False Memory: P300 Amplitude, Topography, and Latency

Antoinette R. Miller, Christopher Baratta, Christine Wynveen, and J. Peter Rosenfeld

Northwestern University

RUNNING HEAD: P300 AND FALSE RECOGNITION

Address correspondence to:

J. Peter Rosenfeld
Psychology Department
Northwestern University
2029 Sheridan Rd.
Evanston, IL 60208-2710
(847) 491-3629
jp-rosenfeld@nwu.edu

Abstract

Two experiments are described in which the P300 component of the Event-Related Potential was recorded in a modification of the Roediger & McDermott (1995) paradigm. P300 amplitudes and topographies were evaluated in both true recognition of previously presented (Old) words and in false recognition of associatively related, never presented (Lure) words. In the first experiment, P300 topographies and amplitudes differed between recognized and rejected items of both types (Old vs. New words, falsely recognized Lures vs. correctly rejected Lures), while P300 did not appear to differ in either amplitude or topography between true and false recognition. However, false recognition of Lures produced substantially shorter P3 latencies than did the true recognition of Old words. This latency difference, as well as most of the topography results, were replicated in the second experiment, and discussed within the Fuzzy Trace Theory framework.

False Memory: P300 Amplitude, Topography, and Latency

False memory, i.e., memory for events that never occurred or distorted memory of actual events, has been actively researched in recent years. The experimental procedure frequently used to model this memory illusion (Deese, 1959; Roediger & McDermott, 1995) employs study lists composed of primary semantic associates (e.g., mad, fear, hate, rage, temper, fury, ire, wrath, fight, hatred, mean, calm, emotion, enrage) of critical non-presented words (Lures, e.g., anger). False recognition (in subsequent test sessions) of Lures is a robust effect and is experienced as subjectively similar to recognition of Old (studied) words, inasmuch as:

1) In the remember/know procedure, participants identify any memory based on conscious recollection (i.e., re-experiencing of some aspect of the item's previous presentation) as a "remember" experience. "Know" experiences are those in which participants are confident that the item had been presented, but are unable to remember details about its presentation (such as list position). This allows for the categorization of recognition as either recollection-based (remember) or familiarity-based (know). Roediger & McDermott (1995) found comparable numbers of remember judgments for both previously studied (Old) and falsely recognized Lures. This was a striking result, since few false alarms in standard word list paradigms are categorized as remember judgments (Rajaram, 1993). Participants often claim to have based their false memory reports on conscious recollection although these items were never presented (Lampinen, Neuschatz, & Payne, 1998).

2) High confidence levels are often reported for false recognition (Roediger & McDermott, 1995). Furthermore, participants are also willing to indicate which of two or three possible speakers read their falsely recalled Lures during the study period (Payne, Elie, Blackwell, & Neuschatz, 1996).

3) Perceptual priming (seen in implicit memory tests, including word-stem completion) of Lures has been observed (McDermott, 1997), as well as conceptual priming (seen in implicit memory tests including category instance production, Norman & Schacter, 1996).

Similar neural substrates may underlie both true and false recognition. Two recent functional neuroimaging studies (Schacter, Reiman, Curran, Yun, Bandy, McDermott, & Roediger, 1996; Schacter, Buckner, Koustaal, Dale, & Rosen, 1997) showed increased cerebral blood flow in similar brain areas in both true recognition of Old words and false recognition of Lures. Two recent Event-Related Potential (ERP) studies (Johnson, Nolde, Mather, Kounios, Schacter, & Curran, 1997; Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997) demonstrated electrophysiological correlates of false recognition. However, neither of these studies investigated P300, nor did they employ a paradigm that is classically used to elicit P300, e.g., an oddball paradigm (Donchin & Coles, 1988). However, in this previous ERP work, Johnson, et al. (1997) found no topographical differences in their electrophysiological responses for true and false recognition except in a condition in which test items were blocked by type (Old or Lure). Duzel, et al. (1997) found ERP differences between "remember" and "know" judgments, but not between true and false recognition.

We here investigate the P300 component of the ERP in an adaptation of the Deese/Roediger & McDermott paradigm. Several studies have demonstrated that studied or "old" words will elicit a large amplitude P300, differentiating them from unstudied or "new" words, which evoke a smaller or no P300 (Karis, Fabiani, & Donchin, 1984; Johnson, Pfefferbaum, & Kopell, 1985). This larger P300 has been interpreted as reflecting processes contributing to categorization of items based on recognition. Also, P300 has been successfully used in various malingered amnesia detection paradigms as an involuntary index of recognition (Rosenfeld, Sweet, Chuang, Ellwanger, & Song 1996; Rosenfeld, Reinhart, Bhatt, Ellwanger, Sekera, & Sweet, 1998; Ellwanger, Rosenfeld, Sweet, & Bhatt 1996). In these studies, the recognition of either autobiographical or recently learned information consistently elicited the P300 component regardless of whether the overt behavioral responses were truthful or deceptive. Further, these studies also demonstrated P300 amplitude and topography differences between honest participants and those simulating amnesia on these tasks (Ellwanger, et al., 1996; Rosenfeld, et al., 1998; Rosenfeld, Ellwanger, Nolan, Wu, Bermann, & Sweet, in press; Miller, 1999). These data suggested that P300 may be sensitive to aspects of pseudomemory phenomena, which can include both simulated amnesia and false memory (Cercy, Schretlen, & Brandt, 1997).

In the present paradigm, words are classified as either "old" or "new", based on a subjective recognition experience ("old") or lack thereof ("new') in response to each word. Both Old and falsely recognized Lure (Lure-Old) words are thus categorized as "old", which then may result in the elicitation of a P300. P300 amplitude is inversely proportional to stimulus probability (lower probability results in larger amplitude) and directly related to stimulus meaning (Johnson, 1988). In an oddball paradigm such as the one used here, both Lure and Old words will have a low probability of occurrence (15%), and recognition (true or false) may increase stimulus meaning associated with recognized vs. rejected words. For these reasons, we expected larger P300 amplitudes in response to recognized (Old) relative to unrecognized (New) words, and also, to falsely recognized Lures (Lure-Old) relative to correctly rejected Lures (Lure-New), i.e., we considered a larger P300 amplitude as an indicator of a recognition experience, or post-recognition categorization processes.

We also examined P300 topography (scalp distribution). According to the Triarchic Model (Johnson, 1993), P300 amplitude is a function of task and stimulus characteristics, which may be processed by different neural generators whose various contributions produce the scalp-recorded voltage patterns. Differing tasks or conditions may produce voltages which are not uniformly distributed across recording locations because of the differential recruitment of their respective "neurogenerator" sets (Johnson, 1993). Both veridical and illusory recognition may change the meaningfulness of recognized words (Old, Lure-Old) relative to comparable unrecognized items (New, Lure-New), requiring recruitment of different neurogenerators for successful stimulus processing. This mechanism may be reflected in differing scalp voltage patterns for recognized vs. unrecognized items.

Although there are many similarities between illusory and true recognition, recent work (Norman & Schacter, 1997; Mather, Henkel, & Johnson, 1997) has indicated qualitative differences between false and true recognition in the Roediger/McDermott paradigm. For example, greater sensory and contextual detail (Norman & Schacter, 1997), and more remembered feelings and reactions (Mather, Henkel, & Johnson, 1997) accompanied true recognition. P300 amplitude and topography have been shown to be useful in the investigation of one type of pseudomemory phenomenon (simulated amnesia), and we thus hypothesized that differences between true and false recognition may also be reflected in P300 topography differences.

In contrast to either of the other ERP studies of false memory (Johnson, et al., 1997; Duzel, et al., 1997), we also had expectations regarding P300 latency based on the following considerations:

1) According to Fuzzy Trace Theory (Reyna & Brainerd, 1995), two kinds of memory representations are constructed in parallel during encoding of an event: Gist comes from the extraction of senses, patterns, and meaning from events as they are encoded. Verbatim is a much

more accurate representation incorporating more precise representations of the event itself. Verbatim and gist representations are functionally dissociated: gist does not depend upon and is not integrated with the respective verbatim traces (Reyna & Brainerd, 1995).

2) True recognition of Old words may be supported by both verbatim and gist traces, since the Roediger/McDermott paradigm employs lists that are constructed around some theme encompassing both Old and Lure words. True recognition may be accompanied by retrieval of information from the two functionally dissociated gist and verbatim sources. However, false recognition of Lures is supported by the gist trace only. There is no study-related verbatim representation of the Lure words (which are never presented during the study episode), other than a possible implicit associative response related to the gist trace (Roediger & McDermott, 1995). Thus information may be retrieved primarily from gist during false recognition.

3) Typically, P300 latency is related to stimulus evaluation processes such as encoding and classification (Donchin, 1979; Donchin, Karis, Bashore, Coles, & Gratton, 1986; Kutas & Dale, 1997). It is also believed that a P300 is not elicited until an item has been evaluated (Donchin & Coles, 1988; Rugg & Coles, 1995), especially when the task relevance of the stimulus depends upon its proper categorization. Retrieval and utilization of information from both verbatim and gist may be necessary for successful categorization of Old words, and require more time than accessing gist only (as in the false recognition of Lures). Also, since verbatim may include memory traces corresponding to the individual study items (Payne, et al., 1996), more time may also be required to search it for the needed information (for true recognition). Gist contains the more general semantic content of the study lists, with no precise representations of the individual items (Payne, et al., 1996), and less time may be needed to find the appropriate information to support the false recognition of Lures. If false recognition requires less stimulus

categorization time than does true recognition, this may be reflected in shorter P300 latencies for Lure-Old-P300 than for Old-P300.

Methods

Participants

16 Northwestern University undergraduates (8 male) volunteered to participate in this study in partial fulfillment of requirements for the introductory psychology course. We excluded four participants from ERP analyses due to high eyeblink (EOG) artifact rates, but all 16 were included in behavioral analyses. All were unaware of Roediger & McDermott (1995) or similar studies. All were right-handed, and had normal or corrected to normal vision.

Materials

For Study, we used 25 study lists, each composed of fourteen strong associates (Keppel & Postman, 1970) of a critical non-presented word (Lure). The strongest semantic associate of each original Lure was not incorporated into the study list, and later used as an additional Lure (as in Israel & Schacter, 1997). Words were presented visually, with a two second stimulus duration and two second interstimulus interval. A 333 item recognition test followed the completion of the Study task, composed of 50 Old words (two from each list), 50 Lures (the critical word used to construct each study list plus the strongest nonstudied associate), and 233 New words (unrelated to either Lures or Old words). This allowed presentation of Old and Lure items with sufficiently low probability (0.15 each) to elicit the classic oddball P300 (Fabiani, Gratton, Karis, & Donchin 1987).

General testing procedure

Study: After signing an informed consent form, participants were seated in a recording room with a computer monitor. Following electrode placement, participants were told that they would see a series of words on the computer monitor, and that they would later be tested for their memory of those words. Robinson & Roediger (1997) cautioned that when preceded by a recall task, the recognition test results should be interpreted with care due to possible contamination by prior recall of Lures. Therefore, participants engaged in a semantic encoding task (rating each word for pleasantness on a 5-point scale) rather than the more typical recall task. No ERPs were recorded during Study.

Test: Participants received five minutes of instruction, in which they were told to identify displayed words as either Old or New with a button press (right button for "Old", left button for "New"). After remaining on the screen for 2048 ms, a test word was cleared from the screen and replaced with the message "RESPOND". While this message was displayed for 1.5 s, participants were required to make old/new judgments. After this time period elapsed, the screen was cleared, remaining blank for 500 ms before the next word appeared.

ERP recording and analysis

Silver-silver chloride electrodes were attached with conductive EEG paste to the following scalp sites: Fz, Cz, Pz, F3, F4, P3, and P4. Linked mastoids served as the reference with the forehead grounded. Electrodes were also placed supra- and sub-orbitally for EOG recording; i.e., eye movement artifacts occurring during the recording epoch were detected and trials containing 80 microvolt or higher deflections were discarded. Signals were amplified 100,000 times by Grass p511-K preamplifiers with 3 dB filters set to pass signals between 0.3

and 30 Hz. Conditioned signals were led to a 12-bit analog/digital converter sampling one point every eight ms, and then to a microcomputer for data storage.

The ERP was recorded from 104 ms prior to the word onset to 1944 ms after its onset. Participants were instructed to delay their responses until the end of the recording epoch (signalled by the “RESPOND” message), in order to minimize movement-related artifacts. Individual ERP averages were calculated for Lure-Old (Lures called "old"), Lure-New (Lures called "new"), Old, and New words. No averages were calculated for Old words called “new”, or New words called “old”, as too few of these responses were available to yield stable averages. Averages were digitally filtered to pass low frequencies; 3 dB point: 4.23 Hz for all analyses.

P300 Analysis

Figures 1, 2 and 3 (grand averages) show the peak Pz P300 latency for Old, Lure-Old, and Lure-New items. The value of the maximally positive 104 ms segment (13 data points) of Pz waveform in the interval from 400 to 900 ms post-stimulus was calculated, and then subtracted from the EEG average of the 104 ms prestimulus period; this was the standard baseline-to peak (b-p) P300 index. The time measured from stimulus onset to the midpoint of this segment was taken as the P300 latency. The subsequent absolute maximum negative 104 ms segment in the interval from the previously found P300 latency to the end of the sweep was then subtracted from the absolute P300 amplitude value (b-p), yielding a peak-to-peak (p-p) P300 index. It was necessary to compare the same process (component) across sites within an item category (Ruchkin, Johnson, Grafman, Canoune, & Ritter, 1992) in subsequent topographical analyses, and to avoid possibly spurious results due to noise-related latency differences across sites within the wide 400-900 ms analysis window. Therefore, a method was used in which the latency at site Pz was calculated, and the surrounding 104 ms latency segment was used as the time window within which to calculate amplitude at all other sites, and the same method was used for calculation of the subsequent negative segment. P-p amplitude values are reported here. Although less frequently used, the p-p method has been a better discriminator of guilty and innocent participants than the b-p measure in previous studies of detection of concealed information (e.g., Ellwanger, et al., 1996; Rosenfeld, et al., 1998). Further, the amplitude of the negative component that follows P300 (used to calculate p-p, see above) has been shown to be less variable than the pre-stimulus EEG baseline average, making the p-p index a more reliable measure of P300 amplitude in the detection of deception (Soskins, Rosenfeld, & Niendam, submitted).

Analysis of P300 Amplitude, Topography, and Latency

For all analyses, within-participant Item by Site repeated measures ANOVAs were utilized, the details of which are described in the Results. To perform unambiguous topographical analyses, ERP data were scaled using the vector length method (McCarthy & Wood, 1985) prior to ANOVA, to ensure that the comparisons were confined to scalp topography differences alone, and not to possibly confounding effects of overall amplitude differences (Johnson, 1993). A significant Item by Site interaction was considered evidence of distinct Item-related topographies. All p values were Huynh-Feldt corrected where df > 1 in within-participant analyses.

Results

Behavioral

Participants falsely recognized 56% of Lures, a significantly greater rate ($F_{1,15} = 151.751$, $p < 0.001$) than the false alarm rate (calling New words "old") of 14%. This false recognition effect is comparable to rates in other studies using multiple Lures per study list (e.g., Israel & Schacter, 1997; Duzel, et al., 1997). It should be noted that we used visual presentation at Study, which Smith & Hunt (1999) showed may result in significantly lower false recognition rates than auditory presentation of the same lists. Artifact-free EEG data were collected for an average of 31 (SEM 2.04) of the 50 Lure trials (17 Lure-Old, SEM 1.33 and 14 Lure-New, SEM 1.35), 28 (SEM 2.1) of the 50 Old trials, and 140 (SEM 9.8) of the 233 New trials across all participants.

Unscaled P300 Amplitude

P300 is defined as a positive component peaking between 300 and 900 ms post-stimulus, with a parietally maximal distribution. As can be seen in Figure 4, amplitude is largest at Pz, smallest at Fz. We conducted a 2 (Item, Old and Lure-Old) by 3 (Site, Fz, Cz, Pz) repeated measures ANOVA on p-p data. A significant Site effect resulted ($F_{2,22} = 35.004$, $p < 0.001$), in which a significant linear component ($F_{1,11} = 40.352$, $p < 0.001$) confirmed the Pz > Cz > Fz amplitude distribution typical of P300.

While all ANOVAs were performed using amplitude data from all 7 sites, mean differences (MD) are reported below only at site Pz, where P300 amplitude is maximal (Fabiani, et al., 1987).

Old vs. New: Figure 1 contains the grand averaged ERPs for Old and New words at all recording sites. A distinctly larger P300 (p-p) is apparent in the waveforms for Old words, particularly at the parietal and Cz recording sites. Figure 5 contains the mean unscaled P300 p-p amplitudes for all four item types (Old, New, Lure-Old, and Lure-New), in which substantially larger amplitudes are apparent for Old than for New items. A 2 (Item) by 7 (Site) repeated measures ANOVA yielded a significant Item effect, confirming that Old words elicited larger P300 amplitudes than did New words ($F_{1,11} = 11.325$, $p < 0.007$, MD = 3.24 microvolts). A significant Site effect resulted ($F_{6,66} = 21.012$, $p < 0.001$), as well as a significant Item by Site interaction ($F_{6,66} = 9.153$, $p < 0.001$), which may reflect the different distributions of amplitudes across recording sites for Old and New words (see Figure 5); but without scaling, this effect could confound amplitude and topography, see McCarthy & Wood (1985).

Lure-Old vs. Lure-New: Figure 2 shows the grand averaged ERPs for falsely recognized Lures (Lure-Old) and correctly rejected Lures (Lure-New), in which a larger p-p P300 appears associated with falsely recognized Lures than with rejected Lures, at the majority of recording sites. Figure 5 also shows the mean P300 p-p amplitudes for Lure-Old and Lure-New items, where it appears that Lure-Old items elicited larger amplitudes than did Lure-New items. In a 2 (Item) by 7 (Site) repeated measures ANOVA, a significant Item effect ($F_{1,11} = 14.043$, $p < 0.004$, MD =2.11 microvolts) confirmed that falsely recognized Lures (Lure-Old) produced larger P300 amplitudes than did rejected Lures (Lure-New). A significant Site effect also resulted ($F_{6,66} = 18.725$, $p < 0.001$), but the Item by Site interaction was not significant ($F_{6,66} = 1.25$, $p > 0.2$).

Lure-Old vs. Old: Figure 3 shows the grand averaged ERPs for Lure-Old and Old words. No consistent substantial difference is apparent across all recording sites in p-p amplitude between Lure-Old and Old items, with the possible exception of site Cz. Figure 5 shows the group mean P300 p-p amplitudes for all four item types, in which Lure-Old and Old amplitudes appear nearly identical. A post-hoc 2 (Item) by 7 (Site) repeated measure ANOVA on Lure-Old vs. Old P300 p-p amplitudes did not reveal significant amplitude differences ($F_{1,11}$ = 0.098, $p > 0.7$, MD = 0.132 microvolts).

Lure-New vs. New: No consistent difference in amplitude is apparent between Lure-New and New items in Figure 5, with the possible exception of sites P3 and Pz. A post-hoc 2 (Item) by 7 (Site) repeated measure ANOVA on Lure-New and New amplitudes did not reveal a significant amplitude difference ($F_{1,11}$ = 0.911, $p > 0.3$, MD = 1.24 microvolts).

Scaled P300 Topography

Figure 6 contains mean scaled P300 p-p amplitudes for Old, Lure-Old, Lure-New, and New items at all seven sites.

Old vs. New: Different topographies are apparent in Figure 6 for Old and New items. A 2 (Item) by 7 (Site) repeated measures ANOVA on scaled amplitude data yielded a significant Item by Site interaction ($F_{6,66}$ = 5.597, $p < 0.01$), indicating distinct Old and New P300 topographies.

Lure-Old vs. Lure-New: A significant Item by Site interaction resulted in the 2 (Item) by 7 (Site) repeated measures ANOVA on scaled amplitudes ($F_{6,66} = 5.161$, $p < 0.004$), indicating different Lure-Old and Lure-New topographies.

Lure-Old vs. Old: Figure 6 shows distinctly that the P300 topographies for both kinds of recognized items (Old and Lure-Old) are quite similar, yet they are different from both kinds of unrecognized items (New and Lure-New). A post-hoc 2 (Item) by 7 (Site) repeated measures ANOVA on Lure-Old vs. Old scaled P300 amplitudes did not yield a significant Item by Site interaction ($F_{6,66} = 1.781$, $p > 0.2$).

Lure-New vs. New: Figure 6 also shows no obvious difference in topography between Lure-New and New items. A post-hoc 2 (Item) by 7 (Site) repeated measures ANOVA on Lure-New and New scaled P300 amplitudes did not yield a significant Item by Site interaction ($F_{6,66} = 2.128$, $p > 0.1$).

P300 Latency

All ANOVAs were done using Pz latency data only, the means of which for each item category are shown in Figure 7. As can be seen in Figure 3 (grand averages in which the Pz-P300 latency is indicated) and Figure 7, the Lure-Old P300 latency is substantially shorter than Old-P300 latency at all recording sites, including Pz. This latency difference was confirmed by a significant Item effect (Lure-Old vs. Old) in a one-way repeated measures ANOVA on Pz latency ($F_{1,11} = 7.242$, $p < 0.03$, MD = 69.3 ms). As also can be seen in Figures 2, 3, and 7, the Lure-Old P300 latency appears substantially shorter than that for the other item types. A post-hoc repeated measures ANOVA, Lure-Old vs. the mean of Old, New, Lure-New, yielded a significant Item effect ($F_{1,11} = 11.59$, $p < 0.01$, MD = 48.6 ms), suggesting that the false-recognition P300 may be of considerably shorter latency than that in response to any other item.

Replication Study

Because the latency results from the initial experiment have important theoretical and practical implications, a second experiment was conducted using similar methods to those previously described. 13 Northwestern University undergraduates (9 male) volunteered to participate in partial fulfillment of requirements for the introductory psychology course. Participants responded by means of a button press. A long rectangular box with two buttons was placed on the right arm of the participants' chair, and they were instructed to press the right button for "Old" and the left button for "New" words, using the index and middle finger of their right hand. The EEG recording parameters were the same as in the initial experiment, except for the use of a nose reference, rather than a linked mastoid reference (this was done to avoid confounds of amplitude asymmetry by electrode impedance asymmetry). These participants falsely recognized 46.5% of Lures, a rate which was significantly greater than the false alarm rate of 16.1% ($F_{1,12} = 42.398$, $p < 0.001$).

P300 Latency

Figure 8 shows the mean Pz-P300 latencies associated with Old, New, Lure-Old, and Lure-New items. As in the initial experiment (see Figure7), the latency of the P300 peak elicited by false recognition (Lure-Old) appears substantially shorter than that for true recognition (Old), and also shorter than that of the other items (Lure-New and New). As before, the difference between Lure-Old and Old latencies proved significant in a one-way ANOVA ($F_{1,12}$ = 11.498, p < 0.006, MD = 93.5 ms). As in the previous experiment, a one-way ANOVA on Lure-Old vs. the mean of the latencies for the other items (Other) also yielded a significant Item effect ($F_{1,12}$ = 13.711, p < 0.004, MD = 76 ms).

P300 Amplitude and Topography

Figure 9 shows the mean unscaled P300 p-p amplitudes for Old, New, Lure-Old and Lure-New items. As in the previous experiment, no significant difference is apparent between true (Old) and false (Lure-Old) recognition. This was confirmed by a Item (Old vs. Lure-Old) by Site ANOVA, in which the Item effect was not significant ($F_{1,12}$ = 0.160, p > 0.6, MD at Pz = 1.44 microvolts). However, Old items elicited larger P300 amplitudes than New items (Old vs. New, $F_{1,12}$ = 17.123, p < 0.002, MD at Pz = 3.66 microvolts), as can be seen in Figure 9. The amplitude difference between Lure-Old and Lure-New items (Figure 8) was not consistent across all sites, reflected by the lack of a significant Item effect in an ANOVA on these items ($F_{1,12}$ = 1.970, p > 0.1, MD at Pz = 2.6 microvolts).

Figure 10 shows the mean scaled p-p amplitudes for Old, New, Lure-Old, and Lure-New items. As in the initial experiment, the topographies for Old and Lure-Old items are quite similar. This was reflected in the lack of a significant Item by Site interaction ($F_{6,72}$ = 1.767, p > 0.18). And, while Old and New items produced distinct topographies (Item by Site interaction: $F_{6,72}$ = 7.389, p < 0.003), the topographies for Lure-Old and Lure-New items do not appear distinct (Item by Site interaction: $F_{6,72}$ = 0.371, p > 0.7).

Replication Summary

In this second experiment, several of the key effects observed in the initial experiment were replicated. As in the first experiment, true (Old) and false (Lure-Old) recognition did not differ in either P300 amplitude or topography, while a substantial latency difference was observed (Lure-Old << Old). Also as in the first experiment, Old words elicited larger P300 amplitudes and a different P300 topography than did New words. However, the P300 amplitude and topography differences between Lure-Old and Lure-New items did not differ significantly.

Discussion and Conclusions

Two experiments were conducted using a modification of the Deese/Roediger & McDermott paradigm. In the first, larger P300 amplitudes were associated with both true and false recognition than with rejection (non-recognition) of comparable items. Post-hoc analysis revealed no difference in either amplitude or topography between the P300 elicited by true recognition of old words and P300 associated with false recognition of Lure words. This apparent similarity of true and false recognition in P300 amplitude and topography was replicated in a subsequent experiment, in which the amplitudes and topographies associated with Old and Lure-Old items also did not differ. These results may be a reflection of the apparent subjective similarity of the false and true recognition experiences in the Deese/Roediger & McDermott paradigm (see Introduction).

However, while the P300 amplitude and topography associated with false recognition of Lures were remarkably similar to that of veridical recognition, we did find a substantial difference in Pz latency. The false recognition-P300 peaked much earlier the P300 in true recognition, a result that proved significant both in the initial experiment and in the replication.

Because P300 often does not appear until an item has been evaluated (Donchin & Coles, 1988; Rugg & Coles, 1995) and P300 peak latency is related to categorization time (Donchin, 1979; Donchin, Karis, Bashore, Coles, & Gratton, 1986), we might infer that the shorter P300 latency for false recognition of Lures indicates that these evaluation and categorization processes take less time for falsely-recognized Lures. The nature of the false recognition phenomenon may account for the possibly shorter evaluation time required for falsely recognized Lures:

As described above, Fuzzy Trace Theory (Reyna & Brainerd, 1995) assumes that two functionally dissociated memory representations (verbatim and gist) are constructed in parallel during encoding of an event. True recognition of Old words may be accompanied by retrieval of both verbatim and gist information. If these representations are not integrated, it may require more time to access the information from these two dissociated sources than to access gist only (as in false recognition of Lures). Also, since verbatim is more detailed, more time may be required to find the appropriate information within it than within the more general gist representation. Both of these possibilities may result in a shorter stimulus evaluation time for falsely recognized Lures, yielding a shorter P300 latency.

False recognition may be based on feelings of familiarity that are produced by the retrieval of gist (Schacter, Verfaellie, & Pradere, 1996) and this gist information can be misidentified as verbatim information. Rejected Lures may thus elicit a weaker sense of familiarity, and the separate verbatim trace may then be consulted to confirm the source of the item's familiarity. This added step might increase the classification time and concomitant P300 latency for Lure-New responses (compared to Lure-Old).

New words are unrelated to the study lists, and so may not elicit the gist experience. Only a possible verbatim trace may be searched for use in their categorization (as "new"). Searching the detailed verbatim trace for the New (nonexistent) word may require more time, and may also result in a longer stimulus evaluation time and a longer New-P300 latency.

P300 latency is not necessarily correlated with reaction time (Donchin & Coles, 1988), which is consistent with our finding a substantial P300 latency difference between true and false recognition while no difference in reaction time was found in previous research (Johnson, et al., 1997). If a dissociation is found directly between P300 latency and reaction times in false recognition, it could provide further evidence that the latency difference is a measure of stimulus evaluation processes, rather than response-related processes. We are currently researching this question.

In conclusion, we have provided evidence for yet another way in which false recognition of lure words is similar to true recognition of studied words. However, our finding of an earlier latency for the false recognition-P300 provides an interesting correlate of false recognition not yet seen in previous work. This latency effect may be a substrate/correlate of unconscious recognition of the true nature of the Lure, or explained in terms of Fuzzy Trace Theory (see above). Regardless, the participants' recognition experience of Lure-Old words is similar to that of Old words, and the P300 amplitudes and topographies are similar in true and false recognition. But, P300 latency distinguishes between Lure and Old words, when both are experienced as "old". Continued investigation of this result may allow further characterization of differences between true and false recognition.

References

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology, 58(1), 17-22.

Cercy, S.P., Schretlen, D.J., & Brandt, J. (1997). Simulated amnesia and the pseudo-memory phenomena. In R. Rogers (Ed.), Clinical Assessment of Malingering and Deception.

Donchin, E. (1979). Event-related brain potentials: a tool in the study of human information processing. In H. Begleiter (Ed.) Evoked potentials and behavior. New York: Plenum Press

Donchin, E., & Coles, M. G. (1988). Is the P3 component a manifestation of context updating? Behavioral and Brain Sciences, 11(3) 357-427.

Donchin, E., Karis, D., Bashore, T. R., Coles, M.G.H., & Gratton, G. (1986). Cognitive psychophysiology and human information processing. In M.G.H. Coles, E. Donchin, & S. W. Porges (Eds.) Psychophysiology: systems, processes, and applications. New York: Plenum Press

Duzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.-J., & Tulving, E. (1997). Event-related potential correlates of two states of conscious awareness in memory. Proceedings of the National Academy of Sciences, USA, 94, 5973-5978.

Ellwanger, J., Rosenfeld, J.P., Sweet, J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P3 event-related potential. International Journal of Psychophysiology, 23, 9-23.

Ellwanger, J., Rosenfeld, J.P., & Sweet, J.J. (1997). P300 event-related potential as an index of recognition response to autobiographical and recently learned information in closed-head injury patients. The Clinical Neuropsychologist, 11(1), 1-5.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). The definition, identification, and reliability of measurement of the P3 component of the event related brain potential. In P.K. Ackles, J.R. Jennings, & M.G.H. Coles (Eds.), Advances in psychophysiology, Vol. 2. Greenwich: JAI Press.

Israel, L., & Schacter, D.L. (1997). Pictorial encoding reduces false recognition of semantic associates. Psychonomic Bulletin and Review, 4(4), 577-581.

Johnson, M.K., Nolde, S.F., Mather, M., Kounios, J., Schacter, D.L., & Curran, T. (1997). The similarity of brain activity associated with true and false recognition memory depends on test format. Psychological Science, 8(3), 250-257.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. Psychophysiology, 30, 90-97.

Johnson R., Pfefferbaum, A., & Kopell, B.S. (1985). P300 and long-term memory: latency predicts recognition performance. Psychophysiology, 22, 497-507.

Karis, D., Fabiani, M., & Donchin, E. (1984). 'P300' and memory: individual differences in the von Restorff effect. Cognitive Psychology, 16, 177-216.

Keppel, G. & Postman, L. (Eds.) (1970) Norms of word association New York: Academic Publishers, Inc.

Kutas, M. & Dale, A. (1997) Electrical and magnetic readings of mental functions. In M.D. Rugg (Ed.) Cognitive Neuroscience. Cambridge, Massachusetts: MIT Press.

Lampinen, J.M., Neuschatz, J.S. & Payne, D.G. (1998). Memory illusions and consciousness: Examining the phenomenology of true and false memories. Current Psychology: Development, Learning, Personality, Social, 16(3-4), 181-194.

Mather, M., Henkel, L.A., & Johnson, M.K. (1997). Evaluating characteristics of false memories: remember/know judgments and memory characteristics questionnaire compared. Memory and Cognition, 25(6), 826-837.

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

McDermott, K.B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. Psychonomic Bulletin and Review, 4(4) 582-586.

Miller, A.R. (1999). P300 amplitude and topography in pseudomemory phenomena. Unpublished doctoral dissertation, Northwestern University.

Norman, K.A. & Schacter, D.L. (1996). Implicit memory, explicit memory, and false recollection: a cognitive neuroscience perspective. In Reder, L.M. (Ed.) Implicit memory and metacognition. Greenwich, New Jersey: Lawrence Erlbaum Associates.

Norman, K.A. & Schacter, D.L. (1997). False recognition in younger and older adults: exploring the characteristics of illusory memories. Memory and Cognition, 25(6), 838-848.

Payne, D.G, Elie, C.J., Blackwell, J. M, &.Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. Journal of Memory and Language.,35(2) 261-285

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. Memory and Cognition, 21, 89-102.

Reyna, V.F. & Brainerd, C.J. (1995). Fuzzy trace theory: an interim synthesis. Learning and Individual Differences, 7(5), 1-75.

Robinson, K.J., & Roediger, H.L (1997). Associative processes in false recall and false recognition. Psychological Science, 8(3), 231-237.

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(4), 803-814.

Rosenfeld, J.P., Ellwanger, J.W., Nolan, K., Wu, S., Bermann, R.G., & Sweet, J. (in press). P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model. International Journal of Psychophysiology.

Rosenfeld, J.P, Reinhart, A.M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., & Sweet, J. (1998). P300 correlates of simulated malingered amnesia in a matching-to-sample task: topographic analyses of deceptive vs. truthtelling responses. International Journal of Psychophysiology, 28(3), 233-249.

Rosenfeld, J.P., Sweet, J.J., Chuang, J., Ellwanger, J. & Song, L. (1996). Detection of simulated malingering using forced choice recognition enhanced with event-related potential recording. The Clinical Neuropsychologist, 10(2), 163-173.

Ruchkin, D.S., Johnson, R., Grafman, J., Canoune, H., & Ritter, W.(1992) Distinctions and similarities among working memory processes: an event-related potential study. Cognitive Brain Research,1, 53-66

Rugg, M.D. & Coles, M.G.H. (1995). The ERP and cognitive psychology: conceptual issues. In Rugg, M.D. (Ed.) Electrophysiology of the Mind. Oxford Press, New York.

Schacter, D.L., Buckner, R.L., Koustaal, W., Dale, A.M., & Rosen, B.R. (1997). Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. Neuroimage, 6, 259-269.

Schacter. D.L., Reiman, E., Curran, T., Yun, L.S., Bandy, D., McDermott, K.B., Roediger, H.L. (1996). Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. Neuron, 17, 267-274.

Schacter, D.L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: false recall and recognition in amnesic patients. Journal of Memory and Language, 35, 319-334.

Smith, R.E. & Hunt, R.R. (in press). Presentation modality affects false memory. Psychonomic Bulletin and Review.

Figure Captions

Figure 1: Grand averaged ERPs, Old (thin line) vs. New (thick line) words. (a) Superimposed over the traces is a vertical line indicating the group mean Pz-Old peak latency. Positive is down.

Figure 2: Grand averaged ERPs, Lure-Old (thin line) vs. Lure-New (thick line) words. Superimposed over the traces are (b) a vertical line indicating the group mean Pz-Lure-Old peak latency, and (c) a vertical line indicating the group mean Pz-Lure-New peak latency. Positive is down.

Figure 3: Grand averaged ERPs, Old (thin line) vs. Lure-Old (thick line) words. Superimposed over the traces are (a) a vertical line indicating the group mean Pz-Old peak latency, and (b) a vertical line indicating the group mean Pz-Lure-Old peak latency. Positive is down.

Figure 4: Mean unscaled p-p P300 amplitudes, Lure-Old and Old words.

Figure 5: Mean unscaled p-p P300 amplitudes, Old, Lure-Old, Lure-New, and New words.

Figure 6: Mean scaled p-p P300 amplitudes, Old, Lure-Old, Lure-New, and New words.

Figure 7: Mean P300 latencies at site Pz.

Figure 8: Mean P300 latencies at site Pz, replication experiment.

Figure 9: Mean unscaled p-p P300 amplitudes, replication experiment.

Figure 10: Mean scaled p-p P300 amplitudes, replication experiment.

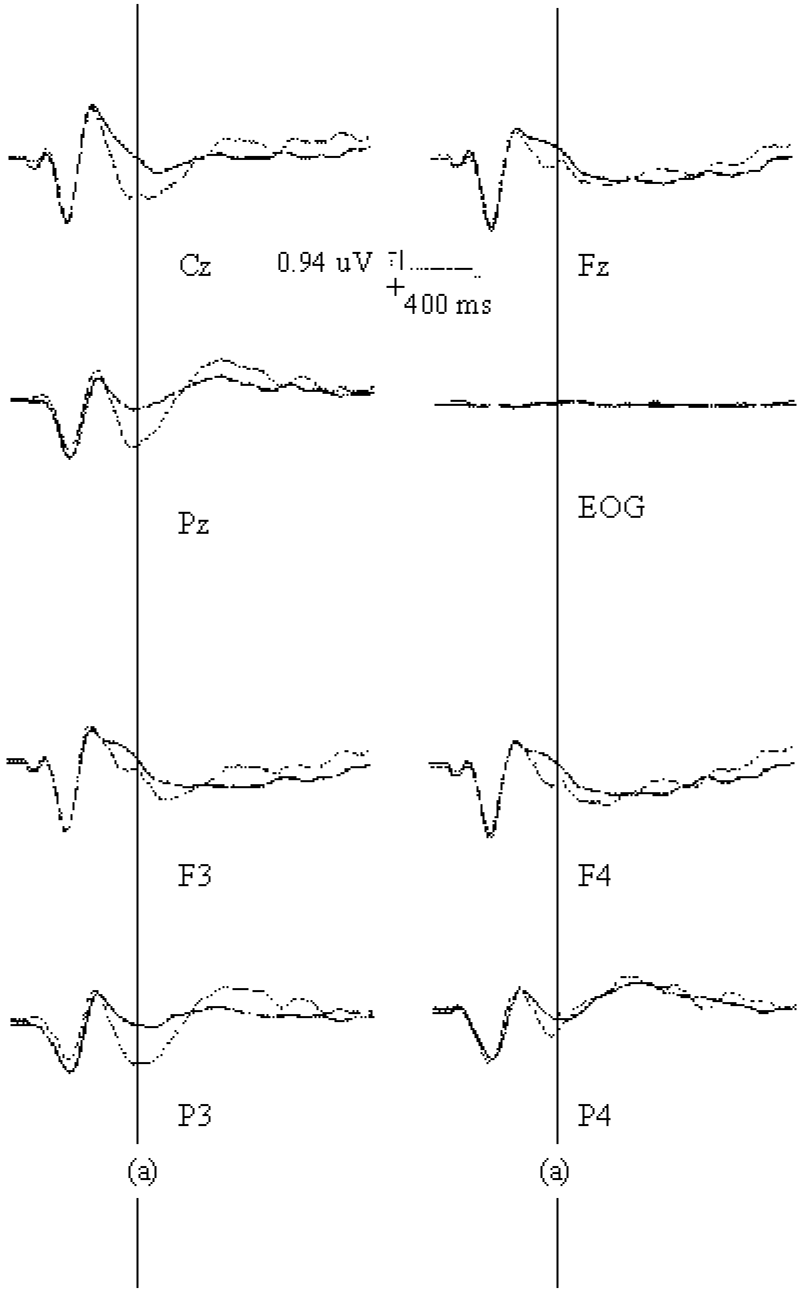


Figure 1

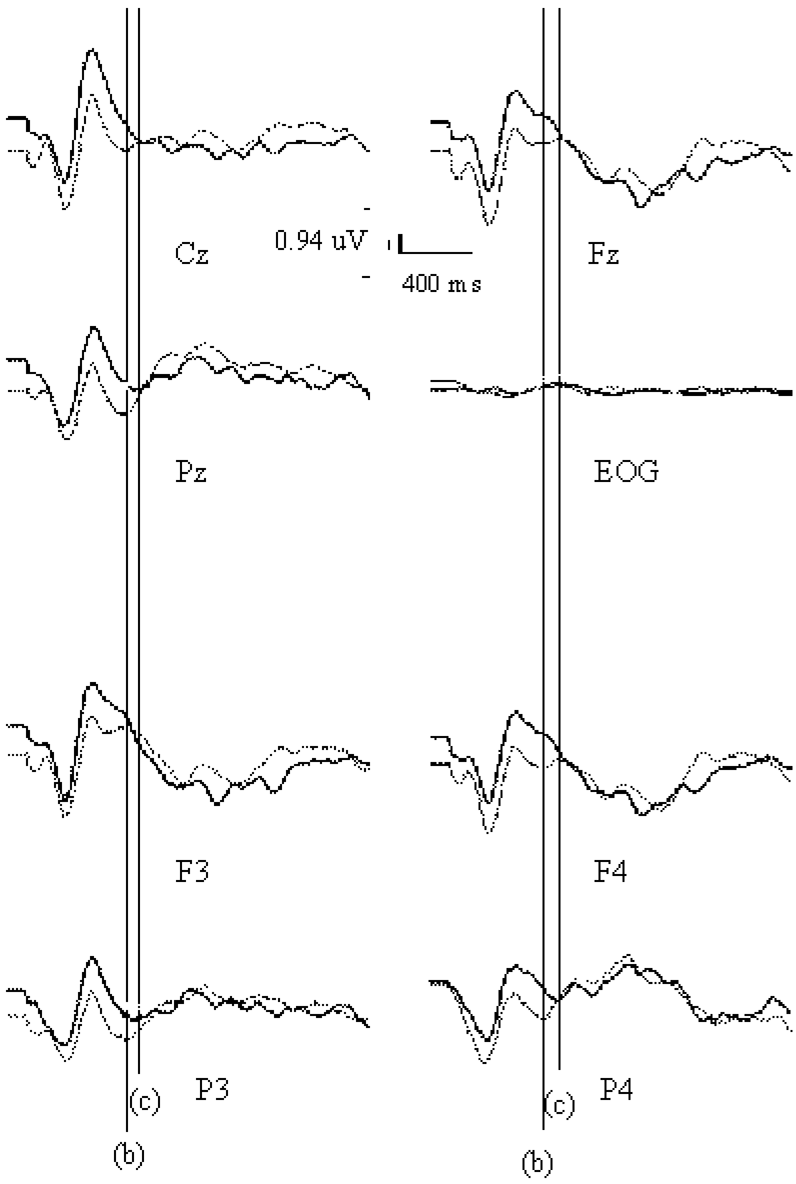


Figure 2

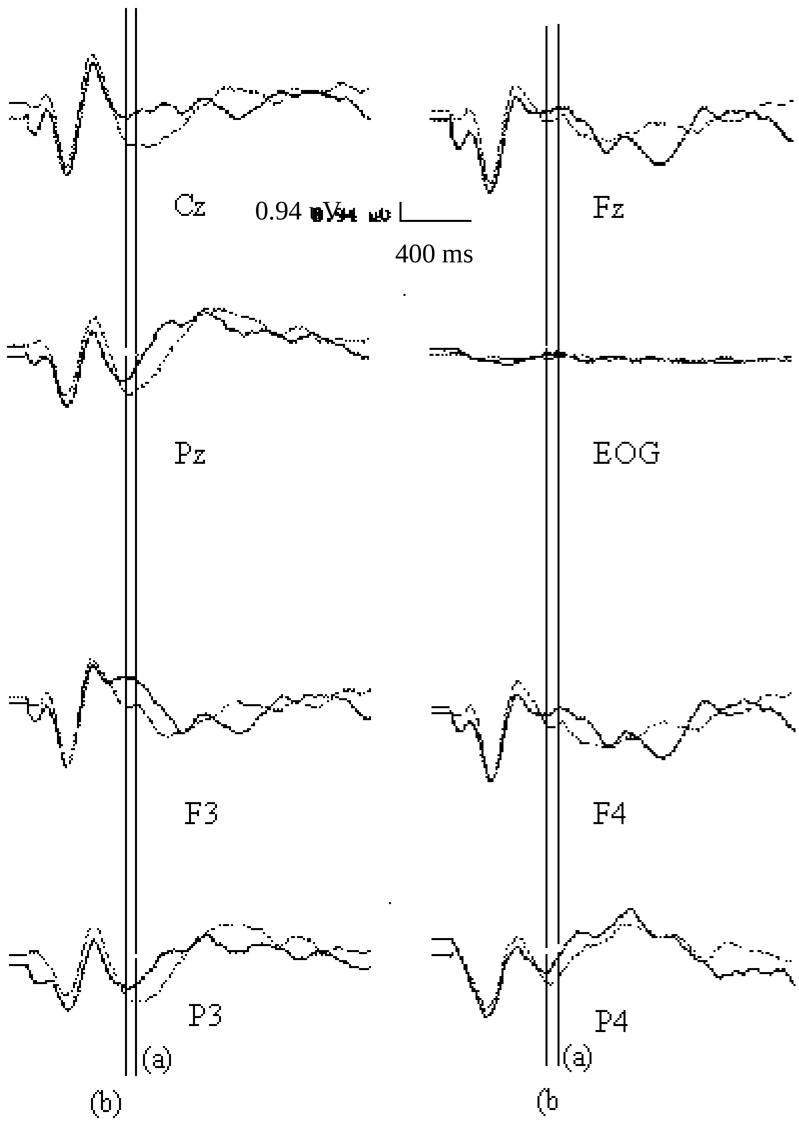


Figure 3

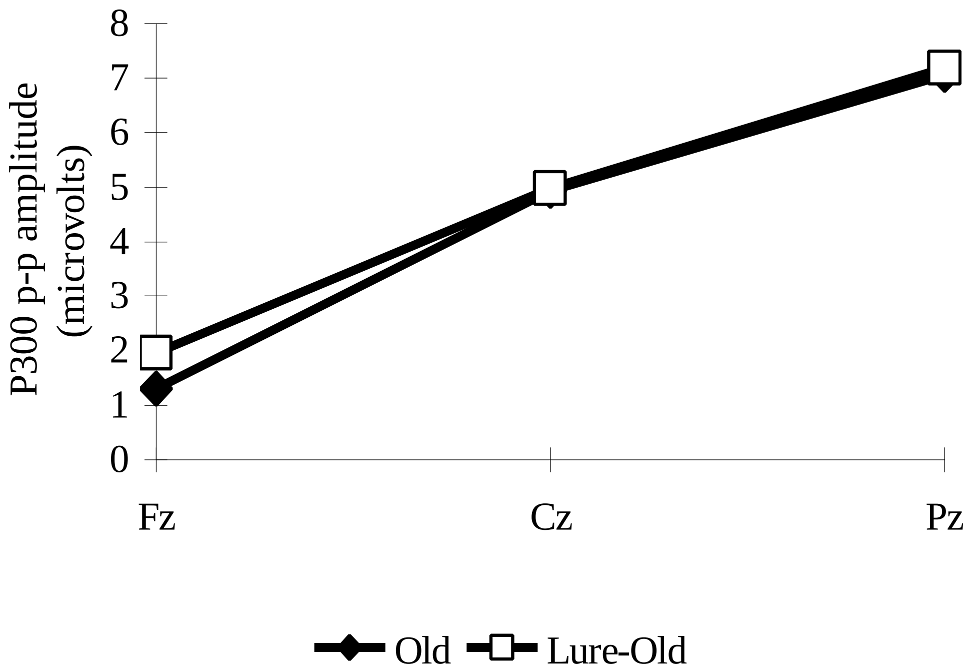

8
7
6
5
4
3
2
1
0
P300 p-p amplitude (microvolts)
Fz
Cz
Pz
Old
Lure-Old

Figure 4

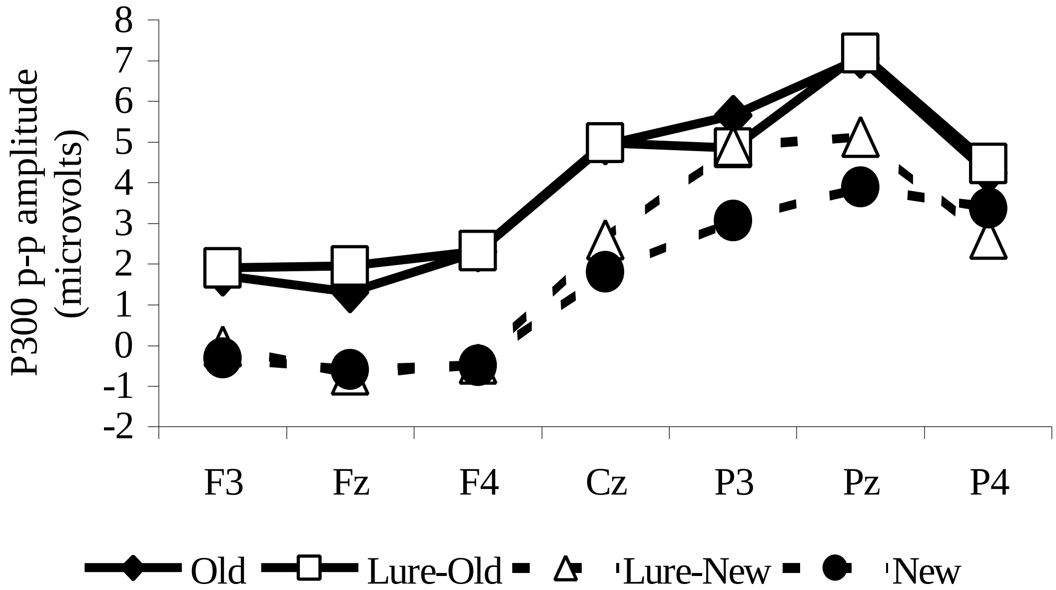

P300 p-p amplitude (microvolts)
8
7
6
5
4
3
2
1
0
-1
-2
F3
Fz
F4
Cz
P3
Pz
P4
Old
Lure-Old
Lure-New
New

Figure 5

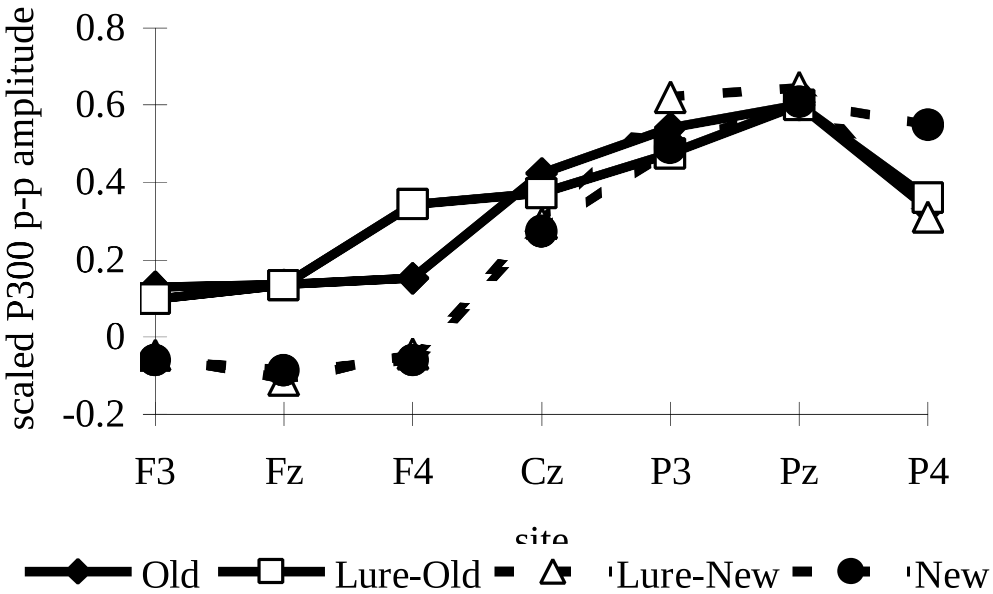

scaled P300 p-p amplitude
0.8
0.6
0.4
0.2
0
-0.2
F3
Fz
F4
Cz
P3
Pz
P4
site
Old
Lure-Old
Lure-New
New

Figure 6

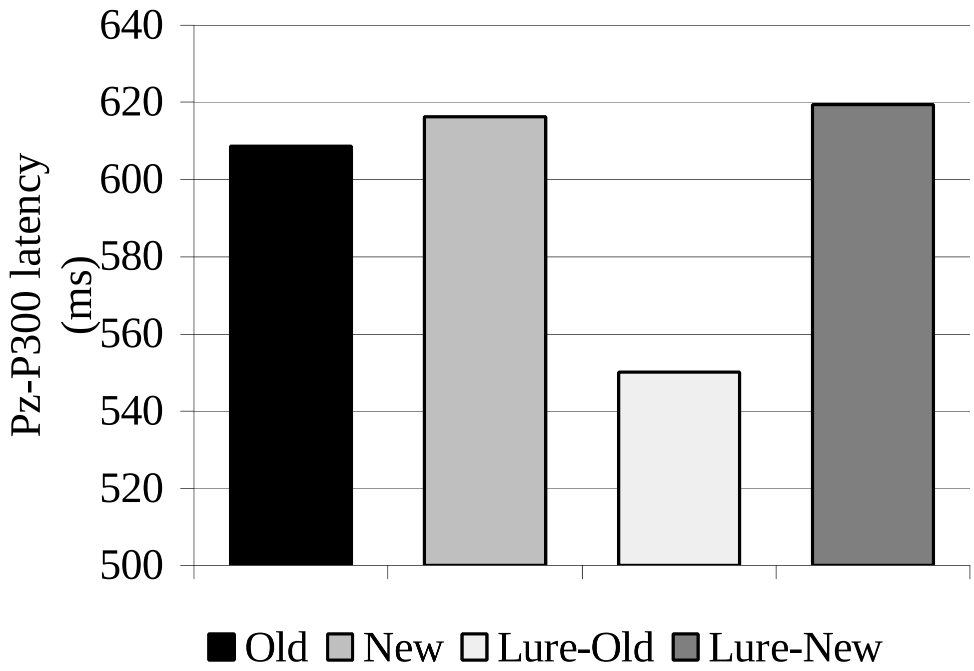

640
620
600
580
560
540
520
500
Pz-P300 latency (ms)
Old
New
Lure-Old
Lure-New

Figure 7

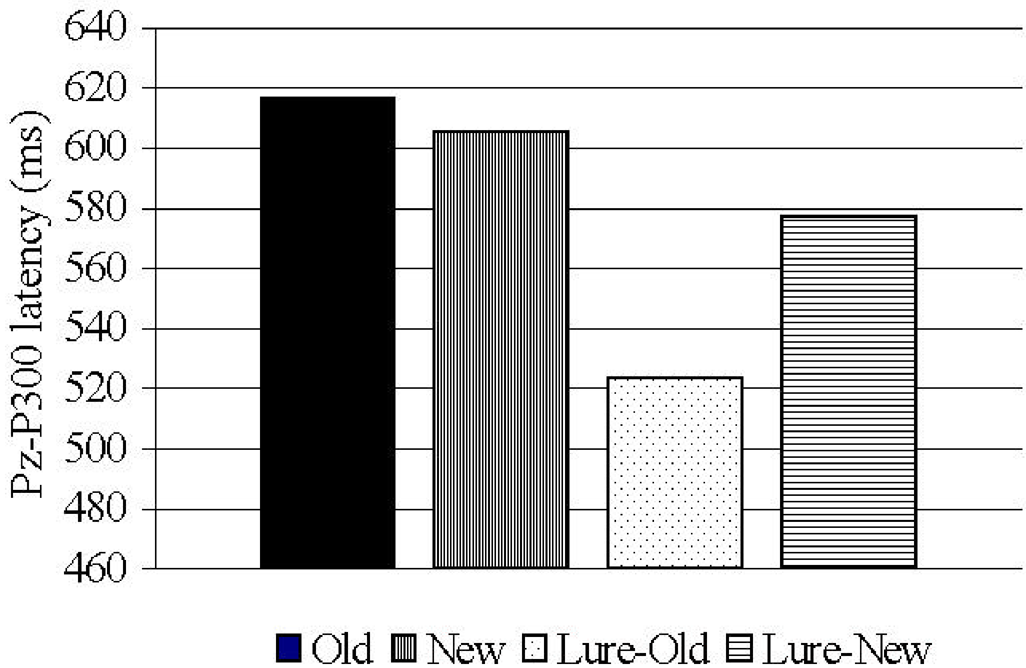
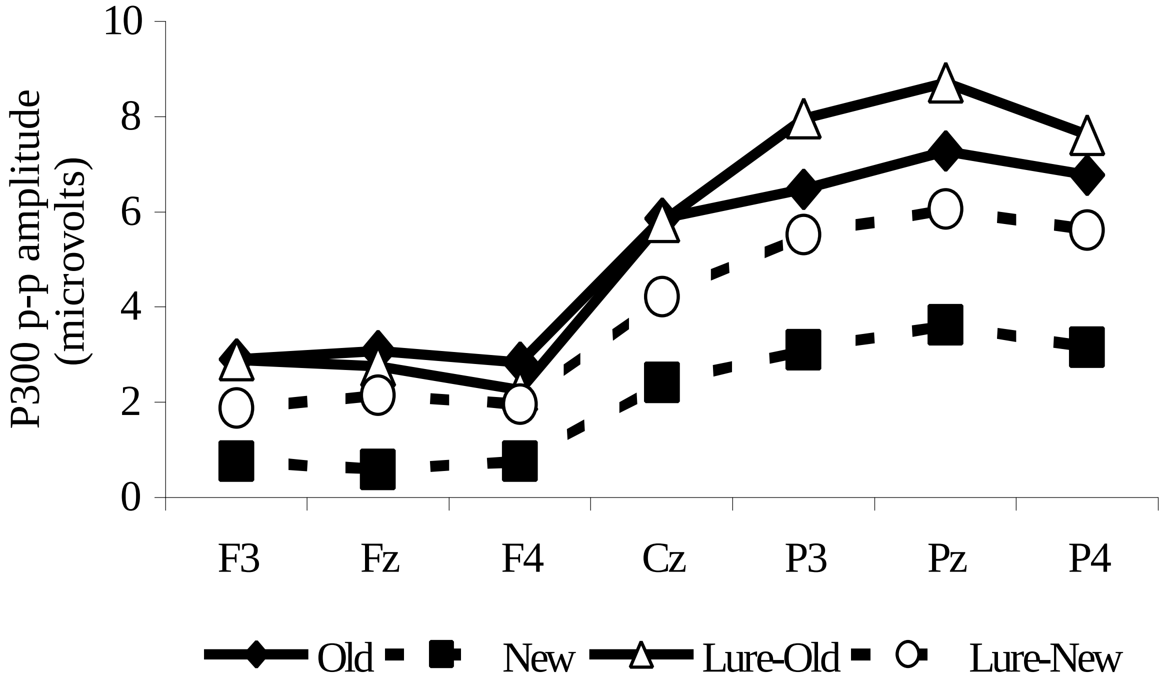
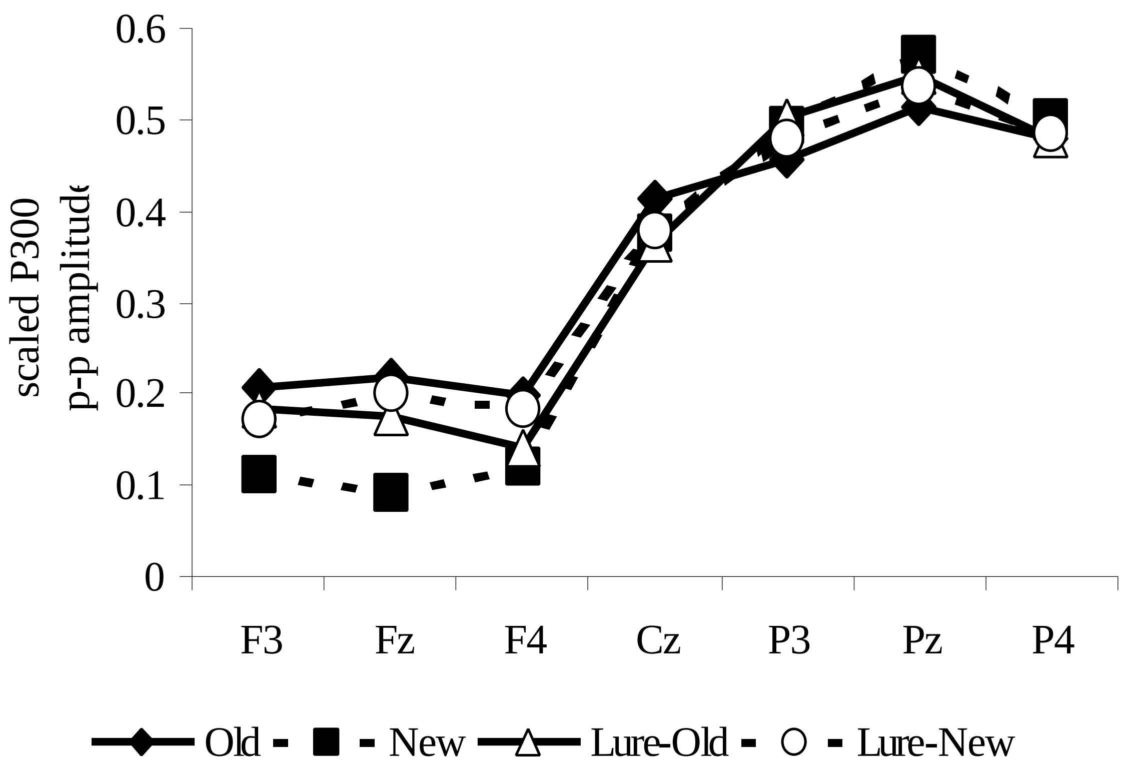

640
620
600
580
560
540
520
500
480
460
Pz-P300 latency (ms)
Old
New
Lure-Old
Lure-New

Figure 8



P300 p-p amplitude (microvolts)
10
8
6
4
2
0
F3
Fz
F4
Cz
P3
Pz
P4
Old
New
Lure-Old
Lure-New

Figure 9



0.6
0.5
0.4
0.3
0.2
0.1
0
scaled P300
p-p amplitude
F3
Fz
F4
Cz
P3
Pz
P4
Old
New
Lure-Old
Lure-New

Figure 10