In what way does the parietal ERP old/new effect index recollection?

Edward L. Wilding*

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

Received 21 May 1999; received in revised form 4 October 1999; accepted 12 October 1999

Abstract

Event-related potentials (ERPs) were recorded while subjects performed a memory retrieval task requiring old/new judgements to visually presented old (previously studied) and new words. For words judged old, subjects made two binary forced-choice context (hereafter source) judgements, denoting the voice (male/female) and task (action/liking) with which the test word had been associated at study. By separating the ERPs according to the accuracy of the voice and task judgements, it was possible to test the prediction that the differences between ERPs to correctly identified old and new words at parietal scalp sites (parietal old/new effects) are sensitive to the amount or quality of information that is retrieved from episodic memory (Rugg, M.D., Cox, C.J.C., Doyle, M.C., Wells, T., 1995. Event-related potentials and the recollection of low and high frequency words. Neuropsychologia 33, 471–484). In keeping with this proposal, the magnitude of the parietal old/new effects co-varied with the number of accurate source judgements. This finding is consistent with proposals that the parietal old/new effect indexes recollection in a graded fashion.

Keywords: Recollection; Source memory; Event-related potentials; Old/new effects; Episodic memory

Event-related potentials (ERPs) acquired during recognition memory tasks differentiate correctly identified old and new items (Johnson, 1995; Rugg, 1995). At parietal electrodes, the ERPs to items judged correctly to be old are more positive than those to items judged correctly to be new, starting 250–300 ms post-stimulus. This relative positivity lasts for up to 1 s, and from approximately 500 ms is larger at left- than at right-parietal sites. This left-lateralised effect likely indexes the process of recollection (Paller and Ku-
According to one proposal, the parietal effect indexes recollection in a graded fashion, whereby the magnitude of the effect is sensitive to either the quality or the quantity of episodic information that is retrieved (Rugg et al., 1995; Wilding and Rugg, 1996). Support for this claim comes from studies that have included comparisons between classes of ERPs that either were or were not associated with accurate memory for source (Smith, 1993; Wilding et al., 1995; Wilding and Rugg, 1996). For example, in the study of Wilding and Rugg (1996), subjects made old/new judgements to visually presented old and new words. For words judged old, subjects made a second judgement, denoting in which voice (male or female) the word had been spoken at study. When the ERPs to correct old judgements were separated according to the accuracy of the subsequent source judgement, reliable parietal old/new effects were evident in both cases, and were larger for words associated with correct source judgements. When the ERPs to correct old judgements were separated according to the accuracy of the subsequent source judgement, reliable parietal old/new effects were evident in both cases, and were larger for words associated with correct source judgements. If recollection is defined operationally as the ability accurately to retrieve source information, these findings are consistent with the view that the parietal ERP old/new effect indexes recollection in a graded fashion.

However, an alternative explanation exists. First, the old/new effect for words associated with incorrect source judgements may have arisen as a result of averaging two types of trials: those not associated with recollection (which do not evoke parietal old/new effects), and those associated with recollection of information that was not diagnostic for the source judgement (voice) that was required in the task. Second, the larger old/new effect for words associated with correct source judgements may have arisen because, in addition to containing trials of the two types described above, on some trials recollection of voice information would have occurred. Thus, it remains a possibility that the ERP signature for recollection is all-or-none, irrespective of what information is in fact recollected, and that the differences between the two critical response categories in the study of Wilding and Rugg (1996) arose because the category associated with incorrect voice judgements contained a higher proportion of trials on which recollection did not occur. Broadly similar arguments can be applied to comparable findings in other ERP studies of memory that have required some form of source memory judgement (Smith, 1993; Wilding et al., 1995, also see Senkfor and Van Petten, 1998). Furthermore, the results of studies in which there has been a positive correlation between recognition memory performance and the magnitude of the old/new effect (e.g. Paller and Kutas, 1992; Johnson et al., 1998) are consistent with an all-or-none account, since the level of recognition performance might well have varied inversely with the proportion of trials on which recollection did not occur.

In order to disentangle the graded and all-or-none accounts of the parietal old/new effect, the experiment described here comprised study phases in which each word was associated with two aspects of source information. These were followed by test phases in which two forced-choice source decisions were required for items judged to be old. In this paradigm, the critical comparison is between the old/new effects associated with either one or two correct source judgements. Given relatively good memory for source, this comparison will be between two classes of ERPs that contain high proportions of trials on which recollection of task-relevant source information occurred. The principal disparity between the two categories will be the degree of recollection, as indexed by the number of correct source judgements (1 vs. 2). If the magnitude of the parietal old/new effect varies positively with the number of correct source judgements, then this would constitute strong evidence in support of the graded hypothesis first proposed by Rugg et al. (1995).

Twenty-one subjects (all right-handed, average age 21 years, 8 female) took part in the experiment. Each was paid at a rate of £5.00 per hour. One female was discarded from the final analyses due to excessive EEG artifact. EEG was recorded from 54 electrode sites that were embedded in an elasticated cap. The data were acquired continuously (250 Hz acquisition rate, bandpass 0.03–30
Hz) and epoched off-line into epochs of 2048 ms duration, which included a 100-ms pre-stimulus baseline. A left-mastoid reference was employed during acquisition, and the data were re-referenced off-line to the average of the left and right mastoids. Additional electrodes placed above and below, as well as to the left and right, of the eyes were used to monitor the EOG. Trials with significant eye blinks or eye movements were rejected prior to averaging. In keeping with previous work (Wilding et al., 1995) subjects were not included in the experiment if they did not contribute at least 16 trials to each of the critical response categories.

Each subject completed one task list, consisting of five study-test cycles. Study lists contained 32 words, an equal number spoken in the male/female voice. Test lists contained the 32 studied words and 32 new words. For counterbalancing purposes, four task lists were constructed. No word appeared in more than one cycle and an equal number of subjects completed each task list. On study trials subjects heard words spoken in either a male or a female voice, and to each they made an active/passive or a pleasant/unpleasant judgement (for a more complete description of the study phase format, see Wilding, 1999). Test trials started with a fixation asterisk (1 s duration), which was removed 200 ms pre-stimulus. Test words were presented visually for 300 ms, to which subjects made a speeded old/new judgement; 1700 ms after this response they were cued to make a task (action/liking) and a voice (male/female) judgement. Cues were VOICE? and TASK? which stayed on the screen until a response was made. Two hundred milliseconds intervened between the offset/onset of these cues. The order of voice/task cue presentation was constant for each subject and varied across subjects. The next trial commenced 1 s after the final source judgement. For trials on which a new response was made, subjects pressed any key when the VOICE? and TASK? cues appeared. Subjects were asked to restrict eye blinks to the period between the appearance of the first source judgement cue and the reappearance of the fixation cue. The average time between study and test phases was 2 min. In order to discourage the use of rehearsal strategies, subjects were asked to count backwards in three’s from 100 during this interval.

Recognition memory and source memory accu-

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Probabilities of correct old/new judgements and of correct source judgements to words judged correctly to be old*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy Old/new judgement</td>
<td>New</td>
</tr>
<tr>
<td>P(correct)</td>
<td>0.93 (0.07)</td>
</tr>
<tr>
<td>Source judgement</td>
<td></td>
</tr>
<tr>
<td>Voice correct</td>
<td>0.67 (0.12)</td>
</tr>
<tr>
<td>Task correct</td>
<td>0.83 (0.09)</td>
</tr>
<tr>
<td>Reaction times (ms)</td>
<td>Old/new judgement</td>
</tr>
<tr>
<td>Correct</td>
<td>1186 (227)</td>
</tr>
<tr>
<td>Incorrect</td>
<td>1446 (308)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes. Old words are separated according to which voice they were spoken in at study (male/female: M/F) and which type of study judgement was required (action/liking: A/L). The lower part of the table shows the reaction times for correct and incorrect old/new judgements to old and new words, as well as the reaction times for correct old judgements separated according to the accuracy of the subsequent source judgements. Standard deviations are shown in brackets.
Discrimination \( [P(\text{hit}) - P(\text{false alarm})] \) was reliably above chance for the four classes of old word (all possible conjunctions of voice and task: \( t(19) > 30, P < 0.001 \)), and a one-way ANOVA comparing the four measures revealed no reliable differences. A second ANOVA on the conditional probabilities of correct source judgements incorporated the factors of test judgement (task and voice) and word type (the four possible voice/task conjunctions). The only reliable effect revealed by this analysis reflected the fact that correct task judgements were more likely than correct voice judgements (\( F_{1,19} = 18.94, P < 0.001 \)). An initial ANOVA of the RT data employed the factors of accuracy (correct vs. incorrect) and word type (new vs. old)—collapsed across the subsequent source judgements. It revealed that correct judgements were faster than incorrect judgements (\( F_{1,19} = 44.82, P < 0.001 \)). Although the lower right portion of Table 1 shows that RT varies inversely with the number of correct source judgements, a one-way ANOVA (both correct vs. voice correct vs. task correct vs. neither correct) revealed no reliable differences.

Overall, the RTs in this task are comparable to those obtained in other tasks where an old/new judgement preceded a forced-choice source judgement (e.g. Wilding and Rugg, 1996). However, the RTs are slower than those obtained in ERP studies where only a binary recognition memory judgement was required (e.g. Rugg and Doyle, 1992; Johnson et al., 1998). This pattern of results suggests that one or more of the processing stages that lead to an old/new judgement are influenced by the requirement to make a subsequent source judgement. The stage (or stages) that may be affected remain unclear.

For the ERP data, the principal question was whether the magnitude of the parietal old/new effect varied with the number of correct source judgements. Consequently, ERPs were formed from trials on which either two or one source judgements were correct (the 2-correct and 1-correct response categories, respectively). Due to the relatively high probabilities of correct source judgements for voice and for task, it was not possible to form reliable averaged ERPs to words that were judged correctly to be old and associated with two incorrect source judgements. Within the 2-correct and 1-correct categories the ERPs were collapsed across study voice and study task. Paired comparisons of the 2-correct, 1-correct and correct rejection ERPs over the 500–800-ms epoch at the left- and right-parietal electrodes P5 and P6 revealed main effects of category (2-cor-

![Graph](image-url)  

Fig. 1. Grand averaged ERPs to the 2-correct (2-Corr), 1-correct (1-Corr) and correct rejection (C Rej) response categories at two frontal and two parietal electrode locations (AF7/AF8, P5/P6).
rect vs. 1-correct: $F_{1,19} = 6.40, P < 0.05$: 2-correct vs. correct rejection: $F_{1,19} = 66.41, P < 0.001$: 1-correct vs. correct rejection: $F_{1,19} = 42.76, P < 0.001$). The latter two comparisons also revealed interactions between category and site (2-correct: $F_{1,19} = 12.42, P < 0.01$: 1-correct: $F_{1,19} = 5.01, P < 0.05$). These effects reflect the fact that the 2-correct ERPs are more positive than those to the other two categories, and that the 1-correct ERPs are more positive than those to correct rejections. The interaction terms reflect the fact that the parietal old/new effects are larger over the left hemisphere than over the right. Fig. 1 shows the ERPs to the 2-correct, 1-correct and correct rejection response categories, while Fig. 2 shows the scalp distribution of the parietal ERP old/new effect.

The directed analysis at parietal sites was followed by analyses that were intended to elucidate how the old/new effects for the 2-correct and 1-correct response categories differ at frontal locations. These analyses were restricted to time windows over which reliable frontal old/new effects have been reported in other ERP studies of source memory (300–500, 500–800, 1100–1400 and 1400–1900 ms, see Allan et al., 1998). Following Wilding (1999) these analyses were restricted to the left- and right-frontal electrode pair AF7/AF8, and consisted of all possible paired comparisons of the 2-correct, 1-correct and correct rejection response categories. Over the three epochs, the ERPs to the 2-correct and 1-correct categories were statistically indistinguishable, while both were reliably more positive than the ERPs to correct rejections (300–500: 2-correct, $F_{1,19} = 6.98, P < 0.05$: 1-correct, $F_{1,19} = 4.77, P < 0.05$: 500–800: 2-correct, $F_{1,19} = 4.43, P < 0.05$: 1-correct, $F_{1,19} = 5.20, P < 0.05$: 1100–1400: 2-correct, $F_{1,19} = 6.64, P < 0.05$: 1-correct, $F_{1,19} = 29.13, P < 0.001$: 1400–1900: 2-correct, $F_{1,19} = 6.38, P < 0.05$: 1-correct, $F_{1,19} = 9.18, P < 0.01$). The analyses of the 2-correct and 1-correct old/new effects over the 1100–1400-ms epoch also revealed interactions between category and site, as did the analysis of the 1-correct effect over the 1400–1900-ms epoch (1100–1400: 2-correct, $F_{1,19} = 6.64, P < 0.05$: 1-correct, $F_{1,19} = 9.95, P < 0.01$: 1400–1900: 1-correct, $F_{1,19} = 6.36, P < 0.05$). The interaction terms reflect the fact that the old/new effects are larger over the right than over the left hemisphere.

The analyses thus revealed reliable old/new effects at parietal and frontal electrode sites. The magnitude of the parietal old/new effect varied with the number of correct source judgements, consistent with the graded hypothesis of Rugg.

Fig. 2. Topographic map showing the scalp distribution of the parietal ERP old/new effect over the 500–700-ms time period. The distribution was interpolated from the difference scores obtained by subtracting the ERPs to correct rejections from a weighted average of the ERPs to the 1-correct and 2-correct response categories.
and colleagues (Rugg et al., 1995; Wilding and Rugg, 1996). However, it is worthwhile considering the potential influence of guessing on the 2-correct and 1-correct ERPs. Given that the paradigm required forced-choice source judgements, a proportion of the trials in the 2-correct and 1-correct categories were likely correct guesses, rather than judgements that were made on the basis of veridical source information (for extended comments see Wilding and Rugg, 1996). Could the influence of guesses on the ERPs to the 2-correct and 1-correct categories explain these data in a way that is consistent with an all-or-none account of the parietal ERP old/new effect? It seems unlikely. For an all-or-none interpretation of the parietal effect to hold, the guesses that might influence the differences between the two critical categories are those on which neither task nor voice information was retrieved. These trials are presumably not associated with old/new effects, and would, in the absence of a significant response bias, be represented in approximately equal numbers in the 2-correct and 1-correct categories. Since the number of trials in the 1-correct category was greater than the number in the 2-correct category (the probability of one correct source judgement was greater than the probability of two correct source judgements), the guesses would result in a relatively greater attenuation of the 2-correct than of the 1-correct old/new effects. Thus, if an all-or-none interpretation of the parietal old/new effect was correct, the effect would, if anything, be larger for the 1-correct than for the 2-correct response category.

The analyses at frontal sites revealed that, while not differing from each other, the two classes of ERPs associated with correct source judgements were reliably more positive than the ERPs to correct rejections. From 300 to 800 ms this relative positivity was distributed bilaterally, while later in the epoch the positivity was larger at right- than at left-frontal electrode sites. This lateralised modulation likely corresponds to the right-frontal old/new effect that has been identified in previous ERP studies of source memory (Wilding and Rugg, 1996). In a recent paper, Wilding and Rugg (1997) reported that the right-frontal effect is neurally and functionally dissociable from an earlier frontal modulation, evident from approximately 400–800 ms post-stimulus. While this earlier effect also takes the form of a greater positivity to words associated with correct source judgements, it does not have the lateralised distribution of its right-frontal counterpart. In ERP studies of associative recognition and recall, Donaldson and Rugg (1998, 1999) have demonstrated that the early-frontal effect is functionally dissociable from the parietal old/new effect. To the extent that the frontal old/new effects over the 300–800-ms epoch in this experiment reflect a modulation of the early-frontal effect, the present findings are consistent with those of Donaldson and Rugg, since only the parietal effect varied in magnitude according to the number of correct source judgements that were made. The early-frontal effect may be related to the frontal modulation observed by Rugg et al. (1998) in an ERP study of recognition memory. Drawing on dual-process accounts of recognition memory, they proposed that the effect indexes the process of familiarity (see Jacoby, 1991). For additional discussions of the functional significance of the early-frontal effects see Wilding and Rugg (1997), Curran (1999), as well as Donaldson and Rugg (1998, 1999).

To summarise, analyses of the ERP data revealed reliable left-parietal, early-frontal and right-frontal old/new effects. The findings build on previous claims that at least three neurally and functionally dissociable processes are engaged during successful retrieval of source information (Wilding and Rugg, 1997). Of primary interest, the magnitude of the parietal old/new effect varied with the number of correct source judgements that were made, providing strong support for the view that the parietal ERP old/new effect indexes recollection in a graded rather than an all-or-none fashion. The immediate corollary to this finding is that the process of recollection is itself graded.

Acknowledgements

This research was completed in the Depart-
ment of Experimental Psychology, University of Oxford, with the support of an MRC fellowship awarded to Ed Wilding. I am very grateful to Kia Nobre for the use of her laboratory facilities, as well as for her comments on earlier versions of this manuscript.

References


