



---

All Theses and Dissertations

---

2011-07-06

# The Neural Correlates of Retrospective Memory Monitoring: Convergent Findings from ERP and fMRI

Jeremy Clark Roper

*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Psychology Commons](#)

---

## BYU ScholarsArchive Citation

Roper, Jeremy Clark, "The Neural Correlates of Retrospective Memory Monitoring: Convergent Findings from ERP and fMRI" (2011). *All Theses and Dissertations*. 3052.  
<https://scholarsarchive.byu.edu/etd/3052>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

The Neural Correlates of Retrospective Memory Monitoring:  
Convergent Findings from ERP and fMRI

Jeremy Roper

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Brock Kirwan, Chair  
Michael Larson  
Erin Bigler

Department of Psychology  
Brigham Young University  
August 2011

Copyright © 2011 Jeremy C. Roper

All Rights Reserved

## ABSTRACT

### The Neural Correlates of Retrospective Memory Monitoring: Convergent Findings from ERP and fMRI

Jeremy C. Roper  
Department of Psychology, BYU  
Master of Science

Monitoring the accuracy of memory is an automatic but essential process of memory encoding and retrieval. Retrospective memory confidence judgments are making effective and efficient decisions based on one's memories. The neural processes involved in retrospective confidence ratings were investigated with EEG and fMRI using a recognition memory task designed such that participants also rated their confidence in their memory response. Correct trials (hits and correct rejections) were examined for differences related to the participants' level of confidence in their response. There were significant differences in electrophysiological activity (in the FN400 and the late parietal component) associated with confidence rating, with mean deflection increasing as confidence decreased. fMRI analysis revealed activity that appeared to be specific to the process of confidence rating. Activity was found to increase in the medial frontal, lateral frontal, and lateral parietal cortices as confidence decreases, but only for hits. In the lateral frontal, lateral parietal, and medial parietal cortices, activity decreased as confidence increased. These data indicate that there are neural mechanisms specifically related to making retrospective memory confidence judgments.

Keywords: memory, long term memory, metamemory, monitoring fMRI, EEG, ERP

## ACKNOWLEDGMENTS

Thanks are due to Brock Kirwan, advisor and mentor, for teaching me to think like a scientist; to Michael Larson and Erin Bigler for unflagging guidance and essential instruction; to research assistants who enabled me to not have to be everywhere at once; to the Squire lab at UCSD; and most importantly to Amy, for enduring the demands of my time as a graduate student.

## Table of Contents

List of Tables and Figures.....	v
Introduction.....	1
Metamemory and Confidence .....	1
Parietal Lobe Contributions to Memory.....	5
Frontal Lobe Contributions to Memory.....	7
Hypotheses .....	10
Experiment 1 – Confidence Rating Reaction Times.....	12
Method.....	12
Results .....	13
Discussion.....	14
Experiment 2 – Event-Related Potentials .....	15
Method.....	15
Results .....	17
Discussion.....	19
Experiment 3 – Functional MRI .....	21
Method.....	21
Results.....	24
Discussion.....	26
General Discussion .....	28
References.....	31
Tables and Figures .....	39

## List of Tables and Figures

<i>Table 1.</i> Size of overlapping regions and degree of overlap .....	39
<i>Figure 1.</i> Response proportions for Experiment 1.....	40
<i>Figure 2.</i> Mean reaction times for memory response and confidence response.....	41
<i>Figure 3.</i> Electrode clusters used in analyses .....	42
<i>Figure 4.</i> Mean distribution of responses for EEG participants .....	43
<i>Figure 5.</i> Mean peak amplitude, left and right FN400 .....	44
<i>Figure 6.</i> FN400 Waveforms for hits and correct rejections .....	45
<i>Figure 7.</i> Mean peak amplitude, left and right parietal effect .....	46
<i>Figure 8.</i> Late parietal component waveforms .....	47
<i>Figure 9.</i> Mean distribution of responses for fMRI participants .....	48
<i>Figure 10.</i> ROIs with stimulus type $\times$ confidence interaction pattern.....	49
<i>Figure 11.</i> Mean activation for ROIs in Figure 10 .....	50
<i>Figure 12.</i> ROIs with increased activation for decreased confidence .....	51
<i>Figure 13.</i> Mean activation for ROIs in Figure 12 .....	52

## **The Neural Correlates of Retrospective Memory Monitoring: Convergent Findings from ERP and fMRI**

Monitoring the accuracy of memory is an automatic but essential process of memory encoding and retrieval (Nelson & Narens, 1990). Confidence in the accuracy of memory can be assessed prior to storage of the memory, after storage of the memory, or after memory retrieval. Though the nature of the memory confidence differs between those stages, it is heavily based on strength and accuracy. Without the ability to make confidence judgments about accuracy, it would be difficult to effectively and efficiently complete even mundane, day-to-day tasks. Important as it is, the neural processes behind memory confidence are not well characterized.

### **Metamemory and Confidence**

Complex cognitive processes such as memory can be thought of as having two levels: the object-level and the meta-level. Within the domain of memory, object-level processes include the storage and retrieval of the memory trace while meta-level processes involve evaluating the ease and accuracy with which the object-level processes are accomplished. Monitoring is an aspect of metamemory in which the metacognitive level is able to continuously attend to the contents of memory (Nelson & Narens, 1990). Monitoring is a way of overseeing all incoming and outgoing memory information and makes contributions to both long-term and short-term memory (Mitchell, Johnson, Raye, & Greene, 2004). The model proposed by Nelson and Narens (1990) includes both prospective memory monitoring (regarding one's predicted ability to remember information in the future) and retrospective memory monitoring (regarding the confidence in the contents of previously-recalled information).

Prospective monitoring has three categories: judgments of learning (JOLs), feeling-of-knowing (FOK), and tip-of-the-tongue (TOT) states. Each presumes different memory processes, and occur at different stages in the encoding/retrieval timeline (Schwartz & Bacon, 2008). JOLs occur earliest of the three, as one is learning something and gauging how well it has been stored in memory. The magnitude of JOLs is commonly higher than the actual degree of learning (Koriat & Bjork, 2005), though the ratio is markedly decreased in older adults (Dunlosky, Kubat-Silman, & Hertzog, 2003; Emanuel Robinson, Hertzog, & Dunlosky, 2006). Accuracy of JOLs is not affected by distraction during the judgment, suggesting that it is separate from attentive processes (Barnes & Dougherty, 2007). Feeling-of-knowing is thought to occur during the retention of knowledge, and participants are generally asked to report it prior to retrieval (Nelson & Narens, 1990). It could be considered an early part of the retrieval process, possibly initiating memory retrieval, and is sometimes reported in combination with retrospective confidence judgments (Chua, Schacter, & Sperling, 2009b). FOKs are generally less accurate than recall performance (Modirrousta & Fellows, 2008), and accuracy decreases with both mild cognitive impairment and acute brain injury (Perrotin, Belleville, & Isingrini, 2007; Schmitter-Edgecombe & Anderson, 2007; Schnyer et al., 2004). The TOT state has been described as a FOK state (Brown, 2000), and is very similar to the concepts of familiarity or recognition in that one clearly has experience with a given stimulus, but is unable to explicitly state that experience (Cleary & Specker, 2007).

Nelson and Narens (1990) propose that retrospective monitoring occurs in two steps. First, at the object-level, the memory itself must be found and retrieved through a

search process. In the second step, at the meta-level, the level of confidence in the memory is judged once the sought-for information is found. High confidence in the accuracy of the memory indicates a high likelihood that the information retrieved is correct, and one's thoughts may move forward. Low confidence stimulates a deeper search for correct information, terminating either when more accurate information is found or when all search options have been exhausted.

While there is an abundance of prospective monitoring research, retrospective monitoring research is lacking. Work that addresses retrospective confidence judgments directly includes a recent series of studies by Chua and colleagues (Chua, Rand-Giovannetti, Schacter, Albert, & Sperling, 2004; Chua, Schacter, Rand-Giovannetti, & Sperling, 2006; Chua, Schacter, & Sperling, 2009a; Chua, et al., 2009b). Much of the research that addresses retrospective confidence judgments does so as a secondary purpose of the study, and thus the construct is given neither extensive attention nor analysis to yield a full understanding of the processes involved.

Some cognitive neuroscience findings support the object- and meta-level model that Nelson and Narens (1990) proposed. In fact, some neuroscientists have suggested that this model is more than just a theoretical approach, but representative of actual behavioral and neurocognitive processes (Schwartz & Bacon, 2008). Much of the current understanding of metacognition in the brain suggests both that metacognitive systems are separate from memory systems, and that there are discrete systems—in discrete brain regions—responsible for the different metacognitive processes. Several regions of the prefrontal cortex are associated with various aspects of metacognition (Shimamura, 2008).

Some recent research clearly dissociates object-level and meta-level processes in memory. Activity in parietal regions has been observed during memory monitoring (Chua, et al., 2006), consistent with neurologically-constrained models of metamemory (Shimamura, 2008). Patients with temporal lobe epilepsy, which is known to commonly affect memory performance, have shown clear deficits in episodic memory accuracy but not in metacognitive performance (Howard et al., 2010). Thus, while the temporal lobe is seen as critical to declarative memory, it appears not to be essential to metamemory. It has also been shown that, while activity in the medial temporal lobe (MTL) may vary with levels of retrospective memory confidence, the medial temporal lobe may perform as a separate system from other regions involved in meta-level processes (Chua, et al., 2006), again suggesting multiple levels of metacognitive processing.

One recent study focused on the interaction of processes related to memory confidence and memory veridicality in the MTL (Kirwan, Shrager, & Squire, 2009). In this study, functional MRI (fMRI) data were collected from participants as they performed a recognition memory task. Activity in the MTL distinguished between old and new stimuli regardless of behavioral response. Activity in the MTL also varied as a function of recognition confidence. These results are consistent with a role of the MTL in object-level memory processes. However, the results reported by these authors were limited in that the analyses reported were restricted to the data from the medial temporal lobe and did not include other brain regions that have a role in metacognition. Also, these authors did not investigate processes related to both the object- and meta-levels of retrospective memory monitoring.

## **Parietal Lobe Contributions to Memory**

The parietal lobe has also been shown to be involved in declarative memory. Parietal lobe activity is related to episodic memory retrieval (Cabeza, 2008; Wagner, Shannon, Kahn, & Buckner, 2005), spatial memory (Bonmassar et al., 2001; Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006), and successful memory retrieval (Donaldson, Wheeler, & Petersen, 2009; Iidaka, et al., 2006). This activity has been observed with fMRI (during retrieval of both visual and auditory memories) in regions throughout the parietal lobe. There is also an ERP component observed over the parietal lobe in the 500-800 ms post-stimulus window, associated with old/new distinction, which is affected differentially by the type of memory process utilized.

This late parietal component is commonly observed over the left hemisphere of the parietal lobe, but can also be observed over the right hemisphere (Rugg & Curran, 2007; Woodruff, Hayama, & Rugg, 2006). Over the left temporo-parietal area, the amplitude of this component is more positive for items correctly classified as old (hits) than for those items correctly classified as new (correct rejections). This effect is not present in incorrect responses, and so does not appear to be a result of general decision processes or whether the stimulus is novel or familiar. This suggests that the parietal old/new effect may be modulated by the accuracy of recognition memory (Allan & Rugg, 1997; Allan, Wilding, & Rugg, 1998).

The late parietal component also appears to be affected by memory monitoring. Studies that examine recognition confidence judgments have shown modulation of this component related to confidence for both correctly and incorrectly classified items (Curran, 2004; Johnson, Kreiter, Russo, & Zhu, 1998).

Functional MRI investigations of metamemory-related parietal lobe activity have also found activity suggesting that the parietal lobe does indeed play a strong role in memory monitoring. Lateral parietal cortex has greater activation during a metamemory task (making FOK and retrospective confidence judgments) than during recognition trials or baseline tasks (Chua, et al., 2009b). These data indicate that recognition processes do not significantly affect confidence processes. As memory confidence is a subset of metamemory, it is unclear whether the parietal activity observed was specific to confidence or rather the more general construct of metamemory. It is possible that this activity is related to general memory monitoring processes (Cabeza, 2008; Cabeza, Ciaramelli, Olson, & Moscovitch, 2008).

However, other findings indicate that there are regions of the parietal lobe that are involved in specific types of confidence rating. Left parietal regions have been shown to have increased fMRI activity during low-confidence ratings of correct “old” responses, as compared to high-confidence ratings of correct “old” responses (Henson, Rugg, Shallice, & Dolan, 2000). Other findings have shown that the left parietal cortex is involved in both high and low confidence conditions (Kim & Cabeza, 2007), with greater activation in the left ventrolateral posterior parietal cortex (PPC) for high confidence than for low confidence, and greater activation in the left dorsal PPC for low confidence recognition memory decisions than for high confidence recognition decisions. (Kim & Cabeza, 2009).

Bilateral parietal activity also has been observed during recognition confidence decisions. In a recognition memory task, bilateral dorsal parietal cortex was active during both high-confidence hits and low-confidence hits (Kim & Cabeza, 2009). This

activity was not seen during the correct rejection trials, and so it may not be so much a matter of confidence, but rather successful recognition. (It should be noted that confidence assessment in those conditions was not differentiated in the analysis.) However, bilateral parietal activation has been found when metamemory processes are specifically addressed. It was shown that the lateral and medial parietal regions were consistently more active for high-confidence correct trials than for low-confidence correct trials in a face-naming paradigm, although no distinction was made for incorrect trials (Chua, et al., 2006), making it clear that there is parietal involvement specifically during recognition confidence decisions.

Neuropsychological investigations with patients with bilateral parietal lesions also indicate a role of parietal cortex in memory confidence decisions. Across a series of three experiments, patients with bilateral parietal lesions performed no worse than controls at source recollection, but exhibited decreased confidence in their memory abilities (Simons, Peers, Mazuz, Berryhill, & Olson, 2010). Thus it appears that the parietal lobe plays an important role in feelings of memory confidence, though perhaps not in memory retrieval *per se*.

### **Frontal Lobe Contributions to Memory**

The frontal lobe, more particularly the prefrontal cortex (PFC) is another region that is involved in declarative memory. Activity in the PFC appears to be associated more with monitoring of memory than with encoding or retrieval, possibly because of the frontal lobe's role in executive control and decision-making. With the commonalities between metamemory and executive control (Nelson & Narens, 1990), it is reasonable to expect to observe memory monitoring activity in the frontal lobe (Shimamura, 2000).

The neural basis of memory monitoring in the frontal lobe has been the subject of a number of studies, with a wide variety of findings.

The FN400 (in the 300-500 ms window) is a well-established memory-related ERP component that also demonstrates an old/new effect. This mid-frontal component has a decrease in amplitude for old items compared to new (Rugg & Curran, 2007), and this effect has been found for items classified as old regardless of response accuracy (Mecklinger, 2006). Using a remember/know paradigm and two possible confidence response levels (high or low), the magnitude of the FN400 old/new effect has been found to vary according to confidence that an item is old, but such a relationship was not found for new items (Woodruff, et al., 2006). The FN400 appears to vary in direct relationship with the “oldness” of an item, as well as the confidence in the classification of items as old.

The higher spatial resolution of fMRI allow for better localization of memory confidence processes within the PFC. Consequently, fMRI is used more frequently for studying memory confidence processes in this area. Prefrontal regions are thought to be involved in memory monitoring processes that are a combination of memory judgment and continuous attention to accuracy of memory (Yonelinas, Otten, Shaw, & Rugg, 2005).

The lateralization of memory confidence processes suggested by some ERP data is confirmed by fMRI research. Activity in the left hemisphere has been associated with high-confidence recognition decisions, with greater activity in the left posterior PFC for higher-confidence ratings (Yonelinas, et al., 2005), and greater activity in the left inferior PFC for high-confidence responses, as opposed to low-confidence responses, regardless

of memory accuracy (Chua, et al., 2004). These results are also supported by findings from a transcranial magnetic stimulation (TMS) study, which found that disruption of the left ventrolateral PFC produced an overall decrease in memory confidence (Kahn et al., 2005).

The right PFC appears to be involved in low-confidence processes. The right dorsolateral PFC consistently shows increased activity when participants make low-confidence judgments as opposed to high-confidence judgments (Chua, et al., 2009b; Fleck, Daselaar, Dobbins, & Cabeza, 2006; Henson, et al., 2000). Higher levels of activity in the right PFC may be indicative of a heavier load placed on monitoring processes during low-confidence ratings. In making low-confidence judgment, there is likely more searching for memory accuracy cues than in high-confidence judgments, leading to more activity. Results of the TMS study show that right ventrolateral disruption leads to an increase in medium-confidence responses, presumably because the low-confidence system has been disrupted (Kahn, et al., 2005).

Bilateral dorsal PFC activity does not appear to be involved so much in memory processes, but rather in general monitoring and metacognitive processes. Research indicates that dorsal PFC activity is associated with the process of confidence rating and not with specific levels of confidence (Achim & Lepage, 2005; Henson, et al., 2000). Ongoing bilateral dorsolateral PFC activity during metamemory tasks supports the theory that bilateral dorsal PFC activity reflects general monitoring and metacognitive processes (Chua, et al., 2009b; Modirrousta & Fellows, 2008).

It is clear that the PFC plays an essential role in memory monitoring processes. Functional MRI studies have added specific localization of function beyond the

capability of ERP methods. Unlike many of the studies demonstrating parietal activity in memory monitoring, some of the studies reporting PFC activity attempt to specifically address the neural processes behind memory confidence rating. Comparing fMRI results with ERP results is still problematic, however, because of the difference in memory tasks used and the problem of source localization inherent in ERP. One way to address this challenge that has not been done previously in metamemory research is to use the same memory task during both ERP and fMRI measurement, thereby allowing direct comparison of the data from each.

### **Hypotheses**

Functional MRI and ERP methods have both served to greatly elucidate neural processes in memory confidence, and will continue to play an important role in memory research. Data from ERP studies do not provide clear information about the source of specific event-related activity, due to the nature of the activity measured with EEG. More specific localization is available with fMRI, though it involves constraints on temporal resolution. The involvement of prefrontal and parietal cortices in memory monitoring and confidence has been established by recent research. However, it is possible that neural activity in these processes moves more rapidly than fMRI can detect on its own, but pinpointing the location of that activity is difficult with ERP alone. The proposed hypotheses regarding these questions are:

**Hypothesis 1.** If the confidence response time is significantly shorter than the memory response time, it will indicate that the confidence and memory decisions do not occur sequentially, but rather simultaneously. If they occur simultaneously, then a six-

point scale can be used (old-new decision with high, medium, and low confidence) that will facilitate an event-related experimental design.

**Hypothesis 2.** Two ERP components, the FN400 and the late parietal old/new effect, will show different responses depending on the level of the participant's confidence in their response. The FN400 magnitude will be mediated by the participant's level of confidence, with increasingly confident responses resulting in an attenuated FN400 waveform. (The effect of response confidence is expected in both old and new words, though the FN400 is also affected by whether a word is old or new.) The amplitude of the late parietal component will increase with the participants' level of confidence, regardless of whether a word is endorsed as old or new.

**Hypothesis 3.** Bilateral parietal fMRI activity is expected. Specifically, left parietal cortical activity proportional to the level of confidence memory responses is predicted for both old and new words, regardless of response accuracy. During this task, the left PFC is also expected to show activity associated with high-confidence responses, and the right PFC is expected to show activity associated with low-confidence responses, the level of activity being mediated by the level of confidence, though not by response accuracy.

## Experiment 1 – Confidence Rating Reaction Times

### Method

**Participants and Materials.** Twenty-one student volunteers were recruited from Brigham Young University to participate in Experiment 1. Participants gave written informed consent before participating and received course credit as compensation for their participation.

The stimuli used were 720 nouns with a mean KF frequency of 27 (range 1-198) and concreteness ratings  $> 500$  (mean = 573) obtained from the MRC Psycholinguistic Database (Wilson, 1988). Half the words were assigned to the study list and half the words were assigned to the foil list for the recognition memory test. The assignment of words to the study and foil conditions was randomized across participants.

**Procedures.** In the study portion, participants were asked to make a pleasant/unpleasant judgment for 360 words (presentation time 2.5 seconds per word) by pressing one of two marked keys on a computer keyboard. Participants were instructed to pay close attention to the words, as their memory for the words would be tested later. The study session was divided into six blocks of 60 trials, and participants were allowed a short break between each block. The delay between study and test was approximately three minutes. A recognition memory test was given, where words were presented one at a time and participants were asked to make a judgment for each word as to whether they had seen the word at study or not. Following each memory response, they made a confidence judgment using one of three possible responses: *maybe* correct, *probably* correct, or *definitely* correct. The task was self-paced, to allow for examination of reaction times for both judgments. Nine blocks of 80 trials were completed.

**Analysis.** For both the memory responses and the confidence ratings, trials were separated into twelve categories (accuracy [2] × old-new response [2] × confidence rating [3]). Mean reaction times were then calculated for each category. A paired-samples t-test was conducted comparing grand mean reaction times for the old-new response and confidence rating. A 2 (stimulus type) × 2 (memory response) × 3 (confidence rating) repeated measures ANOVA was then conducted for reaction times for the memory response and the confidence ratings separately.

## Results

**Behavioral results.** Figure 1 shows response proportions. Mean percent correct (mean ± SD) was  $79.7 \pm 12.7\%$ , and  $d'$  was  $1.99 \pm 0.89$ .

**Reaction times.** Figure 2 shows mean reaction times for both memory responses and confidence ratings. The paired-samples t-test comparing grand mean reaction times for memory response and confidence ratings revealed a significant difference between the two reaction times ( $t[20] = -15.044, p < .001$ ). In the analysis of memory response reaction times, there was a significant effect of confidence level ( $F[2, 19] = 21.33, p < .001$ ), but no significant effects of memory response (old/new) ( $F[1, 20] = 1.28, p = .272$ ) or accuracy ( $F[1, 20] = 0.25, p = .621$ ). There was also a significant linear trend for confidence level ( $F[1, 20] = 40.12, p < .001$ ), with reaction time increasing as confidence decreases. For the confidence response reaction times, there was a significant effect of confidence level ( $F[2, 19] = 21.24, p < .001$ ), a nearly-significant effect of memory response ( $F[1, 20] = 4.30, p = .051$ ), and no significant effect of accuracy ( $F[1, 20] = 0.33, p = .572$ ). A significant linear trend for confidence level was present ( $F[1, 20] = 36.37, p < .001$ ), again with reaction time increasing as confidence decreases.

## Discussion

One fundamental assumption underlying this paradigm is that the memory and confidence judgments occur, roughly, simultaneously. There is a possibility that the judgments and the processes that drive them are temporally distinct. Before extensive experimentation, it was necessary to examine whether the planned paradigm was conceptually sound. To test this, a self-paced recognition memory task was used, where participants were asked to give a memory response and then a confidence rating for each word. It was hypothesized that if the confidence response time was significantly shorter than the memory response time, it would indicate that the decisions do not occur so far apart as to confound the task used for the first two experiments. Results indicated that the type of response (memory or confidence) had a very strong effect on the reaction time. Figure 2 provides a clear illustration of the differences in reaction time between the two response types, showing that the reaction time for the memory responses is consistently longer than for the confidence ratings. These results do seem to confirm the theory regarding the robustness of the planned experimental paradigm.

A consistent trend of decreased reaction time for high confidence trials as opposed to low confidence trials was also observed. This supports the idea that low-confidence responses have a heavier load placed on them, so not only do some regions show increased activity for low-confidence responses, but participants appear to spend more time making low-confidence decisions—both for the memory response and the confidence rating.

## Experiment 2 – Event-Related Potentials

### Method

**Participants and Materials.** Thirty-three healthy, right-handed student volunteers were recruited from Brigham Young University to participate in Experiment 1. Participants gave written informed consent before participation and received either course credit or \$10 for their participation. Participants were free from head injury, neurological insult and major psychiatric disorders. Stimuli used were the same as in Experiment 1.

**Procedures.** The study portion was the same as in Experiment 1. Participants completed the study portion in a separate room before being fitted with the EEG sensor cap, which took approximately 10-15 minutes.

In the test portion, participants were given a recognition memory test for the 360 target words as well as the 360 foil words (word order randomized), divided into nine blocks of 80 trials. Participants were again allowed a short break between each block. Each word was presented for 3.4 seconds, with a 100 ms interstimulus interval. Participants made an old/new recognition judgment using a six-point confidence scale (1 = "sure new"; 2 = "probably new"; 3 = "guess new"; 4 = "guess old"; 5 = "probably old"; and 6 = "sure old"). Participants were encouraged to use the entire scale. A brief practice session was presented before the test to ensure that participants understood the task and the response scale.

**Electrophysiological data recording and processing.** The electroencephalogram (EEG) was recorded from 128 scalp sites using a HydroCel Geodesic Sensor Net™ and an Electrical Geodesics Inc. (EGI; Eugene, Oregon, USA) amplification system (amplification 20K, nominal bandpass 0.10-100 Hz). The EEG was

referenced to the vertex electrode and digitized at 250 Hz. Impedances were maintained below 50 k $\Omega$ . EEG data were processed off-line beginning with a 0.1 Hz first-order highpass filter and 30 Hz lowpass filter. ERPs were segmented based on trial type criteria (specified below). Eye movement and blinking artifacts in EEG data were removed using the algorithm suggested by Gratton, Coles, and Donchin (1983).

Stimulus-locked ERP averages were derived spanning 200 ms pre-stimulus to 1000 ms post-stimulus, and were re-referenced to the average reference. The baseline used was the 200 ms pre-stimulus period, as suggested by Luck (2005). Stimuli were presented in an electrically-shielded testing room, on a 17 inch LCD computer monitor, and responses recorded with a computer keyboard. Electrode clusters of interest were based off of Curran (2004), who used a similar 128-channel recording system to observe the FN400 and late parietal components (See Figure 3 for a map of electrode clusters).

**Analysis.** ERPs were derived for correct-response trials only. Hits (correct responses to old stimuli) were segmented into those with high, medium, and low response confidence. Similarly, correct rejections (correct responses to new stimuli) were also segmented according to confidence. The FN400 peak amplitudes were extracted as the average of 8 ms pre- and post-peak negative amplitude within the 300-500 ms post-stimulus window. Late parietal peak amplitudes were extracted using the same method, using the peak positive amplitude within the 500-800 ms post-stimulus window. Latencies were calculated as the time of the peak amplitude within the specified windows. Each hemisphere of each component was then analyzed separately to examine effects specific to each cluster, as has been done with previous studies (Curran, 2004; Mecklinger, 2006; Rugg & Curran, 2007; Woodruff, et al., 2006), using a 2 (stimulus

type; hits, correct rejections)  $\times$  3 (confidence level; low, medium, high) repeated measures ANOVA for each cluster.

## Results

**Behavioral results.** Figure 4 shows response proportions for both old and new items. Overall, participants demonstrated good memory for the words. Mean percent correct was  $78.6 \pm 8.1\%$ , and  $d'$  was  $1.78 \pm 0.53$ .

**FN400 ERP results.** The FN400 was analyzed over two frontal, dorsal groups of seven electrodes each centered near the standard F3 and F4 locations (see Figure 3). Consistent with the hypothesis, the ANOVA revealed a significant main effect of confidence level for both the left ( $F[2, 31] = 4.43, p < .05$ ), and the right clusters ( $F[2, 31] = 3.49, p < .05$ ). The main effect of stimulus type (old vs. new) was not significant in either hemisphere ( $F$ 's  $< 1, p$ 's  $> 0.10$ ); whether the word was previously studied had no significant effect in itself on the amplitude of the signal. This is somewhat unexpected, as the FN400 was initially conceived as a component that demonstrated an old-new effect. A significant interaction between confidence level and stimulus type was present in both clusters (Left:  $F[2, 31] = 3.55, p < .05$ ; Right:  $F[2, 31] = 3.89, p < .05$ ). Figure 5 illustrates the mean deflection associated with each confidence level of both stimulus types, and Figure 6 shows the observed FN400 waveforms. The left cluster showed a significant linear (decreasing activity as confidence increases) trend for confidence regardless of stimulus type ( $F[1, 32] = 6.21, p < .05$ ); for hits only, the linear trend for confidence in this cluster was also significant ( $F[1, 32] = 5.67, p < .05$ ). There was no significant linear trend for confidence on the right cluster ( $F[1, 32] = 1.22, p > .05$ ); there was, however, a significant linear trend when only hits were examined (again,

decreasing activity as confidence increases) ( $F[1, 32] = 9.85, p < .005$ ). Latency effects failed to reach significance ( $p > .05$ ), and were minimal relative to the window size (approximately 1-3 ms differences, within a window of 200 ms).

**Late parietal ERP results.** The late parietal component was analyzed over two dorsal groups of seven electrodes each over the parietal lobe, slightly anterior to the standard P3 and P4 locations (see Figure 3). The ANOVA revealed a significant main effect of response confidence level for both the left ( $F[2, 31] = 5.03, p < .05$ ) and the right clusters ( $F[2, 31] = 4.54, p < .05$ ). There was a significant effect of stimulus type ( $F[1, 32] = 7.95, p < .01$ ) in the right parietal cluster. Contrary to predictions, peak amplitude *increased* with decreasing confidence, though only for hits. The linear trend for hits was statistically significant on the left side ( $F[1] = 9.57, p < .005$ ), but only approaching significance on the right side ( $F[1] = 3.23, p = 0.082$ ). A significant interaction between confidence level and stimulus type was present for the left side ( $F[2, 31] = 4.37, p < .05$ ), but not the right ( $F[2, 31] = 2.25, p > .05$ ). Neither cluster showed significant linear trends for confidence (when collapsed over stimulus type), possibly because an effect present in one stimulus type could have been canceled out by an effect in the other stimulus type. Figure 7 illustrates the mean deflection associated with each confidence level of both stimulus types for the parietal component, and Figure 8 shows the observed parietal waveforms. Again, latency effects did not reach significance ( $p > .05$ ) and were minimal relative to the window size (approximately 1-3 ms differences, within a window of 300 ms).

## Discussion

A word recognition memory task was used while recording scalp electrical activity with EEG to assess participants' level of confidence in their responses in an attempt to more clearly understand the neural processes of retrospective memory confidence.

For both the FN400 and late parietal components, an effect of confidence on peak amplitude was observed, with the greatest mean peak amplitude for low confidence responses and decreasing mean peak amplitude with increasing confidence. This effect was only apparent for previously studied words. These results were only partially consistent with the predictions, which were that in the FN400, the mean deflection would be decreased for low confidence responses as opposed to high confidence responses and mediated by stimulus type, and the late parietal component would show increased mean deflection with increasing response confidence, regardless of stimulus type.

The FN400 characteristically distinguishes between old and new stimuli, regardless of accuracy, and was thus a useful component to gauge effects of metamemory processes (Mecklinger, 2006; Rugg & Curran, 2007). Consistent with previous findings, there was a relationship between confidence level and the amplitude of the FN400. The significant stimulus type  $\times$  confidence interaction we found in the left cluster also confirmed past findings that the amplitude of the FN400 varies according to confidence, but only for previously studied items (Woodruff, et al., 2006). It was also hypothesized that the late parietal component would show confidence effects on amplitude that occur regardless of whether the word was previously studied, but instead a pattern of a stimulus

type  $\times$  confidence interaction was observed (significant in the left cluster but not the right), similar to that seen in the FN400 activity.

Previous studies using the FN400 to investigate memory confidence have used a two-point confidence scale. A three-point confidence scale was implemented with the goal of understanding recognition confidence processes more fully. Through the use of a three-point scale, it was possible to examine the patterns of activity related to recognition more closely than with a two-point scale. For example, in Figure 5, inspecting the mean amplitude for old words shows that there is a sharper increase in amplitude between high and medium confidence than between medium and low confidence. This may indicate greater similarity between the cognitive processes behind the medium- and low-confidence responses, which is particularly evident in the left FN400 CRs waveform in Figure 6.

Activity mediated by confidence level was expected, and the activity pattern observed in the FN400 was consistent with expectations. Response confidence and mean peak amplitude were inversely related—high confidence is associated with low amplitude, and low confidence is associated with high amplitude. An opposite pattern in the late parietal component was predicted, but activity similar to that in the FN400 was observed. This may indicate that with low confidence responses, the memory may be weaker and therefore more searching is occurring, whereas for high confidence responses, the memory may be stronger and thus less searching for confidence cues.

The activity observed in the FN400 was not consistent with its past conception as a component sensitive to old/new stimuli effects (Curran, 2004; Mecklinger, 2006; Rugg & Curran, 2007). In both clusters of the FN400 there was not a significant main effect

for stimulus type. This difference may stem from a difference in experimental design; most studies that define the FN400 as a component showing an old/new effect use a remember/know paradigm. Perhaps the design used (a recognition memory judgment in combination with a confidence judgment) is responsible. Either way, experimenting to clarify the extent and limits of the FN400 old/new effect would be helpful for future studies.

Next, we were interested in more accurately localizing the elements of post-recognition confidence processes in the brain. Thus, a third experiment using fMRI and a slightly modified version of the same task was carried out. A subset of the fMRI data focusing on the effects in the MTL was reported previously (Kirwan et al., 2009).

### **Experiment 3 – Functional MRI**

#### **Method**

**Participants and Materials.** Thirteen right-handed volunteers were recruited for the fMRI experiment from the University of California at San Diego, and gave written consent before participation. Materials used were the same as those in Experiments 1 and 2.

**Procedures.** The overall task design was similar to that of Experiment 2. Participants first completed a study portion prior to being placed in the MRI scanner for the recognition memory test. The delay between study and test was approximately 15 minutes. Words in the recognition memory test were again presented for 3.5 seconds and participants made a confidence judgment using a 6-point scale. The recognition memory test was divided into nine blocks. Odd/even digit task trials were intermixed in the test blocks to serve as a baseline for estimating the hemodynamic response. In the digit task,

participants saw a digit (1-8), and indicated whether the digit was odd or even by button press. Each scan run began and ended with at least 12 digit trials; digit trials occurred in groups of 2, 4, or 6. Presentation time for each digit trial was 1.75 seconds, and the mean inter-trial interval was 5.1 seconds (range 0-10.5 seconds) (Kirwan, Shrager, & Squire, 2009).

**Functional imaging data recording and processing.** Imaging was conducted at the University of California at San Diego Center for Functional MRI, using a 3T GE scanner. Functional images were acquired using a gradient-echo, echo-planar, T2\*-weighted pulse sequence (TR = 1750 ms; 264 TRs/run; TE = 30ms; flip angle 90 degrees, matrix size = 64 x 64; field of view 22 cm). To allow for T1 calibration, the first five TRs were discarded. Twenty-nine oblique coronal slices (slice thickness = 5 mm) were acquired perpendicular to the long axis of the hippocampus and covering the whole brain (voxel volume = 3.44 x 3.44 x 5 mm). High-resolution structural images were acquired using a T1-weighted IR-SPGR pulse sequence (24 cm field of view; flip angle 10 degrees; TE = 3.7 ms; 166 slices; 1.4 mm slice thickness; matrix size = 256 x 256), following the nine functional runs.

Functional scans were coregistered with three-dimensional whole-brain anatomical data. Data were also slice-time corrected and corrected for the effects of minor head movements. Data during major head movements was excluded from the analysis, in a manner without affecting the temporal aspect of the rest of the data. Spatial normalization was achieved using Advanced Normalization Tools (ANTs; Avants et al., 2008; Klein et al., 2009; Lacy, Yassa, Stark, Muftuler, & Stark, 2011; Yassa et al., 2010).

Functional data were categorized according to the stimulus type (old or new) and behavioral response (old or new with three levels of confidence) for a total of 12 vectors of interest. Due to low response rates, some of the incorrect response types were combined to achieve adequate numbers of trials to estimate the hemodynamic response. Trials in which participants failed to make a response were included in the fMRI model but excluded from further analysis. Behavioral vectors and vectors that coded for motion and rotation were used in a deconvolution analysis (Ward, 2001). Deconvolution is based on multiple linear regression and estimates the shape of the hemodynamic response from the data itself. The resultant fit coefficients (beta-coefficients) represent activity versus baseline in each voxel for a given time point and for each of the stimulus types. This activity was summed over the expected hemodynamic response (0-15.75 s after trial onset) and taken as the estimate of the response to each stimulus type (relative to the digit task baseline). Following the analysis of the ERP data, only correct trials were included in the overall analysis.

Following individual deconvolution analysis, individual subject parameter estimate maps were entered into group-level analyses and thresholded at a voxel-wise p-value of  $p=0.03$ . A cluster correction technique was used to correct for multiple comparisons, and Monte Carlo simulations were used to determine how large a cluster of voxels was needed to be statistically meaningful ( $p < 0.05$ ; minimum cluster extent of 104 voxels) (Forman et al., 1995; Xiong, Gao, Lancaster, & Fox, 1995).

## Results.

**Behavioral results.** Figure 9 shows response proportions for both old and new words. Mean percent correct was (mean  $\pm$  standard deviation)  $78.1 \pm 8.6\%$  and  $d'$  was  $1.69 \pm 0.61$ .

**Functional imaging results.** The first analysis sought to examine the interaction of memory confidence and stimulus type. A regression analysis was conducted that sought regions where activity decreased for hits with increasing confidence and constant activity for CRs (weights=3, 2, 1 for low, medium, and high confidence hits and 1, 1, 1 for low, medium, and high confidence CRs). The regression analysis revealed three regions: a right medial frontal cluster, a right lateral frontal cluster, and a right lateral parietal cluster (Figure 10 depicts these three regions). A 2 (stimulus type)  $\times$  3 (response confidence level) repeated measures ANOVA was used to examine the activity in each region. See Figure 11 for an overview of the patterns of activation within each region.

In the right medial frontal region, the ANOVA revealed a significant main effect of stimulus type ( $F[1, 12] = 5.71, p < .05$ ); the main effect of confidence approached significance ( $F[2, 11] = 3.66, p = .060$ ). There was no significant linear trend for confidence collapsed across old and new stimuli ( $F[1, 12] = 1.99, p = .184$ ), or when hits were considered alone ( $F[1, 12] = 1.81, p = .203$ ). In the right lateral frontal region, a similar pattern of effects was evident. There was a significant effect of stimulus type ( $F[1, 12] = 6.08, p < .05$ ), but no significant effect of confidence level ( $F[2, 11] = 2.24, p = .152$ ). As with the medial frontal region, there was not a significant linear trend for confidence ( $F[1, 12] = 4.34, p = .059$ ); there was, however, a significant linear trend for confidence (increasing activity as confidence decreases) when considering hits alone

( $F[1, 12] = 5.40, p < .05$ ). In the right lateral parietal cluster, however, there were statistically significant main effects of both stimulus type ( $F[1, 12] = 5.46, p < .05$ ) and confidence level ( $F[2, 11] = 4.06, p < .05$ ). There was also a significant linear trend for confidence level ( $F[1] = 7.89, p < .05$ ), again with increasing activity as confidence decreases.

The next analysis examined regions where activity increased linearly as confidence decreased, collapsed across old and new stimuli. Three such regions were found, two of which overlapped with two of the regions found in the previous analysis. See Table 1 for the size (in voxels) of each region and the degree of overlap. Figure 12 depicts the location of the regions, and Figure 13 illustrates mean activation for each category.

In the right lateral frontal region, there was a significant main effect of stimulus type ( $F[1, 12] = 4.97, p < .05$ ) and for confidence level ( $F[2, 11] = 4.12, p < .05$ ), and a significant linear trend for confidence ( $F[1] = 8.97, p < .05$ ) (increasing activity as confidence decreases). The ANOVA revealed no significant main effect of either stimulus type ( $F[1, 12] = .01, p = .93$ ) or confidence level ( $F[2, 11] = 3.38, p = .072$ ) in the right medial parietal region, and there was no significant linear trend for confidence ( $F[1] = 4.20, p = .063$ ). Finally, in the right lateral parietal region there was no significant main effect of stimulus type ( $F[1, 12] = .049, p = .829$ ), but there was a strong main effect of confidence ( $F[2, 11] = 26.91, p < .001$ ) and linear trend for confidence ( $F[1] = 54.81, p < .001$ ) (again, increasing activity as confidence decreases).

## Discussion

Seeking regions of the brain that may be related to the activity observed in Experiment 2, a second experiment was conducted, using a version of the task adapted for fMRI. Two analyses were conducted—first, looking for regions associated with the pattern of activation observed in the ERP results (a confidence level  $\times$  stimulus type interaction) and, second, looking for regions associated with a trend of decreasing activation with increasing confidence, regardless of stimulus type. Two discrete brain regions, all in the right hemisphere, were found to be associated with both patterns of activation (lateral frontal and lateral parietal). Of these two regions, the lateral frontal region showed significant effects of stimulus type and confidence, and both regions exhibited strong linear trends for confidence. In the first analysis a medial frontal region was identified, in which there was not a significant effect of confidence or stimulus type and no significant linear trend for confidence, but there was a significant linear trend when hits alone were examined. High-confidence response activity was not observed in the left frontal cortex, but there was increased activity in the frontal cortex, mediated by confidence level, as predicted based on previous findings. There was not significant bilateral activity in the parietal cortex. However, the observed pattern of results was consistent with predictions about the right parietal cortex, which was expected to show activity mediated by confidence level.

The activity observed in the right lateral frontal cortex—increased activity with low confidence compared to high confidence—is consistent with findings of several previous studies (Chua, et al., 2009b; Fleck, et al., 2006; Henson, et al., 2000), confirming that low-confidence judgments likely place heavier strain on the monitoring

processes occurring in the region. One study observed activity in the same medial frontal region found in the present study, with greater activity for metamemory processes than for non-metamemory processes (Chua, et al., 2009b). The analyses used in this study allowed for further clarification of the nature of this activity: not only is activity greater for metamemory processes, but it is also specifically associated with a pattern of increasing activity for decreasing confidence hits, for hits only.

Several fMRI studies of memory confidence have suggested a strong lateralization in the frontal cortex. Much of the activity previously observed to be associated with confidence processes has been in the left hemisphere. Little research has seen activity in the right frontal cortex. Activity was found that is specifically associated with retrospective confidence judgments in the right medial frontal cortex, and the right lateral frontal cortex. It is possible that right frontal activity has not been observed as frequently because research has not focused as directly on confidence processes.

Lateral parietal activity has also been shown to be associated with metamemory responses more than non-metamemory processes (Chua, et al., 2009b). As with the frontal cortex, activity was found in the lateral parietal lobe that was associated specifically with confidence processes—increasing activity for decreasing confidence, but only for hits. Greater activity was also found in the lateral parietal cortex and medial parietal cortex for low confidence ratings than for high confidence ratings, for both hits and CRs. A number of studies have suggested parietal activity associated with general monitoring processes, but as with the frontal cortex activity we observed, these results clarify that there is activity in specific parietal regions associated with retrospective confidence judgments.

## General Discussion

A task was implemented that was structured to facilitate investigation of how activity in the brain is affected by the retrospective memory confidence judgments, and if there is an effect of the level of confidence for an item. The results indicate that there are clearly specific neural processes involved in retrospective memory confidence judgment—both ERP components and fMRI activity changed as a function of confidence level. Both ERP and fMRI analyses indicate that activity increased as confidence decreased, in some cases dependent on whether a word had been previously studied, and in other cases, regardless of whether a word had been studied. Combined with the increased reaction times for medium and low confidence words compared to high confidence words, these data indicate that when an individual makes a lower-confidence recognition memory judgment, a heavier load is placed on the neural mechanisms involved in the judgment. This increased activity may be due to more searching involved in making the judgment.

These findings may indicate that activity in certain regions of the brain is associated with certain patterns of EEG readings. However, other possibilities cannot conclusively be ruled out. EEG readings were not obtained from the same individuals who were scanned, and so between-subjects variability cannot be ruled out as a possible confound. Ideally, simultaneous EEG and fMRI recordings would allow for ruling out inter-subject variability as a confound. Simultaneous recording is a method becoming used more frequently, and may have provided more information.

Simultaneous EEG-fMRI is not without its limitations, however. Nunez and Silberstein (2000) note several potential concerns with the method. One possibility is

that activity in the cortex could be picked up by fMRI, but because of opposing dipoles in sulci, electrical signal from that region may cancel itself out, and in theory activity in that region could be nullified as far as EEG readings go. Another issue is the timescale difference between electrophysiological data and hemodynamic data. Nunez and Silberstein (2000) suggest that the readings obtained could come from qualitatively different sources. An ERP component can come and go within a window of 200 ms or shorter, perhaps not representing enough activity to actually cause a significant hemodynamic response; thus it is possible that two different mechanisms could be interpreted as the same.

These findings may have relevant implications for patients with injury to the frontal or parietal cortices or disorders such as Alzheimer's disease. In cases where the function of these regions is disrupted, retrospective memory confidence is not likely the first deficit encountered or considered. Impairment in making memory confidence judgments may be less apparent than other memory deficits, but inaccurate memory confidence judgments could play a strong role in behavior that indicates general memory impairment. Lack of confidence or misplaced confidence in a memory could lead one to make decisions that on the surface seem to indicate impaired memories, but may in fact represent a stronger impairment in decision-making abilities or similar executive functions. If the difference could be discerned, or if there was known damage to the regions identified as associated with retrospective memory confidence processes, more effect treatment or compensatory measures could be used. Future research with clinical populations could explore how directly these processes are associated with impaired memory function or damage to the identified regions.

To the knowledge of the researchers, no work has been done specifically addressing memory confidence processes on this scale. A direct relationship between memory confidence and neural activity using was found using a variety of measures. Lower confidence resulted in *more* activity, perhaps reflecting more searching or otherwise greater processing.

## References

- Achim, A. M., & Lepage, M. (2005). Dorsolateral prefrontal cortex involvement in memory post-retrieval monitoring revealed in both item and associative recognition tests. *NeuroImage*, *24*(4), 1113-1121. doi: 10.1016/j.neuroimage.2004.10.036
- Allan, K., & Rugg, M. D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia*, *35*(4), 387-397. doi: S0028-3932(96)00094-2 [pii]
- Allan, K., Wilding, E. L., & Rugg, M. D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychologica*, *98*(2-3), 231-252.
- Avants, B., Duda, J. T., Kim, J., Zhang, H., Pluta, J., Gee, J. C., & Whyte, J. (2008). Multivariate analysis of structural and diffusion imaging in traumatic brain injury. *Academic radiology*, *15*(11), 1360-1375. doi: 10.1016/j.acra.2008.07.007
- Barnes, K. A., & Dougherty, M. R. (2007). The effect of divided attention on global judgment of learning accuracy. *American Journal of Psychology*, *120*(3), 347-359.
- Bonmassar, G., Schwartz, D. P., Liu, A. K., Kwong, K. K., Dale, A. M., & Belliveau, J. W. (2001). Spatiotemporal brain imaging of visual-evoked activity using interleaved EEG and fMRI recordings. *NeuroImage*, *13*(6 Pt 1), 1035-1043. doi: 10.1006/nimg.2001.0754

- Brown, S. R. (2000). Tip-of-the-tongue phenomena: An introductory phenomenological analysis. *Consciousness and Cognition*, 9(4), 516-537. doi: 10.1006/ccog.2000.0421
- Cabeza, R. (2008). Role of parietal regions in episodic memory retrieval: The dual attentional processes hypothesis. *Neuropsychologia*, 46(7), 1813-1827. doi: 10.1016/j.neuropsychologia.2008.03.019
- Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience*, 9(8), 613-625. doi: 10.1038/nrn2459
- Chua, E. F., Rand-Giovannetti, E., Schacter, D. L., Albert, M. S., & Sperling, R. A. (2004). Dissociating confidence and accuracy: Functional magnetic resonance imaging shows origins of the subjective memory experience. *Journal of Cognitive Neuroscience*, 16(7), 1131-1142. doi: 10.1162/0898929041920568
- Chua, E. F., Schacter, D. L., Rand-Giovannetti, E., & Sperling, R. A. (2006). Understanding metamemory: Neural correlates of the cognitive process and subjective level of confidence in recognition memory. *NeuroImage*, 29(4), 1150-1160. doi: 10.1016/j.neuroimage.2005.09.058
- Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009a). Neural basis for recognition confidence in younger and older adults. *Psychology and Aging*, 24(1), 139-153. doi: 10.1037/a0014029
- Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009b). Neural correlates of metamemory: A comparison of feeling-of-knowing and retrospective confidence

- judgments. *Journal of Cognitive Neuroscience*, 21(9), 1751-1765. doi: 10.1162/jocn.2009.21123
- Cleary, A. M., & Specker, L. E. (2007). Recognition without face identification. *Memory and Cognition*, 35(7), 1610-1619.
- Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, 42(8), 1088-1106. doi: 10.1016/j.neuropsychologia.2003.12.011
- Donaldson, D. I., Wheeler, M. E., & Petersen, S. E. (2009). Remember the source: Dissociating frontal and parietal contributions to episodic memory. *Journal of Cognitive Neuroscience*. doi: 10.1162/jocn.2009.21242
- Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Effects of aging on the magnitude and accuracy of quality-of-encoding judgments. *American Journal of Psychology*, 116(3), 431-454.
- Emanuel Robinson, A., Hertzog, C., & Dunlosky, J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition*, 13(3-4), 458-478. doi: 10.1080/13825580600572983
- Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex*, 16(11), 1623-1630. doi: 10.1093/cercor/bhj097
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic

- resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33(5), 636-647.
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4), 468-484.
- Henson, R. N., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience*, 12(6), 913-923.
- Howard, C. E., Andres, P., Broks, P., Noad, R., Sadler, M., Coker, D., & Mazzoni, G. (2010). Memory, metamemory and their dissociation in temporal lobe epilepsy. *Neuropsychologia*, 48(4), 921-932. doi: 10.1016/j.neuropsychologia.2009.11.011
- Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., & Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebral Cortex*, 16(9), 1349-1360. doi: 10.1093/cercor/bhl040
- Johnson, R., Jr., Kreiter, K., Russo, B., & Zhu, J. (1998). A spatio-temporal analysis of recognition-related event-related brain potentials. *International Journal of Psychophysiology*, 29(1), 83-104. doi: 10.1016/S0167-8760(98)00006-3
- Kahn, I., Pascual-Leone, A., Theoret, H., Fregni, F., Clark, D., & Wagner, A. D. (2005). Transient disruption of ventrolateral prefrontal cortex during verbal encoding affects subsequent memory performance. *Journal of Neurophysiology*, 94(1), 688-698. doi: 10.1152/jn.01335.2004

- Kim, H., & Cabeza, R. (2007). Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. *Journal of Neuroscience*, 27(45), 12190-12197. doi: 10.1523/JNEUROSCI.3408-07.2007
- Kim, H., & Cabeza, R. (2009). Common and specific brain regions in high- versus low-confidence recognition memory. *Brain Research*, 1282, 103-113. doi: 10.1016/j.brainres.2009.05.080
- Kirwan, C. B., Shrago, Y., & Squire, L. R. (2009). Medial temporal lobe activity can distinguish between old and new stimuli independently of overt behavioral choice. *Proceedings of the National Academy of Sciences, USA*, 106(34), 14617-14621. doi: 10.1073/pnas.0907624106
- Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M. C., . . . Parsey, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*, 46(3), 786-802. doi: 10.1016/j.neuroimage.2008.12.037
- Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 187-194. doi: 10.1037/0278-7393.31.2.187
- Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory*, 18(1), 15-18. doi: 10.1101/lm.1971111
- Luck, S. J. (2005). *An introduction to the event-related potential technique* (1 ed.). Cambridge, MA: The MIT Press.

- Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience*, 37(4), 292-299.
- Mitchell, K. J., Johnson, M. K., Raye, C. L., & Greene, E. J. (2004). Prefrontal cortex activity associated with source monitoring in a working memory task. *Journal of Cognitive Neuroscience*, 16(6), 921-934. doi: 10.1162/0898929041502724
- Modirrousta, M., & Fellows, L. K. (2008). Medial prefrontal cortex plays a critical and selective role in 'feeling of knowing' meta-memory judgments. *Neuropsychologia*, 46(12), 2958-2965. doi: 10.1016/j.neuropsychologia.2008.06.011
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, 26, 125-173.
- Nunez, P. L., & Silberstein, R. B. (2000). On the relationship of synaptic activity to macroscopic measurements: Does co-registration of EEG with fMRI make sense? *Brain Topography*, 13(2), 79-96.
- Perrotin, A., Belleville, S., & Isingrini, M. (2007). Metamemory monitoring in mild cognitive impairment: Evidence of a less accurate episodic feeling-of-knowing. *Neuropsychologia*, 45(12), 2811-2826. doi: 10.1016/j.neuropsychologia.2007.05.003
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251-257. doi: 10.1016/j.tics.2007.04.004
- Schmitter-Edgecombe, M., & Anderson, J. W. (2007). Feeling of knowing in episodic memory following moderate to severe closed-head injury. *Neuropsychology*, 21(2), 224-234. doi: 10.1037/0894-4105.21.2.224

- Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., & Kaszniak, A. W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgements: Evidence from patients with lesions to frontal cortex. *Neuropsychologia*, *42*(7), 957-966. doi: 10.1016/j.neuropsychologia.2003.11.020
- Schwartz, B. L., & Bacon, E. (2008). Metacognitive neuroscience. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of Metamemory and Memory* (1 ed., pp. 355-371). New York: Psychology Press.
- Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition*, *9*(2 Pt 1), 313-323; discussion 324-316. doi: 10.1006/ccog.2000.0450
- Shimamura, A. P. (2008). A neurocognitive approach to metacognitive monitoring and control. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of Metamemory and Memory* (1 ed., pp. 373-390). New York: Psychology Press.
- Simons, J. S., Peers, P. V., Mazuz, Y. S., Berryhill, M. E., & Olson, I. R. (2010). Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. *Cerebral Cortex*, *20*(2), 479-485. doi: 10.1093/cercor/bhp116
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445-453. doi: 10.1016/j.tics.2005.07.001
- Ward, B. (2001). Deconvolution analysis of fMRI time series data. 109. Retrieved from <http://afni.nimh.nih.gov/pub/dist/doc/manual/Deconvolvem.pdf>

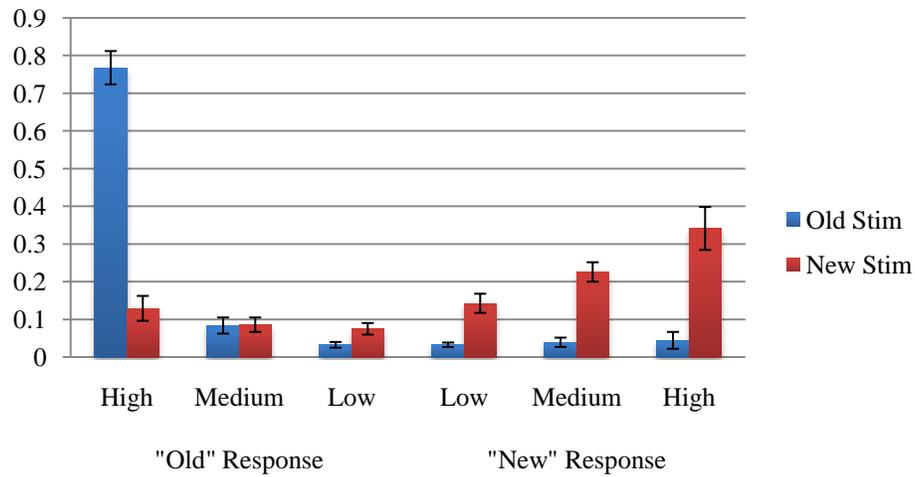
- Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioral Research Methods, Instruments and Computers*, 20, 6-11.
- Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, 1100(1), 125-135. doi: 10.1016/j.brainres.2006.05.019
- Xiong, J., Gao, J.-H., Lancaster, J. L., & Fox, P. T. (1995). Clustered pixels analysis for functional MRI activation studies of the human brain. *Human Brain Mapping*, 3(4), 5. doi: 10.1002/hbm.460030404
- Yassa, M. A., Stark, S. M., Bakker, A., Albert, M. S., Gallagher, M., & Stark, C. E. (2010). High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnesic mild cognitive impairment. *Neuroimage*, 51(3), 1242-1252. doi: 10.1016/j.neuroimage.2010.03.040
- Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, 25(11), 3002-3008. doi: 10.1523/JNEUROSCI.5295-04.2005

### Tables and Figures

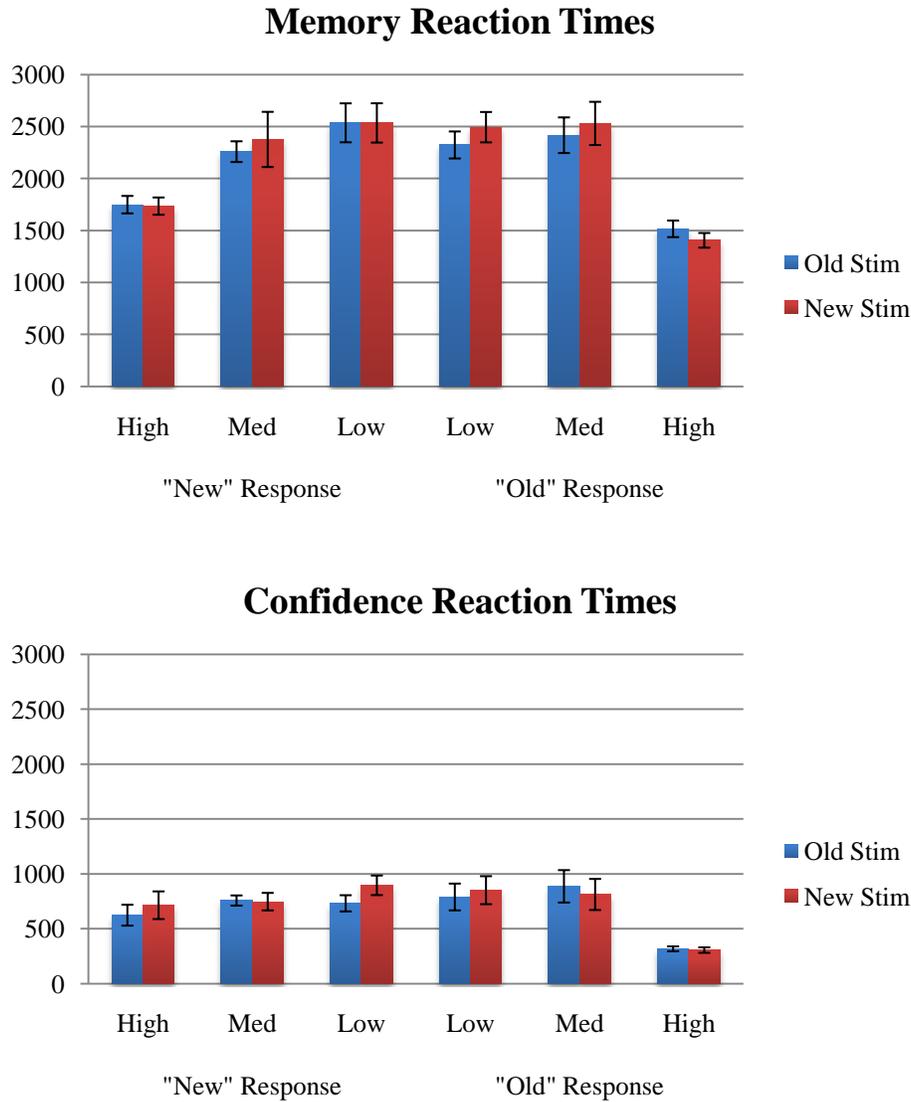
Region	Confidence Regression	“Interaction” Regression	Overlap
Right Lateral Frontal	636	87	29
Right Parietal	67	112	49

*Table 1.* Size of overlapping regions and degree of overlap, in voxels, identified by two regression analyses—seeking regions where activity increased linearly as confidence decreased, collapsed across old and new stimuli, and regions where activity decreased for hits with increasing confidence and constant activity for CRs.

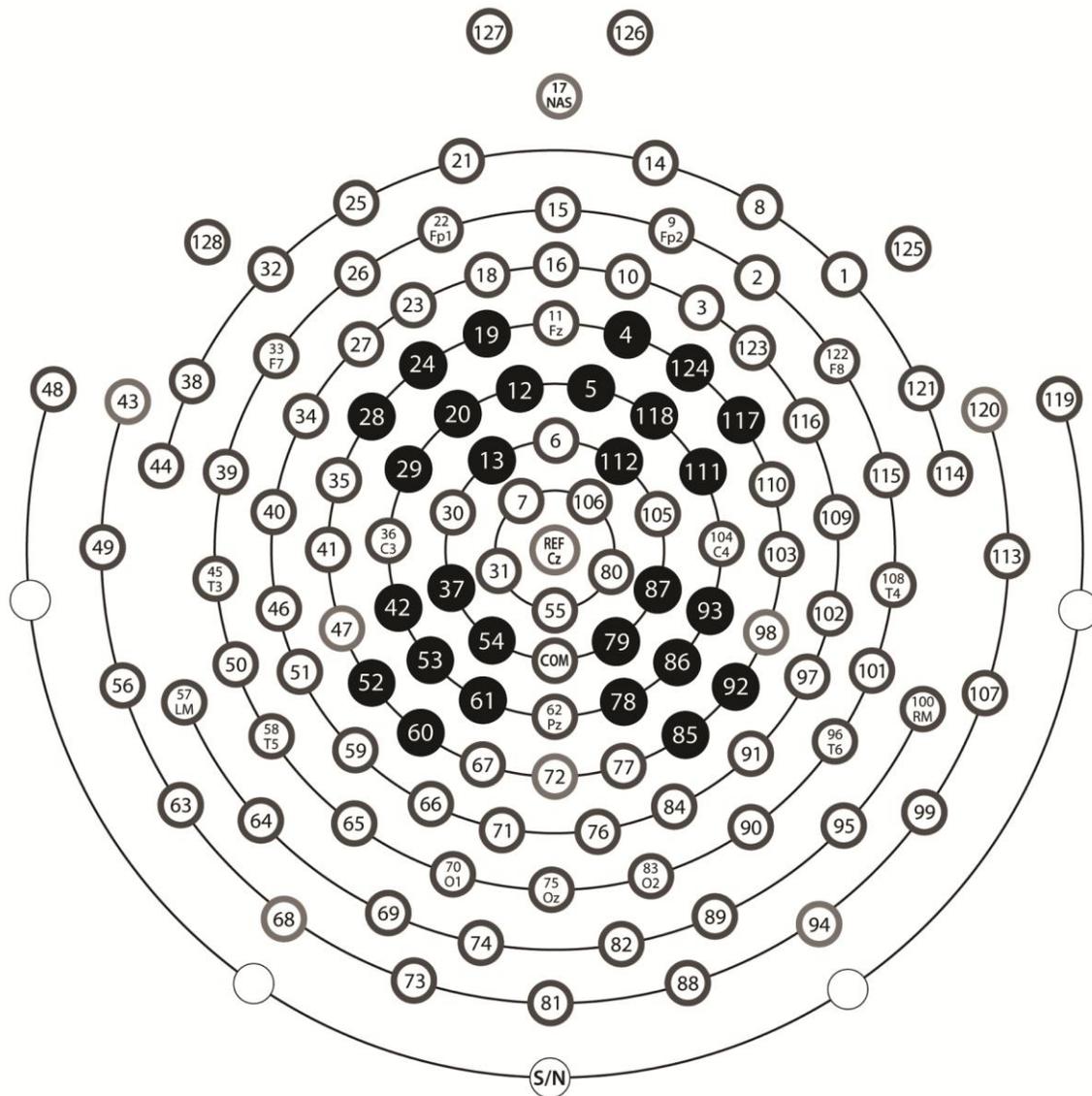
### Experiment 1 Response Proportions



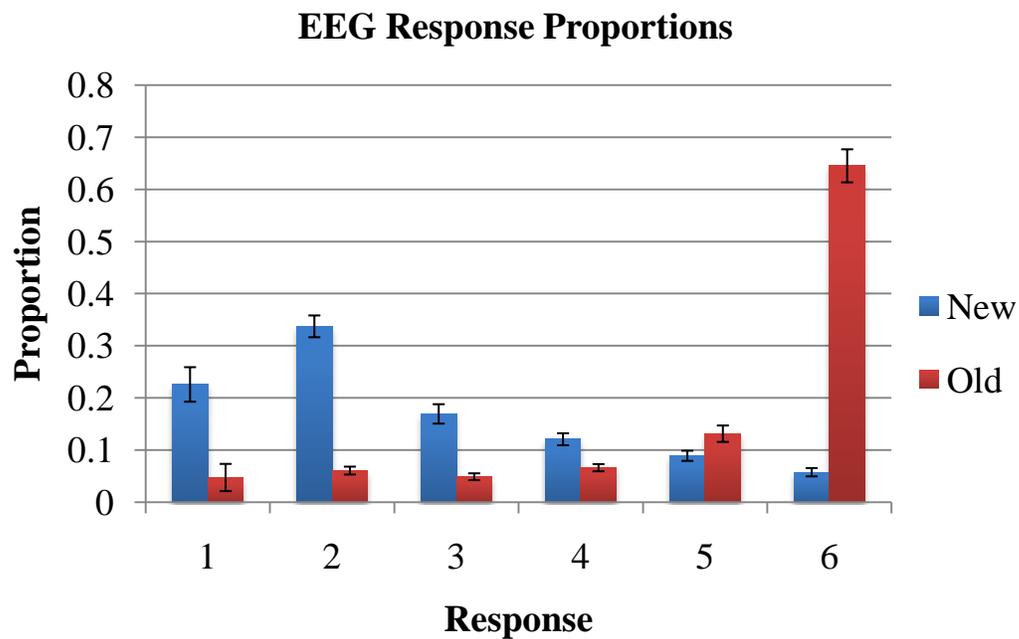
*Figure 1.* Response proportions for Experiment 1. Error bars represent standard error of the mean. (“Old” responses to old stimuli are hits; “new” responses to old stimuli are misses; “old” responses to new stimuli are false alarms; “new” responses to new stimuli are correct rejections.)



*Figure 2.* Mean reaction times for memory response and confidence response. Error bars represent standard error of the mean. (“Old” responses to old stimuli are hits; “new” responses to old stimuli are misses; “old” responses to new stimuli are false alarms; “new” responses to new stimuli are correct rejections.)



*Figure 3.* Electrode clusters used in analyses. FN400 Left: 12, 13, 19, 20, 24, 28, 19;  
 FN400 Right: 4, 5, 111, 112, 117, 118, 124; Late Parietal Left: 37, 42, 52, 53, 54, 60, 61;  
 Late Parietal Right: 78, 79, 85, 86, 87, 92, 93.



*Figure 4.* Mean distribution of responses for EEG participants, for both old and new words. Error bars represent standard error of the mean.

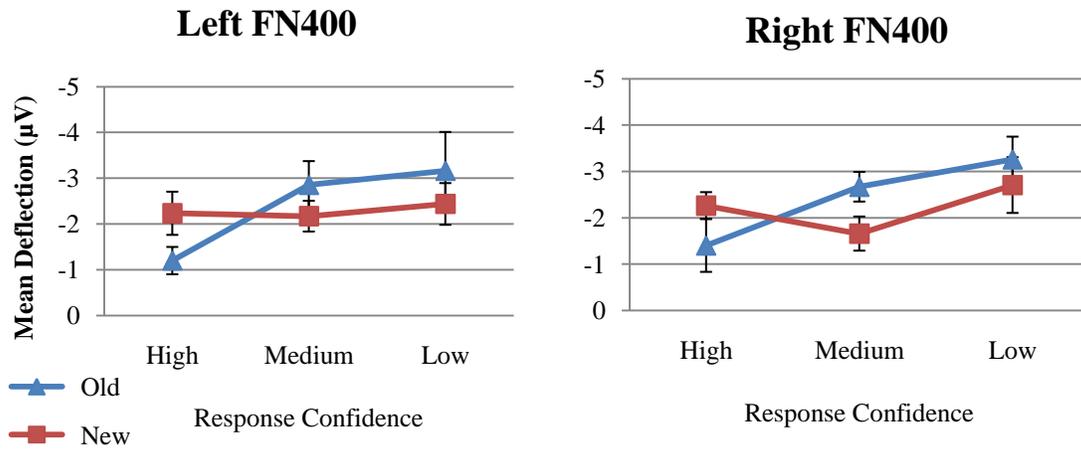


Figure 5. Mean peak amplitude, left and right FN400, by confidence level and stimulus type. Error bars represent standard error of the mean.

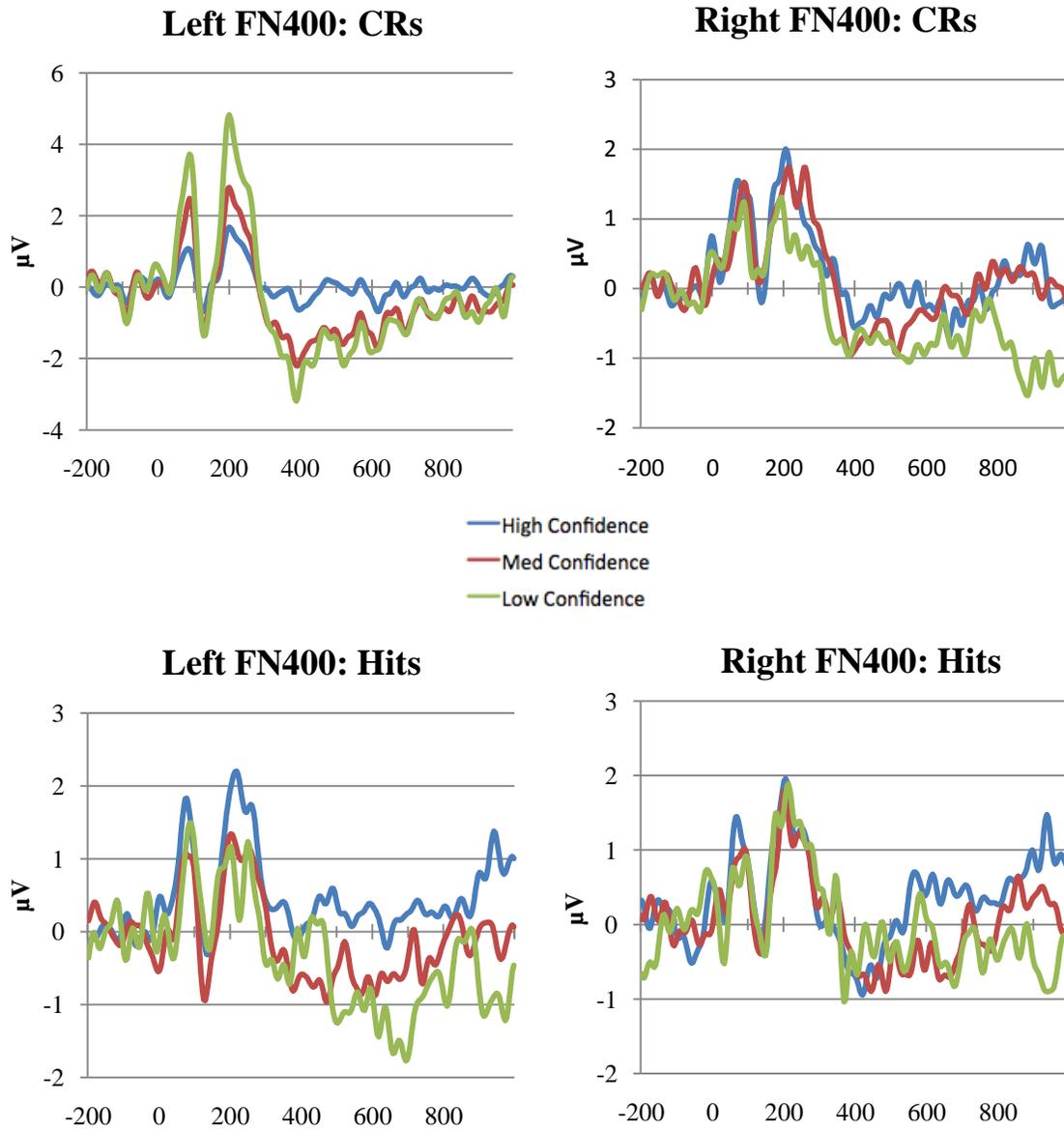
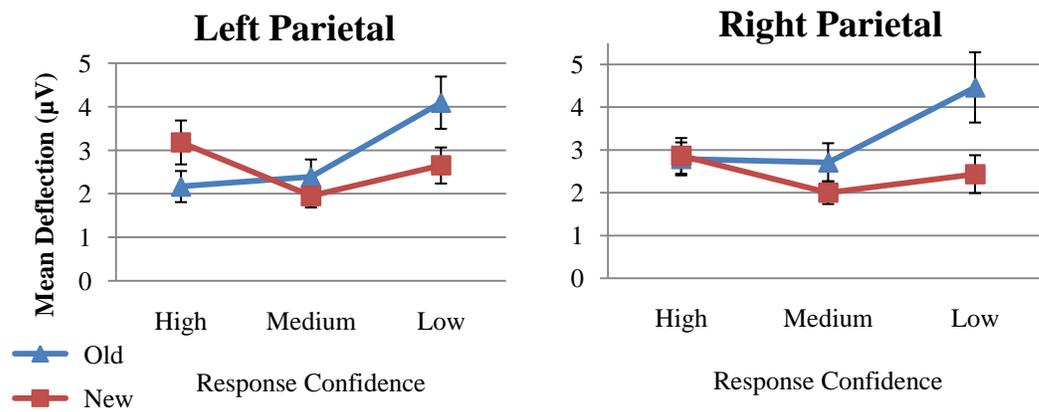


Figure 6. FN400 Waveforms for hits and correct rejections.



*Figure 7.* Mean peak amplitude, left and right parietal effect, by confidence level and stimulus type. Error bars represent standard error of the mean.

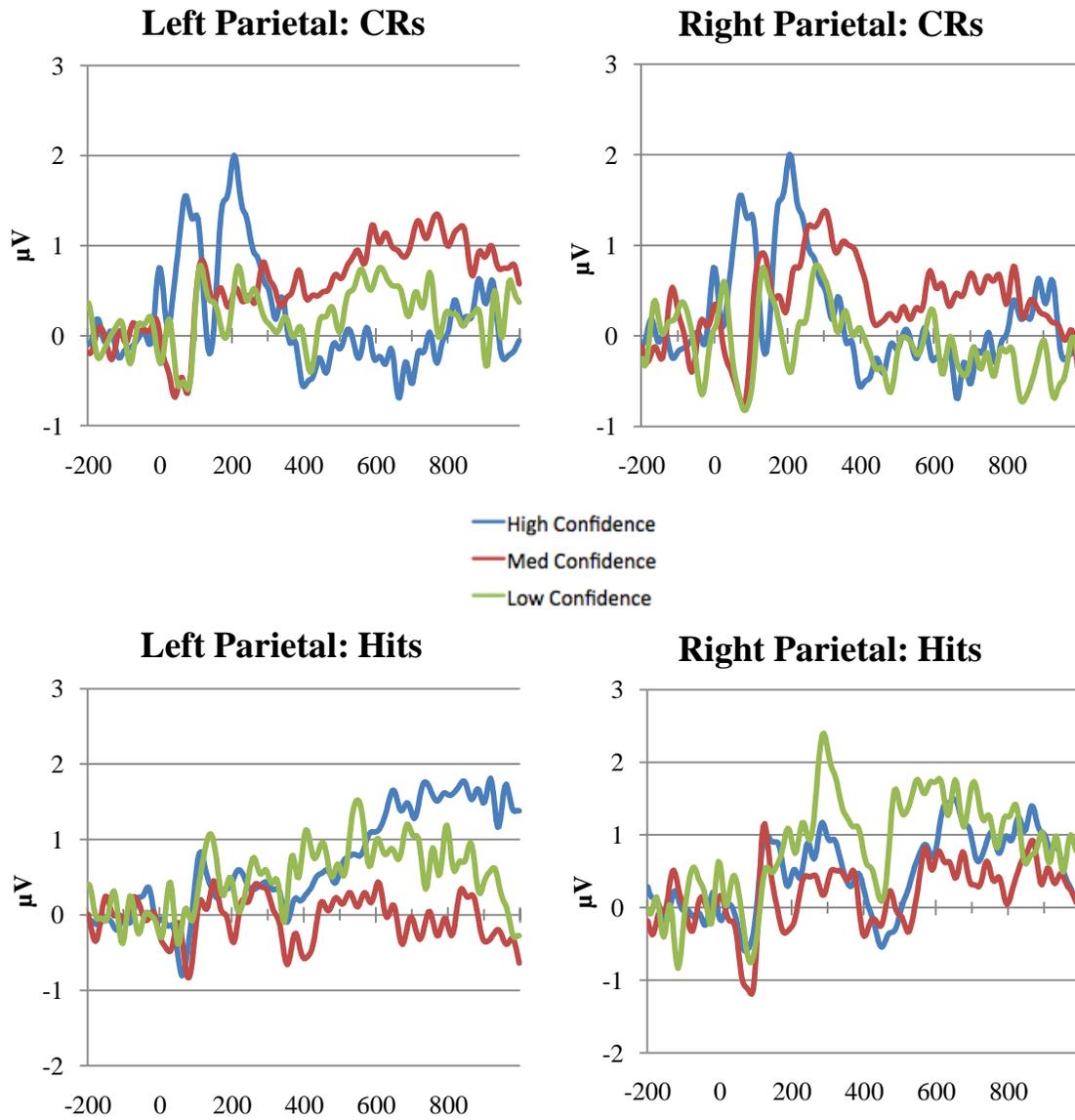
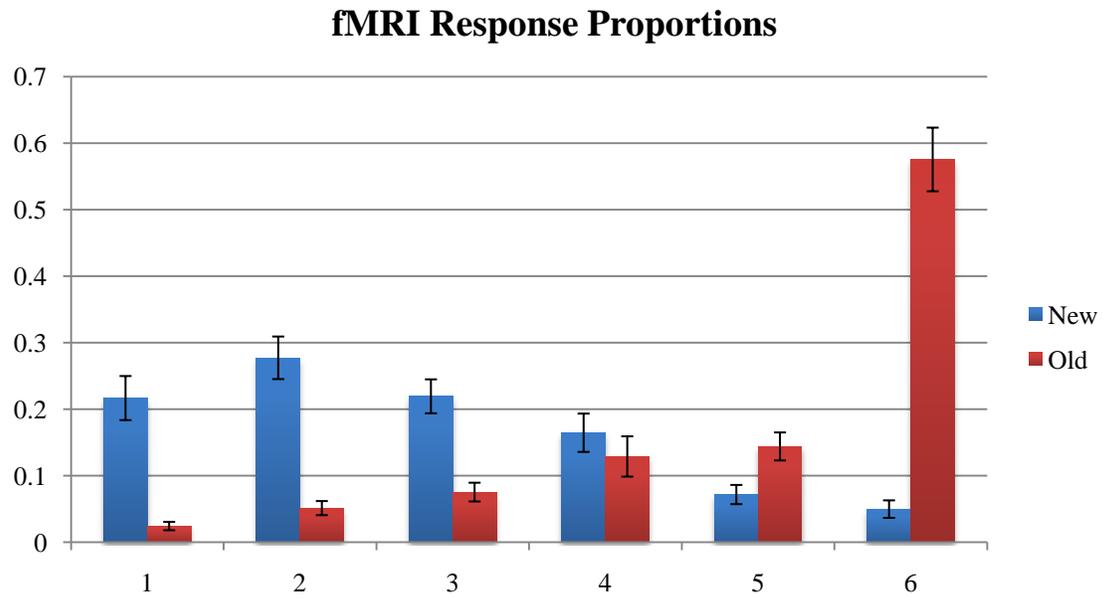
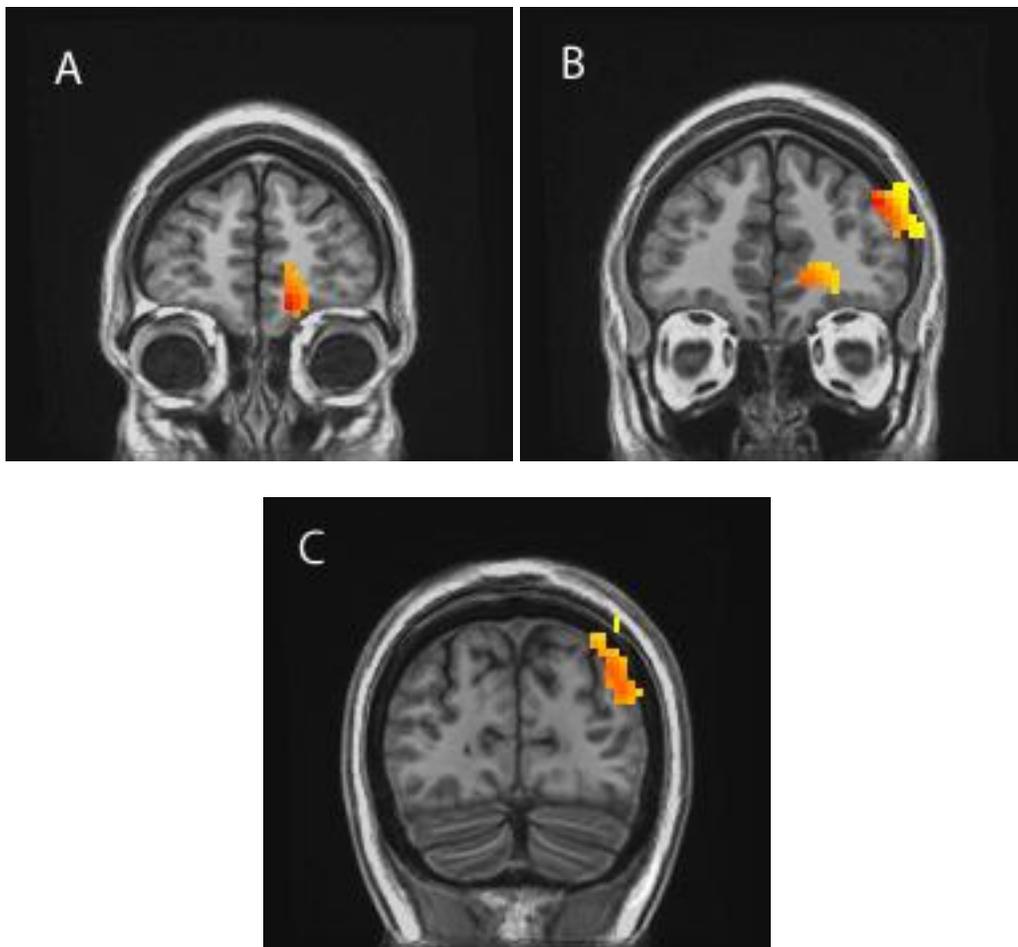


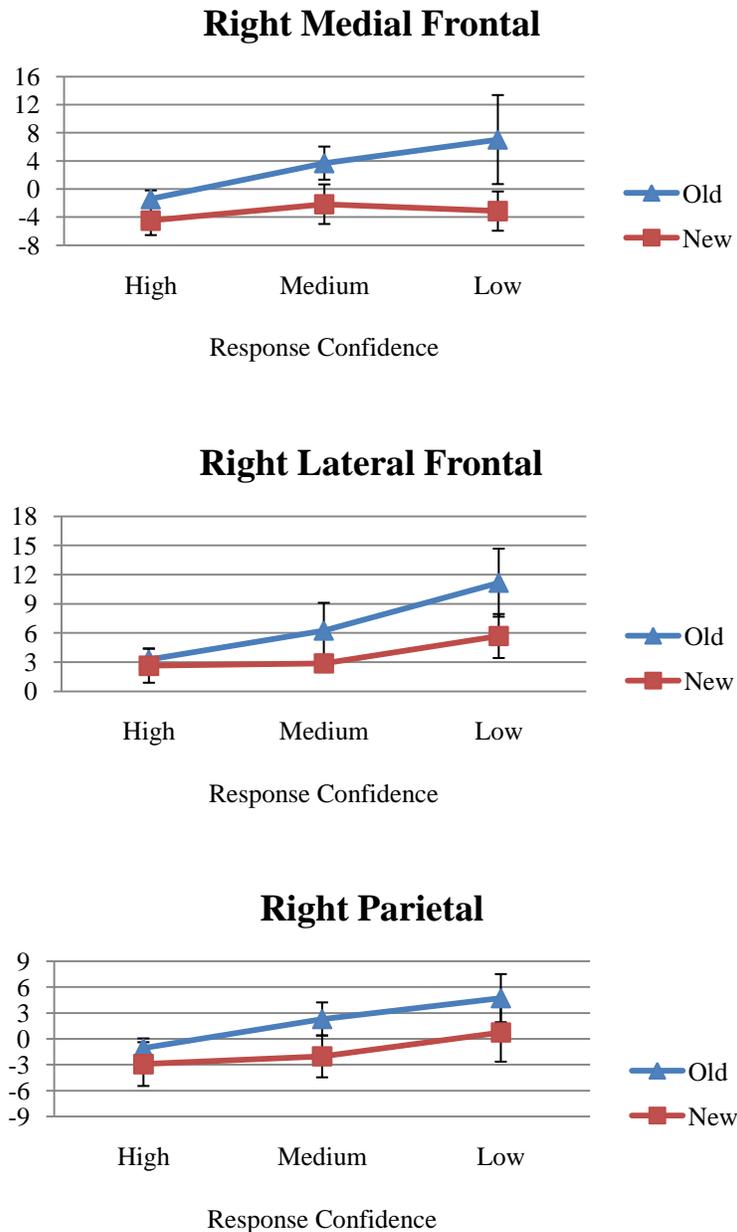
Figure 8. Late parietal component waveforms for hits and correct rejections.



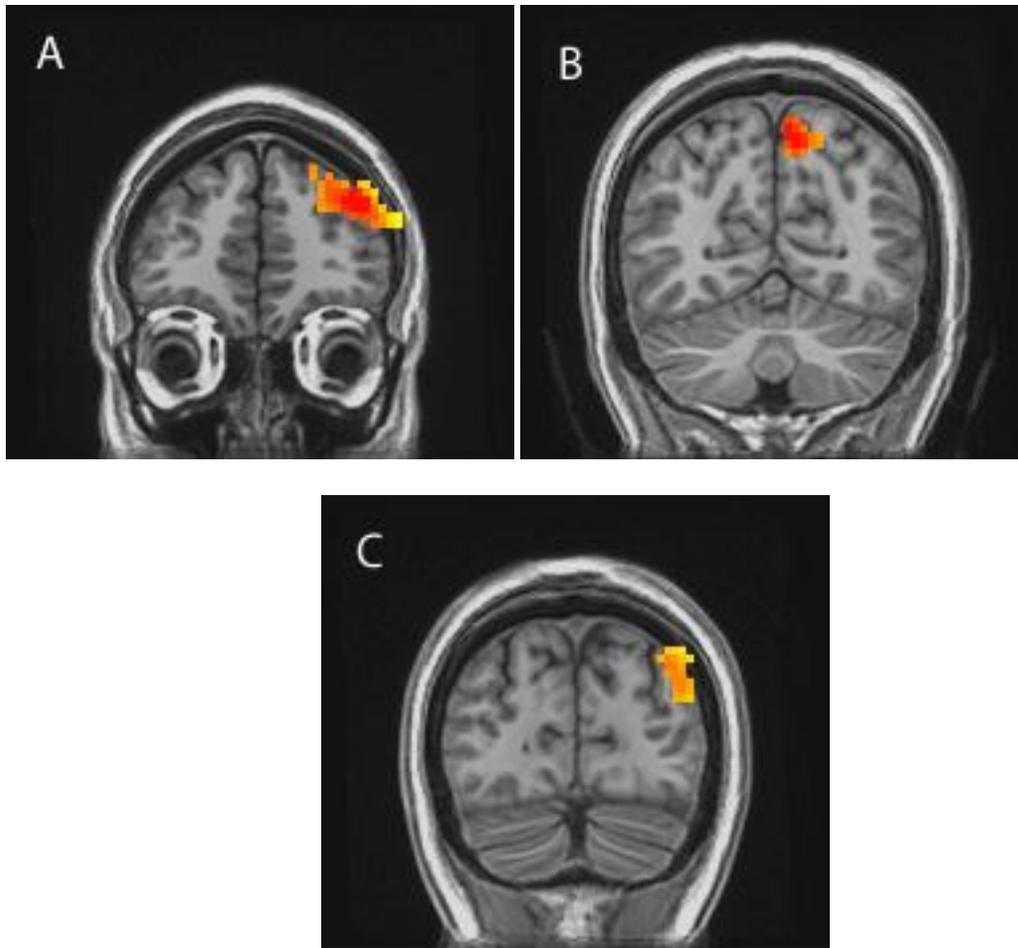
*Figure 9.* Mean distribution of responses for fMRI participants, for both old and new words. Error bars represent standard error of the mean.



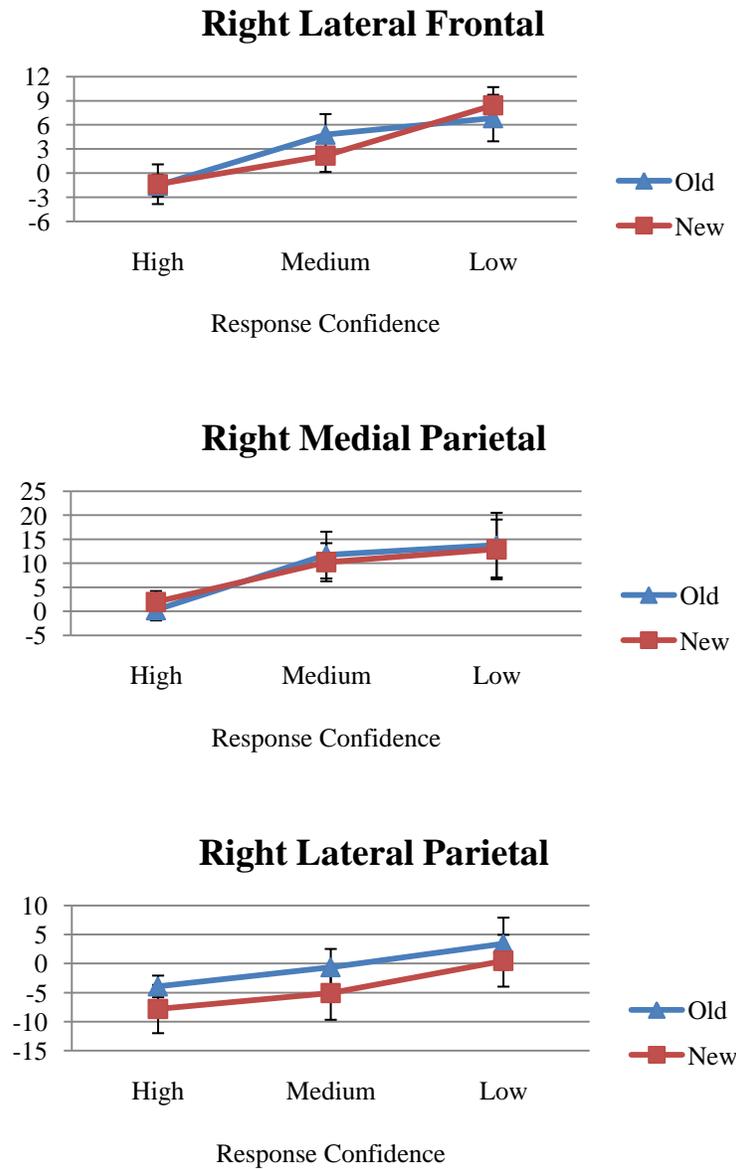
*Figure 10.* ROIs with stimulus type  $\times$  confidence interaction pattern identified by regression analysis: 10a – medial frontal cortex; 10b – medial frontal (cont.) and lateral frontal cortex; 10c – lateral parietal cortex.



*Figure 11.* Mean activation for three ROIs with stimulus type  $\times$  confidence interaction identified by regression analysis. Error bars represent standard error of the mean.



*Figure 12.* ROIs with increased activation for decreased confidence identified by regression analysis: 12a – lateral frontal cortex; 12b – medial parietal cortex; 12c – lateral parietal cortex.



*Figure 13.* Mean activation for three ROIs with increased activation for decreased confidence identified by regression analysis. Error bars represent standard error of the mean.