Research Report

Magnetoencephalographic correlates of processes supporting long-term memory judgments

N.C. Bridson, S.D. Muthukumaraswamy, K.D. Singh, E.L. Wilding*

Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff, CF10 3AT, Wales, UK

ARTICLE INFO

Article history:
Accepted 23 May 2009
Available online 6 June 2009

Keywords:
Familiarity
ERF
ERP
Old/new effect
Recollection
Recognition memory

ABSTRACT

The sensitivity of event-related fields (ERFs) to memory retrieval processes is not well determined. This stands in sharp contrast to event-related potential (ERP) studies, as ERPs have been employed widely to address questions about the functional architecture supporting memory retrieval. Despite their success in this endeavour, however, the sensitivity of ERPs to one retrieval process—familiarity—is somewhat limited. This experiment was designed to determine the sensitivity of ERFs to familiarity, and thus to examine the functional leverage that is available to investigate item familiarity via magnetic means of indexing retrieval processing in real-time. The analyses of the ERF data focused on old/new effects, which are differences between the neural activities associated with old (previously studied) and new test items that attract correct memory judgments. The ERFs showed a level of sensitivity to changes in item familiarity superior to that reported previously in very similar studies where ERPs were acquired. Moreover, analyses of the ERF data revealed four functionally distinct old/new effects. These findings provide strong incentives for employing ERFs in subsequent studies of human memory retrieval processing operations.

1. Introduction

According to dual-process accounts, two independent processes support recognition memory judgments (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992; Mandler, 1980). The two terms that are now employed widely to denote these processes are Recollection and Familiarity (Yonelinas, 2002). Recollection is recovery of qualitative information about an episode. Familiarity provides quantitative information that permits judgments of prior occurrence. The status of this binary distinction, as well as the nature of any relationship between recollection and familiarity, remain controversial.

Behavioural data held to support dual-process accounts have been questioned recently (for relevant arguments, see Rotello et al., 2005; Wais et al., 2008; Yonelinas and Parks, 2007), as have findings in some animal experiments (Eichenbaum et al., 2008; Wixted and Squire, 2008) as well as those in neuropsychological and functional magnetic resonance imaging (fMRI) studies (e.g. Squire et al., 2007).

Another source of evidence that is relevant to these issues is event-related potential (ERP) data. A number of studies have revealed that at least two functionally and electrophysiologically distinct indices of explicit memory processes are engaged during recognition memory tasks (Curran et al., 2006; Donaldson et al., 2003; Friedman and Johnson, 2000; Mecklinger, 2000; Rugg and Curran, 2007; Wilding and Sharpe, 2003). The earlier of these is evident at anterior scalp sites between 300 and 500 ms post-stimulus. It comprises a greater
relative positivity elicited by old (previously studied) items than by new test items that attract correct judgments. This mid-frontal ERP old/new effect has been linked to familiarity (for reviews, see Curran et al., 2006; Rugg and Curran, 2007). The second effect evident primarily between 500 and 800 ms, also comprises a greater relative positivity elicited by old than by new items, and has a posterior scalp maximum which is typically left-lateralised for verbal test stimuli. This left-parietal ERP old/new effect has been linked to recollection (Allan et al., 1998).

If these functional accounts are correct, then the electrophysiological data points summarised above comprise strong evidence in support of dual-process accounts of recognition memory, and in this regard, the link between the left-parietal effect and recollection is strong (for a review, see Wilding and Sharpe, 2003). The link between the mid-frontal old/new effect and familiarity might be seen to be challenged, however, by the failure of the effect to respond to some manipulations that are thought to influence item familiarity. These include depth of processing (Rugg et al., 1998) and divided attention manipulations (Curran, 2004), as well as changes in the confidence associated with recognition memory judgments (Azimian-Faridani and Wilding, 2006; Curran, 2004; although see Woodruff et al., 2006). One plausible explanation for these insensitivities is a lack of power to detect changes in a small electrophysiological effect (see comments by Azimian-Faridani and Wilding, 2006; and power assessments reported by Curran, 2004). This argument may well be correct, but if it is, then it implies that the ERP index of familiarity, while providing data points that in combination with others are broadly consistent with dual-process accounts (e.g. Duzel et al., 2001), may be of only limited use in assessing the relative contributions that familiarity and recollection can make on a given task, or in assessing the conditions under which familiarity changes as a result of variables such as age or focal brain damage (Azimian-Faridani and Wilding, 2006).

Another cognitive neuroscience tool that might provide evidence relevant to these issues is event-related field (ERF) data. While the number of ERF studies of recognition memory is small in comparison to the number of ERP studies, the existing data indicate that ERFs index some retrieval processes, is not well established. This was one of the goals reviewed, see Wilding and Sharpe, 2003). The link between the mid-frontal old/new effect and recollection is strong (for a review, see Duzel et al., 2003, 2004; Guderian and Duzel, 2005; Neufang et al., 2006).1

Walla and colleagues have also acquired ERF data in a series of studies in which old/new recognition memory judgments were required, but in none of these studies were analyses of old/new effects reported (Walla et al., 1999, 2001a, b). Moreover, in the all of the MEG studies described so far, the magnetic contrasts were restricted to data acquired in a pair of conditions. Thus, while the findings described above suggest ERFs index memory processes, the findings in and of themselves provide limited information about the identity of those processes.

ERFs associated with three memory conditions of interest were, however, reported by Staresina et al. (2005). In their study, participants made old/new judgments to words in a modified recognition memory task. For the words judged to be old, participants also made a binary distinction between words for which the ‘old’ judgment was a high or a low confidence response. The analyses comprised paired contrasts between the ERFs associated with incorrect old judgments (misses) and with correct judgments (hits) that were separated according to confidence. The initial analyses revealed no effect that differentiated hits split according to confidence, but there were effects that differentiated these response categories from the ERFs associated with misses between 300 and 500 ms (Staresina et al., 2005). The distributions of these differences bear correspondences with those reported by Duzel and colleagues and while the insensitivity to response confidence makes it difficult to establish a link with either recollection or familiarity, the fact that the differences were revealed in contrasts between hits and misses suggests that in this time window ERFs index processes which contribute to the basis for an explicit memory judgment. It is somewhat surprising, however, that the modulations were restricted to the 300–500 ms time window, a finding which is at odds with the identification of later occurring ERF indices of memory processes in other studies (Duzel et al., 2003, 2005; Tendolkar et al., 2000).

These disparities across studies, and the fact that only paired contrasts have been reported in most ERF memory studies to date, mean that the sensitivity of ERFs to one or more memory processes, as well as the identity of those processes, is not well established. This was one of the goals motivating the experiment described here. A somewhat more specific goal, however, was to determine whether, and if so with what degree of resolution, ERFs index familiarity, in light of the potential sensitivity limitations of the mid-frontal ERP old/new effect described above. Towards this end, ERFs were acquired in a modified memory exclusion task (Jennings and Jacoby, 1993; 1997). Participants read aloud a list of words initially, and in a subsequent test phase these words were represented, interspersed with new (unstudied) words. Some of the new words were repeated at test after a lag of 7–9 intervening words. Participants were asked to respond on one key to old (studied) words and on another to new as well as to repeated test words.

1 In some other MEG studies of memory retrieval, the analyses have been conducted in source space only (Gonsalves et al., 2005; Lee et al., 2005; Dhond et al., 2005).
This task was selected for the ERF experiment because of the findings in ERP studies in which the same task was employed. Functionally dissociable mid-frontal and left-parietal ERP old/new effects have been obtained by contrasting the ERP old/new effects elicited by studied and repeated test words that attracted correct judgments (Bridson et al., 2006; Fraser et al., 2007). While the parietal effect has behaved in a way that is consistent with a recollection account (larger for old than for repeated test words), the mid-frontal effect has not, as it has been of equal magnitude for studied and repeated words. This finding might be seen to be at odds with a familiarity account because, using a continuous recognition memory task, Yonelinas and Levy (2002) showed that item familiarity decreased across conditions under which there were either 8 or 32 intervening items between first and second presentations. The repetition lag for repeated test items in the ERP tasks was 7–9 items, and the gap between presentation and re-presentation was markedly longer for studied than for repeated test words. As a result, it is reasonable to assume that these repeated test items should be more familiar than those presented at study and only once at test. Thus, it may be the case that the equivalent mid-frontal old/new effects for studied and repeated test words reflect limitations in the resolution with which ERPs index familiarity, a possibility that is lent some support by the finding that old/new effects in the ERP tasks was 7–9 items, and the gap between presentation and re-presentation was markedly longer for studied than for repeated test words. As a result, it is reasonable to assume that these repeated test items should be more familiar than those presented at study and only once at test. Thus, it may be

The use of this task with ERFs thus offers two important opportunities. First, an assessment of the number of memory processes that ERFs index. Second, an assessment of the sensitivity of ERFs to item familiarity. According to the foregoing argument, an ERF index of familiarity should be reliable for both old studied and repeated test words, but larger for the latter. Were this pattern to be observed, it would suggest that ERFs provide a more sensitive index of familiarity than do ERPs, and indicate their utility in assessing changes in item familiarity as well as predictions that follow from dual-process accounts of recognition memory.

2. Results

2.1. Behavioural data

Table 1 shows response accuracies and reaction times (RTs) for correct and incorrect responses to the three classes of test stimuli. Two measures of discrimination were calculated using Pr [p(hit) − p(false alarm)]: Snodgrass and Corwin (1988). The first was for discrimination between old (studied) and new test words (mean Pr = 0.60), while the second was for discrimination between old and repeated test words (mean Pr = 0.51). Both Pr values were reliably greater than 0 (t(15) > 15.00, p < 0.001 in each case), and Pr was superior for the old/new discrimination (t(15) = 5.63, p < 0.001). A one-way ANOVA on the RTs for correct responses revealed a main effect of response category (F(1.3, 19.9) = 6.10, p < 0.05). Follow-up paired comparisons revealed only that RTs for correct responses to new words were quicker than those for correct responses to old and repeated test words (t(15) = 2.78, p < 0.05 and t(15) = 3.54, p < 0.01, respectively).

2.2. MEG data

Fig. 1 shows scalp maps based on difference scores that were obtained by subtracting mean signal measures associated with correct rejections from those associated with hits (left-hand side) and repeated hits (right-hand side). The selection of time windows was guided by the findings in previous ERF (Staresina et al., 2005; Duzel et al., 2005) as well as related ERP (Bridson et al., 2006; Fraser et al., 2007) studies. The data are shown for the 300–500 and 500–800 ms epochs. Fig. 2 show ERFs for old and repeated test words averaged across sites for

![Fig. 1 - Spline-interpolated scalp distributions for the 300–500 and 500–800 ms epochs of the differences in signal strength obtained by subtracting mean values associated with ERFs elicited by correct rejections from those associated with correct judgments to old words (left-hand side) and repeated test words (FT = femtotesla). The magnitudes of the effects can be interpreted via recourse to the colour bar in the centre of the figure.](image)
specific clusters at left- and right-hemisphere anterior and posterior scalp regions, as detailed in the figure legend. The figures show that there are old/new effects in both epochs. At anterior locations, the distributions of the old/new effects become somewhat more inferior over time, and reverse polarity across the hemispheres. The old/new effect in the earlier epoch is larger for repeated than for old test words, while the reverse is true for the effect in the later epoch. At posterior locations, there is an old/new effect in the 300–500 ms epoch which is common to old and repeated test words. In the later epoch, there is a left-lateralised old/new effect that is evident primarily for old words.

Table 2 – Outcomes (F-values) of the paired contrasts between the ERFs elicited by correct responses to New, Old and Repeated (Rep) test words for the 300–500 and 500–800 ms epochs at anterior (left-hand side of table) and posterior locations.

<table>
<thead>
<tr>
<th></th>
<th>Anterior</th>
<th>300–500 ms</th>
<th>Posterior</th>
<th>500–800 ms</th>
<th>Posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>Old vs. New ns</td>
<td>Rep vs. New ns</td>
<td>Old vs. Rep ns</td>
<td>Old vs. New ns</td>
<td>Rep vs. New 5.80 * ns</td>
</tr>
<tr>
<td></td>
<td>CAT×HM 12.40 **</td>
<td>45.67 ***</td>
<td>4.94 * 28.06 ***</td>
<td></td>
<td>10.53 ** ns</td>
</tr>
<tr>
<td></td>
<td>CAT×CH ns 4.61 * 59</td>
<td>17.54 ***</td>
<td>ns ns ns</td>
<td>ns ns 7.18 * ns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAT×HM×CH 7.54 ** (69)</td>
<td>17.54 ***</td>
<td>ns ns ns</td>
<td>ns ns 3.44 * 0.90 ns</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CAT</th>
<th>Old vs. New ns</th>
<th>Rep vs. New ns</th>
<th>Old vs. Rep ns</th>
<th>Old vs. New ns</th>
<th>Rep vs. New ns</th>
<th>Old vs. Rep ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT×HM</td>
<td>62.20 *** 56.66 ***</td>
<td>11.80 **</td>
<td>ns ns 4.54 * 0.79</td>
<td>ns 3.44 * 0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT×CH</td>
<td>ns 29.96 *** (59)</td>
<td>4.12 * 0.75</td>
<td>3.62 * 0.96</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT×HM×CH</td>
<td>34.89 ** (0.69)</td>
<td>4.12 * 0.75</td>
<td>3.62 * 0.96</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
</tbody>
</table>

CAT=response category, CH=chain, HM=hemisphere.
*** =p<0.001.
** =p<0.01.
* =p<0.05, ns=non-significant. Epsilon values are bracketed and in subscript.
The ERF mean signal measures were subjected to an initial global ANOVA including the factors of response category (3 levels: correct responses to old, new and repeated test words), epoch (2: 300–500 and 500–800 ms), the anterior/posterior dimension (3: anterior, central, posterior), hemisphere (2: left, right) and chain (3: inferior, mid-lateral and superior regions). This global analysis revealed an interaction between all five factors ($F(4.2, 62.8)=14.07$, $p<0.001$), and the highest-order (four-way) interaction terms remained reliable when analyses including the same response category and location factors were conducted separately for each epoch (300–500 ms: $F(3.4, 51.1)=5.67$, $p<0.01$; 500–800 ms: $F(3.3, 50.2)=13.65$, $p<0.01$).

In light of this pattern of reliable effects, all possible paired contrasts were computed separately at anterior and posterior locations in each epoch, with the results shown in Table 2. These follow-ups do not include analyses for central locations, as Fig. 1 illustrates that the maxima of the effects of interest are at anterior and posterior locations in both epochs.

2.3. Analyses at anterior locations

Table 2 shows that reliable ERF old/new effects were revealed for old words and for repeated test words at anterior locations between 300 and 800 ms. For both the 300–500 and 500–800 ms time periods, the analyses of the old/new effects (Old vs. New and Rep vs. New) revealed reliable three-way interactions between category, hemisphere and chain, which reflect the inferior maxima of the effects, in combination with the polarity reversals across the hemispheres. These differences are clearly evident in the plots of signal strength depicted in Fig. 3. The same interaction term for the repeated vs. old comparison in the 500–800 ms epoch reflects the larger old/new effect for old than for repeated test words. By contrast, the interaction between category and hemisphere for repeated vs. old words in the 300–500 ms epoch is due to the larger old/new effect for repeated test words.

2.4. Analyses at posterior locations

In the 300–500 ms epoch, the repeated and old word old/new effects were statistically equivalent, with the category by hemisphere interactions reflecting the polarity reversal of the effects across hemispheres. In the later epoch, there were reliable old/new effects at posterior locations for old words only, with the interactions involving chain reflecting the left-superior maximum of the effect.

2.5. Topographic analyses

In a final analysis stage, the ERF old/new effects were subjected to analyses of scalp distributions, which were carried out on data rescaled to avoid confounding differences between the amplitudes of old/new effects with differences between their scalp distributions (McCarthy and Wood, 1985). The data from each of the 18 regions were rescaled using the max-min method, with the data submitted to rescaling comprising the difference scores that were obtained by subtracting the mean field strengths for correct rejections from those for correct responses to old and repeated test words, respectively (Urbach and Kutas, 2002, 2006; Wilding, 2006). The initial analysis included the factors of epoch (300–500, 500–800 ms), category (old, repeated), the anterior/posterior dimension, hemisphere and chain. This analysis revealed an interaction between all five factors ($F(2.4, 36.1)=4.95$, $p<0.01$) which moderated a three-way interaction between epoch, category and hemisphere ($F(1, 15)=11.20$, $p<0.01$). Follow-up analyses within each epoch revealed a reliable effect in the 500–800 ms epoch only, which comprised a three-way interaction between category, the anterior/posterior dimension and hemisphere ($F(1, 15)=11.20$, $p<0.01$). Follow-up analyses within each epoch revealed a reliable effect in the 500–800 ms epoch only, which comprised a three-way interaction between category, the anterior/posterior dimension and hemisphere ($F(1, 15)=11.20$, $p<0.01$). This outcome reflects the presence of a posteriorly distributed old/new effect for old words only (see Figs. 1 and 2), and an anteriorly distributed effect that is evident also for

Fig. 3 – Mean signal strength measures (Femtotesla: fT) of the ERF old/new effects for Old and Repeated (Rep) test words for six regions covering anterior and posterior scalp locations for the 300–500 and 500–800 ms epochs. Note the different y-axis strength ranges across the four panels.
repeated test words, but larger for old words. In keeping with this claim, separate analyses at anterior and posterior locations in the 500–800 ms epoch revealed a reliable interaction between category and hemisphere at posterior locations only ($F(1,15)=7.46, p<0.05$).

3. Discussion

Analysis of the event-related field (ERF) data revealed evidence for up to four functionally dissociable processes in the 300–800 ms post-stimulus time period. These were obtained in a task where the pattern of response accuracy mapped closely onto that obtained in previous studies in which ERPs, rather than ERFs, were recorded (Bridson et al., 2006; Fraser et al., 2007). This claim also holds for the reaction time (RT) data, where RTs were fastest for the category associated with the highest level of response accuracy: correct responses to items presented for the first time at test were made quickly than were responses to items that were presented either at study and test, or that were re-presented during the test phase.

The evidence for the existence of four functionally dissociable processes was revealed in analyses of the ERF old/new effects that were elicited by old and repeated test words that attracted correct task judgments. The different ways in which these effects responded to the retrieval task manipulations can be seen clearly in the plot of mean signal strength measures in Fig. 3, and the findings for the 300–500 and 500–800 ms epochs are discussed in turn below.

3.1. 300–500 ms

A specific question motivating this study was the sensitivity of ERFs to familiarity, and the findings in this epoch are relevant to this issue. There were reliable ERF old/new effects at anterior and at posterior locations. The ERF old/new effects distinguished between old and repeated test words at anterior locations only, with larger effects associated with repeated test than with old words. The polarity of this old/repeat difference (as well as the polarity of the old/new difference) reversed over left- and right-frontal scalp regions, suggesting that these left- and right-hemisphere modulations are generated by the same magnetic field as it enters and leaves the head. Thus, the effects over these lateral scalp regions are likely to comprise contributions from the same generator (or set of generators) and index the same process/es. The same argument can be applied to the posteriorly distributed effect in this epoch as well as the frontally distributed effect in the 500–800 ms epoch (see Curran, 1999, for related comments for ERP data acquired in retrieval tasks). In all of these cases, the similarities between the time courses of the lateralised effects also argue for the view that they share the same neural generators. It is not impossible for some or all of these lateralised effects to have different neural sources, but even if this turned out to be the case, their comparable sensitivities to the experiment manipulations leave little room here to argue that they index functionally distinct processes.

The fact that the sensitivity of ERFs to the old/repeated test status of words differed over anterior and posterior scalp in the 300–500 ms is consistent with the view that two functionally distinct memory-related processes were engaged, and a similar claim has been made in ERP studies of memory retrieval: a posterior effect with a similar time course has been shown to be sensitive only to the old/new status of test items, while the anterior effect has varied according to the accuracy of memory judgments (for a visualisation of this pattern of ERP data in the same exclusion task employed in this ERP study, see Fig. 3 in Fraser et al. (2007); for the first demonstration of this dissociation, see Rugg et al. (1998)). In light of these findings it has been proposed that the posterior ERP repetition effect is an index of implicit memory, while the anterior ERP old/new effect indexes familiarity (Rugg et al., 1998).

Can the posterior ERF old/new effect described here be considered as a functional homologue of the posteriorly distributed ERP old/new effects that have been observed previously over this time period? The insensitivity of this ERP effect to the old/repeated status of old words (see the LPM region ERFs in Fig. 1), coupled with its time course, provides some support for the view that this effect is a functional homologue of the repetition effect that has been observed in comparable ERP studies. An important endeavour in subsequent ERP (as well as ERF) studies will be to establish links between this early posterior effect and behavioural indices of implicit memory, such as reaction time changes to repeated stimuli in indirect and/or direct retrieval tasks (Azimian-Faridani and Wilding, 2006; Grove and Wilding, 2009).

In contrast to the posterior ERP old/new effect described above, links between the anterior ERP old/new effect in the 300–500 ms epoch and frontally distributed ERP old/new effects with similar time courses are less immediately apparent. This is because the ERP old/new effect was larger for repeated test than for old words, whereas, as described in the Introduction, this has not been observed in this task for the mid-frontal ERP effect (compare the LAI and RAI regions in Fig. 1 with Bridson et al. (2006); Fraser et al., 2007). While it is possible that this ERP effect and the mid-frontal ERP old/new effect have different functional properties, the correspondences between their time courses and scalp distributions (which were equivalent for old and for repeated test words) provide a reasonable basis for the view that they are functionally homogeneous. In this regard, it is noteworthy that an ERF field acquired with radial gradiometers that shows a polarity reversal across right- and left-hemisphere frontal scalp is consistent with a corresponding electrical field that manifests itself primarily at frontal midline scalp locations (Vrba and Robinson, 2001). This observation supports the view that this early ERP old/new effect is a neural as well as a functional homologue of the mid-frontal ERP old/new effect. Thus, rather than being taken as evidence for the view that these ERP and ERF effects have different functional correlates, the differences across these imaging modalities for measures of neural activity associated with old and repeated test words can be regarded as evidence for greater sensitivity of ERPs than ERFs to the same process.

As noted in the Introduction, the mid-frontal ERP old/new effect has been linked with familiarity, and if this is the case then it is arguably surprising that the ERP correlate of this process is insensitive to differences between the familiarity of old and repeated test items. Yonelinas and Levy (2002) have shown that the familiarity of test items declines over periods
One set of reasons for this advantage in sensitivity for ERFs over ERPs can be developed from consideration of the factors that determine the electrical and magnetic signals that can be detected by sensors located outside the skull. Relative to ERFs, the neural activity indexed at a location (or locations) by ERPs may reflect contributions from a larger number of distinct neural populations (Mosher et al., 1993). If these populations have overlapping time courses and do not support entirely the same cognitive operations, then the possibility that some patterns of neural activity that are of functional interest will be obscured is greater for ERPs than for ERFs. The likelihood of this occurring for ERPs is increased further because of the distortion that ERPs (but not ERFs) are subject to as they pass from source to recording sensor (Hamalainen et al., 1993; Mosher et al., 1993). Another reason why MEG may on occasions provide greater sensitivity than EEG is because of the superior spatial sampling that can be achieved with markedly less preparation time. When coupled with data reduction techniques such as averaging signals across clusters of sensors—the approach adopted here—this may result in superior sensitivity for ERFs than for their electrical counterparts.

It may well turn out to be the case that all of these factors contribute to some degree to the differential sensitivities of ERFs and ERPs to familiarity. Irrespective of the relative contributions of these factors, however, the central point here is the possibility that ERFs may provide a more sensitive index of changes in familiarity than do ERPs. As a result, early frontal ERF old/new effects may offer a broader functional window for investigating the role of familiarity in retrieval processing than is available via the analysis of ERP old/new effects.

These claims only hold, however, if it is reasonable to assume that familiarity should in fact be greater for repeated test words than for old (studied) words. In this regard, there is a confound in the present experiment because, in addition to a different average interval between presentation and re-presentation of old and repeated test words, the cognitive operations to which they were subjected upon first presentation were not equivalent. For old words, the task was simply to read words aloud. For repeated test words, the operations engaged upon first presentation at test were those necessary to make appropriate judgments to test items. There is no basis for asserting confidently how subsequent familiarity would be affected by the different operations engaged during the first presentation of old and repeated test words. The first presentations of old as well as repeated words can be regarded as “intentional” encoding conditions, in so far as encoding information about the context (the study or test list) in which they were encountered would benefit the decision that is required upon re-presentation, but clearly it is very unlikely that they were subjected to entirely the same sets of cognitive operations. It will be important in subsequent experiments to avoid confounds of this kind, and in addition to include a behavioural measure that will permit separation of ERF signals into categories that can be linked more directly to recollection and to familiarity than is the case in the experiment described here.

3.2. 500–800 ms

There was also evidence for two functionally dissociable old/new effects in this later epoch. One indicator is the reliable interaction between response category and location in the analyses of scalp distributions, in combination with the outcomes of the follow-up analyses of field strengths at anterior and posterior locations. These showed that the old/new effects at these different locations were not equivalent for old and for repeated test words. In what is the most directly comparable published ERP exclusion task study (Experiment 2: Fraser et al., 2007), there was evidence for only one old/new effect in this epoch.

Common to the ERF data reported here, and many ERP studies in which old/new effects have been analysed, is a posteriorly distributed old/new effect in the 500–800 ms epoch (see the LPM and LPS region ERFs in Fig. 2). This oft-reported ERP effect comprises a greater relative positivity for old than for new items attracting correct memory judgments, extends from approximately 500 to 800 ms post-stimulus, and is typically largest at left-parietal electrode locations. Critically, in exclusion tasks where the level of response accuracy is relatively high, this effect is markedly larger for old than for repeated test words that attract correct judgments (Czernochowski et al., 2005; Dywan et al., 1998, 2002, 2001; Dzulkifli et al., 2006; Dzulkifli and Wilding, 2005; Herron and Rugg, 2003; Herron and Wilding, 2005; Wilding et al., 2005).

A comparable sensitivity to the old/repeated test status of critical words is evident in the posteriorly distributed ERF old/new effects that were obtained here: there was a reliable posteriorly distributed old/new effect in the 500–800 ms epoch for old words only. The similarities between the time course of this ERF effect and the left-parietal ERP old/new effect, when coupled with their comparable sensitivities to the old/repeated test manipulation, support strongly the claim that they index the same process, and the ERP literature converges on the view that this process is recollection (Rugg and Allan, 2000).2 The claim that a very similar ERF old/new effect indexes recollection has been made before (Tendolkar et al., 2000), and the current findings bolster that proposal by showing for the first time that the ERP and ERF old/new effects respond in the same way to items that are repeated in a retrieval task after different intervals.

Moreover, there is a functional double dissociation between this late parietal ERF old/new effect and the earlier (300–500 ms) frontal effect already described. The parietal 2 The reasons why ERP indices of recollection will be larger for old than for repeated test words in some exclusion tasks have been articulated in detail elsewhere and are not recapitulated here (see Dzulkifli et al., 2006; Herron and Rugg, 2003; Wilding and Herron, 2006). The critical points for present purposes are the correspondences between the sensitivities of ERP and ERF old/new effects to experiment manipulations, alongside considerations of their temporal and spatial similarities.
effect was reliable only for old words, while the earlier frontal effect was reliable for old as well as repeated test words, and reliably larger for the latter. Because these two effects behave as correlates of recollection and of familiarity, respectively, these data points provide strong evidence in support of the claim that these two temporally and functionally distinct processes are engaged during exclusion tasks, which is consistent with dual-process accounts of recognition memory. It is also notable that, despite the large number of ERP studies of memory retrieval, there are few demonstrations that putative ERP correlates of recollection and familiarity can be doubly dissociated in the same experiment (Jager et al., 2006; Woodruff et al., 2006; Stenberg et al., 2009).

In addition to this putative index of recollection, a second old/new effect with different functional properties was evident at anterior locations in the 500–800 ms epoch (see the LAI and RAI region ERPs in Fig. 1). This effect was reliable for old as well as for repeated test words. It was, however, reliably larger for the former, thereby separating it from the posteriorly distributed effect in the same epoch, as well as doubly dissociating it from the frontally distributed ERP effect in the preceding epoch, where greater relative field strength was associated with repeated test rather than with old words. The anterior distribution of this effect encourages attempts to link this ERP old/new effect to retrieval processing operations supported by prefrontal cortex, of which processes that operate on the products of retrieval are one set of candidates. An accurate functional characterisation, however, awaits further information about the ways in which this effect is influenced by task and stimulus manipulations.

In summary, these findings in an ERF study of memory retrieval provide some support for dual-process accounts of recognition memory. They also provide a strong incentive to acquire ERPs in other retrieval tasks. First, to determine whether there is converging evidence for the links between the memory-related ERP modulations described here and the processes of recollection, familiarity, implicit memory and post-retrieval operations. Second, to explore further the possibility that employing MEG in studies of memory retrieval will provide functional leverage regarding changes in item familiarity (and possibly other retrieval processes) over and above that which can be obtained when ERPs are acquired.

4. Experimental procedures

4.1. Participants

17 completed the experiment. The data set from one was discarded because of failure to follow the task instructions. The average age of the remaining participants was 21 years (range 18 to 23, 13 female). All were right-handed native English speakers, and none were taking neuroleptic medication at the time of testing or reported a history of mental illness. All were paid at a rate of £7.50/hour and gave informed consent.

4.2. Stimuli

Words were taken from the MRC psycholinguistic database (frequency range = 1–7 occurrences per million, length = 4–9 letters: Coltheart, 1981). They were presented in white letters on a black background, projected onto a flat screen placed 0.5 m from participants. The words subtended a maximum of 3° of visual angle horizontally and 1° vertically.

4.3. Design

400 words were split into sixteen equal groups of 25 words. These groups were assigned randomly into four sets, each containing 4 word groups. Each set comprised the stimuli for one of four study-test cycles. One complete experiment list comprised all four study-test cycles. For each cycle, there were 25 words (one whole group) on each study list. There were 125 words on each test list. These comprised the 25 words presented at study along with 50 unstudied words (2 word groups), and a further 25 unstudied words which then repeated once after 7–9 intervening words. No word appeared in more than one test list. The groups within each cycle were rotated so that, across lists, all words appeared at test as a new word, an old word, and a repeated word. This procedure resulted in the creation of four complete task lists, and a further four were created by changing the order in which the study-test cycles were presented. The order of presentation of words in each of the study lists was determined randomly for each participant. The order of presentation of test words was determined pseudo-randomly, and the numbers of words represented after lags of 7, 8, and 9 words were approximately equal. In total, each participant saw 600 presentations of words (100 study words, 500 test words).

4.4. Procedure

Prior to the experiment, participants were informed that they would see words on a computer screen that would be shown one at a time. They were informed that they would complete four study-test cycles, and that words encountered in each cycle would not appear in another cycle. At study, participants were asked to read each word aloud. Each study trial began with an asterisk (*), which was displayed for 1000 ms and followed by a blank screen (100 ms), after which the study word was presented for 300 ms. The asterisk signalling the start of the next trial appeared 2000 ms later. Each test trial started with a ‘Blink Now’ screen (1500 ms duration) followed by a fixation asterisk (500 ms), which was removed from the screen 100 ms prior to presentation of a test word (300 ms). The screen was then blanked until the participant responded, and the next trial started 1700 ms afterwards.

Participants were asked to press one response button to words that were seen at study (old words), and another to unstudied test words (new words), as well as repeated test words. Responses were made on a key-pad with the index and middle fingers of the same hand. The fingers and hands used for responses were balanced across participants, who were all asked to restrict their blinking to the period when the ‘Blink Now’ instruction was visible. The RT criterion for trial inclusion in behavioural and ERP analyses was that responses were slower than 300 ms and faster than 4000 ms. No trial was rejected according to this criterion. There was a 1 minute interval between study and test phases within each cycle, and a 3 minute interval between cycles. During both of these
intervals, participants were reminded of the response requirements in the following study or test phase.

4.5. MEG acquisition

Whole head MEG recordings were made using a 275-channel radial gradiometer system (VSM MedTech, Canada). An additional 29 reference channels were recorded for noise cancellation purposes and the primary sensors were analysed as synthetic third order gradiometers (Vrba and Robinson, 2001). 4 of the 275 channels were removed due to excessive sensor noise. The sampling rate was 1200 Hz and recordings were filtered off-line with a bandpass of 0.03 to 40 Hz. Intra-individual head movement was kept to a minimum, and head position was localised at the start and finish of each study-test block. ERFs were epoched off-line into 1100 ms epochs, with a 200 ms pre-stimulus baseline, relative to which all mean amplitudes were computed. Trials containing large signal artefact and those containing obvious blink-related activity were rejected (mean number of trials rejected per participant=8%, range=1–14%). Baseline corrected averaged ERFs were formed for correct judgments to old, new and repeated test words. Participants were excluded if they did not contribute at least 16 artefact free trials to each of the response categories associated with correct responses. There were insufficient trials for some participants to permit formation of averaged ERFs associated with incorrect test responses. Mean number of trials was 63 (range 40–83), 267 (range 190–293) and 79 (range 56–91), for correct responses to old, new and repeated test words.

4.6. MEG data reduction

The ERFs recorded from clusters of adjacent sites were averaged together, forming 18 regions in a 6×3 grid covering anterior (A), central (C) and posterior (P) scalp over the left (L) and right (R) hemispheres at inferior (I), mid-lateral (M) and superior (S) regions (for a similar approach for multi-channel ERP data, see Curran, 1999, 2000). Triplets of these letter combinations (e.g. left anterior superior=LAS) are employed in the remainder of this manuscript to denote specific scalp regions. The data points in each cluster represent the average signal from 6 MEG sensors, and a complete list of the sensors in each cluster is provided in the Appendix. Averaging ERFs across clusters of proximal sensors also ameliorates concerns about variability in fixed individual sensor locations across participants relative to the underlying generators.

4.7. Behavioural and MEG data analysis strategy

Measures of response accuracy and discrimination were assessed using t-tests. Reaction times and MEG data were assessed using repeated measures ANOVAs, with follow-up analyses comprising subsidiary ANOVAs and paired t-tests as appropriate. The repeated measures ANOVAs incorporated the Greenhouse–Geisser correction when necessary (Greenhouse and Geisser, 1959; Winer, 1971), and corrected degrees of freedom are shown for the ANOVAs reported in the text. In tabulated outcomes, epsilon values are reported alongside full degrees of freedom. For all ERF analyses, only effects involving the factor of category are reported as they are of principal interest here. The time courses of the ERF and ERP old/new effects described in the Introduction guided the partitioning of the ERF data into the 300–500 and 500–800 ms post-stimulus epochs. The outcomes of somewhat more fine-grained analyses from 0 to 800 ms with data averaged for successive 100 ms epochs confirmed the validity of this approach.

Acknowledgments

This work was completed with support from the Wales Institute of Cognitive Neuroscience (WICN) and the UK Biotechnology and Biological Sciences Research Council (BBSRC). The authors would like to thank Emma Bridger, Damian Cruse, Yee Ying Yick, Rachael Elward and Lisa Evans for comments on drafts of this paper.

Appendix

Individual MEG sensor locations contributing to each of the 18 regions over which individual sensor data were averaged. L=left, R=right, A=anterior, C=central, P=posterior, I=inferior, M=mid-lateral, S=superior.

LAI: LT22, LT31, LT32, LT41, LT42, LT51
LAM: LF34, LF45, LF55, LT11, LF35, LT21
LAS: LF42, LF52, LF62, LS32, LF53
RAS: RF42, RF52, RF62, RF32, RF53
RAM: RF43, RF45, RF55, RT11, RF35, RT21
RAI: RT22, RT31, RT32, RT41, RT42, RT51
LCA: LT23, LT24, LT34, LT35, LT43, LT44
LCM: LT13, LT14, LF66, LC16, LP57, LP45
LCS: LC24, LC31, LC53, LP35, LP23, LP12
RCM: RT13, RT14, RF66, RC16, RP57, RP45
RCI: RT23, RT24, RT34, RT35, RT43, RT44
LPI: LT37, LT46, LT47, LT55, LT56, LT57
LPM: LT27, LO14, LO34, LO24, LO23, LP54
LPS: LO11, LP52, LF41, LP53, LO12, LO13
RPS: RO11, RP52, RP41, RP53, RO12, RO13
RPM: RT27, RO14, RO34, RO24, RO23, RP54
RPI: RT37, RT46, RT47, RT55, RT56, RT57

References


