ELSEVIER

Neuropsychologia 44 (2006) 1120–1130

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Memory retrieval processing: Neural indices of processes supporting episodic retrieval

M.A. Dzulkifli, J.E. Herron, E.L. Wilding *

School of Psychology, Cardiff University, Cardiff CF10 3AT, Wales, UK

Received 5 April 2005; received in revised form 12 October 2005; accepted 21 October 2005
Available online 5 December 2005

Abstract

Event-related potentials (ERPs) were acquired during separate test phases of a verbal recognition memory exclusion task in order to contribute to current understanding of the functional significance of differences between ERPs elicited by new (unstudied) test words, which are assumed to index processes engaged in pursuit of task-relevant information. Participants were asked to endorse old words from one study task (targets), and to reject new test words as well as those from a second study task (non-targets). The study task designated as the target category varied across test phases. The left-parietal ERP old/new effect – the electrophysiological signature of recollection – was reliable for targets and for non-targets in all test phases, consistent with the view that participants recollected information about both of these classes of test word. The contrast between the ERPs evoked by new test words separated according to target designation revealed no reliable differences. These findings contrast with those in a recent study in which the same tasks were used, but in which the accuracy of task judgments was markedly higher (Dzulkifli, M.A., & Wilding, E. L. (2005). Electrophysiological indices of strategic episodic retrieval processing. *Neuropsychologia, 43*, 1152–1162). In that study, there were reliable differences between the ERPs evoked by the two classes of new words, but reliable left-parietal ERP old/new effects for targets only. In combination, the findings suggest that differences between ERPs evoked by new test words can reflect processes that are important for controlling what kinds of information will and will not be recollected.

Keywords: Episodic memory; Recollection; Event-related potentials; Retrieval orientation; Retrieval processing

1. Introduction

According to the principle of transfer appropriate processing, the likelihood of successful retrieval increases along with the extent to which the processes engaged during retrieval recapitulate those that were engaged during encoding (Lockhart, 2002; Morris, Bransford, & Franks, 1977). For example, Morris et al. (1977) required participants to complete encoding tasks in which the focus was on either the semantic or phonemic properties of words. Superior retrieval for words encoded semantically was observed when the retrieval task was old/new recognition memory, a finding consistent with the levels of processing framework (Craik & Lockhart, 1972). The opposite pattern of findings – superior retrieval for words encoded with respect to their phonological properties – was observed when the task required participants to determine whether test words rhymed with those presented during encoding.

These findings (for a careful commentary, see Nairne, 2002) emphasise that the success or failure of retrieval is determined at least in part by the processes that are engaged at the time of a retrieval attempt, and this assumption is central to accounts of the way in which memory retrieval can be influenced by processes that operate during retrieval tasks. For example, on the basis of a detailed protocol analysis Burgess and Shallice (1996) identified cue-specification and elaboration as processes that are engaged in pursuit of task-relevant memories. These processes operate directly upon retrieval cues in order to influence the likelihood of recovery of task-relevant information. Broadly in keeping with the transfer appropriate processing principle (Morris et al., 1977), one way in which they might accomplish this is by maximising overlap between a retrieval cue and a target memory.[1]

* Corresponding author. Tel.: +44 2920875048; fax: +44 2920874858.
E-mail address: wildinge@cardiff.ac.uk (E.L. Wilding).
URL: http://www.cardiff.ac.uk/psych/home/wildinge/index.html.

0028-3932/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2005.10.021

[1] The concepts of cue-bias (Anderson & Bjork, 1994) and focusing (Schacter, Norman, & Koutstaal, 1998) encapsulate broadly similar ideas.

Rugg and Wilding (2000) proposed that the engagement of retrieval cue processing is a consequence of the adoption of an appropriate retrieval orientation—a task set that determines the processes that will be set in train when a retrieval cue is encountered (see also Donaldson, Wilding, & Allan, 2003; Wilding, 1999). Rugg and Wilding (2000) proposed that task-specific indices of cue-processing – the outcome of the successful adoption of an appropriate orientation – can be measured by contrasting the neural activity that is elicited by unstudied test items in retrieval tasks having different retrieval demands. Other factors being equal, such contrasts should reveal processes that are engaged around the time of retrieval and which are not contaminated by activity that occurs as a result of successful retrieval (Rugg & Wilding, 2000).

With few exceptions (e.g. Ranganath, Johnson, & D'Esposito, 2000), the studies in which these contrasts have been made have employed event-related potentials (ERPs). The findings in these studies provide support for the concept of orientation in so far as the divergences between ERPs evoked by classes of new items have differed across studies (Dzulkifli, Sharpe, & Wilding, 2004; Johnson, Kounios, & Nolde, 1997; Ranganath & Paller, 1999, 2000; Robb & Rugg, 2002; Rugg, Allan, & Birch, 2000; Wilding, 1999; Wilding & Nobre, 2001). This statement remains accurate, moreover, for those studies in which retrieval effort has been ruled out – at least to a reasonable degree – as an explanation for the differences between the critical classes of ERPs (Dzulkifli et al., 2004; Ranganath & Paller, 2000; Robb & Rugg, 2002).

In one recent study, Dzulkifli and Wilding (2005) reported differences between classes of ERPs evoked by new items that were acquired during two different retrieval tasks. Participants studied words initially, all of which were concrete nouns. For 50% of the words, the task was to decide how difficult the object denoted by each word would be to draw. For the remainder, the task was to generate uses for the object denoted by each word. Old and new (unstudied) words were presented at test, and participants completed exclusion tasks (Jacoby, 1991). In this task, one class of studied item – for example, words encoded in the function task – is designated as belonging to the *target* category. The other class of old item is designated as belonging to the *non-target* category. On each trial of an exclusion task, participants are asked to respond on the same key to new items and to old items designated as non-targets, while responding on a different key to old items designated as targets.

In this experiment, there were separate blocks. In each of these, words encoded under either function or drawing task instructions were designated as targets. This meant that it was possible to contrast the ERPs evoked by new test words in separate test blocks where all that differed was whether words encoded under function or drawing instructions were designated as targets. The new test words elicited in the function target designation condition were more positive-going from 500 to 1000 ms, particularly at frontal and central scalp locations. There was a tendency for these differences to be larger over the right than the left hemisphere. Dzulkifli and Wilding (2005) noted that these differences were unlikely to be due to differences in task difficulty, because of the very similar pattern of behavioural data that was obtained in the two target designations. In light of this, they interpreted these differences as ERP indices of retrieval orientations. More specifically, they proposed that these indices of orientation reflected processes that were important for restricting recollection to task-relevant information. This specific claim concerning selective recollection – see also Herron and Rugg (2003a) – was based on consideration of the ERPs evoked by new items alongside the pattern of left-parietal ERP old/new effects that was obtained for targets and for non-targets.

The left-parietal ERP old/new effect is typically largest at parietal scalp locations over the left hemisphere. It is evident primarily from 500 to 800 ms, and comprises a relatively greater positivity for correct memory judgments to old compared to new items (Friedman & Johnson, 2000; Rugg, Herron, & Morcom, 2002). The weight of evidence suggests that the effect is an electrophysiological index of recollection (Wilding & Sharpe, 2003), and the pattern of ERP old/new effects observed by Dzulkifli and Wilding formed the basis for their functional claim concerning the differences between ERPs evoked by new test items. They observed reliable left-parietal ERP old/new effects for targets only, which suggests that participants prioritised recollection of information about targets over information about non-targets. This data formed the basis for the proposal that the indices of retrieval orientation they obtained reflected retrieval processing that permitted selective recollection of task-relevant information.

An important question, however, is why a strategy of prioritising recollection of information about targets might be adopted during completion of exclusion tasks. In a number of recent exclusion studies, the parietal ERP old/new effects for targets have been markedly larger than the effects for non-targets (Dzulkifli & Wilding, 2005; Herron & Rugg, 2003a; Herron & Wilding, 2005). Herron and Rugg (2003b) have suggested that this comes about because when the likelihood of recollecting targets is high, participants rely primarily on the success or failure of recollection of information about targets to make task judgments. They proposed that all items failing to elicit target recollection (non-targets as well as new items) are given a 'new' response. Herron and Rugg also suggested that the efficacy of this strategy diminishes as the likelihood of recollecting information about targets decreases, and these proposals were motivated by their finding that left-parietal ERP old/new effects for non-targets were reliably larger in an experiment where the likelihood of target recollection was low than in an experiment where the likelihood was somewhat higher. Critically, the task associated with non-targets in these two experiments was the same, as was accuracy of non-target judgments.

The findings in the study of Dzulkifli and Wilding (2005) described earlier are in line with this account, since the accuracy of target judgments was high (>0.80) in both target designations, and reliable left-parietal ERP old/new effects were revealed for targets only. These findings are therefore in keeping with the view that participants prioritised recollection of target information over non-target information, and it was this aspect of the data that Dzulkifli and Wilding (2005) relied upon in order to support their proposal that the differences between the ERPs evoked by new test words and separated according to target designation in their experiment indexed processes that were important for the selective retrieval of contextual information associated with targets.

The study described here builds on the previous work of Dzulkifli and Wilding (2005) as well as that of Herron and Rugg (2003a). It was designed in order to test the proposal that the differences between the ERPs evoked by new items in our previous study index processes that are responsible for selective recollection of different kinds of contextual information. The proposal we offered was that the differences between the ERPs evoked by new items in the function and drawing target designation conditions reflected the fact that in one case participants were prioritising recollection of information associated with the function task, and in the other case they were prioritising recollection associated with the drawing task. The critical observation is that if this account is correct then the differences between the ERPs evoked by new items reported by Dzulkifli and Wilding (2005) should diminish when the same task is completed by relying in all cases upon recollection of information from both the function and the drawing task to a greater degree than was the case in our previous experiment.

This prediction was tested by maintaining the same encoding tasks employed by Dzulkifli and Wilding (2005), while lowering the accuracy of target judgments by reducing the number of study–test cycles and increasing the lengths of study–test intervals. According to the account offered by Herron and Rugg (2003b), reducing the likelihood of recollecting information about targets should attenuate the extent to which participants prioritise recollection of one form of information over another, as the efficacy of relying on the presence or absence of recollection of information associated with targets as a basis for task judgments diminishes as the likelihood of recollecting information about targets decreases. Thus, separating the ERPs elicited by new test words according to target designation in this experiment should reveal smaller differences than in our previous study.

In addition, it is of course vital to provide: (1) indicators that the likelihood of recollecting information about targets is lower in this experiment than in our previous experiment, and (2) evidence consistent with the view that participants did in fact rely upon recollection of information about non-targets as well as targets in the current experiment. The first of these two criteria will be assessed by contrasting directly the accuracy of target memory judgments across experiments. The second criterion will be assessed by analyzing the ERP data with a view to determining whether there are reliable left-parietal ERP old/new effects for non-targets as well as for targets. As already stated, the left-parietal ERP old/new effect has been linked to the process of recollection, thus reliable old/new effects for targets and non-targets would suggest that participants relied upon recollection of information associated with both of these classes of test stimulus in order to complete the memory task. In combination, lower target accuracy in this experiment than in our previous experiment, coupled with the presence of reliable target and non-target old/new effects as well as the absence of marked retrieval orientation effects (differences between ERPs evoked by new test words) would provide strong support for the claim that ERP indices of retrieval orientations reflect processes that enable selective recollection of task-relevant information.

2. Materials and methods

Twenty-two right-handed participants (15 female) were paid at the rate of £7.50 per hour for taking part in the experiment. The data from four participants (two female) was discarded due to excessive EOG artefact (see below). The average age of the remaining participants was 21 years (standard deviation (S.D.) = 1.5 years). All had a minimum of 6 years of secondary school education (mean = 6.9, S.D. = 0.6 years) and between 6 and 30 months of higher education (mean = 16 months, S.D. = 10 months). Participants reported no history of psychological of neurological illness and were not taking neuroleptic medication at the time of participation. All participants gave informed consent prior to completing the experiment.

2.1. Stimuli and design

Three hundred and sixty critical words from the MRC psycholinguistic database (www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm) were presented in white letters on a black background on a computer monitor placed 1 m from participants (frequency 1–7 million^{-1}, 4–9 letters in length). Maximum horizontal and vertical visual angles were 2.4° and 1.4°. One complete 360-word task list comprised one study list and two test lists. The 360 words were split into 6 equal groups. Words appeared in only one group. The study list comprised 4 of the 6 groups of words (240 words in total). An asterisk preceded two groups of study words, a plus sign the other two. These cues signalled the task participants should complete for each word (see Section 2.2). Each test list comprised 3 of the 6 groups of words (180 words in total, 2 groups of which were also on the study lists). No words appeared in both test lists. Rotating the groups of words across study and test lists so that across lists all words appeared after an asterisk and a plus sign, and all were presented at study and test as well as at test only, resulted in the creation of three complete task lists. The order of presentation of words in the study and the test lists was determined randomly for each participant. Twenty filler words were placed at the beginning and the end of each study list. These filler words did not appear in either test list. One filler word was added to the beginning of each test list. In total, each participant saw 642 words (240 study words + 40 fillers, 360 test words + 2 fillers).

2.2. Procedure

In each study phase, participants completed one of two tasks on each word. In the function task, they were asked to say aloud a suitable function for the object denoted by the word. In the drawing task, they were asked to rate verbally the difficulty of drawing the object denoted by the word on a five point scale: 1—'very easy'; 5—'very difficult' (for similar task requirements, see Johnson et al., 1997). For half of the participants, an asterisk before study words signalled that a function judgment should be made, and a plus sign signalled that a drawing judgment should be made. This correspondence was reversed for the remaining participants. One of these two cues initiated each study trial and remained on the screen for 1000 ms. The screen was then blanked (100 ms) before the study word was presented for 300 ms. After a 1000 ms gap, the message PLEASE SPEAK NOW appeared. Participants were asked to withhold their response until this message appeared. The message was removed when participants pressed a key. The next trial started 1000 ms later.

The test phase was preceded by a 40 min period during which participants were fitted with an electrode cap (see below). Study and test phases were completed in the same testing chamber. Each test trial started with a fixation asterisk (500 ms duration), which was removed from the screen 100 ms prior to presentation of a test word (300 ms duration). The screen was then blanked until the participant responded, and the next trial started 1200 ms after the response. Participants were asked to balance response speed and accuracy equally. For each test phase, participants responded with one hand to words from the function/drawing study task (targets), and with the other to words from the alternate task (non-targets), as well as to unstudied test words. Responses were made on a key-pad with the left and right thumbs. The thumbs used for responses were balanced across participants, and participants were informed of target designation for each test phase only at the start of that phase. An equal number of participants completed the function/drawing target designation condition first. Participants were informed prior to the experiment that target designation would not necessarily differ across test phases and they were not informed of the number of test phases. A short break was given after each phase.

2.3. ERP recording

Twenty-five recording locations from the International 10–20 system (Jasper, 1958) comprised midline (Fz, Cz, Pz), left and right hemisphere sites (FP1/FP2, F7/F8, F5/F6, F3/F4, T7/T8, C5/C6, C3/C4, P7/P8, P5/P6, P3/P4, O1/O2). Additional electrodes were located on the mastoid processes. EEG was acquired continuously (6 ms/point) over a frequency band of 0.03–40 Hz with Fz as reference. Vertical and horizontal EOG were recorded bipolarly from electrodes placed above and below the right eye, and on the outer canthi of the eyes. ERPs were re-referenced off-line to linked mastoids and the data from Fz was recovered. Data were epoched off-line (1536 ms (256 point) epochs, with a 102 ms pre-stimulus baseline, relative to which all mean amplitudes were computed). Trials containing large EOG artefact and those containing A/D saturation or baseline drift exceeding ±80 μV were rejected. Other EOG blink artefacts were corrected using a linear regression estimate (Semlitsch, Anderer, Schuster, & Presslich, 1986).

3. Results

3.1. Behavioural data

Table 1 displays the probabilities of correct responses to each class of test word in the function and drawing target designations. In order to determine that participants were able to discriminate between targets and non-targets, and between targets and new words; values of Pr (Snodgrass & Corwin, 1988) were computed where Pr = p(hit) − p(false alarm). Pr was calculated separately for each target designation (function/drawing). For all measures of discrimination, p(hit) was the likelihood of a correct response to a target. For target/non-target discrimination, p(false alarm) was the likelihood of an incorrect (target) response to a non-target, while for target/new discrimination p(false alarm) was the likelihood of an incorrect (target) response to a new word. In all cases, these discrimination measures were reliably above zero, indicating that in both target designations participants were able to discriminate between the critical classes of test stimuli (in each case $t(17) > 10.40$, $p < 0.001$). The likelihoods of correct responses were subjected to ANOVA with target designation (function/drawing) and response category (correct response to target, new, and non-target words) as factors. The analysis revealed only a main effect of category ($F(1.9,31.9) = 33.25$, $p < 0.01$).[2] Follow-up analyses were conducted by collapsing data across target designation, and comprised all possible paired comparisons of the likelihood of correct responses to target, new and non-target words. Bonferroni corrected t-tests (adjusted alpha level = 0.017) indicated that judgments were more accurate for new than for old words (targets: 0.67 versus 0.92; $t(17) = 7.22$, $p < 0.001$; non-targets: 0.81 versus 0.92; $t(17) = 4.05$, $p < 0.01$), and more accurate for non-targets than for targets (0.67 versus 0.81; $t(17) = 4.56$, $p < 0.001$).

As outlined in Section 1, an important element of the behavioural data in this experiment is that the ability of participants to make correct judgments to targets should be lower than in our previous study (Dzulkifli & Wilding, 2005). In order to ascertain that this was indeed the case, the likelihoods of correct target judgments were contrasted across the two experiments. This measure was collapsed across function/drawing target designation as in neither experiment was the accuracy of target judgments moderated by which class of studied items was designated as the target category. The mean probabilities of a correct judgement to targets in the previous and the current experiment were 0.82 and 0.67, respectively. An unpaired t-test revealed that there was superior accuracy for target judgments in our previous than in our present experiment ($t(34) = 2.99$, $p < 0.01$).

The reaction times for correct responses to target, new and non-target words in the two target designation tasks are also shown in Table 1. ANOVA of RTs incorporated the same factors as above and again revealed a main effect of category only ($F(1.4,24.6) = 19.74$, $p < 0.001$). Bonferroni corrected t-tests (adjusted alpha level = 0.017) conducted on data collapsed across the factor of designation indicated only that new words were associated with significantly faster RTs than old words (targets: 1374 versus 1179; $t(17) = 3.65$, $p < 0.001$; non-targets: 1422 versus 1179; $t(17) = 6.60$, $p < 0.001$).

Table 1
Probabilities of correct responses (p(correct)) and reaction times (RT) to target, non-target and new words in the function and drawing target designation conditions

Target designation	Word type		
	New	Target	Non-Target
Function			
p(correct)	0.89 (0.12)	0.69 (0.09)	0.82 (0.08)
RT	1178 (308)	1352 (349)	1408 (323)
Drawing			
p(correct)	0.94 (0.04)	0.65 (0.12)	0.80 (0.09)
RT	1181 (273)	1396 (446)	1437 (364)

S.D.s are in parentheses.

3.2. ERP analyses

3.2.1. ERPs evoked by correct rejections

Fig. 1 shows the ERPs evoked by correct rejections, separated according to target designation. The figure shows that there are few differences between these two classes of ERPs, with some suggestion of divergences at right frontal and left-parietal electrode locations from approximately 700–800 ms onwards, where those evoked in the function target designation tend to be more positive-going at anterior sites and more negative-going at posterior sites than those evoked in the drawing target designation.

The analyses of the ERPs evoked by new words were guided by the findings in our previous study (Dzulkifli & Wilding, 2005), where reliable differences between these classes of test word were restricted to the 500–900 ms time period, with some suggestion of changes in the distribution of the effects over

[2] This and all subsequent ANOVAs incorporated Geisser-Greenhouse corrections when necessary (Greenhouse & Geisser, 1959), and corrected degrees of freedom are shown where appropriate.

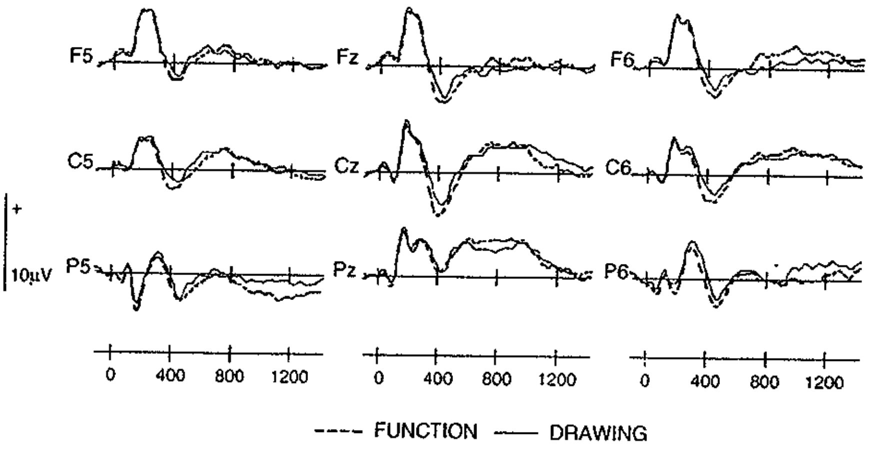

Fig. 1. Grand average ERPs evoked by new words in the two target designation conditions. The data are shown for nine locations at midline as well as left and right hemisphere sites over anterior (F5, Fz, F6), central (C5, Cz, C6) and posterior scalp (P5, Pz, P6).

the 500–700 and 700–900 ms time windows. Accordingly, the data were analysed from 500 to 700, 700 to 900 and 900 to 1400 ms, the final time window covering the remaining part of the recording epoch in which inspection of Fig. 1 suggests that there are some divergences according to designation. In keeping with the analysis strategy adopted in our previous study, the ERPs were analysed using a 3 × 3 grid of electrode locations, comprising the locations shown in Fig. 1. The ANOVAs

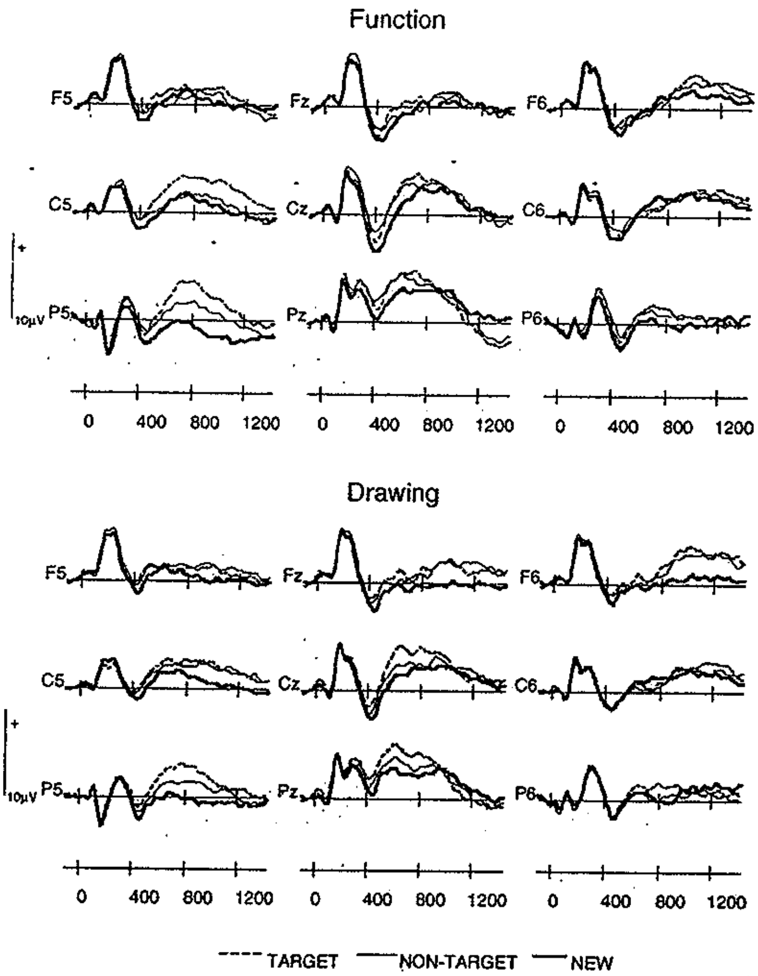

Fig. 2. Grand average ERPs evoked by correct judgments to target, non-target and new words in the function target (upper panel) and drawing target designation conditions. Electrode locations as for Fig. 1.

included the factors of target designation (function/drawing), the anterior–posterior dimension (AP: anterior, central, posterior) and the left-right dimension (LR: left hemisphere, midline, right hemisphere). The dependent measure in these and all subsequent analyses of the ERP data comprised the mean amplitude measures for the time periods designated in each case. No reliable effects involving designation were obtained in any of the time windows, although for the 900–1400 ms epoch the three-way interaction between designation, AP and LR approached significance ($F(2.7,46.3)=2.73, p=0.06$). The principal contributor to this interaction is the relatively greater positivity associated with the function designation at right-frontal sites and the drawing designation at left posterior sites (see Fig. 1).

3.2.2. ERP old/new effects

Fig. 2 shows the ERP old/new effects that were obtained in the two target designations. The figure shows that at posterior sites the ERPs evoked by old words are more positive-going than those evoked by new words at left-hemisphere and midline scalp locations from approximately 400–1100 ms. The ERPs evoked by targets are also more positive-going than those evoked by non-targets during this period. At anterior locations from 900 ms onwards, the ERPs evoked by old words in the drawing designation are more positive-going than those evoked by new words, particularly at right-frontal scalp sites. This relative positivity is markedly less evident in the function designation condition.

The initial analysis of the ERP old/new effects comprised a directed analysis of the effects that were obtained at parietal electrode locations. This analysis was completed in order to determine the relationship between this effect for targets and for non-targets and was restricted to sites P5 and P6, as the effect is typically largest at left-hemisphere parietal electrode locations (Rugg & Allan, 2000). The initial analysis (factors of designation, condition and site) revealed a main effect of condition ($F(1.5,24.8)=8.12, p<0.01$) as well as an interaction between this factor and site ($F(2.0,34.0)=11.23, p<0.001$). In the absence of reliable effects involving designation, follow-up analyses were run on data collapsed across this factor and comprised all possible comparisons of the ERPs evoked by old and new words. All three contrasts revealed reliable effects of condition as well as interactions between condition and site (condition: target versus new $F(1,17)=10.06, p<0.01$; non-target versus new $F(1,17)=5.18, p<0.05$; target versus non-target $F(1,17)=6.53, p<0.05$: condition × site: target versus new $F(1,17)=23.26, p<0.001$; non-target versus new $F(1,17)=6.38, p<0.05$; target versus non-target $F(1,17)=4.69, p<0.05$). The outcomes reflect the fact that the ERP old/new effects are larger at P5 than at P6, with the effects being larger for targets than for non-targets.

3.3. Global analyses

In addition to these hypothesis driven analyses, the ERP old/new effects were subjected to a series of global analyses in order to ascertain the correspondence between the effects obtained here and those in other ERP studies in which recollection has been required for at least some task judgments. The effects were analysed using data from the 3 × 3 montage described above over four epochs: 300–500, 500–800, 800–1100 and 1100–1400 ms. These epochs correspond to those employed in our previous study (Dzulkifli & Wilding, 2005), as well as in other previous related ERP memory studies (for reviews, see Friedman & Johnson, 2000; Rugg et al., 2002). The initial analyses included the factors of designation, condition, AP and LR. Follow-up ANOVAs comprising paired contrasts were employed in order to determine the reasons for effects involving condition. Reports of the outcomes of the follow-up analyses are restricted to those effects that were revealed in the initial analysis. Where three-way interactions involving condition, AP and LR were obtained in the paired contrasts they were followed up with post hoc analyses (Newman–Keuls) at each of the nine electrode locations. In each time window, no reliable effects of designation were obtained, and all follow up analyses (paired contrasts and post hoc analyses by site) are collapsed across this factor.

3.3.1. 300–500 ms

The initial analysis revealed a main effect of condition ($F(1.8,30.9)=9.85, p<0.01$) as well as a CC × LR interaction ($F(3.4,58.1)=4.42, p<0.01$). The same two interaction terms were revealed by the paired contrasts between the ERPs evoked by old and new words (collapsed across designation), while the contrast between the ERPs evoked by the two classes of old words revealed no reliable effects (targets versus new; condition $F(1,17)=11.79, p<0.01$; condition × LR $F(2.0,33.4)=7.65, p<0.01$: non-target versus new; condition $F(1,17)=18.50, p<0.01$; condition × LR $F(1.7,29.4)=5.13, p<0.05$). These effects reflect the fact that overall the greater relative positivity for the ERPs evoked by old compared to new words is larger at midline than at lateral sites. Inspection of Fig. 3 suggests an interaction involving designation and target/non-target status, but the relevant effects were not statistically significant.

3.3.2. 500–800 ms

The initial analysis revealed the same two effects as in the previous epoch (condition $F(1.8,30.2)=6.25, p<0.01$; condition × LR $F(2.9,48.8)=5.80, p<0.01$), which were moderated by an interaction between these two factors and AP ($F(4.9,82.7)=7.29, p<0.001$). Follow up analyses revealed the same three-way interaction for each paired contrast (target versus new $F(3.1,53.2)=10.70, p<0.001$; non-target versus new $F(3.4,57.1)=6.40, p<0.01$; target versus non-target $F(2.5,42.3)=3.11, p<0.05$). Both contrasts involving new words revealed condition × LR interactions (target versus new $F(1.6,26.7)=8.55, p<0.01$; non-target versus new $F(2.0,33.7)=4.67, p<0.05$). The target versus new contrast also revealed a main effect of condition ($F(1,17)=9.48, p<0.01$) and a condition × AP interaction ($F(1.4,24.5)=4.45, p<0.05$). As Fig. 3 shows, these effects come about primarily because of the left-lateralisation of the target and non-target ERP old/new effects at posterior sites, which is less evident at central and

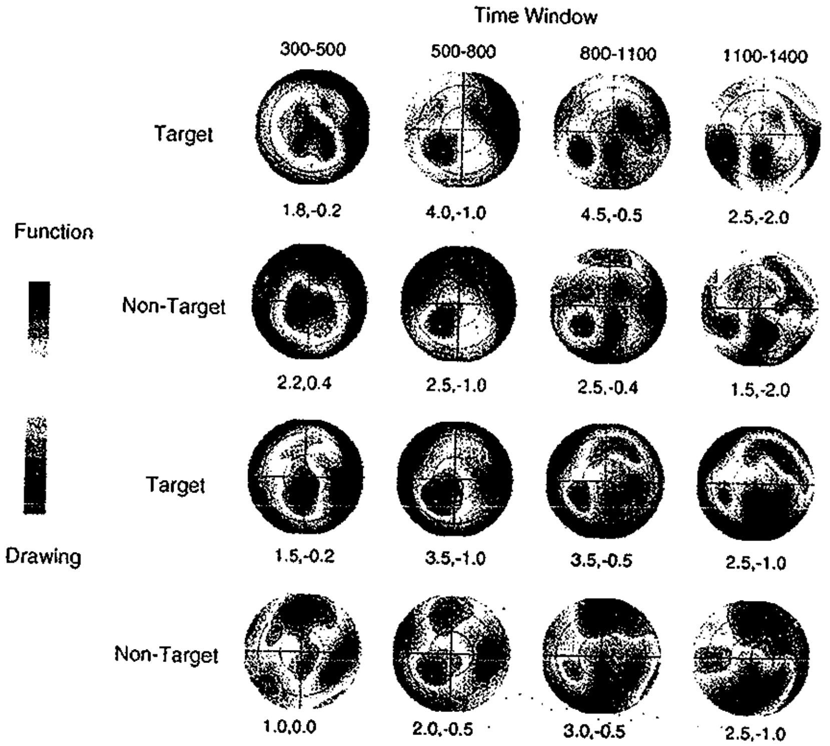

Fig. 3. Topographic maps depicting the scalp distributions of the old/new effects for targets and for non-targets, separated according to target designation (function/drawing) and epoch (300–500, 500–800, 800–1100, 1100–1400 ms). The maps were computed on difference scores obtained by subtracting the mean amplitudes for correct rejections from those for targets and non-targets in the two target designation conditions. The paired values below each map denote the maxima and minima of the amplitude differences between conditions, and can be interpreted relative to the colour bar on the left-hand side of the figure.

anterior locations. In keeping with this description, the post hoc analyses revealed reliable non-target old/new effects at the P5 electrode only, and target old/new effects were reliable at P5 as well as at Cz. The interaction between targets and non-targets is most likely due to the fact that while sharing the same distribution the non-target old/new effects are smaller than the target effects. While the post hoc analyses revealed no reliable effects at individual sites, this quantitative interpretation is supported by the outcome of an additional analysis on data rescaled using the vector length method described by McCarthy and Wood (1985). The difference scores obtained by subtracting the mean amplitudes of the ERPs evoked by new words from those evoked by both classes of old words were rescaled and submitted to ANOVA with all other factors as described above. The analysis did not reveal an interaction between condition and site—the statistical signature of qualitative differences between the scalp distributions of ERP effects.

3.3.3. 800–1100 ms

For the initial analysis over this epoch, a main effect of condition ($F(1.7,29.3)=3.87$, $p<0.05$) was accompanied by an interaction between this factor, AP and LR ($F(4.2,71.8)=9.71$, $p<0.001$). The same interaction term was reliable for the paired contrasts (collapsed across designation) involving new words, and approached significance ($p=0.06$) for the contrast between the two classes of old words (target versus new $F(3.1,53.5)=16.18$, $p<0.001$; non-target versus new $F(2.7,46.7)=8.25$, $p<0.001$). The target versus new contrast also revealed a main effect of condition ($F(1,17)=5.71$, $p<0.05$). Post hoc analyses revealed reliable old/new effects at P5 only, thus in keeping with the previous epoch, these three-way interactions likely reflect the fact that the old/new differences at posterior sites are smallest over the right hemisphere. At anterior sites, there is a trend towards the opposite asymmetry, but the individual site post hoc analyses did not reveal reliable differences at the individual scalp locations.

3.3.4. 1100–1400 ms

The initial analysis revealed a pair of two-way interactions (condition × AP $F(2.1,35.6)=3.95$, $p<0.05$; condition × LR $F(3.7,62.1)=3.87$, $p<0.01$) that were both moderated by a three-way interaction involving these factors ($F(3.6,61.0)=7.70$, $p<0.001$). Inspection of Fig. 2 suggests an interaction that involves condition and task, given the markedly larger old/new effects in the drawing designation at right-frontal electrode locations. No interactions with task were revealed, however, and this remained true when an ANOVA restricted to sites F5, Fz and F6 was conducted (all other factors as above). The follow up analyses collapsed across the factor of task revealed significant effects only in the contrasts involving old words, which included three-way interactions between condition, AP and LR (targets $F(3.3,56.1)=13.98$, $p<0.001$; non-targets $F(2.5,41.9)=9.60$, $p<0.001$). The target versus new contrast also revealed a condition × LR interaction ($F(1.9,31.8)=7.42$, $p<0.01$), while for the non-target versus new contrast the condition × AP interaction was significant ($F(1.4,23.1)=7.27$, $p<0.01$). Fig. 3 shows that the right-lateralised old/new effects at anterior locations are accompanied by a midline maximum posterior negativity, but the individual site post hoc analyses revealed reliable differences at Pz only, where the ERPs evoked by new words were more positive-going than those evoked by old words.

4. Discussion

This experiment was designed in order to test a claim regarding the functional significance of the differences between ERPs elicited by new test items that we reported in a previous study (Dzulkifli & Wilding, 2005). Our previous proposal was that differences between ERPs elicited by new items and separated according to target designation indexed processes responsible for the selective recovery of task-relevant information. That is, in the two target designations the differences between the ERPs elicited by new test words arose because participants were prioritising recollection of different kinds of information, and the ERPs indexed processes important for accomplishing that. The way this claim was assessed in this experiment was by acquiring the relevant ERP data in a task where, in comparison to our previous study, there was less incentive to adopt such a selective retrieval orientation and instead an incentive to rely on more similar kinds of information in the two different target designation conditions. Attenuation of the differences between the ERPs elicited by new items in this study would thus support the functional account we offered previously.

There were no reliable differences between the ERPs elicited by new words and separated according to target designation, a finding consistent with our pre-experimental hypothesis. There are also two important aspects of the data that were assessed in order to provide support for our functional account. The first was the accuracy of target memory judgments across this experiment and our previous experiment. As reported in the Results section, the probability of a correct target judgment was reliably lower in the present than in our previous experiment (0.67 versus 0.82, collapsed across target designation). According to Herron and Rugg (2003a, 2003b), the extent to which recollection of target over non-target information will be prioritised diminishes as the likelihood of recollecting information about targets diminishes. Further evidence germane to this issue, moreover, is the fact that there were reliable left-parietal ERP old/new effects for targets as well as for non-targets in the current experiment. These findings differ from those in our previous study (Dzulkifli & Wilding, 2005), where left-parietal ERP old/new effects were reliable for targets only.[3]

To the extent that the left-parietal ERP old/new effect indexes recollection, these data are therefore consistent with the claim that participants prioritised recollection of target over non-target information to a lesser degree in the present than in the previous study. In the previous study only, moreover, there were reliable indices of retrieval orientations–differences between the ERPs that were elicited by new test words and separated according to target designation. In combination, the absence of reliable differences between the ERPs elicited by new items in the present study, alongside the presence of reliable left-parietal ERP old/new effects for non-targets as well as for targets, supports strongly the claim that the ERP orientation effects in the study due to Dzulkifli and Wilding (2005) index processes important for selective recollection of task-relevant information. The critical data from our present and previous study is summarised in Fig. 4, which shows ERP old/new effects at a left-parietal (P5) electrode and orientation effects at a mid-frontal (Fz) electrode in the two experiments.

One possible objection to this account, however, is that the smaller differences between the ERPs evoked by new test words in the present experiment came about simply because of the lower probabilities of correct judgments to new words in comparison to our previous study. This objection is based on the assumption that in forced-choice tasks the neural signatures of some memory-related processes become less separable as the accuracy of task judgments decreases. This would in part be due to the fact that as accuracy decreases the proportion of responses made on the basis of impoverished information increases. One possible outcome of this set of circumstances is a greater degree of similarity between the neural activity associated with test items that are separated into different response categories. Thus, the likelihood of observing differences – for example, between the neural activity associated with hits and with correct rejections – falls along with reductions in the accuracy of task judgments.

While this argument holds for contrasts involving old test items, however, it is less applicable in the case of comparisons between classes of new test items. This is because the critical contrast here is between test items with identical study histories separated solely according to the target/non-target distinction. Since all that differs is the test context, there is little reason to assume that indices of orientation will be attenuated in all circumstances according to the accuracy of task judgments. Consistent this view, Robb and Rugg (2002) manipulated retrieval orientation and the level of response accuracy orthogonally in four separate recognition memory tasks. The indices of orientation they identified were invariant across changes in accuracy. Dzulkifli et al. (2004), moreover, demonstrated that in some circumstances the magnitude of indices of orientation can increase as the accuracy of task judgments decreases. These observations and data argue against this competing account of the current findings.

Issues concerning the impact of factors that may attenuate the magnitude of memory-related effects are also relevant to consideration of the differences between the amplitudes of the target and non-target ERP old/new effects in the present experiment. The non-target old/new effects, while reliable, were smaller than the target old/new effects. These data suggest that, although to a markedly lesser degree than in our previous study, participants still prioritised recollection of targets over non-targets in the present study. There are a number of ways in which this could have come about and the data that are presented here do not

[3] The pattern of ERP old/new effects across these two studies also mirrors that reported by Herron and Rugg (see Section 1). In their experiments, the accuracy of task judgments was higher in the experiment where reliable left-parietal ERP old/new effects were obtained for targets only.

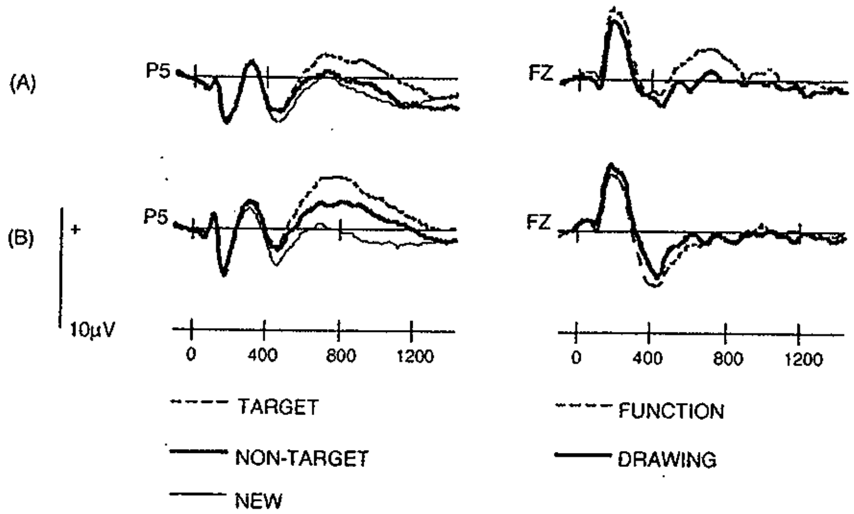

Fig. 4. (Row A) ERP old/new effects at P5 and retrieval orientation effects at Fz in the study of Dzulkifli and Wilding (2005). (Row B) The same effects in the present study.

provide a means of distinguishing between them. The results may reflect a general predisposition for participants to prioritise recollection of targets to some degree. Alternatively, they may reflect the fact that a proportion of participants on a proportion of test trials adopted this strategy (Wilding & Sharpe, 2004). Irrespective of which if either of these possibilities is correct, however, the fact that there were reliable ERP old/new effects for targets as well as for non-targets suggests that participants employed recollection of targets as well as non-targets as a basis for task judgments to a greater degree in this experiment than in our previous experiment (Dzulkifli & Wilding, 2005).

It is also important to note that there is no identifiable 'non-target miss' in an exclusion task: the design of the task is such that a correct non-target response can be made either when a non-target is recognised as an old item, or when it is forgotten, because responses to non-targets and to new words are made on the same response key (Wilding & Rugg, 1997). This is not the case for targets, thus there will be one kind of trial – forgotten words – contributing to the non-target averaged ERPs that contributes markedly less to the averaged target ERPs. The results in a number of previous studies have shown that the ERPs evoked by misses are more similar to those evoked by correct rejections than those evoked by correct responses to old items (Neville, Kutas, Chesney, & Schmidt, 1986; Smith, 1993; Wilding & Rugg, 1996, 1997). Of particular note here is the work due to Wilding and Rugg (1997), who demonstrated that this pattern of activity for misses obtains when ERPs are acquired while participants complete an exclusion task. The likely impact on the magnitude of ERP old/new effects, therefore, is a reduction for non-targets relative to targets.

According to this account, non-target ERP old/new effects should in general be somewhat smaller than target ERP old/new effects (Wilding & Rugg, 1997; Wilding & Sharpe, 2004). It is important to note, however, that these observations cannot provide a complete account of the patterns of differences between target and non-target ERP old/new effects that have been reviewed here. The reason for this is that if this factor were the sole contributor then the disparity between the sizes of target and non-target ERP old/new effects would decrease as the accuracy of task judgments increases, since the proportion of 'forgotten' non-targets will diminish as task accuracy increases. The opposite pattern of data was reported by Herron and Rugg (2003b), and can also be seen in the data from our present and previous study: reliable non-target left-parietal ERP old/new effects were evident only in the experiment where the accuracy of task judgments was lower (see Fig. 4).

Another finding in the present study was that the patterns of ERP old/new effects that were obtained during the post-stimulus epoch did not vary qualitatively according to target designation. This finding is in keeping with the majority of ERP studies of memory retrieval, in which there has been little evidence for content-specific retrieval (although see Senkfor, Van Petten, & Kutas, 2002). These findings support the view that, at least to a first approximation, ERP old/new effects index core retrieval processes that are engaged (albeit to greater or lesser degrees) irrespective of the content of information that is to be retrieved (Allan, Robb, & Rugg, 2000). It is of course important to note that this inference is based upon a series of null results, but the consistency with which this general pattern of effects has been observed is also of note.

While not varying according to target designation, however, the ERP old/new effects for targets and for non-targets did vary according to epoch, and the changes in scalp distributions with time that can be seen in Fig. 4 are broadly similar to those that have been reported in previous studies (Friedman & Johnson, 2000; Wilding & Sharpe, 2003). In particular, the primarily left centro-parietal distribution of the old/new effects from approximately 500–1000 ms post-stimulus shifts to a combination of this effect and an anterior right-sided distribution later in the recording epoch (for similar distributions, see Rugg, Allan, & Birch, 2000; Wilding & Rugg, 1996). Fig. 3 shows that in addition there is also a relatively greater negativity associated with correct old judgments than with correct rejections from approximately 1000 ms onwards, which is most prominent at midline posterior scalp sites. In the case of both the right-frontal and the posterior negative effect the present data add nothing to functional interpretations that have been offered previously (Johannsen & Mecklinger, 2003; Rugg & Allan, 2000).

Finally, we turn to the question of the mechanisms that may be responsible for the selective control of recollection that is indexed by the patterns of ERP old/new effects for targets and for non-targets, and hence what processes are indexed by the differences between the ERPs evoked by correct rejections in our previous study but not in the present study. According to one framework, there are at least three classes of bias operations that may support selective recollection (Anderson & Bjork, 1994; Levy & Anderson, 2002). One possibility is that target bias operations are engaged that operate directly upon memory representations and which influence the likelihood that they will interact with retrieval cues. While it is possible that this class of operations is engaged on a trial-by-trial basis, it has been argued that target bias operations would likely be sustained throughout a given retrieval task and would not therefore be manifest as differences between ERPs evoked by classes of new items (Herron & Rugg, 2003a).

Two further classes of bias mechanism are cue-bias and attention bias. The former influences the ways in which retrieval cues can be manipulated in order to influence the likelihood of the cue interacting with some memory contents rather than others. The latter influences the ways in which processing resources will be allocated preferentially to some contents rather than others. Herron and Rugg (2003a) suggested that differences between ERPs evoked by new items likely index cue bias processes. The data consistent with this view came from a study in which the encoding phase comprised presentation of words and pictures. Test stimuli were words, and the old words were either representations of words, or of words corresponding to the objects shown in the pictures. In separate retrieval phases, targets were designated as old words encountered either as words or pictures at encoding.

There were reliable parietal old/new effects for targets in both target designations, but reliable parietal effects for non-targets only when pictures were designated as targets. Herron and Rugg (2003a) argued that the most parsimonious explanation for this pattern of data was that: (1) in the word target condition participants were able to process cues sufficiently selectively to restrict recollection to information associated with studied words, perhaps because of the perceptual match between target stimuli at study and at test, while (2) in the picture target condition the absence of this perceptual match meant that processes engaged to recover information diagnostic for judgments about studied pictures involved cue processing operations that also resulted in recollection of information about studied words.

For present purposes, the important point is that these data are not straightforward to account for in terms of an attention bias account, thus the weight of evidence, although not conclusive, favours a cue bias explanation for the mechanisms by which selective recollection is accomplished, at least in the study due to Herron and Rugg (2003a). It should also be noted, however, that there is no reason in principle why a combination of these bias mechanisms is not typically responsible for selective episodic retrieval. A related question is whether the extent to which these classes of process might act in concert varies according to factors such as the content that is to be retrieved and the precise structure of retrieval tasks. The data from this experiment do not in and of themselves permit strong claims to be made regarding the mechanisms that are responsible for the selective control of memory retrieval, but they do point to the potential utility of the combined use of behavioural and electrophysiological data in order to address this issue.

In summary, in a previous study reliable indices of retrieval orientation were observed alongside data indicating that participants restricted recollection almost wholly to only certain memory contents (Dzulkifli & Wilding, 2005). This finding was the basis for the proposal that the indices of orientation indexed processes responsible for selective retrieval. In the present study, the design of the task encouraged participants to be less selective with respect to what was retrieved from memory, and under these circumstances the indices of orientation reported in the earlier study were attenuated markedly. This combination of findings provides strong support for the claim that ERP indices of retrieval orientation index processes that are influential in the selective recovery of information from episodic memory. An important goal for future research is an accurate delineation of the mechanisms that are responsible for this selectivity.

Acknowledgements

This research was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) and the Wellcome Trust.

References

Allan, K., Robb, W. G. K., & Rugg, M. D. (2000). The effect of encoding manipulations on neural correlates of episodic retrieval. *Neuropsychologia, 38*, 1188–1205.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long term memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). San Diego, California: Academic Press.

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

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour, 16*, 519–533.

Donaldson, D. I., Wilding, E. L., & Allan, K. (2003). Fractionating retrieval from episodic memory using event-related potentials. In A. E. Parker, E. L. Wilding, & T. J. Bussey (Eds.), *The cognitive neuroscience of memory: Episodic encoding and retrieval* (pp. 39–58). Hove: Psychology Press.

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.

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Techniques, 51*, 6–28.

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

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

Herron, J. E., & Rugg, M. D. (2003b). Strategic influences on recollection in the exclusion task: Electrophysiological evidence. *Psychonomic Bulletin and Review, 10*, 703–710.

Herron, J. E., & Wilding, E. L. (2005). An electrophysiological investigation of factors facilitating strategic recollection. *Journal of Cognitive Neuroscience, 17*, 1–11.

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.

Johannsen, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology, 64*, 91–117.

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

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6*, 299–305.

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory, 10*, 397–403.

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

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.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10*, 389–395.

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. L. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language, 25*, 75–92.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience, 20*(RC108), 1–6.

Ranganath, C., & Paller, K. A. (1999). Frontal brain potentials during recognition are modulated by requirements to retrieve perceptual details. *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. K., & Rugg, M. D. (2002). Electrophysiological dissociation of retrieval orientation and retrieval effort. *Psychonomic Bulletin and Review, 9*, 583–589.

Rugg, M. D., & Allan, K. (2000). Event-related potential studies of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 521–537). Oxford: Oxford University Press.

Rugg, M. D., Allan, K., & Birch, C. S. (2000). Electrophysiological evidence for the modulation of retrieval orientation by depth of study processing. *Journal of Cognitive Neuroscience, 12*, 664–678.

Rugg, M. D., Herron, J., & Morcom, A. (2002). Electrophysiological studies of retrieval processing. In L. R. Squire & D. L. Schacter (Eds.), *Neuropsychology of memory* (3rd ed., pp. 154–165). New York: Guilford Publications.

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.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology, 23*, 695–703.

Senkfor, A. J., Van Petten, C., & Kutas, M. (2002). Episodic action memory for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neurosicence, 14*, 402–419.

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgements. *Journal of Cognitive Neuroscience, 5*, 1–13.

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

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., & Nobre, A. C. (2001). Task-switching and memory-retrieval processing: Electrophysiological evidence. *NeuroReport, 12*, 3613–3617.

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain, 119*, 889–905.

Wilding, E. L., & Rugg, M. D. (1997). Event-related potentials and the recognition memory exclusion task. *Neuropsychologia, 35*, 119–128.

Wilding, E. L., & Sharpe, H. (2003). Episodic memory encoding and retrieval: Recent insights from event-related potentials. In A. Zani & A. Mado-Proverbio (Eds.), *The cognitive electrophysiology of mind and brain* (pp. 169–196). San Diego: Academic Press.

Wilding, E. L., & Sharpe, H. L. (2004). The influence of response-time demands on electrophysiological correlates of successful episodic retrieval. *Cognitive Brain Research, 18*, 185–195.