree than responses based upon recollection.

The critical ERP analyses were those for the 300–500 msec time window. This is the time period in which the putative ERP index of familiarity—the mid-frontal ERP old/new effect—occurs. The analyses of the data from the conservative condition demonstrated that there are two functionally distinct memory-related effects in this time period. At posterior electrode locations, the ERPs evoked by misses and hits were reliably more positive-going than those evoked by correct rejections, but were not reliably different from each other. At anterior locations, meanwhile, hits were reliably more positive-going than both of the other response categories.

On the basis of a qualitatively similar pattern of findings, Rugg, Mark, et al. (1998) suggested that the posteriorly distributed effect (see Figure 2A and B) is an index of implicit memory, although to our knowledge...
there are no published studies to date in which this modulation has been correlated with a behavioral index of implicit processes (for further comments relevant to this repetition effect, see Rugg, Allan, & Birch, 2000). For present purposes, however, the important point is that the outcomes of the contrasts involving misses indicate that the greater relative positivity for hits than for correct rejections at anterior and posterior sites comprises two distinct modulations. The absence of reliable differences between misses and correct rejections at anterior sites, moreover, is broadly consistent with the view that the mid-frontal ERP old/new effect indexes familiarity. Markedly stronger evidence in support of this claim stems from the outcome of the direct contrast between the ERPs elicited by hits and by correct rejections in the conservative and the liberal conditions at mid-frontal electrode locations. Previous proposals that the mid-frontal ERP old/new effect indexes familiarity imply that the relatively greater positivity for hits than for correct rejections is a consequence of the greater familiarity of old items (Curran, 2000; Mecklinger, 2000), presumably because of a time-limited increase in their relative familiarity by virtue of their presentation in a prior study phase. If this account is correct, then the consequences of adopting different response criteria should be indexed by the mid-frontal effect in the following way: Test items in the conservative condition should be more positive-going than those from the liberal condition. The reason for this is illustrated in the top section of Figure 4, where it can be seen that as response criteria become more conservative (move to the right), there is an increase in the mean level of familiarity associated with old and new items attracting correct judgments. The correspondence between these predictions and the critical ERPs acquired in this experiment is summarized in the middle and bottom sections of Figure 4, which shows that ERPs evoked by test words in the conservative condition are more positive-going than those evoked in the liberal condition. The data therefore provide strong evidence in support of the claim that the mid-frontal ERP old/new effect indexes familiarity.

To our knowledge, the only other ERP study in which old/new effects have been analyzed when separated according to criterion is due to Windmann, Urbach, and Kutas (2002). They created two groups of participants on the basis of a median split of the response criterion that was adopted (see also Windmann, Sakhavat, & Kutas, 2002; Windmann & Kutas, 2001). ERP old/new effects between 300 and 500 msec were evident primarily at anterior scalp locations, with a tendency to be larger in the conservative than in the liberal condition. There were no overall group differences comparable to those described in this article. This disparity may, in part, reflect a power issue because of the between-participants design, and the number of participants per group (n = 15). Of perhaps greater importance, when participants are separated according to criterion, it is not possible to determine whether differences in criterion come about because of changes in overall familiarity strength (a relative shift in the familiarity distributions), changes in criterion, or some combination of the two (Azimian-Faridani & Wilding, 2004; Verde & Rotello, 2003).

The findings for the putative index of recollection in this experiment are less straightforward to interpret than those for the mid-frontal effect. As Figure 3 shows and the analyses confirmed, at posterior electrodes there is evidence for a degree of left lateralization of the ERP old/new effects in the 500–700 msec time window in both the conservative and the liberal response conditions. This pattern of lateralization is part of the signature of
the left parietal ERP old/new effect and suggests that recollection was engaged, at least to some extent, in both conditions. Questions concerning the extent to which recollection was engaged in the two conditions are hard to address, however, because of the response probability confound across the conservative and liberal conditions, and the subsequent influence of the P300 potential on the ERPs evoked by test words.

In the liberal condition, approximately 61% of test responses were old; in the conservative condition, the figure was 43%. The P300 component typically peaks between 300 and 800 msec poststimulus, and has a Pz maximum with little, if any, hemisphere lateralization (Donchin, 1981; Sutton, Braren, Zubin, & John, 1965). This component is particularly sensitive to response probability, being larger for low probability response classes (Donchin & Coles, 1988). Thus, any differences between the ERPs evoked by old and new items in the two conditions may be contaminated by overlapping P300 activity. In particular, the fact that the P300 is sensitive to response probability leads to the prediction that a relatively larger P300 will be evoked by hits in the conservative than in the liberal condition, whereas the reverse is true in the case of correct rejections. The data in the 500–900 msec time window are broadly consistent with this account. This can be seen in the markedly larger differences between the ERPs evoked by hits and by correct rejections in the conservative case—where hits are the lower probability category—than in the liberal case—where the reverse is true. For the liberal condition, it is reasonable to assume that the relatively small left parietal ERP old/new effect is a consequence of greater P300 activity for correct rejections offset to some degree by a left-lateralized recollection-related positivity for hits. The greater relative positivity for correct rejections in the liberal than in the conservative condition from 500 msec onwards, which is most pronounced at Pz (see Figure 1) and shows little indication of hemisphere lateralization, is also consistent with this account.

For these reasons, however, it is difficult to employ the ERP data in order to make inferences concerning whether participants relied differentially upon recollection in the conservative and liberal conditions. Nonetheless, as noted earlier, the existing behavioral data suggest that criterion manipulations exert their influence almost wholly on familiarity rather than on recollection (Yonelinas, 2002). In light of this, perhaps a more important question for the current data is whether the response probability confound can explain the relatively greater positivity for old words in the conservative condition at mid-frontal electrode sites between 300 and 500 msec poststimulus. Is it possible that the condition-related differences are a reflection of the anterior projection of P300-related activity? The reason why this is unlikely is because although P300 activity should be greater for hits in the conservative than the liberal case, the opposite is true for correct rejections.

Given that the analyses of the mid-frontal effect revealed a greater relative positivity for old as well as for new test words in the conservative condition, it is difficult to conceive how these aspects of the ERP data can be explained via recourse to probability-related modulations of the P300 component.

Another aspect of the ERP data that is worthy of comment is the fact that the ERPs elicited by false alarms in the liberal condition did not differ reliably from those elicited by correct rejections in the 300–500 msec period. A similar finding can be seen in the data of Windmann, Urbach, et al. (2002), but these findings are seemingly at odds with those in previous studies, where the ERPs associated with false alarms have elicited a mid-frontal effect of comparable magnitude to that elicited by hits. In those studies, however, the false alarms were to lures that bore a close resemblance to studied items—singular/plural reversals (Curran, 2000), semantically related words (Nessler, Mecklinger, et al., 2001), and physically similar pictures (Curran & Cleary, 2003; Curran, Tanaka, et al., 2002). In all of these studies, a proportion of new items were less similar to test items than the similar lures, but the levels of behavioral performance precluded formation of reliable ERPs to false alarms for these items. Thus, the generality of the finding for false alarms in this experiment awaits further work. It is also worth emphasizing that the false alarm data were acquired in the liberal response condition. This may be important, as in this condition participants are likely to have accepted a new word as old on the basis of comparatively low levels of familiarity.

These previous findings therefore provide strong evidence in support of the view that the mid-frontal ERP old/new effect indexes familiarity, in keeping with previous proposals (Nessler, Mecklinger, et al., 2005; Curran, 2000; Mecklinger, 2000; Rugg, Mark, et al., 1998). An alternative account, however, is that this aspect of the electrical record is in fact an index of conceptual priming (Yovel & Paller, 2004). The basis for this proposal was the absence of a reliable mid-frontal ERP old/new effect in a study where the stimuli were unfamiliar faces (for other unanticipated null results and comments, see Curran & Friedman, 2004; Nessler & Mecklinger, 2003; Curran, Schacter, et al., 2001). Because there was adequate discrimination between old and new faces in this experiment, a mid-frontal old/new effect should have been observed for these stimuli if the effect is an index of a relatively amodal familiarity signal (Curran & Dien, 2003). The picture is complicated, however, by the fact that reliable mid-frontal ERP old/new effects have been obtained in another study in which the stimulus set comprised nonfamous faces (Nessler, Mecklinger, et al., 2005). There are also some data suggesting that a comparable effect can be obtained using impossible figures (Penney et al., 2001).

Setting these inconsistencies aside, it remains a possibility that in tasks where the stimulus sets are associated
with conceptual information, the mid-frontal ERP old/new effect indexes conceptual priming. According to one dual-process account, familiarity can arise from conceptual and/or perceptual processing fluency, followed by an attribution of that fluency to a prior encounter (Jacoby & Kelley, 1992; Jacoby & Dallas, 1981). Because this model allows for conceptual priming to contribute to familiarity, then that would suggest the effect is associated more generally with familiarity than with one process which may contribute to that basis for recognition memory judgments. Curran, Tepe, and Piatt (in press) provide a detailed overview of these issues and also provide some comments on regional brain activity that may generate the mid-frontal ERP old/new effect. The data that are presented here add nothing new to this issue, but the functional claims made in this article do not depend upon identification of the brain regions that are responsible for old/new effects.

Finally, what implications do these findings have for dual-process accounts of recognition memory? The ERP data acquired from mid-frontal electrodes support the claim that one process that is engaged during recognition memory tasks varies in a continuous fashion. The ERP data acquired from posterior electrodes post-500 msec provide some evidence for the engagement of a second process. The scalp distribution of this effect is similar to that observed in previous studies, and the data to date support strongly the claim that this effect is tied closely to recollection (Rugg, 2004; Friedman & Johnson, 2000; Wilding & Rugg, 1996). The data do not speak to questions concerning the functional properties of recollection—for example, whether it can be modeled as a threshold process (Yonelinas, 1994, 1997). Nor, for the reasons discussed above, is it possible to employ the ERP data to make strong inferences about the extent to which recollection was engaged in the liberal and conservative response conditions. In combination with findings in previous ERP studies, however, the data support strongly the claim that there are two processes which contribute to recognition memory judgments, and that one of the two—indexed by the mid-frontal ERP old/new effect—is best characterized as a process that acts as a continuous index of memory strength.

Acknowledgments

We thank Tim Curran and Jane Herron for their helpful comments on earlier versions of this manuscript. This research was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) and the Wellcome Trust.

Reprint requests should be sent to Nazanin Azimian-Faridani, School of Psychology, Cardiff University, Cardiff, CF10 3AT, UK, or via e-mail: nazanin@cardiff.ac.uk.

Notes

1. In a series of papers, Curran and colleagues have reported data from a dense electrode array and an average reference. Using this approach, differences at mid-frontal electrodes are accompanied by polarity-reversed differences (hits < correct rejections) over posterior scalp sites. The hits versus correct rejections differences are commonly restricted to mid-frontal scalp locations when a linked-mastoid reference is used, as in the current experiment (for review, see Curran, Tepe, et al., in press).

2. The labels “conservative” and “liberal” are relative terms here. It is irrelevant for present purposes whether participants adopted criteria falling either side of a designated neutral point. What is important is that the likelihood of an old response is greater in the liberal than in the conservative condition.

3. All analyses incorporated the Greenhouse–Geisser correction for violations of sphericity where appropriate (Greenhouse & Geisser, 1959).

4. The measures $d_a$ and $c_a$ take into account the fact that old item familiarity distributions can be characterized as having greater variability in familiarity strength than new item distributions (Hirshman & Hostetter, 2000; Ratcliff, Sheu, & Gronlund, 1992; Glanzer, Kim, Halford, & Adams, 1991; Macmillan & Creelman, 1991). The estimated slope of 0.8 was employed on the basis of the values obtained in other published studies (c.f. Verde & Rotello, 2003) and in an unpublished experiment in our laboratory where confidence rating data was obtained in a recognition memory task with the same encoding operations as in the present study as well as similar levels of discrimination.

5. ANOVA of the response accuracy and RT data for these 19 participants revealed a qualitatively similar pattern of reliable effects as for the 24 participants reported above, with the exception that in the analysis of the accuracy data the interaction between status and condition only approached significance ($p = .08$). Mean probabilities of old responses for this subset of participants were: conservative: old words .74, new .07; liberal: old .91, new .30.

REFERENCES


