PERGAMON

Neuropsychologia 40 (2002) 2288–2304

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

ERP 'old/new' effects: memory strength and decisional factor(s)

Simon Finnigan [a,*], Michael S. Humphreys [b], Simon Dennis [b], Gina Geffen [c]

[a] School of Psychology, The University of Queensland, Brisbane, Qld 4072, Australia
[b] The ARC Key Center for Human Factors and Applied Cognitive Psychology, The University of Queensland, Brisbane, Qld, Australia
[c] Cognitive Psychophysiology Laboratory, The University of Queensland, Brisbane, Qld, Australia

Received 4 June 2001; received in revised form 2 July 2002; accepted 2 July 2002

Abstract

Event-related potentials (ERPs) were recorded while subjects made old/new recognition judgments on new unstudied words and old words which had been presented at study either once ('weak') or three times ('strong'). The probability of an 'old' response was significantly higher for strong than weak words and significantly higher for weak than new words. Comparisons were made initially between ERPs to new, weak and strong words, and subsequently between ERPs associated with six strength-by-response conditions. The N400 component was found to be modulated by memory trace strength in a graded manner. Its amplitude was most negative in new word ERPs and most positive in strong word ERPs. This 'N400 strength effect' was largest at the left parietal electrode (in ear-referenced ERPs). The amplitude of the late positive complex (LPC) effect was sensitive to decision accuracy (and perhaps confidence). Its amplitude was larger in ERPs evoked by words attracting correct versus incorrect recognition decisions. The LPC effect had a left > right, centro-parietal scalp topography (in ear-referenced ERPs). Hence, whereas, the majority of previous ERP studies of episodic recognition have interpreted results from the perspective of dual-process models, we provide alternative interpretations of N400 and LPC old/new effects in terms of memory strength and decisional factor(s).

Keywords: Recognition; N400; LPC; Single-process models; Familiarity; Recollection

1. Introduction

A common finding of many event-related potential (ERP) studies of recognition memory is that presentation of old/repeated items elicits more positive-going ERPs than does presentation of new/unrepeated items (reviewed in [23,34,35]). Such ERP 'old/new effects' typically onset approximately 300–400 ms post-stimulus, last 300–600 ms and, when words are used as stimuli, are generally of greatest magnitude at left parietal and adjacent centro-parietal electrodes. On the basis of differing scalp topographies (e.g. [43]) and differential sensitivities to manipulation of a number of experimental variables (e.g. word frequency and repetition lag [33]), ERP old/new effects are now assumed to comprise dissociable early and late effects which reflect the modulation of the N400 and a subsequent late positive component (or complex). There are incongruities in the literature over the name given to this late positive component of the ERP old/new effect with P3, P300, P600 and late positive complex (LPC) having been used by different authors. Without attempting to resolve the issue of whether or not these labels all represent the exact same entity, the term LPC will be used here in reference to this ERP component. Further support for the N400-LPC old/new effect distinction is provided by evidence from a number of intracranial ERP studies (e.g. [13]) which indicate that the N400 and LPC are generated by different neural populations.

Over the past decade most authors have interpreted ERP old/new effects from the perspective of dual-process models (e.g. [19,27]) of recognition memory. Generally, these models stipulate that recognition comprises familiarity which is often assumed to be a context-insensitive, automatic process bereft of the phenomenological experience of remembering; and recollection, a context-sensitive, strategic, recall-like process involving the conscious retrieval of specific information about the encoding episode. Attempts to relate components of ERP old/new effects to putative familiarity and recollection processes have sometimes possessed low discriminatory power, and must still be considered speculative until further supporting evidence is obtained. It is also noteworthy that some authors have concluded that results from a number of ERP studies of episodic memory provide scant support for dual-process models [35]. Furthermore, these

* Corresponding author. Tel.: +61-7-3365-4100; fax: +61-7-3365-3383.
E-mail address: finnigan@cmr.uq.edu.au (S. Finnigan).

0028-3932/02/$ – see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0028-3932(02)00113-6

models have been found to be unsatisfactory on a number of grounds (e.g. [17]) and alternative approaches have been posited (e.g. [10,18]). Thus, while the results of many ERP studies of episodic recognition have been interpreted in the context of dual-process models, alternative memory models warrant further consideration in the interpretation of ERP old/new effects.

It has been proposed that a N400 component in ERPs recorded at parietal electrodes is sensitive to implicit memory processes [41], and also that it is unlikely that the N400 old/new effect reflects activity associated with the explicit discrimination of old from new items [23,34]. The latter proposal was made on the basis of the failure of a number of investigations to obtain an N400 old/new effect when the study-test interval exceeds some time between 2 and 15 min (e.g. [25,37]). In addition, a N400-like component with a more frontal distribution, the 'FN400' (following [6]), has been proposed to index familiarity [6,7,41]. However, as noted above, ERP data need not be associated with approaches that distinguish between putative familiarity and recollection processes. For example, an alternative framework is provided by global matching models [16] which assume that the recognition decision is based on the summed strength of all matches between the cue(s) and all traces in a single global memory system. Separate episodic and non-episodic systems are not proposed. Instead the memory system can be cued with both item and contextual information so in contrast to the concept of familiarity the matching operation is generally thought to be context-dependent. Some memory models propose a relatively rapid change of contextual representations over time [15]. Thus, the observed susceptibilities to temporal decay of single item recognition performance (e.g. [14]) and of the N400 old/new effect, may be the consequences of such contextual change. The strength of the global match is a continuous variable, and the same is often thought to be true of dual-process familiarity (e.g. [55]). We will hereafter refer to any such continuous or graded variable (as distinct from a categorical variable, such as recollection) as strength. Differences between this and other notions of familiarity will be highlighted where appropriate. Within the global matching framework a decision criterion is set, as in signal detection theory (SDT) [45], such that strengths above the criterion will result in an old response whereas those below will produce a new response. Proponents of global matching models would thus predict that at least some component of the ERP old/new effect reflects the strength of the memory trace and of the matching operation.

In contrast to the N400, the LPC old/new effect has reliably been observed after longer study-test intervals and is broadly considered to be reflective of long-term episodic recognition processes [23,34]. Indeed the LPC old/new effect has been linked to the recollection component of dual-process recognition models (reviewed in [2]). Alternative interpretations of the cognitive processes reflected in the LPC old/new effect have also been posited. Some have proposed that LPC peak amplitude reflects memory trace strength [5], others discriminability [24], or decision accuracy and/or confidence (e.g. [24,26,28,30,32,46]). Within the SDT framework confidence is conceptualized as distance from the response criterion. Hence it appears that LPC amplitude is potentially influenced not purely by items' old/new status but also by a number of variables incorporating strength, the position of the response criterion, and consequently recognition performance. So a simple old/new (much less, hit-correct rejection) ERP comparison cannot adequately address this issue, as old/new effects may reflect differing levels of item strength or confidence. The effect may also potentially reflect response-related processes when ERPs are conditionalized on response and ERPs associated with differing responses are compared (although there is some evidence that this is not the case, e.g. [38]). Moreover an old/new ERP comparison wherein only correct response trials are chosen for averaging is prone to contamination by item selection artifacts, as only a subset of all available old/new items are considered for comparison.

Thus, we employed a strength manipulation: half of the study words were presented once ('weak') and half presented three times ('strong'), a procedure which provides three distinct levels of strength (new, weak, strong) at test. Excluding repetitions in continuous tasks, the only previous reported examinations of ERPs evoked by words which had already been repeated within a different phase of the experiment were those of Bentin and co-workers [3,5]. However, in those experiments, words were presented first in a study list (and categorical decisions made), second during an episodic recognition task, and for a third time in a subsequent implicit memory task. Here we compare ERPs elicited during an explicit episodic recognition test by words which were new, or which were presented once or thrice in a previous study list. For this new–weak–strong ERP comparison, all available items are considered so this comparison is not subject to item selection artifacts.

Despite the large body of literature investigating ERP old/new effects over the past two decades, most studies have not examined in detail ERPs accompanying erroneous new and old responses (misses and false alarms, respectively). Indeed some authors have concluded that the effects are not found for incorrect recognition judgments to either old or new words (e.g. [1,53,54]). However, visual inspection of ERPs in those papers which have illustrated error response ERPs suggests that this may be an oversimplification, with different ERP patterns evidently obtained between N400 and LPC old/new effects, and between frontal and posterior electrodes. For example, in the results of Rugg et al. [41] (Fig. 1), Wilding et al. [51] (Fig. 3) and Wilding and Rugg [53] (Fig. 2) ERPs for misses exhibited an apparent N400 (but not LPC) old/new effect, at parietal (but not at frontal) electrodes. Van Petten and Senkfor [48] (Figs. 5 and 6) obtained false alarm ERPs which exhibited an apparent N400 old/new effect relative to correct rejection ERPs. Whereas, LPC amplitudes in correct rejection ERPs have been obtained which were evidently larger than those in false alarm ERPs in the data of Van Petten and Senkfor [48] and of Rubin et al. [32] (Figs. 2 and 3). Hence, a further aim of this study was to systematically compare false alarm and miss ERPs, with those associated with hits and correct rejections. Despite the fact that by conditionalizing ERPs on response our comparisons are potentially subject to item selection artifacts, if results from those converge with data from the initial strength comparisons, then we have a stronger position from which to interpret the data. Moreover, our conclusions will be based upon a greater number of independent variables and degrees of freedom than have typically been utilized, in an attempt to further determine the neurocognitive processes underlying ERP old/new effects.

2. Methods

Two experiments were conducted, each on a different group of subjects. As outlined below, the experimental procedures of these were very similar, but the test lists (and consequently, mean study-test repetition lag times) were shorter in the second experiment. Data for each experiment was analyzed separately, for a number of reasons: first, as summarized in Section 1, N400 old/new effects have been shown to be sensitive to repetition lag, hence those effects may significantly differ between experiments. In addition, we feel that a replication of results (e.g. similar effects during a given latency interval at particular electrodes) between experiments would strengthen our claims and, given that we approach this research from a different theoretical framework to the majority of previous related studies, such a replication would be especially beneficial.

2.1. Subjects

2.1.1. Experiment 1

All subjects for both experiments were recruited from the University of Queensland School of Psychology first year subject pool. All received psychology course credit for their participation. A total of 22 subjects participated in experiment 1. Behavioral and electrophysiological data from five subjects were discarded due to insufficient numbers of trials with which to form reliable ERPs for response-related conditions following artifact rejection (see below for rejection criteria). The remaining 17 subjects included in the analyses had a mean age of 23.0 years, 7 were female and 15 were right-handed.

2.1.2. Experiment 2

A total of 23 subjects participated in experiment 2. Behavioral and electrophysiological data from four subjects were discarded due to insufficient numbers of trials with which to form reliable ERPs for response-related conditions following artifact rejection. The remaining 19 subjects included in the analyses had a mean age of 19.1 years, 9 were female and 17 were right-handed.

2.2. Procedure

2.2.1. Experiment 1

Each subject participated in three study-test cycles. At study subjects, viewed words (five letters, frequency 10–30 per million, source The Sydney Morning Herald Word Database [9]) presented sequentially, centrally on a computer monitor and were instructed to memorize them for an immediate subsequent episodic recognition test. Each word was presented for 400 ms, with an interval of 200 ms between words. Each study list comprised 120 words total. Within each third of a study list, 10 words were presented once (‘weak’ words) and 10 words presented three times (‘strong’ words), making a total of 40 word presentations (20 unique words) presented randomly within each third. A test list immediately followed each study list. At test words that were old (weak and strong) and new (not in the study list) were presented one at a time, and subjects made an old/new episodic recognition judgment for each test word. Each third of a test list comprised of 10 new, 10 weak and 10 strong words, randomly intermixed. Thus, the total length of each test list was 90 words. All old words were presented in the same third of the test list as at study in order to minimize study-test repetition lag variability. Average study-test repetition lag for items from the first third of the study and test lists was 138 s, and for items from the third of the lists 347 s. Each test word was preceded by the appearance of a fixation asterisk for 600 ms, a blank screen for 200 ms, then the word appeared for 400 ms. Following presentation of each test word the screen was blanked for 1500 ms and subjects were instructed to withhold their response during this interval. Next, the words ‘old’ and ‘new’ appeared on each side of the screen center for 1500 ms as a prompt to respond. Subjects’ index fingers were each resting on one of two response buttons throughout the test phase and subjects responded by pressing the button corresponding to their recognition judgment. Hands used for a given response were fully counterbalanced across subjects.

2.2.2. Experiment 2

The sole procedural difference between experiments 1 and 2 was that in experiment 2, the total length of the recognition test lists was reduced from 90 to 60 words. As was the case in experiment 1, total study list length was 120 words in experiment 2. However, each test list in experiment 2 comprised 20 weak, 20 strong, and 20 new, words (whereas 30 of each were tested in experiment 1). Again each third of a test list comprised approximately equal numbers of weak, strong and new words, and each old test word was presented in the same third of the test list as at study in order to minimize study-test repetition lag variability. Average study-test repetition lag for items from the first third of the study and test lists was 117 s, and for items from the third of the lists 237 s. The 10 weak and 10 strong words from each study list that were not immediately tested for recognition, were presented along with an equal number of new words in a further test subsequent to the three study-test blocks. The final test required a frequency of study presentation judgment, the response options being 0, 1, 2, or 3 presentations. Subjects were not instructed about the requirements of the frequency judgment test until after completing the third recognition test. Whilst EEG was recorded during the frequency judgment test, that data is not presented here.

2.3. Electrophysiological recording and analysis

Identical electrophysiological data acquisition and analysis procedures and parameters were used in both experiments. Scalp EEG was recorded from 30 tin electrodes embedded in an elastic cap. Electrode locations corresponded to the following sites of the International 10–20 system: FP1, FP2, FPZ, F3, F4, F7, F8, FZ, FC3, FC4, FCZ, FT7, FT8, C3, C4, CZ, CP3, CP4, CPZ, P3, P4, PZ, T3, T4, T5, T6, TP7, TP8, O1, O2 [22]. All EEG electrodes were referenced to a linked pair of electrodes, one positioned on each earlobe. Electrode impedance was predominantly 10–15 kΩ or less. Vertical eye movements and blinks were monitored via electrodes placed on the supraorbital ridge of, and below, the left eye. Horizontal eye movements were monitored via two electrodes, one on the outer canthus of each eye. EEG, recorded continuously, was filtered (bandpass 0.01–100 Hz) and digitized at a sampling rate of 500 Hz. Vertical and horizontal electrooculogram (EOG) artifacts were reduced using the procedure of Semlitsch et al. [42]. Continuous EEG data were later divided into epochs beginning 200 ms before, and ending 1200 ms after, each word presentation, baseline corrected over a 200 ms pre-stimulus interval, and low-pass filtered at 40 Hz. Epochs in which EEG amplitude exceeded criteria of $\pm 100\ \mu V$ were rejected prior to averaging. Separate grand average ERPs were initially computed for strong hits and misses, weak hits and misses, correct rejections and false alarms. Data was discarded from subjects for whom every individual response-conditionalized ERP average did not comprise at least 10 artifact-free trials. Very similar results were obtained in each experiment and we report this ERP data collapsed across both studies. Furthermore, in each experiment's data, grand average ERPs associated with error conditions (false alarms, strong and weak misses) were relatively noisy so by collapsing the data across both experiments, signal-to-noise ratios are enhanced. Across both experiments, the mean numbers of ERP trials per response condition per subject were: correct rejections, 36.83; false alarms, 11.78; weak hits, 28.67; weak misses, 20; strong hits, 39.11; strong misses, 11.86. ERPs (for each subject) for new, weak and strong words were formed by computing weighted averages of the two respective response-conditionalized ERPs for each strength level. In all cases, each subjects' ERP for a given condition was weighted equally into the grand average ERP. Given the relatively large numbers of trials contributing to each strength condition, separate grand average strength-conditionalized ERPs were computed and analyzed for each experiment. For experiment 1, the mean numbers of ERP trials per strength condition per subject were: new, 53.94; weak, 56.12; strong, 55.82. For experiment 2, the mean numbers of ERP trials per strength condition per subject were: new, 43.84; weak, 42.26; strong, 44.84.

3. Results

An alpha level of 0.05 was used for all statistical tests. For all ANOVAs reported hereafter, violations of the heterogeneity of covariance assumption were corrected using the Greenhouse–Geisser procedure [12] and corresponding F ratios are, where appropriate, reported with corrected degrees of freedom. In addition all sets of planned-comparison t-tests employed a Bonferroni adjustment (according to the number of comparisons in the set, at a given electrode, in the case of ERP data) in order to maintain the family-wise Type I error rate at a sufficiently low level.

3.1. Behavioral data

3.1.1. Experiment 1

Averaged across the three study-test blocks, hit rates (and, in brackets, the corresponding standard deviations) for strong and weak words were 0.744 (0.126) and 0.563 (0.091), respectively, and the false alarm rate was 0.229 (0.095). Recognition performance data were submitted to within-subjects ANOVA with strength and block as factors and probability of an old response as the dependent variable. As expected, a significant strength effect was obtained ($F(1.4, 22.5) = 152.25$, $P < 0.001$). Planned pairwise comparisons (Bonferroni-corrected, two comparisons) revealed that hit rates to strong words ($M = 0.744$, S.D. $= 0.126$) were significantly greater than weak hit rates ($M = 0.563$, S.D. $= 0.091$; $t(16) = 9.84$, $P < 0.001$) which were in turn significantly greater than false alarm rates ($M = 0.229$, S.D. $= 0.095$; $t(16) = 10.39$, $P < 0.001$). Identical graded simple effects of strength were observed for all three blocks, and no significant block effect ($F(2, 32) = 0.589$, $P = 0.561$) nor any strength × block interaction ($F(4, 64) = 1.529$, $P = 0.204$), were obtained.

3.1.2. Experiment 2

Averaged across the three study-test blocks, hit rates (and, in brackets, the corresponding standard deviations) for strong and weak words were 0.783 (0.094) and 0.596 (0.105), respectively, and the false alarm rate was 0.248 (0.093). Recognition performance data were again submitted to within-subjects ANOVA with strength and block as factors and probability of an old response as the dependent variable. As in experiment 1, a significant strength effect was obtained ($F(2, 36) = 195.866$, $P < 0.001$).

As for experiment 1, planned pairwise comparisons (Bonferroni-corrected, two comparisons) revealed that hit rates to strong words ($M = 0.783$, S.D. $= 0.094$) were significantly greater than weak hit rates ($M = 0.596$, S.D. $= 0.105$; $t(18) = 7.797$, $P < 0.001$) which were in turn significantly greater than false alarm rates ($M = 0.248$, S.D. $= 0.093$; $t(18) = 11.371$, $P < 0.001$). Also in accordance with the experiment 1 data, identical graded simple effects of strength were observed for all blocks, and no significant block effect ($F(1.4, 23.5) = 0.527$, $P = 0.532$) nor strength × block interaction ($F(2.5, 43.3) = 2.163$, $P = 0.115$) were obtained.

3.2. Electrophysiological data

3.2.1. Experiment 1

Grand average ERP waveforms evoked by strong, weak and new words (collapsed across both old and new responses) in experiment 1 are displayed in Fig. 1 for electrodes C3, C4, CZ, F3, F4, F7, F8, FZ, P3, P4, PZ, T3, T4, T5, T6. Given that these 15 scalp electrode sites provide a well-distributed and representative sample of all recording sites utilized and all ERP effects apparent to visual inspection at any electrode in the present study, and for the purpose of brevity, ERPs from these electrodes alone were submitted to ANOVA. It appears that between approximately 300 and 600 ms post-stimulus onset, ERP amplitudes at parietal (and in some cases, adjacent) electrodes are positively correlated with memory strength (strong positive relative to weak, weak positive relative to new). This apparent ERP 'strength effect' onsets at approximately 300 ms with the peak of a positive component which is apparently the P2, incorporates the entire N400 and persists until approximately 650 ms at which time the LPC approaches peak amplitude. Thus we have evidently obtained an N400 old/new effect when the study-test repetition lag for the majority of old words was greater than 2 min, and for some words, was approximately 6 min. This strength-related pattern is diminished to a certain extent at central, and to a greater extent at frontal electrodes. No graded strength effect was apparent for the LPC. Rather the amplitude of this component was apparently identical for new and weak words at parietal electrodes. Moreover, new word ERPs' LPC peak amplitude was noticeably greater than that of weak ERPs at electrodes T3, C3, CZ and C4, and was in fact equivalent to that of strong ERPs at electrode CZ.

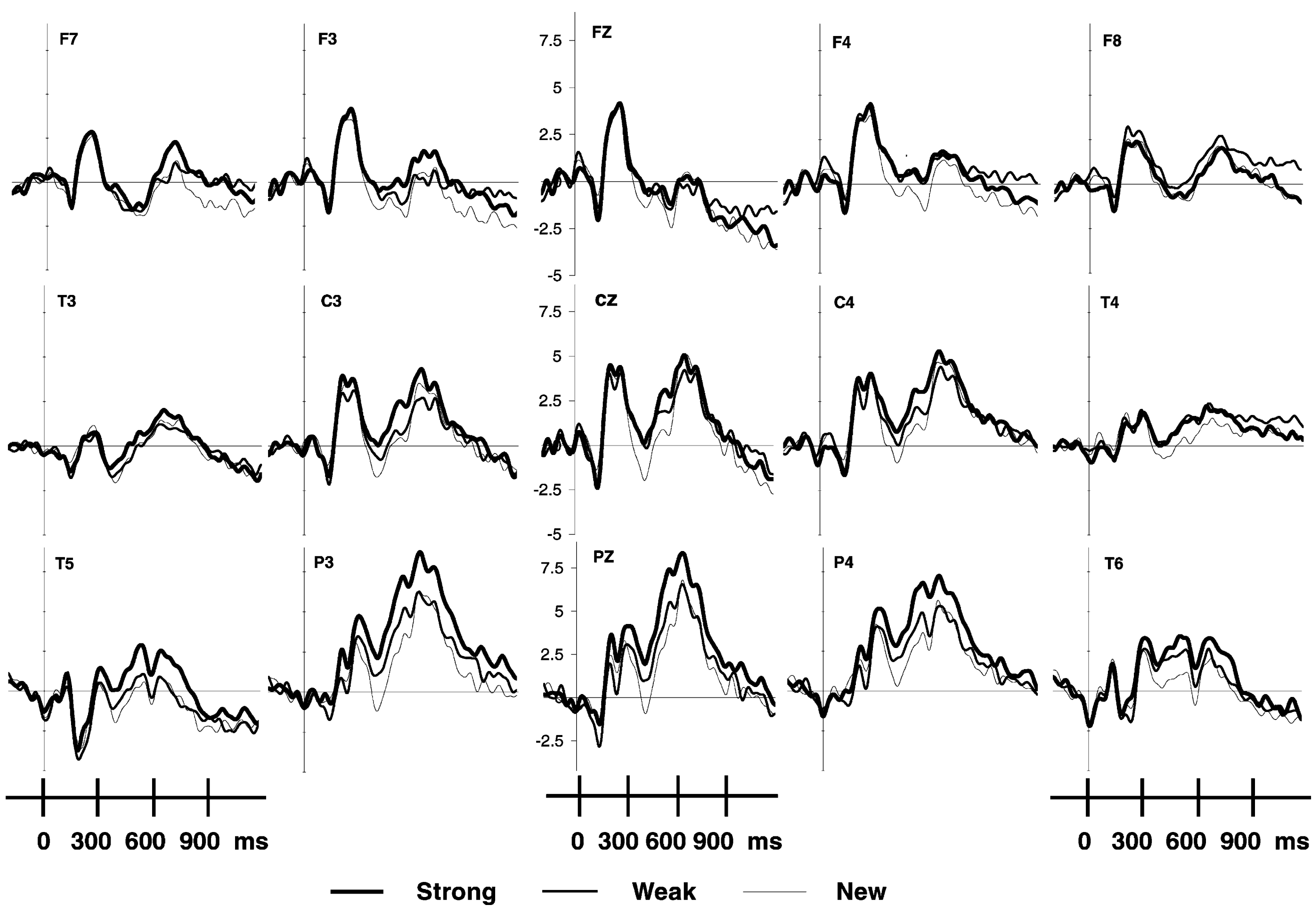

Fig. 1. Grand average ($N = 17$) ERPs evoked by new, weak (presented once at study) and strong (presented thrice at study) words in experiment 1. The mean numbers of individual ERP trials per strength condition per subject were: new, 53.94; weak, 56.12; strong, 55.82.

ERPs were quantified for analysis by computing the mean amplitudes (relative to the mean of the 200 ms pre-stimulus baseline interval) of the 300–500 and 500–800 ms intervals, which have been utilized in a number of previous ERP studies (e.g. [33,36,41,50]) to represent the N400 and LPC components, respectively. Consistent with previous studies, separate within-subjects ANOVAs were conducted for each of these time intervals. These employed the factors of strength (new, weak, strong) and electrode site (C3, C4, CZ, F3, F4, F7, F8, FZ, P3, P4, PZ, T3, T4, T5, T6).

3.2.1.1. The 300–500 ms interval. At this latency range, significant strength ($F(2, 32) = 8.291$, $P < 0.005$) and electrode ($F(3.5, 56.4) = 3.951$, $P < 0.01$) effects, and a significant strength × electrode interaction ($F(5, 79.4) = 2.898$, $P < 0.05$) were obtained. Considering the strength effect (Bonferroni-corrected, two comparisons), both strong word ($M = 1.28$, S.D. $= 2.561$; $t(16) = 3.597$, $P < 0.01$) and weak word ERPs ($M = 0.805$, S.D. $= 2.277$; $t(16) = 2.954$, $P < 0.05$) were positive relative to those evoked by new words ($M = 0.021$, S.D. $= 2.895$). With regard to the electrode effect, average amplitudes were relatively most negative at electrodes T3 and F7, but were relatively most positive at electrode P4, followed by electrodes PZ and P3. Considering the interaction, amplitude differences between conditions were largest at electrode P3 (then at PZ and P4), and were least at electrode F7. Given several lines of evidence in the ERP literature which indicate that old/new effect(s) centered around left parietal electrodes may be sensitive to memory trace strength, the following planned comparisons (Bonferroni-corrected, two comparisons) were conducted subsequent to the above ANOVA. At electrode P3 strong ERPs ($M = 3.553$, S.D. $= 3.587$) were positive relative to weak ($M = 2.129$, S.D. $= 3.063$; $t(16) = 2.69$, $P < 0.05$), which were in turn positive relative to new, ERPs ($M = 0.647$, S.D. $= 3.502$; $t(16) = 3.418$, $P < 0.01$). Given proposals that the frontal FN400 reflects familiarity, the following planned comparisons (Bonferroni-corrected, two comparisons) were carried out on the electrode FZ data: strong ($M = 0.012$, S.D. $= 4.151$) were not significantly different to weak word ERPs ($M = 0.253$, S.D. $= 4.140$; $t(16) = -0.399$, $P = 0.695$); and there was no significant amplitude difference between the former and new, ERPs ($M = -0.465$, S.D. $= 4.596$; $t(16) = 0.700$, $P = 0.494$).

3.2.1.2. The 500–800 ms interval. ANOVA at this latency range revealed a strength effect which approached significance ($F(2, 32) = 3.083$, $P = 0.072$). In contrast to the outcomes obtained for the 300–500 ms data, there were no significant amplitude differences between new ($M = 1.507$, S.D. $= 3.600$) and weak ($M = 1.815$, S.D. $= 3.206$; $t(16) = 0.755$, $P = 0.461$), nor between the latter and strong ($M = 2.691$, S.D. $= 3.624$; $t(16) = 1.466$, $P = 0.162$) ERPs. A significant electrode effect ($F(4.6, 73.4) = 7.715$, $P < 0.001$) was obtained. Collapsed across strength, the average amplitude was generally greatest at frontal and central electrode sites, but was also of a similar magnitude at electrode P3, and was least at electrode T5. In contrast to the 300–500 ms data, no significant strength × electrode interaction was obtained over the 500–800 ms interval ($F(5.2, 83.5) = 1.756$, $P = 0.128$). In order to assess the proposal that the LPC effect is sensitive to strength [4], the following planned comparisons (Bonferroni-corrected, two comparisons) were computed. At electrode P3, average amplitudes of strong ($M = 7.212$, S.D. $= 4.570$) were positive relative to those of weak, words' ERPs ($M = 4.994$, S.D. $= 4.034$; $t(16) = 2.782$, $P < 0.05$); But there was no significant difference between average amplitudes of weak versus new words' ERPs ($M = 4.429$, S.D. $= 5.009$; $t(16) = 0.799$, $P = 0.436$).

3.2.2. Experiment 2

Grand average ERP waveforms evoked by strong, weak and new words (collapsed across both old and new responses) in experiment 2 are displayed in Fig. 2 for electrodes C3, C4, CZ, F3, F4, F7, F8, FZ, P3, P4, PZ, T3, T4, T5, T6. Between approximately 300 and 700 ms post-stimulus onset, ERP amplitudes at parietal electrodes appear to exhibit a graded strength effect as was obtained in experiment 1. A briefer strength effect is apparent at central and frontal electrodes between approximately 300 and 500 ms. As in experiment 1, LPC peak amplitude at central electrodes is evidently identical for weak and new ERPs. ERPs were again quantified for analysis by computing mean amplitudes (relative to the mean of the 200 ms pre-stimulus baseline interval) of the 300–500 and 500–800 ms intervals, which are representative of the N400 and LPC components, respectively. Separate within-subjects ANOVAs were again conducted for each of these time intervals, employing the factors of strength (new, weak, strong) and electrode site (C3, C4, CZ, F3, F4, F7, F8, FZ, P3, P4, PZ, T3, T4, TS, T6).

3.2.2.1. The 300–500 ms interval. Over this interval significant strength ($F(2, 36) = 19.66$, $P < 0.001$) and electrode ($F(2.8, 50.2) = 4.417$, $P < 0.01$) effects were obtained. However, no significant strength × electrode interaction ($F(5.3, 94.8) = 1.646$, $P = 0.152$) was obtained. With regard to the strength effect, strong word ($M = 0.945$, S.D. $= 2.123$) were positive relative to weak word ERPs ($M = 0.020$, S.D. $= 2.456$; $t(18) = 3.447$, $P < 0.01$) which were in turn positive relative to those evoked by new words ($M = -1.034$, S.D. $= 2.672$; $t(18) = 3.426$, $P < 0.01$). Considering the electrode effect, average amplitude was relatively most negative at electrodes F7 and F3, and most positive at parietal electrodes and T6. As was the case in the corresponding experiment 1 data, amplitude differences between conditions were generally largest at parietal, and least at frontal, electrodes. However, the lack of an interaction in the experiment 2 data appears to be associated with the more pronounced (relative to those obtained in experiment 1) graded N400 strength effects at frontal electrodes. As in experiment 1, the following planned

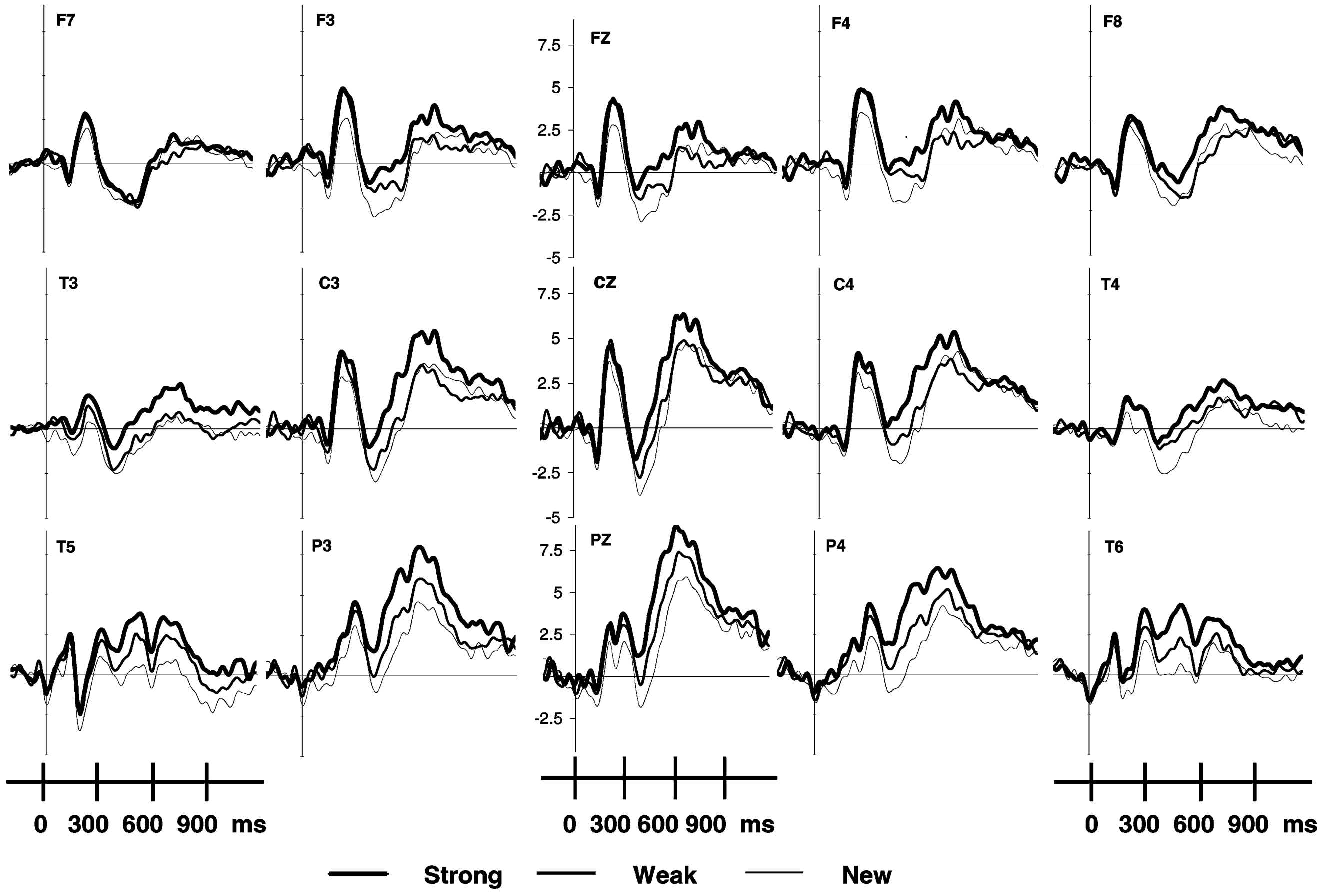

Fig. 2. Grand average ($N = 19$) ERPs evoked by new, weak (presented once at study) and strong (presented thrice at study) words in experiment 2. The mean numbers of individual ERP trials per strength condition per subject were: new, 43.84; weak, 42.26; strong, 44.84.

comparisons (Bonferroni-corrected, two comparisons per set) were conducted subsequent to the above ANOVA. At electrode P3 strong ERPs ($M = 3.066$, S.D. $= 3.245$) were positive relative to weak ($M = 1.632$, S.D. $= 3.511$; $t(18) = 3.107$, $P < 0.05$) which were in turn positive relative to new ERPs ($M = 0.518$, S.D. $= 3.259$; $t(18) = 2.999$, $P < 0.05$); in electrode FZ recordings, strong ($M = 0.008$, S.D. $= 3.436$) were not significantly different to weak word ERPs ($M = -0.705$, S.D. $= 4.111$); $t(18) = 1.635$, $P = 0.119$), whereas the latter were positive relative to new ERPs ($M = -1.966$, S.D. $= 4.469$; $t(18) = 2.705$, $P < 0.05$). Thus the graded N400 strength effect at electrode P3, and the non-significant weak versus strong electrode FZ N400 comparison, obtained in experiment 1, have both been replicated in experiment 2. In addition, a graded N400 strength effect was obtained in the data collapsed across all 15 electrodes, and a significant N400 old/new difference was obtained at electrode FZ. The latter two results, which were obtained only in the experiment 2 data, may be a consequence of the shorter repetition lags between study and test in this experiment.

3.2.2.2. The 500–800 ms interval. ANOVA at this latency range revealed a significant strength effect ($F(2, 36) = 11.243$, $P < 0.001$). Strong ($M = 3.428$, S.D. $= 2.474$) were positive relative to weak word ERPs ($M = 1.959$, S.D. $= 3.019$; $t(18) = 4.472$, $P < 0.001$) which were not significantly different in amplitude to ERPs evoked by new words ($M = 1.480$, S.D. $= 3.209$; $t(18) = 1.130$, $P = 0.273$). A significant electrode effect ($F(2.9, 52.9) = 8.985$, $P < 0.001$) and a significant strength $\times$ electrode interaction ($F(4.9, 88.4) = 2.360$, $P < 0.05$) were obtained. Average amplitude over this latency interval was largest at parietal electrodes (left > right), and least at electrodes F7 and T3. The presence of an interaction evidently reflects the apparent differences in LPC amplitude patterns between, for example, parietal and frontal electrodes. That is, the range of amplitudes across conditions, was larger at parietal than frontal electrodes. At electrode P3 (Bonferroni-corrected, two comparisons), average amplitudes of strong ($M = 6.716$, S.D. $= 4.305$) were positive relative to those of weak, words' ERPs ($M = 4.823$, S.D. $= 4.912$; $t(16) = 3.292$, $P < 0.01$); But there was no significant difference between mean amplitudes of weak versus new words' ERPs ($M = 3.365$, S.D. $= 4.493$; $t(16) = 2.295$, $P = 0.097$). Both outcomes were the same as those obtained in the corresponding comparisons of the experiment 1 data, and indicate that the LPC effect is not sensitive to strength in the same graded fashion as is the N400 effect.

3.2.3. Experiments 1 and 2: response-conditionalized electrophysiological data

In addition to comparing ERPs for the three different levels of strength, comparisons were also made between ERPs conditionalized on response. As noted above, very similar results were obtained in each experiment and for this reason, plus in order to obtain optimal signal-to-noise ratios, we report the response-conditionalized ERP data collapsed across both studies. Grand average ($N = 36$) ERP waveforms for strong and weak hits, strong and weak misses, correct rejections and false alarms, are displayed for midline parietal, central and frontal electrodes in Fig. 3. Amplitudes of strong and weak hit ERPs are predominantly the more positive of all conditions from approximately 200 ms until the end of the epoch at all electrodes. At electrodes CZ and PZ (and adjacent sites), differing patterns are apparent between N400 and LPC peak amplitude orders. Most notably, whereas, N400 amplitude is generally largest (most negative) for correct rejection ERPs, they undertake a steep incline to reach an LPC peak amplitude which is greater (more positive) than that of false alarm and miss ERPs and indeed, at electrode CZ, on a par with that of weak hit ERPs. (The different amplitude order patterns obtained for N400 versus LPC, are examined further via supplementary analyses at the end of this section.) Hence, centro-parietal LPC amplitude is apparently greater for the three correct than the three incorrect, response conditions. These patterns of centro-parietal LPC amplitude are evidently consistent with the view that this ERP component indexes decisional factor(s), such as accuracy, and perhaps—assuming correct recognition decisions are on average generally made with higher confidence than are erroneous decisions (e.g. [32])—confidence.

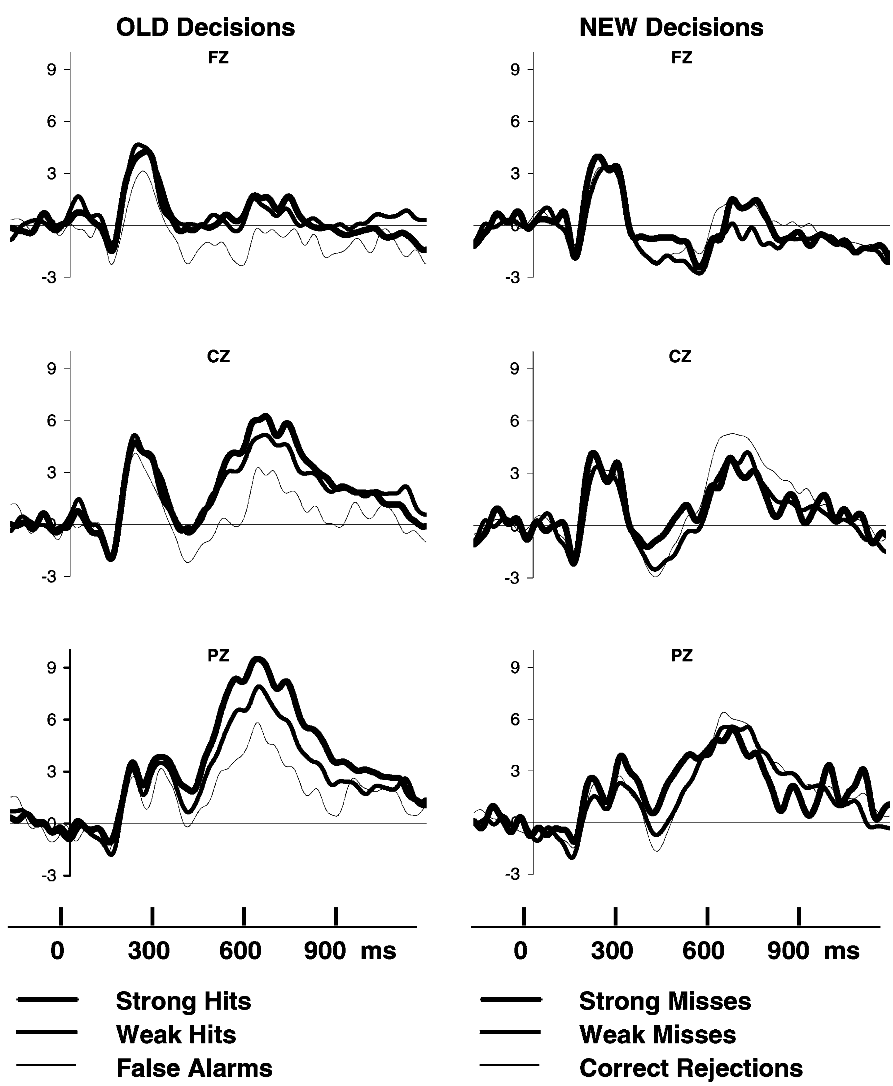

Fig. 3. Grand average ($N = 36$) ERPs evoked by strong hits and misses, weak hits and misses, correct rejections and false alarms, averaged across experiments 1 and 2. The mean numbers of individual ERP trials per response condition per subject were: correct rejections, 36.83; false alarms, 11.78; weak hits, 28.67; weak misses, 20; strong hits, 39.11; strong misses, 11.86.

As for strength-conditionalized ERPs, those associated with response conditions were prepared for analysis by computing mean amplitudes over the 300–500 and 500–800 ms intervals. Separate within-subjects ANOVAs were conducted for each of these time intervals, employing the factors of condition (strong hit, weak hit, false alarm, strong miss, weak miss, correct rejection) and electrode site (C3, C4, CZ, F3, F4, F7, F8, FZ, P3, P4, PZ, T3, T4, T5, T6). Considering previous ERP literature and following inspection of ERP data from experiment 1, the following comparisons were planned for the data from both experiments combined: (1) old (strong and weak hits and false alarms) versus new (strong and weak misses and correct rejections) responses; (2) strong hits versus weak hits; (3) strong and weak misses versus correct rejections; (4) false alarms versus correct rejections; (5) correct (strong and weak hits and correct rejections) versus incorrect (strong and weak misses and false alarms) responses. These comparisons were chosen in an endeavor to obtain evidence which might converge with that from the ERP strength comparisons and thereby provide further support for ERP strength and 'decisional' effects. In particular, comparisons (3) and (4) could be most diagnostic of a relationship between LPC amplitude and decisional factor(s), because given the directions of those differences they would not be consistent with the views that LPC indexes strength. And, given evidence that false alarms are often made on the basis of 'false recollection' (e.g. [31] they would also argue against a sensitivity of the LPC to recollection.

ERP data recorded from electrode P3, over both the 300–500 and 500–800 ms time intervals, were thus compared. Comparison (1) was computed only on the 300–500 ms data, whereas comparison (5) was computed only on the 500–800 ms data. This was done because those comparisons are not fully orthogonal with one another, and the first comparison further assesses a sensitivity to strength, whereas the latter further assesses a sensitivity to decisional factor(s). Otherwise, we chose to do very similar comparisons on data from both time intervals in order to determine whether or not dissociations in the outcomes were obtained between the N400 and LPC components, at a given electrode. Interestingly, such dissociations—which generally become more evident as the number of compared ERP conditions increases—have indeed been reported in recent literature (e.g. [41]), and such outcomes highlight the value of performing separate analyses of the N400 versus LPC latency intervals. Some authors have previously performed analyses of left parietal electrode ERP data over just one time interval. For example, Rugg et al. [40] computed mean amplitudes over the 400–800 ms window alone. Such an approach seems to indicate—at least implicitly—an assumption that 'the left parietal old/new effect' comprises one sole component but, on the basis of a number of lines of evidence, such an assumption does not appear to hold. This issue is further explored by both these comparisons and, furthermore, by the supplementary analyses reported at the end of this section (on amplitude differences between the 300–500 and 500–800 ms intervals). Electrode P3 was selected, as that is the closest electrode site in the present study, to that at which the left parietal ERP old/new effect has in the past been reliably observed to be of greatest magnitude. In addition, it was at this electrode that, during the 300–500 ms interval, the strength effects were largest in both experiments' data, and this was also evident in the response-conditionalized data collapsed across the two studies. During the 500–800 ms interval, average amplitudes were generally greatest at electrode P3. Moreover, in both experiments, the range of amplitudes across conditions, and similarly, the sizes of some effects which are of particular interest to the current research were largest at electrode P3 (and, in some cases, adjacent parietal and central electrodes). (Note that while electrode P3 ERP data are not illustrated in Fig. 3, that data was almost identical to the PZ data.)

Electrode CZ was also chosen for comparisons of 500–800 ms amplitudes, because in the data from experiment 1, the decision accuracy-related LPC amplitude difference (correct > incorrect conditions) was of at least equal (if not greater) magnitude at CZ versus P3. The response-conditionalized data collapsed across both experiments, also exhibit a similar outcome. Moreover, by comparing the directions of LPC amplitude differences between various response conditions at electrodes P3 versus CZ, it is possible to further consider whether or not the effects at these electrodes, should be assumed to reflect the modulation of distinct ERP components (or indeed, complexes of components). That is, it is not fully clear as to whether the LPC effects at P3 versus CZ should somehow be considered distinct. On the other hand, it may be the case that a significant distinction between them is not justified but instead, that they both reflect a common LPC effect which exhibits a centro-parietal scalp distribution in the current ear-referenced data. Hence, the following comparisons, performed on the 500–800 ms data at each electrode, permit further evaluation of these alternative scenarios.

3.2.3.1. The 300–500 ms interval. ANOVA on this interval obtained a significant condition effect ($F(3.4, 117.5) = 6.379$, $P < 0.001$), a significant electrode effect ($F(3.5, 121) = 7.905$, $P < 0.001$), and a significant condition × electrode interaction ($F(11, 385.3) = 1.982$, $P < 0.05$). Ranging from relatively most to least positive, the obtained pattern of average amplitudes, collapsed across electrodes, was as follows: strong hits ($M = 1.202$, S.D. $= 2.347$), weak hits ($M = 0.937$, S.D. $= 2.496$), false alarms ($M = 0.707$, S.D. $= 3.195$), strong misses ($M = 0.044$, S.D. $= 3.529$), weak misses ($M = -0.511$, S.D. $= 2.891$), correct rejections ($M = -0.671$, S.D. $= 2.921$). Collapsed over conditions, amplitudes were relatively most positive at parietal and adjacent temporal electrodes (right > left), and most negative at electrodes F7, T3, and F3. Considering the interaction, the range of amplitudes obtained across conditions, was largest at parietal sites, and least at frontal, sites.

At electrode P3 (Bonferroni-corrected, four comparisons), mean 300–500 ms amplitudes of ERPs associated with old responses ($M = 2.555$, S.D. $= 3.381$) was significantly positive relative to that of ERPs associated with new responses ($M = 1.275$, S.D. $= 3.357$; $t(35) = 3.690$, $P < 0.01$). Strong hit electrode P3 mean ERP amplitude over the 300–500 ms interval ($M = 3.464$, S.D. $= 3.377$) was significantly greater than that of ERPs evoked by weak hits ($M = 2.449$, S.D. $= 3.369$; $t(35) = 2.980$, $P < 0.05$). Mean false alarm electrode P3 ERP amplitude over the same interval ($M = 1.753$, S.D. $= 4.893$) was significantly positive relative to that of correct rejection ERPs ($M = 0.318$, S.D. $= 3.390$; $t(35) = 2.734$, $P < 0.05$), as was the average amplitude of strong and weak miss ERPs ($M = 1.753$, S.D. $= 3.740$; $t(35) = 2.958$, $P < 0.05$).

3.2.3.2. The 500–800 ms interval. ANOVA on this interval also revealed a significant condition effect ($F(4.1, 142.2) = 7.106$, $P < 0.001$), a significant electrode effect ($F(4.2, 146.7) = 14.429$, $P < 0.001$), and a significant condition × electrode interaction ($F(13, 454.4) = 3.192$, $P < 0.001$). Ranging from relatively most to least positive, the obtained pattern of average amplitudes, collapsed across electrodes, was as follows: strong hits ($M = 3.326$, S.D. $= 3.014$), weak hits ($M = 2.345$, S.D. $= 3.096$), strong misses ($M = 1.784$, S.D. $= 3.459$), correct rejections ($M = 1.656$, S.D. $= 3.445$), false alarms ($M = 1.057$, S.D. $= 3.940$), weak misses ($M = 0.959$, S.D. $= 3.809$). Hence, in comparison to the corresponding pattern for the 300–500 ms interval, the position of correct rejection amplitude within this range has risen above that of weak misses and false alarms, and is very similar to that of strong misses. Collapsed over conditions, average amplitudes were relatively most positive at parietal electrodes (left > right), and most negative at electrodes FZ, F3, and F7. With regard to the interaction, the range of amplitudes across conditions obtained at parietal (and central) sites, was relatively broader than that at frontal sites. This is similar to the outcome obtained for the 300–500 ms latency interval. Mean 500–800 ms electrode P3 amplitude (Bonferroni-corrected, four comparisons) of ERPs associated with correct decisions ($M = 5.825$, S.D. $= 4.042$) was significantly positive relative to that of ERPs associated with incorrect decisions ($M = 3.652$, S.D. $= 4.940$; $t(35) = 5.037$, $P < 0.01$). Mean electrode P3 amplitude over the 500–800 ms interval was also significantly greater for strong hit ($M = 7.628$, S.D. $= 4.426$) relative to weak hit ERPs ($M = 5.828$, S.D. $= 4.276$; $t(35) = 3.592$, $P < 0.01$). In contrast to the previous interval, however, no significant mean 500–800 ms electrode P3 amplitude differences were observed between false alarm ($M = 3.530$, S.D. $= 5.825$) and correct rejection ERPs ($M = 4.020$, S.D. $= 4.756$; $t(35) = -0.755$, $P = 0.456$) nor between the average of strong and weak miss ERPs ($M = 3.713$, S.D. $= 5.058$), and those accompanying correct rejections ($M = 4.020$, S.D. $= 4.756$; $t(35) = -0.437$, $P = 0.665$). However, it is indeed noteworthy that the directions of both differences were reversed, relative to the corresponding differences obtained for the 300–500 ms interval. That is, average amplitudes were relatively more positive in correct rejection versus both false alarm and miss, ERPs, and such differences were also obtained at CZ. In ERPs recorded at electrode CZ (Bonferroni-corrected, four comparisons), mean 500–800 ms amplitudes of ERPs accompanying correct decisions ($M = 3.723$, S.D. $= 4.509$) were positive relative to those for incorrect decisions ($M = 1.961$, S.D. $= 4.745$; $t(35) = 3.594$, $P < 0.01$). Mean electrode CZ 500–800 ms amplitude was significantly greater for strong hit ($M = 4.905$, S.D. $= 5.198$) than weak hit ERPs ($M = 2.832$, S.D. $= 4.286$; $t(35) = 3.420$, $P < 0.01$). Furthermore, mean CZ 500–800 ms amplitude was significantly greater in correct rejection ($M = 3.432$, S.D. $= 5.374$) than in both false alarm ($M = 1.584$, S.D. $= 6.070$; $t(35) = 2.811$, $P < 0.05$) and averaged-miss ERPs ($M = 2.15$, S.D. $= 4.723$; $t(35) = 2.649$, $P < 0.05$).

The following significant differences obtained in comparisons of the 300–500 ms ERP data, provide further evidence for a modulation by strength of the amplitude of the N400 old/new effect: strong hit > weak hit; false alarm > correct rejection; miss > correct rejection; old > new response ERPs. For the 500–800 ms data, significant differences obtained in the following comparisons further indicate that the amplitude of the LPC effect is influenced by decisional factor(s), such as accuracy and/or confidence: correct > incorrect decisions; strong > weak hits; false alarm < correct rejection; misses < correct rejection. Furthermore, the direction of the latter two differences, argue against a significant or direct impact of strength on LPC amplitude: rather, as noted above, those differences indicate a significant influence of decisional factor(s), on LPC amplitude. Notably, between electrodes CZ and P3, all amplitude differences were in the same directions. However, in the electrode P3 data, only the first comparison achieved statistical significance (following Bonferroni correction). Hence, while it may indeed be the case that the effects occurring during the LPC latency interval a t electrodes CZ versus P3, reflect the modulation of a common or overlapping component (or complex of components) with a centro-parietal distribution. It appears that this issue is not fully resolved on the basis of the comparisons reported above.

As described earlier, different patterns of ERP amplitudes were obtained between the 300–500 and 500–800 ms intervals. Most saliently, correct rejection ERPs rose very sharply over time relative to other conditions', particularly all error conditions', ERPs. Such a change (and the concomitant slope of the ERP curve) between the N400 and LPC peaks, is an important result but is barely reflected in the analyses reported above and in the literature (i.e. separate analyses for different latency intervals). Indeed, the relationship between N400 and LPC outcomes has been considered by previous ERP investigators (cf. [25,49]), but this has not been explored with additional statistical analyses. Hence, it is worthwhile to consider a supplementary form of analysis, the outcomes of which would be expected to better index the different ERP patterns obtained between the N400 and LPC components. Comparisons of LPC − N400 amplitude difference scores (i.e. average 500–800 ms minus average 300–500 ms, amplitudes) between conditions, appear to be a sufficient means by which to further investigate those differences. It is important to note that this analysis is not motivated by an assumption that the neural generator(s) of the N400 suddenly become inactive, nor by an assumption that they do not in any way modulate the LPC component. That is, they do not reflect an assumption that there is no overlap between the networks of generators of the N400 and LPC effects. Rather, comparisons of LPC − N400 difference scores are proposed as a supplement to the analyses reported above, which might provide converging evidence, and also perhaps better assess the influence of one or more LPC generator(s) on the effects of interest. Such comparisons are evidently of particular relevance to ERPs recorded at electrode P3, because it was at this electrode that amplitude differences between correct rejections' versus other conditions' ERPs (over 300–500 ms) were largest. Hence, amplitudes over 500–800 ms, especially those in correct rejection ERPs, would apparently be most affected by the carryover of N400 effects, at this electrode in particular. Therefore, it may be the case that such carryover yielded a diminution in correct rejection > false alarm and correct rejection > miss, LPC amplitude differences at electrode P3; resulting in the non-significance of those effects at this electrode. Such effects (between correct rejection versus error LPC amplitudes) are the most diagnostic of a sensitivity of the LPC effect to decisional factor(s) and did not appear to significantly differ between ERPs recorded at electrode CZ versus P3, but—perhaps due to the carryover of N400 effects—were of greater magnitude at CZ. Given the anticipated impact of such carryover, and the obtainment of the predicted effects in the experiment 1 data, these two comparisons (between correct rejection versus error LPC − N400 amplitude differences) were planned for the electrode P3 response-conditionalized data, collapsed across the two experiments (Bonferroni-corrected, two comparisons). LPC − N400 difference scores at electrode P3 were as follows: strong hit ($M = 4.164$, S.D. $= 2.462$), correct rejection ($M = 3.702$, S.D. $= 3.358$), weak hit ($M = 3.379$, S.D. $= 2.674$), weak miss ($M = 2.558$, S.D. $= 3.820$), false alarm ($M = 1.777$, S.D. $= 3.375$), strong miss ($M = 1.362$, S.D. $= 3.990$). It is noteworthy that this order is noticeably different from that obtained for the amplitudes over the 500–800 ms interval per se (e.g. correct rejection amplitudes are now substantially greater whereas strong misses are lower, relative to other conditions), and is evidently more correlated with recognition decision accuracy. Moreover, difference scores were significantly greater in correct rejection than in false alarm ERPs ($t(16) = 4.462$, $P < 0.01$). Similarly the former ($M = 3.702$, S.D. $= 3.358$) were significantly greater than the average of LPC − N400 difference scores in both strong and weak miss ERPs ($M = 1.960$, S.D. $= 3.526$; $t(16) = 3.387$, $P < 0.05$). In addition, the six conditions' mean difference scores noted above, taken together with the significance of these two comparisons, indicate that a comparison of correct versus incorrect LPC − N400 scores would be highly significant. Hence, on the basis of the outcomes of these comparisons, it is evident that the LPC effect at electrode P3 is sensitive to decisional factor(s), as is the LPC effect at electrode CZ. That is, it is evidently not the case that the LPCs at these two electrodes are distinct electrophysiological entities, but rather that they can both be considered constituents of a unitary LPC effect which exhibits a centro-parietal scalp topography in the current (ear-referenced) data, which is analogous to the LPC effect reported in previous studies, and which evidently indexes decisional factor(s), such as accuracy.

4. Discussion

An effect during the N400 interval was found to be sensitive to strength in a graded manner (new < weak < strong; old > new, decision ERPs). The other principal result, was that an LPC effect at centro-parietal electrode sites (left > right hemisphere) was found to be related to recognition decision accuracy (correct > incorrect, decision ERPs).

The temporal window during which the graded strength effect achieved statistical significance, predominantly incorporated the N400 component. The graded (new < weak < strong) 'N400 strength effect' was of greatest magnitude (and significant) at the left parietal electrode (P3), and similar results were also apparent at other parietal electrodes. In some cases, ERPs at neighboring (central and temporal) electrode sites exhibited apparently similar, but diminished, effects. These results clearly indicate that the left parietal N400 effect indexes memory trace strength. Converging evidence for the influence of strength on this ERP effect was provided when we examined separate ERPs conditionalized on both strength and decision (Fig. 3). The following significant differences were obtained in electrode P3 mean 300–500 ms amplitude: old > new, decision ERPs: strong > weak, hit ERPs; false alarm > correct rejection ERPs; miss > correct rejection ERPs. Together with the aforementioned strong > weak > new differences, these results provide solid converging evidence that the magnitude of the parietal N400 effect is directly sensitive to memory trace strength.

It has been proposed that the processes underlying the N400 old/new effect are not necessary for the explicit discrimination of old and new items [23,34]. This assertion was made on the basis of observations that the N400 old/new effect dissipates when the study-test repetition lag exceeds some threshold between 2 and 15 min (e.g. [25,37]). However, Duzel et al. [11] obtained an N400 old/new effect in an episodic recognition study in which the study-test repetition lag was at least 30 min. Here we have obtained significant

N400 old/new effects for both weak and strong words in a design where study-test repetition lags ranged from approximately 2–6 min. Rugg et al. [41] observed an N400 in parietal ERPs evoked by (shallowly-studied) unrecognized words which was equivalent in amplitude to that of hit ERPs, whereas both were relatively more positive than was the N400 evoked by correctly-rejected words. Rugg and co-workers thus declared this effect to be a neural correlate of implicit memory in the absence of conscious recognition. The additional result that in ERPs recorded during an implicit (semantic judgment) task both shallowly- and deeply-studied words' N400 amplitudes resembled those evoked by old words in the explicit task, was interpreted by the authors as converging evidence [41]. However, that data can also be accounted for by the proposal that the N400 old/new effect reflects memory trace strength, which would of course be greater for old words than for new words, regardless of whether the old words were recognized or unrecognized. Moreover, the strength associated with hits should generally be greater than that of misses: consistent with this, and in contrast to the results of Rugg and co-workers, both strong and weak (approximately analogous to deep and shallow study, respectively) hits' N400 amplitudes were positive in relation to the corresponding miss amplitudes at parietal electrodes, in the present data. (Note that this hit > miss difference was of the same magnitude as the significant old > new decision N400 amplitude difference.) Rugg et al. did not illustrate ERPs associated with recognized and unrecognized deeply-studied words so it is unclear whether or not an N400 difference was obtained there. However, given that strength is a continuous variable, one would expect the hit-miss N400 differences obtained in the present study if indeed this component reflects episodic strength. Finally, while the observed dissipation over time of the N400 old/new effect may indeed be due to changing contextual representations (cf. [10,15]), it is not possible at this point to determine whether this ERP component reflects an implicit memory process or a contextually-dependent memory process.

In addition to the N400 which typically has a centro-parietal maximum, a frontally-distributed 'FN400' effect [6] has also been reported recent papers. It is unclear whether or not the centro-parietal N400 and the FN400 are distinct components [7], although dissociations have been observed between N400-like old/new effects at parietal versus frontal electrodes [41]. The FN400 effect has been proposed to index familiarity [6,7,41]. However, there appear to be inconsistencies between different authors' concepts of familiarity. For example, Curran's view of familiarity as an 'assessment of global similarity' [7] is quite consistent with our notion of strength; But it appears that the same cannot definitively be said of the view of Rugg et al. [41] who—in relation to familiarity—cite a number of authors, whose own views of familiarity are sometimes disparate from one another, and which are not all consistent with our concept of strength. Moreover, Curran et al. [8] recently reported data that indicated to these authors that the FN400 effect may instead index perceptual fluency or perceptual familiarity. Significant old/new differences were obtained in mean electrode FZ ERP amplitudes over the 300–500 ms interval, only in one of the current experiments. The corresponding weak–strong comparisons did not even approach significance, in either experiment. Within the dual-process framework strong words should certainly engender higher levels of familiarity than weak words (e.g. [20]). So the lack of a significant FN400 strong–weak difference does not appear to be consistent with the hypothesized link between the FN400 and familiarity, regardless of which view of familiarity is adopted. For example, Mandler's [27] concept of familiarity involving the interactive activation of perceptual features would not be expected to asymptote after only a single study presentation. On another note Humphreys et al. [18] found that the second presentation of a target in a list produced no discernible effect on free association (a conceptual implicit memory task) in five different experiments, whereas it did have a large effect on cued recall and recognition. Hence, the replicated lack of a significant difference between weak and strong words' FN400 amplitudes in the present study is more consistent with the idea that the FN400 may reflect some specific form of implicit memory.

Another important point must be noted in relation to the above discussion: the FN400 patterns attributed by Curran [6,7] to familiarity were generally apparent only in average-referenced ERPs (collapsed over groups of neighboring electrodes), and not in ERPs from standard electrode sites (i.e. of the International 10–20 system [22]) which had been re-referenced to electrodes placed on the mastoids. Hence the present ear-referenced frontal ERP data may not be particularly analogous to those FN400 patterns.[1] In order to investigate whether or not the frontal topography of FN400 effects reported by Curran might be due—at least in part—to the employment of the average-reference technique, this transformation was also computed on the current data. Fig. 4 displays ERPs associated with new, weak, and strong, words (averaged across both experiments), and compares those ERPs according to ear versus average-reference techniques. It is evident that a key effect of the average-reference transformation on the current data, is to diminish the magnitude and/or topographical extent of the N400 strength effect and, moreover, that no FN400 effect is evident. However, it is also noteworthy that, relative to the current data, the data of Curran were acquired from many more (128 versus 32), and a markedly broader spatial range of, electrode sites. Consequently, the effects of the average-reference transformation on the current data versus that of Curran, cannot be expected to be identical. Thus, there can ultimately be no simple nor concise comparison between these two data sets. In summary, the current data

[1] We thank Tim Curran for highlighting both the similarity between his concept of familiarity and ours of strength, and the lack of familiarity-related FN400 patterns in his mastoid-referenced ERPs (personal communication, 21 May 2001).

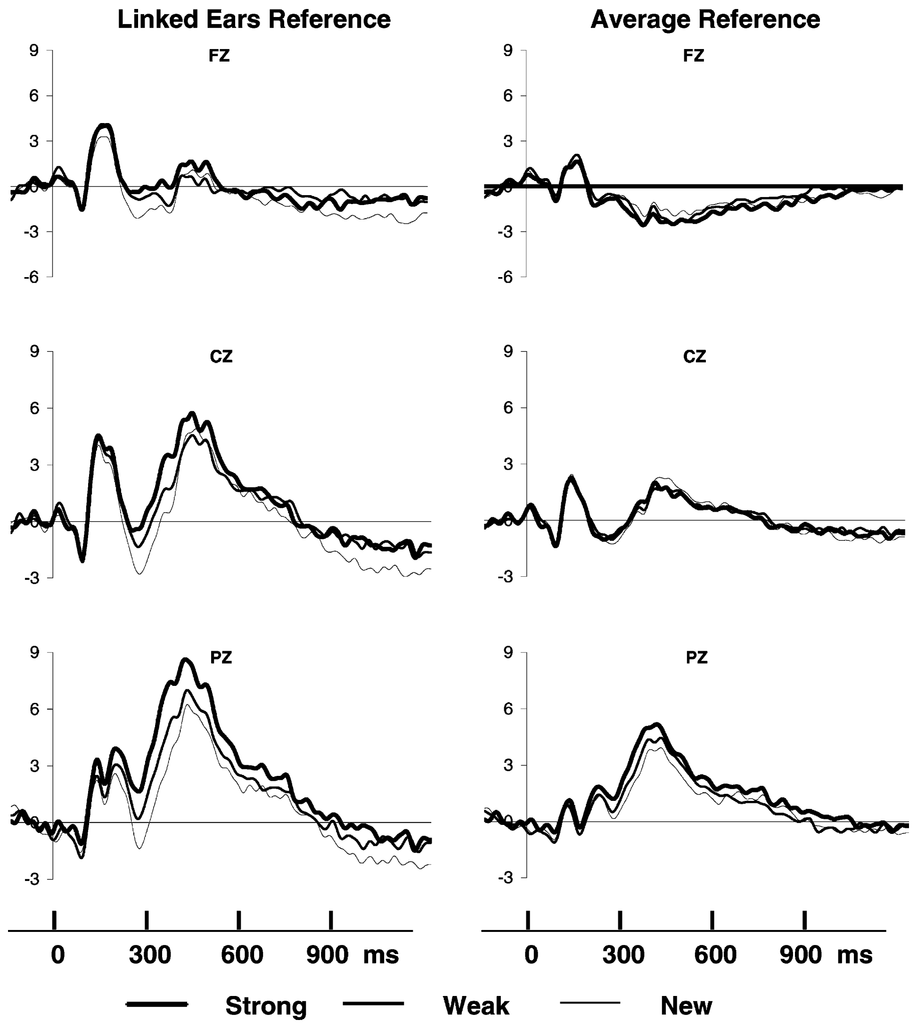

Fig. 4. Grand average ($N = 36$) ERPs evoked by new, weak (presented once at study) and strong (presented thrice at study) words, averaged across experiments 1 and 2. The ERPs on the left are referenced to linked ear electrodes, whereas those on the right represent the same data following an average-reference transformation.

should not be interpreted as being inconsistent with the hypothesized [6,7] link between the FN400 component and familiarity (as conceptualized by Curran). Clearly, the precise nature of the process(es) reflected in the FN400 effect, and of the relationship between this and the (posterior) N400 strength effect obtained in the current data, remain unresolved. Further exploration of these questions represents an important goal for future research, and will probably require (at least) a comparison of ear- versus average-referenced data acquired from a high-density electrode array.

As noted previously, Bentin et al. [5] found LPC amplitude to be sensitive to both lag and recency of repetition and suggested that it is modulated by strength. However, whereas the N400 effect is evidently sensitive to memory strength in a direct or graded manner in the current data, the same does not apply to the LPC. At centro-parietal electrodes, LPC amplitude in new ERPs was generally equivalent to, or in some cases even greater than, that of weak ERPs. These results (and indeed, others described below) evidently refute the existence of a direct relationship between the LPC effect and strength (although, note that this is not to say that strength has no impact whatsoever, on this effect). The three-level strength ERP comparison made here is clearly capable of such a refutation, whereas, a simple old/new comparison is not, for any ERP difference (in either direction) observed for the latter could potentially be erroneously interpreted as a strength-related difference.

The outcomes of the comparisons performed on the 500–800 ms data indicate that the centro-parietal LPC effect indexes decisional factor(s) including accuracy. LPC amplitude was found to be greater in ERPs accompanying correct (strong and weak hits, correct rejections) versus incorrect (strong and weak misses, false alarms) decisions, at both electrodes P3 and CZ; This difference was largest at P3, followed by PZ and CZ. In fact a considerable divergence was even apparent between the LPC amplitude for these two groups, at most centro-parietal electrodes. This was also reflected in the significant correct rejection > false alarm, and correct rejection > miss, 500–800 ms amplitude differences at electrode CZ (and the similar, albeit non-significant, differences at electrode P3). Supplementary analyses were also performed in order to examine the topographical extent of this decision-related LPC effect. That is, those analyses were primarily designed in order to address the question as to whether or not the LPC effect at electrode CZ, was critically different from that obtained at electrode P3. This was done via a form of baseline correction of each condition's ERP over the interval preceding 500 ms: N400 (300–500 ms) amplitudes were subtracted from LPC (500–800 ms) amplitudes, in order to minimize any carryover of effects from the former to the latter, latency interval. Such carryover would be expected to impact most upon correction rejections' LPC amplitudes, as N400 amplitudes in those ERPS were reliably the most negative of all conditions. In addition, this would be particularly problematic at parietal electrodes, as the differences between correct rejection and other conditions' N400 amplitudes, were relatively greatest at those sites. Hence, the aforementioned lack of statistical significance in the correct rejection > error 500–800 ms amplitude comparisons at electrode P3, may be due—at least in part—to a carryover of N400 strength effects. The supplementary comparisons of LPC − N400 difference scores, on electrode P3 data, indicate that this was indeed the case. Both the correct rejection > false alarm and correct rejection > miss differences achieved significance in the latter comparisons of electrode P3 data. In summary, the results summarized above converge to indicate that decision accuracy is indexed by the LPC effect, which has a centro-parietal (left > right) scalp distribution in the current data. To the extent that subjects were generally more confident of correct than incorrect decisions (e.g. [32]), the results of those comparisons are also compatible with the proposal that the LPC effect is modulated by decision confidence. In addition, strong hits and correct rejections are assumed to be at opposite ends of the strength continuum but—assuming there is a relationship between decision accuracy and confidence, and considering the current performance data—were likely to generally be associated with similar mean confidence levels (when confidence is conceptualized within the SDT context as distance from the response criterion). Hence, the relative similarity of LPC amplitudes in ERPs associated with these two conditions, further argues for an LPC sensitivity to decisional factor(s) rather, than to strength.

While the LPC effect had a centro-parietal scalp distribution in the current ear-referenced data, the effect was clearly more centered around parietal (and not central) scalp regions in the average-referenced data (Fig. 4). Given this result, together with the outcomes of the supplementary analyses summarized above, it is evident that this LPC effect is analogous to the LPC reported in previous studies. Consistent with the current results, the authors of a number of ERP studies of memory have proposed that the LPC effect is sensitive to decision accuracy and/or confidence (e.g. [24,26,28,30,32,46]). For example, Johnson et at. [24] repeated the same list of studied words across four study-test blocks and found that, independent of the old/new effect, LPC amplitude increased as the number of correctly classified old and new items increased. The authors suggested that parietal LPC activity reflects increased discriminability, which is concomitant with increasing strength of the memory trace. Again assuming that decision confidence increased as did accuracy with repetition across the four blocks, these data are also consistent with the idea that LPC amplitude is an index of decision accuracy and confidence. In addition, Rubin et al. [32] obtained confidence ratings associated with old/new recognition judgments and found significantly larger LPC amplitudes for high than low confidence hit ERPs, and a corresponding trend for correct rejections. Those authors also obtained significant differences incorporating the centro-parietal LPC between hit and false alarm ERPs, which were obviously much larger and much more sustained than the hit-correct rejection old/new effect. These observations of high > low confidence, and correct rejection > false alarm, LPC amplitude differences also indicate a sensitivity of LPC to decision accuracy and confidence [32], as do the observed LPC amplitudes in both our strength and strength-by-response ERP comparisons. Thus, we have obtained converging evidence which supports a previously-proposed relationship between LPC amplitude, and decisional factor(s) including accuracy and perhaps confidence. More generally, these conclusions are based largely upon comparisons involving ERPs associated with error responses, and this emphasizes the value of computing and comparing such ERPs.

In addition to decisional factor(s), the LPC effect may also be sensitive to other factor(s) or process(es), and this idea warrants further consideration. Similarly, it is worthwhile to consider the potential impact of latency jitter, or variability in the latency of a given ERP component, on LPC amplitude. For example, Spencer et al. [44] have proposed that an LPC amplitude difference obtained between ERPs associated with two different recognition memory conditions, was predominantly a consequence of different amounts of latency jitter between those conditions. If this proposal is accurate, and the amplitude of the LPC effect does in fact index factor(s) associated with the recognition decision, then it may be the case that the current LPC effects are due—at least in part—to different decision latencies between conditions (if such differences exist).[2] Resolution of this issue would require further examination of both single-trial ERPs and the decision latencies associated with each of the conditions of interest. However, in the current experiments, the execution of the recognition response was delayed until after the end of the ERP epoch, which clearly makes the decision latency data less interpretable. This issue and the question as to whether or not the making of the response contaminates ERP memory effects should be explored by future studies that do not require a delayed response.

[2] We thank an anonymous reviewer for highlighting the possible impact of decision latency on the LPC effect.

As noted in Section 1, previous investigators (e.g. [2,6,7,50]) have linked the LPC old/new effect to the recollection component of dual-process recognition theories. Were one to take the dual-process perspective one must assume that correct rejection decisions are consistently made in the absence of the conscious experience of recollection. Hence the current data (i.e. the relatively large LPC amplitudes obtained in correct rejection ERPs) are problematic for the idea that the LPC effect indexes recollection. Some theorists (e.g. [21,32]) assume that false alarms are also associated with an absence of recollection. If this were the case, and the LPC effect does reflect recollection, there should be no significant LPC amplitude difference between these conditions. In fact, there does exist evidence that false alarms can sometimes be associated with some form of 'false recollection' (e.g. [31]), and this also challenges the view that the LPC effect indexes recollection. That is, according to that view, such false recollection should yield greater LPC amplitudes in false alarm versus correct rejection, ERPs, but a significant difference in the opposite direction was obtained in the present data, and this result is completely consistent with the aforementioned data of Rubin et al. [32].

Paller and Kutas [29] were the first authors to propose that the LPC effect indexes recollection. However, this proposal relies on the assumption that depth of study processing selectively influences recollection, whereas more recent evidence indicates that this is not the case [47,56]. The position that the LPC old/new effect reflects a recollection or recall-like process was also adopted by Rugg and co-workers [2,39,40,51–54]. However, it is important to note that these authors' view is that recollection is graded or continuous (and that the LPC effect is sensitive to the amount of contextual information that is retrieved, e.g. [35,50,52]), whereas some key dual-process theorists (e.g. [55]) assume that recollection is a threshold or all-or-none process. In addition, some ERP data from source memory tasks reported by the former group, have substantially challenged some central assumptions of the dual process approach. For example, the data reported by Wilding and Rugg [52] do not support the dual process view that familiarity and recollection are separate cognitive processes associated with distinct patterns of neural activity (see also [35], for review), but rather that data are consistent with the idea that episodic recognition can generally involve only a single process. The question as to whether or not a recall-like process is necessarily involved in—and reflected in ERPs recorded during—single-item episodic recognition, can also be explored via a comparison of ERP effects recorded during recognition versus recall tasks. Allan and Rugg [1] have reported such a comparison, within-subjects. The recognition ERPs exhibited an old/new effect, primarily incorporating the LPC component, which was largest at the left parietal electrode. The cued recall ERPs, however, contained no such distinct LPC peak nor effect, but rather an old/new effect which commenced at approximately 400 ms and was sustained until the end of the ERP epoch (1200 ms). In addition, this cued recall ERP effect differed from the LPC effect not only morphologically but also topographically. The absence of the cued recall effect in ERPs recorded during standard episodic recognition tests, evidently challenges those theories which assume that a recall-like component contributes to performance on—and LPCs recorded during—such tests. In addition, perhaps the absence of a distinct LPC effect in ERPs recorded during cued recall tasks, is indicative of a sensitivity of this effect to recognition decision factor(s). Hence a number of lines of converging evidence challenge the idea that the LPC old/new effect indexes a recall-like or recollection process.

Finally, in relation to the above discussion, it is salient to note the caveat that we do not necessarily assume that a recall-like process never transpires during recognition. For example, such a process may occur during the performance of more complex or demanding recognition (or recognition-like) tasks, such as list discrimination (e.g. [19–21]), source memory (e.g. [50–54]), or study plurality (e.g. [7]), judgments, depending on factors, such as task requirements, subjective strategies, and consequently, the nature of the cues applied to memory. However, the current research is more concerned with standard or single-item recognition tasks, and we posit that only a single memory access process is fundamental to performance of such tasks (cf. [8,14]).

In summary, the amplitude of the N400 effect is modulated not solely by words' old/new status, but more generally, by memory trace strength in a graded manner. However, on the basis of the current data, it remains unresolved as to whether or not this N400 strength effect—which was of greatest magnitude at the parietal electrodes (left $>$ right)—indexes a context-dependent memory access process. In contrast, the LPC effect is sensitive to recognition decision accuracy and, by inference, confidence. This effect had a left $>$ right centro-parietal scalp distribution (in ear-referenced ERPs). Although it is posited that the N400 and LPC effects index strength and decisional factor(s), respectively, is it not proposed that these ERP effects are sensitive to all neurocognitive activity correlated with these factors. Nor is it assumed that these are the only processes or factors reflected in such ERP effects. However, on the basis of the current ERP data, it appears that these factors (at least) do significantly modulate those corresponding ERP effects. Thus, ERP old/new effects can be interpreted within the theoretical framework of single-process, as opposed to dual-process, models of episodic recognition. These inferences are made on the basis of results from this and other studies which have utilized more independent variables (and/or levels of those variables) and degrees of freedom, than have many previous ERP studies of recognition memory, and which consequently have greater analytical power with which to examine the putative influence of a number of neurocognitive processes on the ERP effects of interest.

Acknowledgements

Thanks to Darren Weber, Rohan Clarke, and David Smyth for technical support, and to Tim Curran, Michael Rugg, and two anonymous reviewers for helpful comments on earlier drafts.

References

[1] Allan K, Rugg MD. An event-related potential study of explicit memory on tests of cued recall and recognition. Neuropsychologia 1997;35:387–97.

[2] Allan K, Wilding EL, Rugg MD. Electrophysiological evidence for dissociable processes contributing to recollection. Acta Psychologica 1998;98:231–52.

[3] Bentin S, Moscovitch M. Psychophysiological indices of implicit memory performance. Bulletin of the Psychonomic Society 1990;28:346–52.

[4] Bentin S, Peled B-S. The contribution of task-related factors to ERP repetition effects at short and long lags. Memory and Cognition 1990;18:359–66.

[5] Bentin S, Moscovitch M, Heth I. Memory with and without awareness: performance and electrophysiological evidence of savings. Journal of Experimental Psychology: Learning, Memory and Cognition 1992;18:1270–83.

[6] Curran T. The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. Neuropsychologia 1999;37:771–85.

[7] Curran T. Brain potentials of recollection and familiarity. Memory and Cognition 2000;28:923–38.

[8] Curran T, Schacter DL, Johnson MK, Spinks R. Brain potentials reflect behavioural differences in true and false recognition. Journal of Cognitive Neuroscience 2001;13:201–16.

[9] Dennis S. The Sydney Morning Herald Word Database. Noetica: Open Forum 1995;1 [http://psy.uq.edu.au/CogPsych/Noetica].

[10] Dennis S, Humphreys MS. A context noise model of episodic word recognition. Psychological Review 2001;108:452–78.

[11] Duzel E, Yonelinas AP, Mangun G, Heinze H-J, Tulving E. Event-related brain potential correlates of two states of conscious awareness in memory. Proceedings of the National Academy of Sciences of the United States of America 1997;94:5973–8.

[12] Greenhouse SW, Geisser S. On methods in the analysis of profile data. Psychometrika 1959;24:95–111.

[13] Guillem F, N'Kaoua B, Rougier A, Claverie B. Intracranial topography of event-related potentials (N400/P600) elicited during a continuous recognition memory task. Psychophysiology 1995;32:382–92.

[14] Hockley WE. Item versus associative information: Further comparisons of forgeling rates. Journal of Experimental Psychology: Learning, Memory and Cognition 1992;18:1321–30.

[15] Howard MW, Kahana MJ. Contextual variability and serial position effects in free recall. Journal of Experimental Psychology: Learning, Memory and Cognition 1999;25:923–41.

[16] Humphreys MS, Pike R, Bain JD, Tehan G. Global matching: a comparison of the SAM, Minerva II, Matrix and TODAM models. Journal of Mathematical Psychology 1989;33:36–67.

[17] Humphreys MS, Dennis S, Chalmers K, Finnigan S. Dual processes in recognition: does a focus on measurement operations provide a sufficient foundation? Psychonomic Bulletin and Review 2000;7:593–603.

[18] Humphreys MS, Tehan G, O'Shea A, Bolland SW. Target similarity effects: support for the parallel distributed processing assumptions. Memory and Cognition 2000;28:798–811.

[19] Jacoby LL. A process dissociation framework: separating automatic from intentional uses of memory. Journal of Memory and Language 1991;30:513–41.

[20] Jacoby LL, Jones TC, Dolan PO. Two effects of repetition: support for a dual-process model of know judgments and exclusion errors. Psychonomic Bulletin and Review 1998;5:705–9.

[21] Jacoby LL, Toth JP, Yonelinas AP. Separating conscious and unconscious influences of memory: measuring recollection. Journal of Experimental Psychology: General 1993;122:139–54.

[22] Jasper HA. The ten–twenty system of the international federation. Electroencephalography and Clinical Neurophysiology 1958;10:371–5.

[23] Johnson Jr R. Event-related potential insights into the neurobiology of memory systems. In: Boller F, Grafman J, editors. Handbook of neuropsychology, vol. 10. Amsterdam: Elsevier; 1995. p. 135–63.

[24] Johnson Jr R, Pfefferbaum A, Kopell BS. P300 and long-term memory: latency predicts recognition performance. Psychophysiology 1985;22:497–507.

[25] Karayanidis F, Andrews S, Ward PB, McConaghy N. Effects of inter-item lag on word repetition: an event-related potential study. Psychophysiology 1991;28:307–18.

[26] Karis D, Fabiani M, Donchin E. "P300" and memory: individual differences in the von Restorff effect. Cognitive Psychology 1984;16:177–216.

[27] Mandler G. Recognizing: the judgment of previous occurrence. Psychological Review 1980;87:252–71.

[28] Neville HJ, Kutas M, Chesney G, Schmidt AL. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. Journal of Memory and Language 1986;25:75–92.

[29] Paller KA, Kutas M. Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. Journal of Cognitive Neuroscience 1992;4:375–91.

[30] Paller KA, Kutas M, Mayes AR. Neural correlates of encoding in an incidental learning paradigm. Electroencephalography and Clinical Neurophysiology 1987;67:360–71.

[31] Roediger III HL, McDermott KB. Creating false memories: remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory and Cognition 1995;21:803–14.

[32] Rubin SR, Van Petten C, Glisky EL, Newberg WM. Memory conjunction errors in younger and older adults: event-related potential and neuropsychological data. Cognitive Neuropsychology 1999;16:459–88.

[33] Rugg MD. Event-related brain potentials dissociate repetition effects of high and low frequency words. Memory and Cognition 1990;18:367–79.

[34] Rugg MD. ERP studies of memory. In: Rugg MD, Coles MGH, editors. Electrophysiology of mind: event-related brain potentials and cognition. Oxford: Oxford University Press; 1995. p. 113–27.

[35] Rugg MD, Allan K. Event-related potential studies of memory. In: Tulving E, Craik FIM, editors. Oxford handbook of memory. New York: Oxford University Press; 2000. p. 521–37.

[36] Rugg MD, Doyle MC. Event-related potentials and recognition of low- and high-frequency words. Journal of Cognitive Neuroscience 1992;4:69–79.

[37] Rugg MD, Nagy ME. Event-related potentials and recognition memory for words. Electroencephalography and Clinical Neurophysiology 1989;72:395–406.

[38] Rugg MD, Brovedani P, Doyle MC. Modulation of event-related potentials by word repetition in a task with inconsistent mapping between repetition and response. Electroencephalography and Clinical Neurophysiology 1992;84:521–31.

[39] Rugg MD, Schloerscheidt AM, Mark RE. An electrophysiological comparison of two indices of recollection. Journal of Memory and Language 1998;39:47–69.

[40] Rugg MD, Cox CJC, Doyle MC, Wells T. Event-related potentials and the recollection of low and high frequency words. Neuropsychologia 1995;33:471–84.
[41] Rugg MD, Mark RE, Walla P, Schloerscheidt AM, Birch CS, Allan K. Dissociation of the neural correlates of implicit and explicit memory. Nature 1998;392:595–7.
[42] Semlitsch HV, Anderer P, Schuster P, Presslich OA. Solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. Psychophysiology 1986;23:695–703.
[43] Smith ME, Halgren E. Dissociation of recognition memory components following temporal lobe lesions. Journal of Experimental Psychology: Learning, Memory and Cognition 1989;15:50–60.
[44] Spencer KM, Vila Abad E, Donchin E. On the search for the neurophysiological manifestation of recollective experience. Psychophysiology 2000;37:494–506.
[45] Swets JA, Tannner WP, Birdsall TG. Decision processes in perception. Psychological Review 1961;68:301–40.
[46] Swick D, Knight RT. Contributions of prefrontal cortex to recognition memory: electrophysiological and behavioural evidence. Neuropsychology 1999;13:155–70.
[47] Toth JP. Conceptual automaticity in recognition memory: levels of processing effects on familiarity. Canadian Journal of Experimental Psychology 1996;50:123–38.
[48] Van Petten C, Senkfor J. Memory for words and novel visual patterns: repetition, recognition, and encoding effects in the event-related brain potential. Psychophysiology 1996;33:491–506.
[49] Van Petten C, Kutas M, Kluender R, Mitchiner M, McIsaac H. Fractionating the word repetition effect with event-related potentials. Journal of Cognitive Neuroscience 1991;3:131–50.
[50] Wilding EL. In what way does the parietal ERP old/new effect index recollection? International Journal of Psychophysiology 2000;35:81–7.
[51] Wilding EL, Doyle MC, Rugg MD. Recognition memory with and without retrieval of context: an event-related potential study. Neuropsychologia 1995;33:743–67.
[52] Wilding EL, Rugg MD. An event-related potential study of recognition memory with and without retrieval of source. Brain 1996;119:889–905.
[53] Wilding EL, Rugg MD. Event-related potentials and the recognition memory exclusion task. Neuropsychologia 1997;35:119–28.
[54] Wilding EL, Rugg MD. An event-related potential study of memory for words spoken aloud or heard. Neuropsychologia 1997;35:1185–95.
[55] Yonelinas AP. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. Journal of Experimental Psychology: Learning, Memory and Cognition 1994;20:1341–54.
[56] Yonelinas AP, Kroll NEA, Dobbins L, Lazzara M. Recollection and familiarity deficits in amnesia: convergence of remember/know, process dissociation and ROC data. Neuropsychology 1998;12:323–39.