Published in final edited form as:
*Brain Res*. 2010 March 4; 1317: 180–191. doi:10.1016/j.brainres.2009.12.074.

# Event-related potential correlates of item and source memory strength

**Brion Woroch**[1] and **Brian D. Gonsalves**[1,2]

[1] Department of Psychology, University of Illinois

[2] Beckman Institute for Advanced Science and Technology, University of Illinois

## Abstract

Event-related potential (ERP) studies of recognition memory have shown dissociations between item recognition and source memory, wherein item recognition is associated with the mid-frontal FN400 component, which varies continuously with item memory strength, while source memory is associated with the late parietal effect (LPC). There is current debate about whether source memory can vary along a continuum of memory strength or is a threshold process. The LPC has been shown to be generally sensitive to correct versus incorrect source judgments, but varying levels of "source strength" along a single dimension of source evidence have not been tested. The current experiment had participants encode novel visual objects in one of two different task contexts by performing either a conceptual or perceptual judgment about the object. On a subsequent memory test, participants made an old/new decision on a 4-point confidence scale followed by a source memory confidence judgment, in which they indicated their confidence about which task they had performed with the object at encoding. ERPs from the memory test were examined for electrophysiological correlates of both item and source memory strength. Item memory was associated with differences in the 300–500ms time window, consistent with the timing of the FN400. Differences in the amplitude of the LPC were observed between correct and incorrect source decisions, consistent with previous findings. Comparing low and high confidence source decisions also revealed differences, suggesting that the LPC is also sensitive to variations in the strength of source memory.

### Keywords

Event-related potentials; recognition; source memory

## Introduction

Recognition memory is our ability to discriminate previously encountered stimuli from those that are novel. Theoretical debate about the nature of recognition has largely centered around the debate between single and dual-process models. Dual-process models suppose that recognition can be decomposed into two distinct processes that reflect separate mnemonic mechanisms, known as familiarity and recollection (Jacoby, 1991; Mandler, 1980). Familiarity is the process by which an individual item may be recognized in the absence of the retrieval of information about the source or context in which the item was originally encountered. Recollection is the process of retrieving of these specific contextual details about the encoding episode. Single process models, on the other hand, suppose that the distinction between familiarity and recollection capture instead varying degrees of memory strength along a single

Address Correspondence to: Brion Woroch, Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820, Telephone: (217) 244-1713, Fax: (217) 333-2922, bworoch2@illinois.edu.

continuum, rather than reflecting fundamentally different kinds of memory (Donaldson, 1996; Dunn, 2004; Parks and Yonelinas, 2007; Wixted, 2007a; Wixted, 2007b). In both kinds of models, recognition is often modeled as a signal detection process (SDT), with individual items varying in their degree of memory strength, which is often assumed to be based on the match between the to-be-remembered item and the stored memory representation. For recognition decisions, each item associated with memory strength above a decision threshold is classified as ”old”, and those that fall below are classified as ”new”. Recollection, however, is handled differently by the two classes of models. In the most common instantiation of dual-process models (Yonelinas, 1994; Yonelinas, 2002) recollection is conceptualized as a threshold process, wherein recollection of source or contextual information either occurs or does not, with some probability. One implication of this version of dual-process theory is that items that are recollected are often assumed to be associated with only the highest confidence level of an old/new recognition decision. An alternative interpretation is that both familiarity and recollection are continuous processes, each associated with varying degrees of strength that are summed together prior to a recognition decision. This memory strength signal used to make recognition decisions and assign confidence ratings is thus the combination of both familiarity and recollection processes (Wixted and Stretch, 2004). Thus recollection is not only associated with high confidence recognition decisions, but may contribute to recognition decisions at lower levels of confidence as well. Additionally, high confidence old/new decisions may be associated with varying degrees of recollective strength, though this last assumption is also compatible with the Yonelinas dual process model (Parks & Yonelinas, 2007). These assumptions have important implications for the relationship between old-new recognition confidence and source accuracy, in that models in which recollection is a graded process predict that accurate source retrieval should contribute to recognition memory decisions at all levels of confidence, whereas threshold models predict that source accuracy should be at chance for all but the highest level of recognition confidence. As we outline below, the assumptions that go into behavioral measures of these two kinds of memory also have important implications for attempts to demonstrate dissociable brain mechanisms of recollection and familiarity.

Event-related potentials (ERPs) have been used extensively to adjudicate between single and dual process accounts of recognition memory. On the whole, results from this research seems to support dual-process models by showing that recollection and familiarity tend to be associated with distinct components of the ERP with different timing and scalp distributions (Rugg and Curran, 2007). More specifically, task manipulations that are thought to affect memory for individual items based on familiarity modulate the FN400, a mid-frontally distributed component that peaks from 300–500ms after stimulus onset, with less familiar items having a more negative peak. In contrast, manipulations of recollection affect a separate component often termed the parietal old/new effect or LPC that is localized over parietal regions of the head, peaks from 400–800ms, and is more positive to recollected items than old items for which recollection did not occur. (Curran, 2000; Düzel et al., 1997; Rugg et al., 1998; Rugg and Curran, 2007). Other ERP studies of source memory, which is often used as a way to measure recollection, have also shown the amplitude of the LPC is sensitive to the amount of information or number of sources recalled (e.g., Vilberg, et al., 2006; Wilding, 2000). Typically, however, these studies only examine recollection in a way that is more consistent with threshold models, namely, whether source information is present or absent (which includes whether or not multiple sources have been remembered), and levels of confidence along a single dimension of source evidence are not measured. The possibility that confidence for retrieval of an individual source, and therefore the amplitude of the LPC, may vary has generally not been explored in ERP studies of source memory. One exception is a recent study by Leynes & Phillips (2008), which showed that the LPC was larger for “Remember” than “Know” responses made following accurate source judgments. These responses may serve as a proxy for confidence in the source decision, though the subjects in that experiment were not specifically instructed to respond based on their confidence in their source decision.

Despite the apparent links between these two ERP components and the two memory processes, the validity of the link between the ERP correlates of recollection and familiarity and the specific decision processes associated with them in single vs. dual process models has been challenged in two important ways. First, it has been challenged on the grounds that the FN400 component may not reflect familiarity, but rather conceptual priming, and that measures should be taken to control for conceptual priming when investigating ERP correlates of familiarity (Paller et al., 2007). A second challenge, which is the focus of the current study, has been that the behavioral measures used to identify neural correlates of recollection and familiarity may not be process-pure, such that neural measures will ultimately reflect some combination of the two processes, or different aspects of a single underlying process. For example, most models of recognition memory agree that item recognition based upon familiarity occurs in a continuous manner in accordance with SDT. Thus an important prediction is that the neural correlates of familiarity, such as the FN400, are expected to be graded as a function of changes in the strength of the familiarity signal to test items (Curran, 2004; Azimian-Faridani et al, 2006; see also Gonsalves et al 2005). A recent study by Woodruff and colleagues (2006) showed that the amplitude of the FN400 was indeed modulated in a graded fashion with item confidence judgments made by participants (an index of item memory strength). The LPC effects were not modulated by item confidence, but were instead modulated by a “Remember” response. This was taken as evidence that the FN400 reflects familiarity-based recognition, based on item memory strength, and the LPC reflects recollection.

However, that study assumed a threshold dual-process model of recognition in which recollection would only be associated with the highest levels of confidence, those given a “Remember” response, and confidence ratings for recollection decisions were not queried, though it was generally clear that the results were inconsistent with a unidimensional single-process model. Recent behavioral studies, however, have shown that recollection is not strictly associated with the highest item confidence ratings, but lower confidence ratings as well (Mickes et al., 2009), and that above-chance source memory performance is associated with both “remember” and “know” responses (Hicks, et al., 2002), suggesting that recollection may instead be a continuous process, like familiarity, a notion that is broadly consistent with the source monitoring framework (Dodson et al., 1998; Johnson et al., 1993). Slotnick and Dodson (2005) showed that by modeling source decisions as a strength continuum in multidimensional decision space, they could account for the behavioral evidence originally generated in support of the dual-process threshold model. If recollection is indeed a continuous process, variations in recollection strength and differences in recollective confidence may also be reflected by a graded ERP response. Observations of graded levels of activity in the brain as a function of recognition confidence are often assumed to be markers of item memory strength on the assumption that familiarity is graded but recollection is not. However, such effects may be contaminated by graded source memory, such that the observed neural effects reflect a combination of recollection and familiarity. Similarly, correlates of source memory, typically measured as source hits versus source misses, may be confounded by differences in item memory strength between these two conditions, such that source hits may be associated with higher overall item strength than source misses. Since the LPC has been shown to index to presence or absence of recollection, we hypothesize that it may be modulated by changes in the strength of source memory if confidence ratings are collected for source decisions.

The current study sought to directly examine the ERP correlates of both item and source memory strength, with the goal of obtaining ERP markers of memory strength on one dimension while holding strength on the other relatively constant, to avoid some of the issues discussed above, using logic previously applied in an fMRI study (Kirwan, et al., 2008). We conducted a source memory experiment with novel visual objects in which participants encoded the objects in one of two encoding contexts, conceptual or perceptual. We queried participant's memory confidence during both the old/new and source decisions. This design allows us to measure the neural response to item or source strength independently, while accounting for possible contributions of the other. We first sought to replicate the findings of Woodruff et al. (2006) that showed the modulation of the FN400 to variations in item confidence, but doing so while accounting for possible contributions of source memory to old/new confidence ratings. We sought to isolate the changes in confidence based on item memory by looking for variations in confidence in an old/new recognition decision only on trials when recollection did not occur, as measured by the fact that source was not accurately retrieved (Source Miss). We also queried confidence in source decisions to index source memory strength. This enabled us to examine the ERP correlates of changes in source memory strength while keeping item confidence at a set level, by only considering trials on which subjects gave a high confidence item response. High-confidence item responses were chosen because accurate source memory tends to be associated with higher confidence item decisions, such that there would be too few low-confidence item decisions followed by accurate source memory to generate usable ERP averages. Finally, we used novel visual stimuli that should be minimal in their conceptual content in an attempt to mitigate some of the potential influences of conceptual priming on ERP measures of item memory.

## Results

### Behavioral Results

Behavioral data from the test phase are shown in Tables 1 and 2. The levels of chance were defined as the probability of making a correct response if one were selecting randomly. For item memory, this was 50% (2 out of 4); for source memory it was 40% (2 out of 5), due to the availability of a $5^{th}$ response option ("unsure").

The overall hit rate for the Item test was 80%, the correct rejection rate was 83%. When broken down by confidence, accuracy was higher for high versus low confident hits [97% vs. 70%; $t(20)=12.37$, $p<0.001$], and the hit rate for low confidence responses was significantly greater than chance [50%; ($t(20)=3.81$, $p<0.001$]. For computing source accuracy, only trials on which the item decision was correct were considered. The overall accuracy from the Source test was 67%. Sorting source accuracy by confidence of the preceding item decision revealed significantly greater source accuracy for High Confident (HC) Item Hits (76%) than Low Confident (LC) Item Hits [43%; $t(20)=11.86$, $p< 0.001$]. However source accuracy for LC Item Hits was not significantly above chance [40%; $t(20)=0.88$, $p>0.1$]. When source accuracy was broken down by source confidence, accuracy was higher for high versus low confidence [91% >67%, $t(20)=19.02$, $p<0.001$]. Additionally low confident source accuracy was significantly greater than chance [40%; $t(20)=9.58$, $p<0.001$]. There were differences in item and source accuracy depending upon encoding task. Item accuracy was higher for the conceptual than perceptual tasks, 82% and 78% respectively [$t(20)=3.30$, $p<.005$]. However source accuracy was higher for the perceptual than conceptual tasks, 70% and 65% respectively [$t(20)=2.25$, $p<.05$]. In summary, high confidence was associated with greater accuracy than low for both item and source judgments, and the accuracy of low confident item and source decisions was greater than chance. However, when subjects made a low confidence item decision, their subsequent source accuracy was at chance.

### ERP Results

All ERP analyses for this study were conducted on ERPs time-locked to the presentation of the to-be-remembered object, under the assumption that relevant item and source memory processes should be occurring at this time, despite the fact that responses were delayed. Initial

ERP analyses focused on identifying ERP correlates of item and source memory, considering all item or all source trials regardless of performance on the other test. For the ERP correlates of item memory, the focus was on replication of the Woodruff et al. (2006) item confidence analysis, looking for a graded pattern in the ERPs as a function of item memory strength (high confident hits > low confident hits > low confident correct rejections > high confident correct rejections). The second was a traditional source memory analysis, comparing ERPs to items followed by correct versus incorrect source decisions. For these two analyses, the electrodes were grouped into 4 regions of four channels each reflecting frontal and parietal scalp locations (shown in Fig. 1); left anterior superior (LAS) consisting of F1,F3,FC1,&FC3, right anterior superior (RAS) F2,F4,FC2,FC4, left posterior superior (LPS) CP1,CP3,P1,P3, and right posterior superior (RAS) CP2,CP4,P2,P4. The mean amplitude of the waveform was calculated and averaged over the chosen electrodes for each region for both an Early and a Late time window (300–500 ms & 600–900 ms respectively). A 3-way repeated-measures analysis of variance (ANOVA) was run on these mean amplitude measurements for each time window, with Condition, Hemisphere and Anterior-posterior (AP) axis as factors. A Hyun-Feldt correction for non-sphericity was used to adjust the appropriate degrees of freedom when applicable, and adjusted $p$ values are shown when reporting statistics. Effects of Hemisphere and AP axis are not meaningful unless they interacted with Condition, and thus main effects of these factors are not reported.

## Item Confidence Effects

**Early Time Window (300–500ms):** The first analysis examined ERP correlates of item confidence, regardless of subsequent source performance. The four conditions of interest were low confidence (LC) and high confidence (HC) hits and LC and HC correct rejections (Fig. 2). There was a significant main effect of Condition [$F(3,60)=6.49$, $p<.001$], but no significant interactions of Condition with either Hemisphere or AP axis [all $F<1.67$, all $p>1.84$], indicating a broadly distributed effect of Condition. To further characterize this main effect, a post-hoc polynomial contrast was performed, which showed a significant linear trend across the item confidence conditions: HC hit > LC hit > LC correct rejection > HC correct rejection [$F(1,20)=24.76$, $p<.001$], but no significant quadratic trend [$F(1,20)=0.36$, $p=.56$]. Planned pairwise comparisons showed significant differences between HC correct rejections and all conditions as well as between HC hits and LC correct rejections [all $p<.05$]. There was a marginal difference between HC and LC hits [$p=.059$], and no significant difference between LC hits and LC correct rejections [$p=.85$].

**Late Time Window (600–900ms):** In the late time window, the same ANOVA with Condition, Hemisphere, and AP axis as factors revealed a main effect of Condition [$F(3,60)=14.37$, $p<.001$]. Post-hoc polynomial contrasts also showed a significant linear trend across item confidence [$F(1,20)=18.87$, $p<.001$], as well as a significant quadratic trend [$F(1,20)=21.88$, $p<.001$]. However, this effect differed from the effect observed in the early time window in that it appeared to be carried by the positivity in the waveform for HC hits and did not show the same graded pattern across conditions. Planned comparisons confirmed this impression, revealing significant differences between HC hits and all other conditions [all $p<.001$], but no significant difference amongst the other conditions [all $p>.90$].

**Source Accuracy Effects—**This analysis compared ERPs to item hits with correct versus incorrect source decisions regardless of item or source confidence (Fig. 3) Mean amplitudes from each time window were submitted to an ANOVA with Condition (Source Hit, Source Miss), Hemisphere, and AP axis as factors.

**Early Time Window (300–500ms):** There was no significant difference between conditions in the mean amplitude of the ERP between Source Hits and Source Misses in this time window

[F(1,20)=.389, p=.54]. There was no interaction of Condition with the other factors [F(1,20) =1.03, p<.32].

**Late Time Window (600–900ms):** In the late window, there was a significant main effect of Condition [F(1,20)=6.94, p=.16]. Additionally, there was a 3-way interaction of Condition x Hemisphere x AP axis [F(1,20)=5.00, p<.05], reflecting the fact that the Condition by Hemisphere interaction was significant in the Posterior electrodes [F(1,20)=6.16, p<.05) but not the Anterior electrodes [F(1,20)=.679, p=.42]. The difference between source hits and misses on the Posterior channels was greater on the left than the right [t(20)=2.48, p<.05]. Finally, to test whether these source accuracy effects differed as a function of the specific source, we compared the Source Hit – Source Miss difference waves between trials for items encoded in the Meaning task versus the Complexity task at 600–900ms using the same four-location analysis. There was no significant difference between these difference waves for source type [F(1,20)=0.093,p=.764] as well as no interactions with source type [all F<3.60 all p< .07] indicating that the observed source effects did not differ as a function of encoding task.

To summarize this initial set of analyses, the Item memory comparisons replicated the finding of Woodruff et al. (2006) of a monotonic relationship between item confidence and ERP amplitude during the 300–500 ms time window, which is typically associated with the FN400 familiarity effect. The Source memory comparisons also replicate previous findings by showing a greater amplitude for items accompanied by a correct versus incorrect source decision during the typical LPC time window, and the difference was greater over the left than right hemispheres.

The second set of analyses sought to build on these results by assessing the independent contributions of item and source memory strength to the ERP effects, since, as discussed above, both may be contributing to the two effects that we observed in the initial, more traditional analyses. To assess the unique contributions of each type of memory, we held one type of memory at a set level and looked the differences between high and low confidence in the other. Thus there are two comparisons: item confidence effects including only hits that were followed by a Source Miss (Fig. 4); and HC versus LC Source Hits that were preceded by a HC Item hit (Fig. 5). The Item Confidence analysis was run similar to the initial analyses described above, with Condition (HC Hit, LC Hit, LC CR, HC CR), Hemisphere (L,R), and AP axis as factors. However, visual inspection of the waveforms for the source confidence comparisons revealed that the effects were maximally distributed over central scalp regions, so a single central electrode cluster was chosen over the left (C1,C3,CP1,CP3) and right (C2,C4,CP2,CP4) hemispheres (Fig. 5). Again, the mean amplitude was calculated for early and late time windows, 300–500ms and 600–900ms. A repeated-measures ANOVA was run on the mean amplitude for each time window with Condition and Hemisphere as factors.

### Item Strength, with Source Strength Held Constant

**Early Time Window (300–500ms):** In the Item Confidence comparison, there was a main effect of Condition [F(3,60)=5.70, p<.005], but no interactions between Condition and the other Factors [all F<.77, all p>.56]. A follow-up polynomial trend contrast on the main effect of Condition revealed a significant linear trend as a function of item confidence [F(1,20)=13.74, p<.001], but no quadratic trend [F(1,20)=0.48, p=.50]. Planned comparisons showed a difference between HC hits and all other conditions, and between HC and LC correct rejections [all p<.05]. There was no statistically significant difference between LC hits and LC or HC correct rejections [all p>.15].

**Late Time Window (600–900ms):** In the late time window, there was again a main effect of Condition [F(3,60)=7.68, p<.001], but no Condition x Factor interactions [all F<1.01, all p>.

33]. A polynomial contrast on the main effect of Condition showed a significant linear trend as a function of item confidence [$F(1,20)=12.25$, $p<.01$], as well as a significant quadratic trend [$F(1,20)=8.11$, $p<.01$]. Planned mean comparisons in this time window revealed a significant pairwise comparison between HC hits and all other conditions [all $p<.002$]. There were no other pairwise differences [all $p>.59$]. Overall, the results from this restricted analysis, assessing ERP correlates of item confidence while controlling for source confidence, largely converged with the results observed in the unrestricted analysis. In both sets of analyses, ERPs were graded in amplitude as a function of item memory strength from 300–500ms, consistent with the interpretation that brain activity in this time window reflects processes associated with item memory strength.

### Source Strength, with Item Strength Held Constant

**Early Time Window (300–500ms):** There were no significant main effect of Condition or interactions in the early time window [all $F<.93$, $p>.35$].

**Late Time Window (600–900ms):** In the late time window, there was a significant main effect of Condition [$F(1,20)=4.89$, $p<.05$], as well as a significant Condition by Hemisphere interaction [$F(1,20)=5.47$, $p<.05$]. Follow-up tests revealed that the difference in amplitude between conditions was greater in the right hemisphere than left [$t(20)=2.34$, $p<.05$]. Overall, then, the ERP effects associated with source confidence were restricted to the late time window and were right lateralized, unlike in the comparison of source hits and source misses, which was associated with a late, left-lateralized effect.

**Topographic comparisons—**The analyses of item and source confidence revealed that ERP effects of item confidence are apparent earlier than those of source confidence (300–600ms versus 600–900ms). Given, however, that item confidence effects were apparent in the later time window, we compared scalp topographies of item and source effects in this time window, for the restricted analyses. An ANOVA was run on the differences in scalp topography between item (HC Hits - LC Hits) and source confidence in the 600–900ms time window using the scaling procedure outlined by (McCarthy and Wood, 1985). This comparison revealed no significant interaction between electrode and confidence type with the Hyun-Feldt correction for non-sphericity [$F(63,1260)=1.14$, $p>.19$].

We were also interested in the scalp distributions of the unrestricted and restricted Item and Source ERP effects, and whether the distributions differed between the unrestricted and restricted analyses (scalp topographies shown in Fig. 6). The same analysis procedure as above was used. The comparison of the source topographies was not significant [$F(63,1260)=0.987$, $p=.461$]. Because the critical aspect of the item effects was the graded ERP amplitudes across four conditions, rather than amplitude differences between two conditions, we used a different method to more accurately reflect the scalp distribution of this graded effect across conditions. To plot and compare the item confidence topographies, a four-point linear regression was run at each electrode site on amplitudes from the four conditions (HC hit > LC hit > LC correct rejection > HC correct rejection). Topographic plots show the distribution across the scalp of $r$ values obtained from these analyses. The statistical comparison of the restricted and unrestricted (incorrect source memory) distributions of $r$ values was not significant [$F(63,1260)=1.542$ $p=.092$].

## Discussion

The goal of this experiment was to identify distinct neural markers of item and source memory using ERPs, while controlling for a potentially confounding influence of the lack of process-purity in behavioral measures of item and source memory. Before proceeding to discussion of the ERP findings, some discussion of the behavioral results is warranted. From the perspective of competing models of recognition memory, the behavioral data could be consistent with dual process models of recognition that argue for recollection as a threshold process, but they could also be consistent with multidimensional signal detection models. Specifically, we observed that accuracy of source memory decisions was not significantly different from chance for low confidence item recognition decisions. This is may be construed as inconsistent with the predictions of models that suppose recollection is a continuous process that contributes to item recognition decisions at all levels of confidence (Slotnick and Dodson, 2005; Wixted and Stretch, 2004). However, given that we had only two levels of item confidence to which hits could be assigned, the possibility remains that a third, intermediate level of item confidence may have been associated with above-chance source memory, had such an option been provided. Thus, our behavioral data lack the resolution to argue in favor of one model versus the other, though it is important to note that our study was not designed to do this, but rather to clear up some potential ambiguities about the putative ERP correlates of item and source memory.

Some, though not all, of the focus of ERP studies of recognition memory has been of adjudicating between single and dual process models of recognition, with the idea that demonstrating that recollection and familiarity are associated with distinct neural signatures should argue clearly in favor of dual process models. This effort has been complicated, however, by the lack of a simple mapping between the utilized measures of recognition and the supposed underlying processes. In the current experiment, we attempted to separate the contributions of item and source memory to ERP recognition effects by examining how ERP signatures varied as a function of increasing confidence in one kind of memory while holding confidence in the other kind of memory relatively constant. A similar logic has been applied recently in an fMRI experiment (Kirwan et al., 2008) to look at brain activity at encoding that is predictive of later item and source memory. That study found that later item and source memory are indeed associated with different encoding activation, though not within the medial temporal lobes, which were the theoretical focus of that study. In the current data, we cannot definitively pinpoint the neural sources of the observed electrophysiological indicies of item and source memory, thus we cannot inform the debate as to whether subregions of the medial temporal lobes are functionally heterogeneous with respect or item and source memory, or whether medial temporal lobe activity generally reflects the strength of a memory, irrespective of item versus source distinctions (Wais, 2008).

Our data can, however, inform the interpretations of the putative ERP markers of item and source memory. Our initial ERP analyses focused on characterizing ERP signals associated with item and source memory using the typical sorts of analyses that are performed in the literature. For item memory, we sought to identify portions of the ERP that were graded in amplitude as a function of memory strength, as measured by HC hits, LC Hits, LC Correct Rejections, and HC CRs (Woodruff et al., 2006). As has been found in prior studies, we found an early effect (300–500ms) that showed such a graded effect, with more positive ERPs for HC hits, followed by LC Hits, LC CRs, and finally HC CRs. Under the assumptions of the Yonelinas dual process model, a graded effect such as this is interpreted as reflecting item memory or familiarity-based recognition, since such a model assumes that only familiarity should be graded across these conditions, whereas recollection should occur only for high confidence hits (and certainly not for any correct rejections). However, more recent dual-process signal detection models that assume that both recollection and familiarity are graded, and sum together to form an aggregated strength signal (Wixted and Stretch, 2004) do not allow such a straightforward assignment of a graded ERP effect to item memory/familiarity. Rather, such an effect may reflect contributions of both item and source memory under this sort of model, leaving the functional interpretation of the ERP effect more ambiguous.

Similar problems may cloud the interpretation of traditional ERP source memory effects as well. Our source memory effects obtained using the standard comparison (Source hits vs. Source misses) revealed an effect that largely converged with what is typically observed – more positive ERPs for source hits compared to source misses in a later (600–900ms) time window. However, this comparison may also be confounded, in that ERP differences between source hits and source misses may also reflect differences in item strength between these two response categories, under the assumption that source hits will tend to be associated with higher item confidence, and thus may not reflect neural activity that is exclusively associated with memory for source, though some prior dissociations seem to argue against this interpretation (e.g. Woodruff et al., 2006).

To address these issues, we conducted additional ERP analyses that attempted to better separate brain activity associated with item and source memory. To do so, we collected confidence ratings for both item and source responses during the memory test, with the goal of being able to identify ERP correlates of the strength of one type of memory (item or source) while holding the strength of the other relatively constant. For the item memory effect, we conducted a similar analysis to the one described above, looking for graded ERP effects as a function of item memory strength. However, for this analysis, we only used trials on which subjects had been inaccurate in their source judgment, with the idea that this should limit the contributions of source memory to the observed ERP effects. The results from this restricted analysis largely converged with the results observed in the unrestricted analysis, in that in both analyses, ERPs were graded in amplitude as a function of item memory strength from 300–500ms. In the later time window (600–900ms), the pattern of activity changed, such that high confident hits were associated with a sustained positivity throughout this time window, while the ERPs to the other three conditions (LC hits, LC CRs, HC CRs) were relatively similar. At first glance, it seems that this late effect may be consistent with threshold models of recollection, in that high confident “old” responses may diverge from the other response categories due to the presence of recollection for only these high confident hits. However, for this analysis we looked at only trials where the source was not correctly recalled, which means for these high confident item hits, correct source information was not retrieved. One possibility is that the late effect does indeed reflect recollection, just not recollection of the relevant source information; so-called non-criterial recollection (Yonelinas and Jacoby, 1996). Under this account, the late positivity observed for HC items reflects the retrieval of additional contextual detail associated with the test item, but not the relevant information needed for accurate source memory. Further studies would be needed to explore this possibility.

In addition to this restricted analysis for item memory, we also conducted a restricted analysis to better identify ERP markers of source memory strength. For this analysis, we compared ERPs associated with accurate high confidence source decisions to those associated with accurate low confidence source decisions, restricted to those trials on which subjects had made a high-confidence old/new response. As with the item memory ERP effects, the results of this restricted source analysis largely converged with the results obtained using the unrestricted analysis, with ERPs to high confidence source more positive than those to low confidence source in the late (600–900ms) time window only. These results are generally consistent with the idea that the LPC is modulated by the amount or quality of source information retrieved (Vilberg et al., 2006; Wilding and Rugg, 1996), and furthermore that this is true even when minimizing the potential contributions of item memory to the LPC. It should be noted that the scalp distribution of the source strength effects was different than what is typically labeled as the LPC, in that the LPC typically has a left parietal distribution, while our restricted effect tended to have a right central distribution, though our unrestricted source memory effect did not significantly differ in scalp distribution from the restricted analysis. It is not immediately clear what accounts for this apparent bilaterality/right laterality in the current study, though it should be noted that the stimuli used were different from many studies in that they were novel visual objects that were difficult to verbalize. Studies of source memory that have used similar visual stimuli have also found ERP source memory effects with a bilateral or right-lateralized scalp distribution (Van Petten and Senkfor, 1996; Voss and Paller, 2007; Voss and Paller, 2009). Our observed ERP difference between different levels of confidence in an accurate source decision also converges with results from a recent study (Leynes and Phillips, 2008) which compared ERPs to Remember and Know judgments following an accurate source decision, and found an LPC difference between these two response types, which they took as evidence for variations in the amount or quality of retrieved information associated with an accurate source judgment.

To summarize, the current ERP experiment sought to identify ERP markers of item and source memory, while accounting for some of the potential problems with typical ERP recognition memory effects that are due to ambiguities about how subjects' responses map to underlying memory processes or kinds of memory. The results of this experiment largely converged with what has been obtained using these potentially problematic analyses, indicating that item and source memory, which are often used as proxies for familiarity and recollection, are indeed associated with dissociable ERP effects. Item memory was associated with an earlier (300–500ms) effect that was graded as a function of item memory strength, while source memory was associated with a later (600–900ms) effect that varied with confidence in the source decision, but not with item confidence. These results lend support to the idea that item and source memory rely on dissociable neural mechanisms, and are inconsistent with the notion that these kinds of memory differ only in their relative strength on a single dimension of evidence. Furthermore, the Source confidence ERP effects add to the existing literature showing that the later ERP effects vary with the amount or number of bits of information recollected, by showing that this ERP correlate of recollection and source memory also can vary as a function of subjects' confidence in their source memory decision.

# Experimental Procedure

## Participants

Twenty-one subjects (15 F; ages 18–25 years old) were included in this experiment. An additional 23 subjects were excluded from analysis due to having low numbers of trials for ERP analysis (<15) in some response conditions. It should be noted that participants could be excluded in this way either in cases where their memory performance was poor, or when their performance was exceptionally good, such that there were not enough trials to look at, for example, source misses. Overall performance in item recognition as measured by Hits minus FA was similar between the included [mean = .63, SE = .03] and excluded [mean = .57, SE = .07] subjects [$t(30)=1.01$, $p=.30$] as was overall source memory accuracy [.67 (.03) vs .65 (.05); $t(37) = .64$, $p=.53$] indicating that the included sample was generally representative of the group as a whole. All participants were right-handed by self-report, had no history of psychiatric or neuropsychological disorders, and were not currently taking any psychotropic medications.

## Procedure

The participants performed a recognition memory experiment using novel visual objects (Warren, D. & Cohen, N.J., in preparation). The objects were created using Bryce software and were varied in shape, color, pattern, and texture. The images extended 6.5 degrees of visual angle from top to bottom and left to right. The task was split into 11 encoding/test blocks. All stimuli were counterbalanced across condition and block. During each encoding block, participants were shown 26 objects for 3 seconds each. The first and last objects from each encoding list were not tested to mitigate the effects of primacy and recency. For each object a subjective meaningfulness or complexity judgment was made on a 4-point scale. Prior to each object, participants were cued for 2 seconds with the word “meaning” or “complex” indicating which judgment should be made for the following object. Each cue and object was followed by 1 second of central fixation (white cross on a black background). For the meaningful task, subjects were instructed to rate how much the object looked like something meaningful, like looking at clouds or an ink-blot test. For the complex task, subjects were simply instructed to rate the more varied objects as more complex.

During the subsequent recognition test, participants were presented with 40 objects (24 objects from the encoding phase and 16 novel objects) for two seconds each followed by a response cue. Participants then made an old/new judgment crossed with confidence level (sure old, think old, think new, or sure new). If the object was indicated as “old”, they made an additional source decision, deciding which task, complexity or meaningfulness, they had performed on the object at encoding, again crossed with confidence on a 5-point scale (sure meaning, think meaning, unsure, think complex, or sure complex). The option of indicating “unsure” for the source decision was given to mitigate the effects of guessing on low-confidence responses. Both the old/new and source decisions were self-paced and trials were separated by 1 second of fixation. Objects given a “new”, response were followed by 1 additional second of fixation, to equate the length of new and old response trials, based on response times from pilot data. An example trial is shown in Fig. 7.

### EEG recording, data processing, and ERP analysis

The electroencephalogram was recorded with the Active Two active electrode system from Bio-Semi (www.Biosemi.com). Sixty-four Ag/AgCl electrodes were positioned in a nylon cap according to an extension of the international 10–20 system (Chatrian, et al., 1988). Five additional electrodes were positioned on the mastoids, to the outside of each eye, and under the left eye. The EEG and electroocculogram (EOG) were continuously recorded using a reference-free procedure, amplified between 0.16 and 100Hz, and sampled at a rate of 512Hz. Offline, the data were re-referenced to an average of the left and right mastoid recordings and low-pass filtered at 50Hz. Epochs were created by taking 200ms prior to the onset of the object stimuli to 1500ms after, and baseline corrected to the 200ms pre-stimulus interval. Epochs containing extreme artifacts (changes of more than 1000mv from baseline) were discarded. Next, EOG artifacts were removed using an automated correction procedure (Gratton, et al., 1983), and epochs that continued to show more than a 200mv change from baseline after the correction were discarded. ERPs were generated by selectively averaging the epochs time-locked to the onset of the object stimuli together for each condition of interest. All conditions analyzed were required to have a minimum of 15 artifact free trials; subjects who had less than 15 trials in any condition of interest were excluded from all analyses. The resulting trial counts in the Unrestricted analyses were as follows (mean, range): HCItem (150, 101–216) LCItem (56, 29–108) LCCR (79, 33–130) HCCR (61, 21–125) Source Hit (141, 76–202) Source Miss (65, 37–97); and for the Restricted analyses were: HC Source (79, 34–139); LC Source (37, 17–57); HC Item (35, 16–58); LC Item (31, 15–52). As can be seen from these numbers, the mean trial counts were at least 30 in all conditions of interest.

## Acknowledgments

The authors wish to thank Rick Lawley for assistance with data collection, and Joel Voss and Kyle Matthewson for assistance with aspects of the ERP data analysis.

## References

Azimian-Faridani N, Wilding EL. The influence of criterion shifts on electrophysiological correlates of recognition memory. Journal of Cognitive Neuroscience 2006;18(7):1075–86. [PubMed: 16839282]

Chatrian G, Lettich E, Nelson PL. Modified nomenclature for the “10%” electrode System1. Journal of Clinical Neurophysiology 1988;5(2):183–186. [PubMed: 3250964]

Curran T. Brain potentials of recollection and familiarity. Memory & Cognition 2000;28(6):923–938.

Curran T. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. Neuropsychologia 2004;42(8):1088–106. [PubMed: 15093148]

Dodson CS, Holland PW, Shimamura AP. On the recollection of specific- and partial-source information. Journal of Experimental Psychology: Learning, Memory, and Cognition 1998;24(5):1121–1136.

Donaldson W. The role of decision processes in remembering and knowing. Memory & Cognition 1996;24(4):523–533.

Dunn JC. Remember-know: A matter of confidence. Psychological Review 2004;111(2):524–542. [PubMed: 15065921]

Düzel E, Yonelinas AP, Mangun GR, Heinze H, 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(11):5973–5978. [PubMed: 9159185]

Gonsalves BD, Kahn I, Curran T, Norman KA, Wagner AD. Memory strength and repetition suppression: Multimodal imaging of medial temporal cortical contributions to recognition. Neuron 2005;47:751–761. [PubMed: 16129403]

Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Electroencephalography & Clinical Neurophysiology 1983;55(4):468–484. [PubMed: 6187540]

Hicks JL, Marsh RL, Ritschel L. The role of recollection and partial information in source monitoring. Journal of Experimental Psychology: Learning Memory and Cognition 2002;28(3):503–8.

Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language 1991;30(5):513–541.

Johnson MK, Hashtroudi S, Lindsay DS. Source monitoring. Psychological Bulletin 1993;114(1):3–28. [PubMed: 8346328]

Kirwan CB, Wixted JT, Squire LR. Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. Journal of Neuroscience 2008;28(42): 10541–10548. [PubMed: 18923030]

Leynes PA, Phillips MC. Event-related potential (ERP) evidence for varied recollection during source monitoring. Journal of Experimental Psychology: Learning, Memory, and Cognition 2008;34(4): 741–751.

Mandler G. Recognizing: The judgment of previous occurrence. Psychological Review 1980;87(3):252–271.

McCarthy G, Wood CC. Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. Electroencephalography and Clinical Neurophysiology 1985;62(3): 203–208. [PubMed: 2581760]

Mickes L, Wais PE, Wixted JT. Recollection is a continuous process: Implications for dual-process theories of recognition memory. Psychological Science 2009;20(4):509–515. [PubMed: 19320859]

Paller KA, Voss JL, Boehm SG. Validating neural correlates of familiarity. Trends in Cognitive Sciences 2007;11(6):243–250. [PubMed: 17475539]

Parks CM, Yonelinas AP. Moving beyond pure signal-detection models: Comment on wixted (2007). Psychological Review 2007;114(1):188–201. [PubMed: 17227187]

Rugg MD, Curran T. Event-related potentials and recognition memory. Trends in Cognitive Sciences 2007;11(6):251–257. [PubMed: 17481940]

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(6676):595–598. [PubMed: 9560154]

Slotnick SD, Dodson CS. Support for a continuous (single-process) model of recognition memory and source memory. Memory & Cognition 2005;33(1):151–170.

Van Petten C, Senkfor AJ. Memory for words and novel visual patterns: Repetition, recognition, and encoding effects in the event-related brain potential. Psychophysiology 1996;33(5):491–506. [PubMed: 8854736]

NIH-PA Author Manuscript

NIH-PA Author Manuscript


Vilberg KL, Moosavi RF, Rugg MD. The relationship between electrophysiological correlates of recollection and amount of information retrieved. Brain Research 2006;1122(2):161–170. [PubMed: 17027673]

Voss JL, Paller KA. Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. Learning & Memory 2007;14(4):259–267. [PubMed: 17412965]

Voss JL, Paller KA. An electrophysiological signature of unconscious recognition memory. Nature Neuroscience 2009;12(3):349–355.

Wais PE. fMRI signals associated with memory strength in the medial temporal lobes: A meta-analysis. Neuropsychologia 2008;46(14):3185–3196. [PubMed: 18817791]

Wilding EL. In what way does the parietal ERP old/new effect index recollection? International Journal of Psychophysiology 2000;35(1):81–87. [PubMed: 10683669]

Wilding EL, Rugg MD. An event-related potential study of recognition memory with and without retrieval of source. Brain: A Journal of Neurology 1996;119(3):889–905. [PubMed: 8673500]

Wixted JT. Dual-process theory and signal-detection theory of recognition memory. Psychological Review 2007a;114(1):152–176. [PubMed: 17227185]

Wixted JT. Spotlighting the probative findings: Reply to parks and yonelinas (2007). Psychological Review 2007b;114(1):203–209.

Wixted JT, Stretch V. In defense of the signal detection interpretation of remember/know judgments. Psychonomic Bulletin & Review 2004;11(4):616–641. [PubMed: 15581116]

Woodruff CC, Hayama HR, Rugg MD. Electrophysiological dissociation of the neural correlates of recollection and familiarity. Brain Research 2006;1100(1):125–135. [PubMed: 16774746]

Yonelinas AP. Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. Journal of Experimental Psychology: Learning, Memory, and Cognition 1994;20(6):1341–1354.

Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. Journal of Memory and Language 2002;46(3):441–517.

Yonelinas AP, Jacoby LL. Noncriterial recollection: Familiarity as automatic, irrelevant recollection. Consciousness and Cognition: An International Journal 1996;5(1–2):131–141.

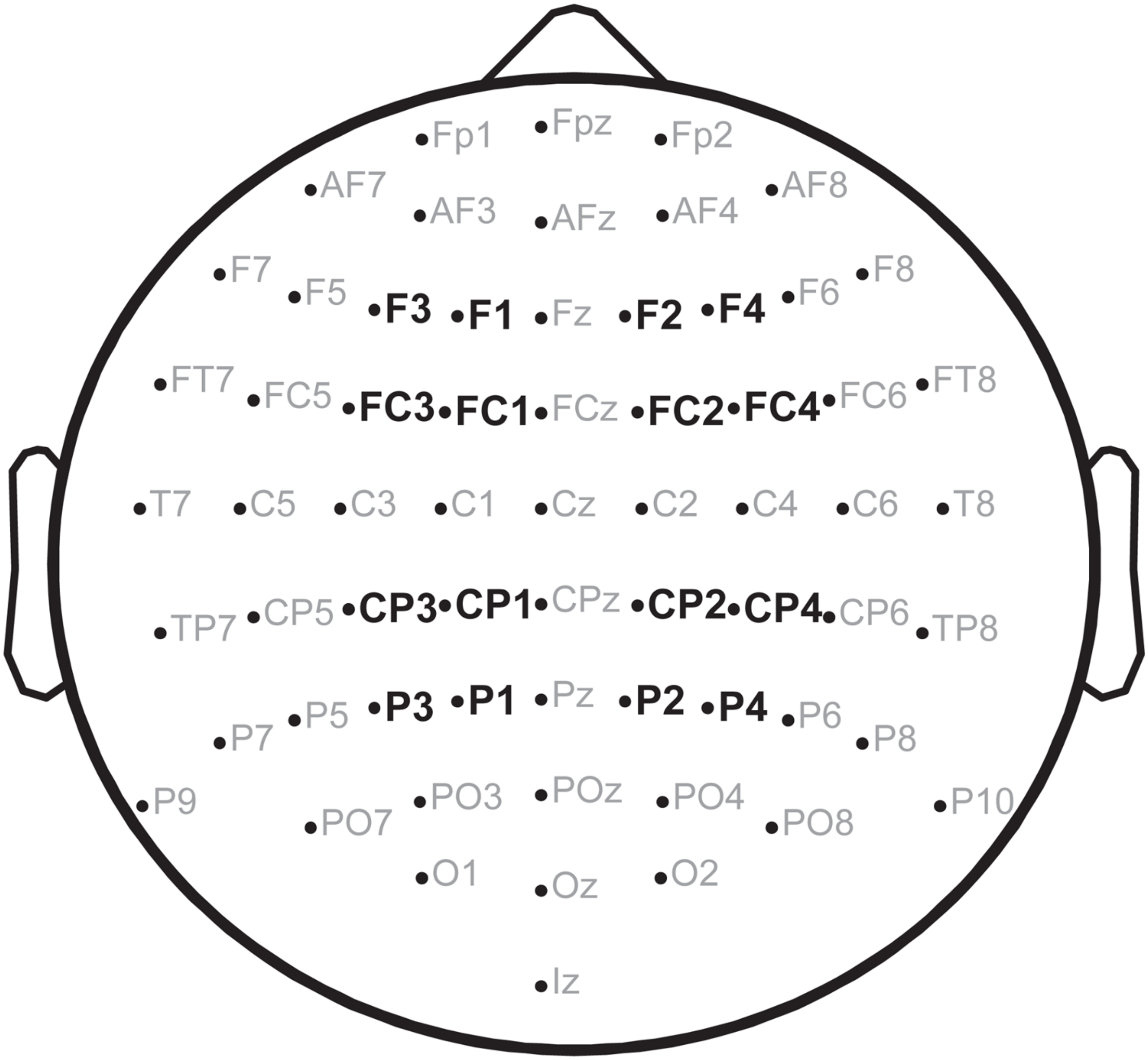


**Figure 1.**

Electrode montage, with the four clusters of electrode channels used for statistical analyses shown in bold.

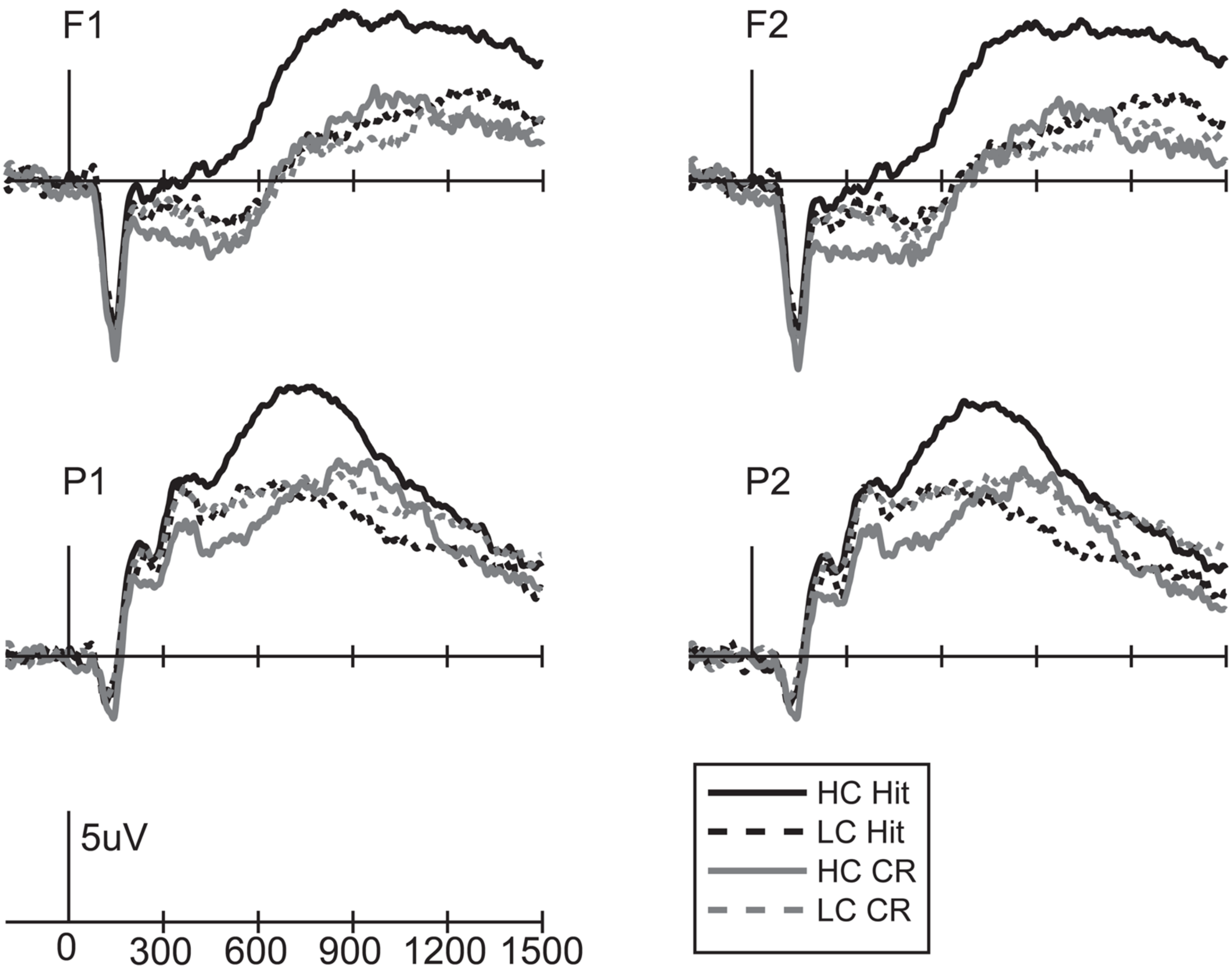


**Figure 2.**
The ERPs from four representative channels from each of the four electrode clusters, showing waveforms sorted by item (Old/New) confidence, regardless of the subsequent source memory decision.

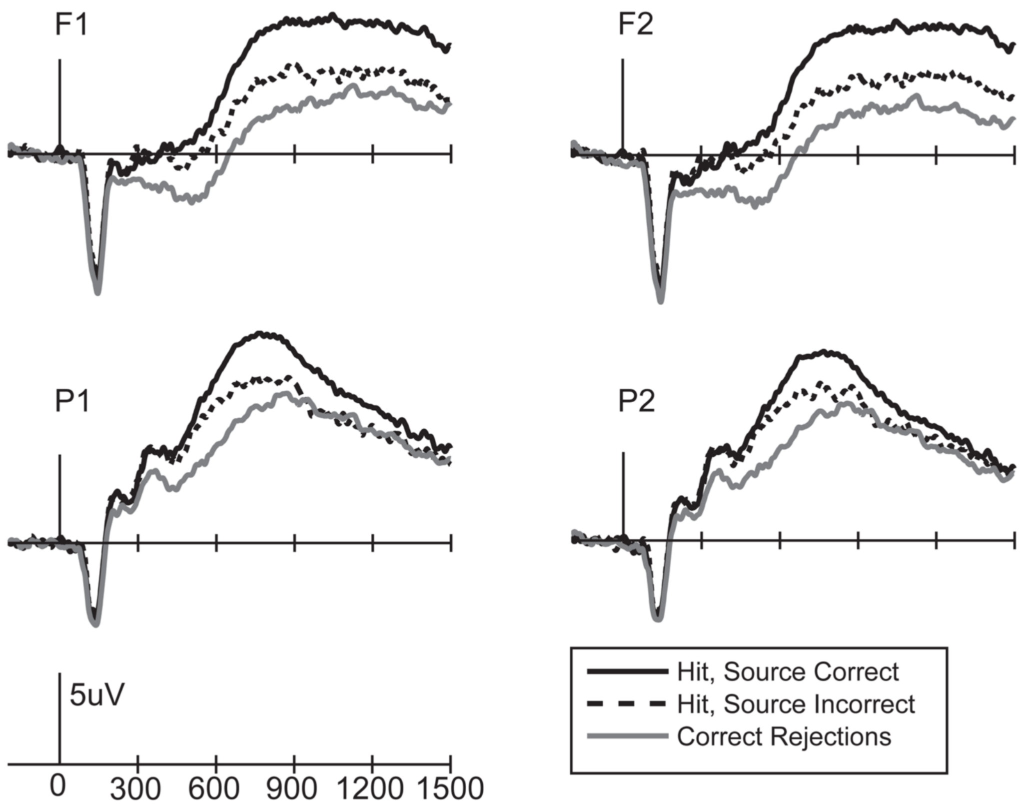


**Figure 3.**

The ERPs from correct item trials, separated by subsequent source memory accuracy, regardless of item confidence. The ERP from correct rejection trials is shown for comparison.

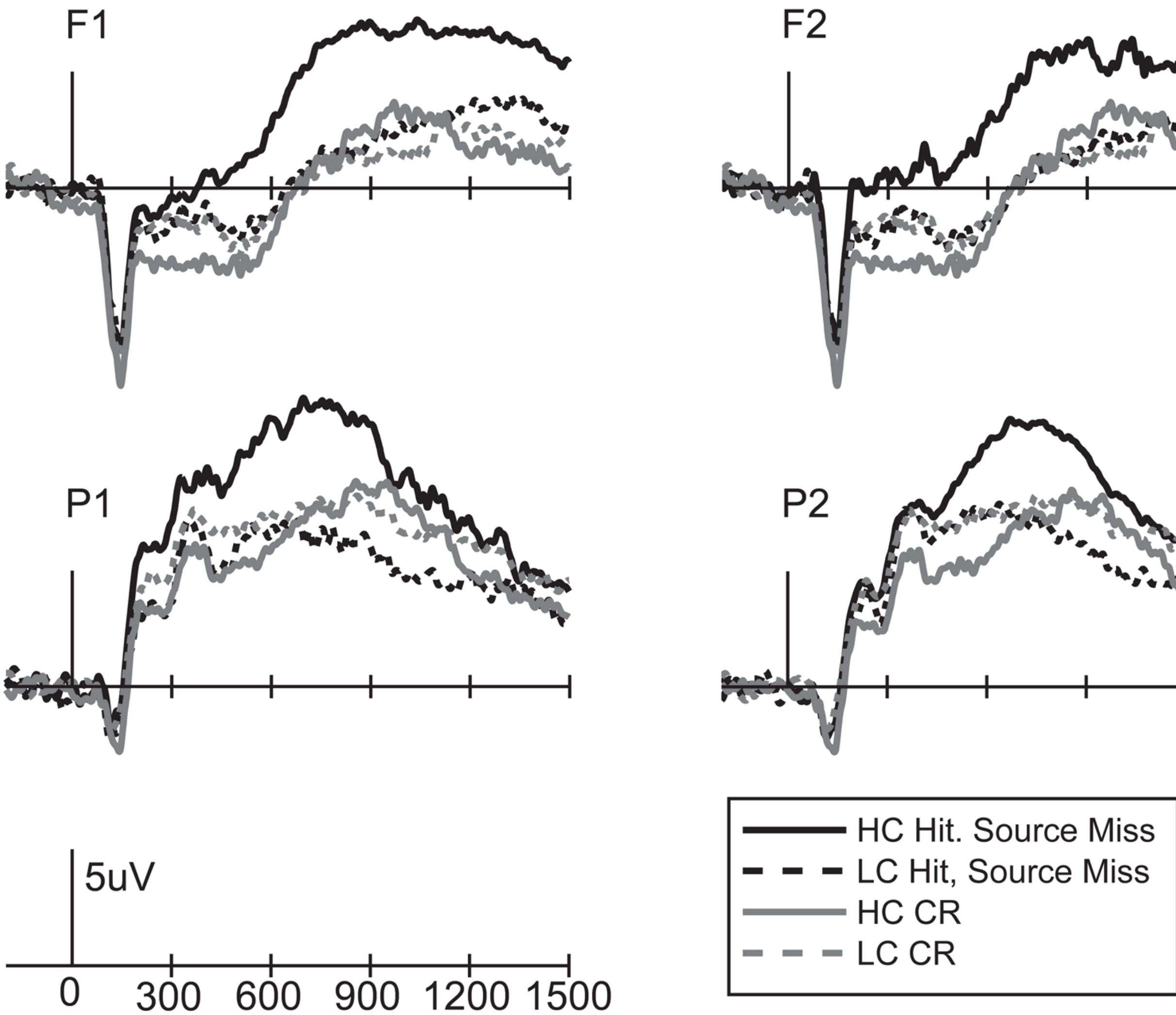


**Figure 4.**

The ERPs from trials separated by item confidence, given a subsequent incorrect source decision. The correct rejections are not conditionalized upon source memory, since no source decision was collected on trials when the participant responded "New".

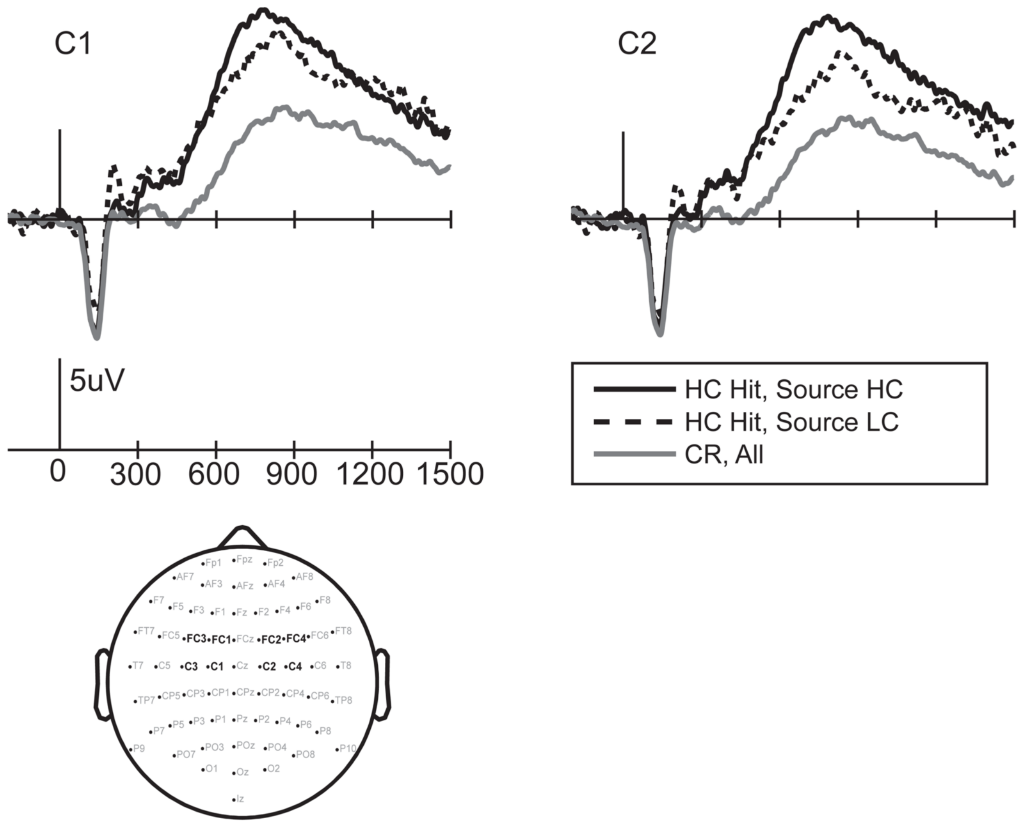


**Figure 5.**
The ERPs from trials sorted based on source confidence, with item memory held constant. These are trials on which the subject responded with high confidence to the item decision, separated by confidence in the subsequent source confidence decision. The ERP from correct rejection trials is shown for comparison. The scalp topography suggested a more central distribution, the electrode locations used in this analysis are bolded,

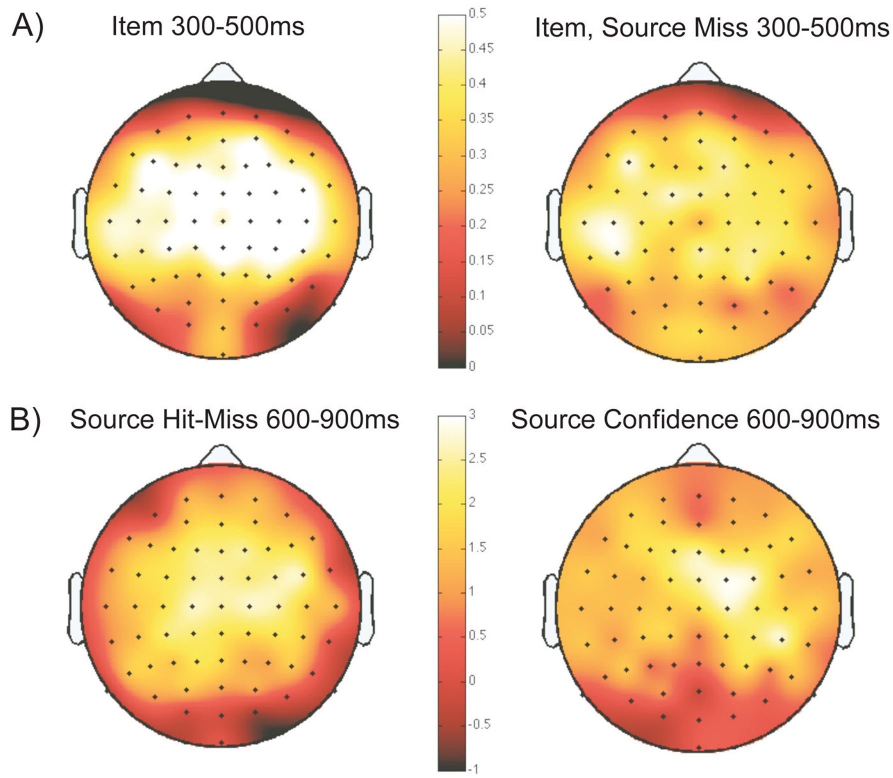


**Figure 6.**

Scalp topographies from both the traditional and restricted confidence analyses. A) Topographic maps created from the mean R values from a four-point linear regression of mean amplitudes from 300–500ms across confidence levels at each electrode site. Shown are these topographies for the unrestricted Item strength analysis (left) and for the Item effects restricted to source misses (right). B) Topographies of the mean difference between items remembered with and without source (left) and the difference between high and low confidence source trials restricted to HC item trials (right), from 600–900ms.

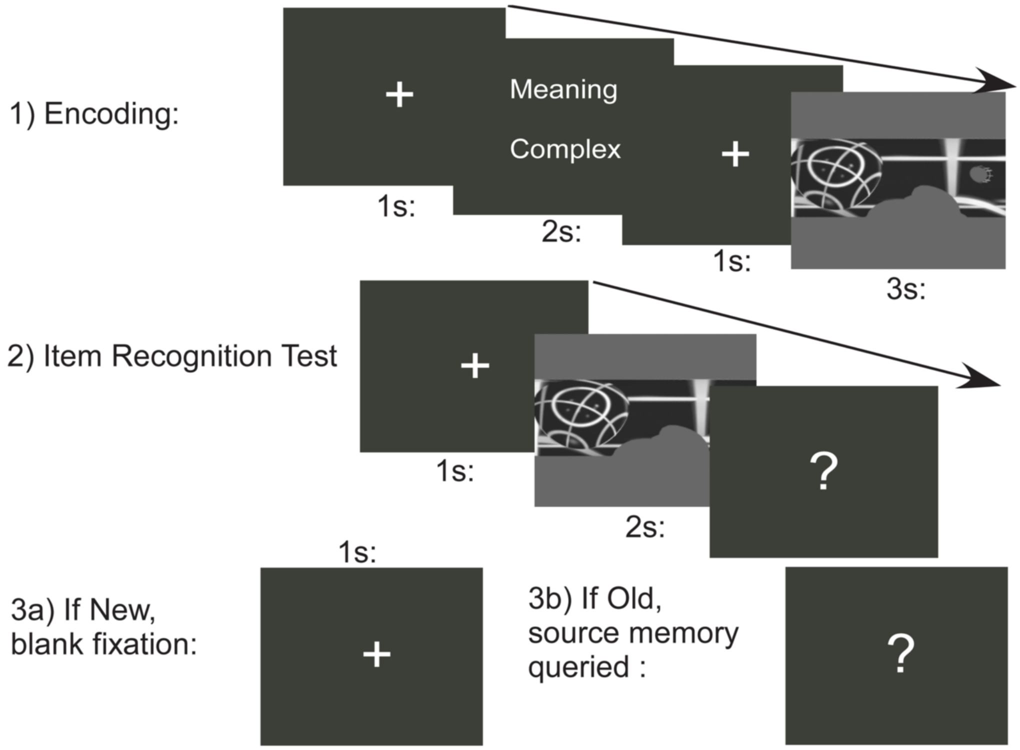


**Figure 7.**

A schematic of the experimental paradigm. 1) During the study phase, the word “meaning” or “complex” was shown prior to the object to indicate the encoding task to be performed on that object. 2) At test, item confidence was assessed with a delayed response on a 4-point confidence scale. 3a) If the item was judged “new”, a fixation cross was shown until the next trial. 3b) If the item was judged “old”, source confidence was assessed on a 5-point confidence scale.

## Table 1

Mean (SE) proportion of confidence responses to old/new item recognition query.

| Study Status | HC Old | LC Old | LC New | HC New |
|---|---|---|---|---|
| Old | 0.58(0.03) | 0.22(0.02) | 0.16(0.02) | 0.03(0.01) |
| New | 0.03(0.01) | 0.14(0.02) | 0.47(0.04) | 0.36(0.04) |

**Table 2**

Mean (SE) proportion of responses to source memory query, broken down by item confidence.

| Item Memory | Unsure | HC Correct | LC Correct | LC Miss | HC Miss |
|---|---|---|---|---|---|
| Overall | 0.14(0.02) | 0.39(0.03) | 0.29(0.02) | 0.15(0.02) | 0.04(0.01) |
| LC Item | 0.32(0.05) | 0.05(0.01) | 0.38(0.03) | 0.23(0.03) | 0.02(0.01) |
| HC Item | 0.07(0.02) | 0.51(0.03) | 0.26(0.02) | 0.12(0.01) | 0.05(0.01) |