Conceptual Priming and Familiarity: Different Expressions of Memory during Recognition Testing with Distinct Neurophysiological Correlates

Joel L. Voss[1,2], **Heather D. Lucas**[1], **and Ken A. Paller**[1]

Abstract

■ Familiarity and recollection are qualitatively different explicit-memory phenomena evident during recognition testing. Investigations of the neurocognitive substrates of familiarity and recollection, however, have typically disregarded implicit-memory processes likely to be engaged during recognition tests. We reasoned that differential neural responses to old and new items in a recognition test may reflect either explicit or implicit memory. Putative neural correlates of familiarity in prior experiments, for example, may actually reflect contamination by implicit memory. In two experiments, we used obscure words that subjects could not formally define to tease apart electrophysiological correlates of familiarity and one form of implicit memory, conceptual priming. In Experiment 1, conceptual priming was observed for words only if they elicited meaningful associations. In Experiment 2, two distinct neural signals were observed in conjunction with familiarity-based recognition: late posterior potentials for words that both did and did not elicit meaningful associations and FN400 potentials only for the former. Given that symbolic meaning is a prerequisite for conceptual priming, the combined results specifically link late posterior potentials and FN400 potentials with familiarity and conceptual priming, respectively. These findings contradict previous interpretations of FN400 potentials as generic signals of familiarity and show that repeated stimuli in recognition tests can engender facilitated processing of conceptual information in addition to retrieval processing that leads to the awareness of memory retrieval. The different characteristics of the electrical markers of these two types of process further underscore the biological validity of the distinction between implicit memory and explicit memory. ■

INTRODUCTION

Explicit memory refers to a class of phenomena that entail the awareness of memory retrieval, such as when one recalls a past event or recognizes a familiar face. In contrast, *implicit memory* involves a change in stimulus processing because of previous experience that can be operative either without or in addition to explicit memory. These distinct memory expressions can be dissociated on psychological and neurobiological grounds (Squire, 2004; Gabrieli, 1998). Explicit recognition memory can be further subdivided into two phenomena termed *recollection* and *familiarity* (Eichenbaum, Yonelinas, & Ranganath, 2007; Aggleton & Brown, 2006; Yonelinas, 2002; Mandler, 1980). Recollection can occur with the recognition of a stimulus or event when contextual details regarding the initial encounter are recalled. In contrast, familiarity-based recognition is unsubstantiated by the retrieval of any pertinent detail, such as when a woman's name cannot be recalled despite the conviction that her face had been previously encountered.

The notion that distinct processes underlie these two types of recognition memory expression has become a dominant theoretical framework guiding research into neural substrates of explicit memory. Nonetheless, some investigators have argued that recognition is best described as the result of a single retrieval process and that the nature of the retrieved information can lead to the phenomenological experience of either recollection or familiarity (e.g., Squire, Wixted, & Clark, 2007; Wixted, 2007; Wais, Wixted, Hopkins, & Squire, 2006). Many brain imaging experiments have supported the dual-process account by identifying distinct neural correlates of recollection and familiarity, and neural evidence from ERPs has been particularly influential in this regard (reviewed in Eichenbaum et al., 2007; Rugg & Curran, 2007; Rugg & Yonelinas, 2003). Recollection has been consistently associated with ERPs at posterior scalp locations from approximately 500–800 msec, sometimes designated the "late positive complex" (LPC) in old/new ERP contrasts (reviewed in Voss & Paller, 2008b; Rugg & Curran, 2007; Paller, 2004; Friedman & Johnson, 2000). In contradistinction, brain potentials designated as FN400 (or midfrontal old/new effects) have been frequently accepted as unique neural correlates of familiarity, thus providing pivotal support for the dual-process notion that recollection and familiarity arise from distinct neural substrates (reviewed in Rugg & Curran, 2007; Mecklinger, 2006).

A challenge facing this research, however, lies in validating neural signals of familiarity (Paller, Voss, & Boehm,

---

[1]Northwestern University, Evanston, IL, [2]University of Illinois at Urbana-Champaign

2007). Until this challenge is surmounted, the relevance of these neural data for clarifying the recollection/familiarity distinction remains in serious doubt (for a similar perspective, see Mandler, 2008). One important difficulty is that stimuli presented in a memory test may elicit neural signals of both familiarity and conceptual priming. Conceptual priming, which is thought to arise due to altered processing of the meaning of a stimulus as a consequence of recent experience, has been assessed using various specialized tests (Schacter & Buckner, 1998). However, the neurocognitive processing responsible for conceptual priming would presumably play out even in the absence of relevant behavioral measurements, as long as some items prime the meaning of subsequently presented items. This position is analogous to the notion that explicit retrieval occurs incidentally during implicit tests of memory (Richardson-Klavehn & Bjork, 1988). Given that recognition testing typically involves meaningful stimuli shown repeatedly (first during an encoding session and again when tested), neural correlates of the two kinds of memory could easily be conflated. In other words, when familiarity and conceptual priming co-occur, their respective neural signatures may be difficult to disentangle in comparisons between old and new items. In particular, FN400 effects that have been observed during prior recognition studies could index conceptual priming rather than familiarity.

In addition, an intriguing but controversial speculation is that conceptual priming reflects an early stage of the processing that ultimately produces familiarity (Yonelinas, 2002; Rajaram & Geraci, 2000; Whittlesea & Williams, 1998; Wagner, Gabrieli, & Verfaellie, 1997; Jacoby & Dallas, 1981). Indeed, one way to foster progress in addressing this issue, while also seeking to validate neural correlates of familiarity, would be to characterize neural signals of familiarity and conceptual priming.

We therefore sought to directly compare neural correlates of conceptual priming and familiarity. In our experiments, people viewed obscure, preexperimentally novel words in a study task in which meaningfulness was rated, followed by either implicit- or explicit-memory testing. Each individual's study-phase ratings were used to designate each word as either high or low in meaningful associations (including only those that the individual claimed they could not define). Results from one group of subjects confirmed that conceptual priming occurs only for subjectively meaningful words (Experiment 1). In another group, ERPs recorded during recognition testing were used to contrast neural correlates of recognition as a function of the potential for concomitant conceptual priming (Experiment 2).

EXPERIMENT 1: CONCEPTUAL PRIMING

Materials and Methods

Visual stimuli were 400 extremely uncommon English words between 4 and 10 letters in length (Appendix A). Words were presented in black Arial font on a white background at central fixation and subtended approximate visual angles of 0.5° vertically and between 1.5° and 3.5° horizontally. A central fixation cross was present during each ISI.

Two conceptual priming tests included a randomly selected set of 240 of the 400 words. Subjects ($n = 12$, ages 18–26 years, one male) performed meaningfulness ratings as a study task, followed at test by either a lexical-decision task or a liking-rating task. There were two study-test blocks. All subjects performed both test-phase tasks, one in each study-test block, with the task order counterbalanced across subjects.

During the study task for each test, 60 uncommon words were allocated into meaningfulness categories (words selected at random for each subject, 2000-msec presentation, ISI randomized between 2000 and 3500 msec). Ratings were made using a 5-point scale, with 1 = *real word*, 2 = *high meaningfulness*, 3 = *medium meaningfulness*, 4 = *low meaningfulness*, and 5 = *negligible meaningfulness*. Instructions were to press 1 if an item was thought to be a word irrespective of whether the precise definition was known. Subjects were thus encouraged to adopt liberal criteria for judging items as real words. For items not endorsed as real words, instructions were to rate a stimulus as *high* if it immediately invoked a concrete meaning, *medium* if it immediately invoked an intangible meaning or connotation, *low* if it was possible to attribute minimal meaning to the stimulus with effort, and *negligible* if the stimulus invoked no meaning. Ratings were thus based on the degree to which any aspect of the stimulus, such as orthographic or phonemic structure, cued associations with some meaning or engendered nonspecific feelings of meaningfulness. Operationally, M+ words were those that received high and medium ratings, and M− words were those that received low and negligible ratings. Subjects were also instructed to distribute their ratings approximately evenly across the four meaningfulness categories.

Pilot testing indicated that the assignment of words to meaningfulness categories was highly idiosyncratic, such that ratings from each individual were far superior to group norms, which would not accurately conform to subjective item meaningfulness (as in Voss & Paller, 2007). As expected, word ratings in Experiment 1 were found to be inconsistent across subjects (average $SD = 1.9$ on the 5-point scale). Thus, many words were assigned to the M+ category by some subjects and to the M− category by other subjects. For instance, 71.3% of the words were designated M+ by at least one third of the subjects while also designated M− by at least one third of the subjects. Rating variability thus led to an approximate counterbalancing of stimuli to the M+ and M− categories across subjects.

In the lexical-decision task, subjects rapidly categorized repeat and novel words into "word" and "nonword" categories. Subjects were misinformed that the stimulus set included both uncommon words and pseudowords (in reality, all were uncommon words). All 60 uncommon words from the prior study task were presented again along with 60 novel uncommon words (selected at random for each subject) and 60 common, high-frequency words (e.g., table, lamp, church) in randomized order (1000-msec presentation, ISI randomized between 2000 and 3000 msec). Subjects pressed one button to indicate that the item on the screen was a word and another button to indicate that the item was not a word. Speed and accuracy were emphasized.

In the liking-rating task, all 60 uncommon words from the prior study task were presented again along with 60 novel uncommon words (selected at random for each subject) in randomized order (1000-msec presentation, ISI randomized between 2000 and 3000 msec). Subjects rated each stimulus using a 4-point scale, with the following response categories: "dislike the most," "like the least," "like a little more," and "like the most." Adjacent ratings were assigned to adjacent fingers, and the button assigned to the lowest rating on the liking scale was the button assigned to the highest rating on the meaningfulness scale, such that high meaningfulness ratings would not lead to higher liking ratings by virtue of response priming. Subjects were encouraged to make ratings at a gut level, given that there was no right or wrong answer, and response speed was not emphasized.

One additional phase followed the second study-test block. All uncommon words that had appeared in the two tasks were shown and rated for meaningfulness using the same scale as in the study task.

Results and Discussion

In the study phase, subjects assigned words fairly evenly into the chief two categories, M+ (combining the high and medium meaningfulness ratings, 50.2% on average, *SE* = 2.2%) and M− (combining the low and negligible meaningfulness ratings, 43.9% on average, *SE* = 2.4%). If a subject claimed to know the definition of a word (indicated by rating 1), that specific word was excluded from all analyses for that subject (5.9%, *SE* = 2.2%). These initial meaningfulness ratings were compared with meaningfulness ratings for the same words obtained in the final phase of the experiment. Initial and repeat ratings were highly consistent, in that 94% (*SE* = 4.3%) of the stimuli were categorized as M+ for both the initial and the repeat ratings or as M− for both the initial and the repeat ratings. The initial ratings were used to categorize stimuli for analyses. New words in the final phase were categorized fairly evenly into M+ and M− conditions (48.3%, *SE* = 3.1% and 44.7%, *SE* = 4.0%, respectively). New words that were in a subject's lexicon (7.0%, *SE* = 2.8%) were excluded from all analyses.

Both the lexical-decision and the liking-rating tasks yielded indications of conceptual priming, as shown in Table 1. In the former task, subjects rapidly assigned each word to the word or nonword category. Given the inclusion of very common words in this test, along with equal numbers of old (i.e., studied) and new obscure words, the correct response was taken to be nonword for the obscure words. Indeed, most of the obscure words were endorsed as nonwords. We hypothesized that classification speed for old words could be enhanced because of prior conceptual analysis and that evidence for conceptual priming would be obtained selectively for old M+ words. Conceptual priming was identified selectively for stimuli with meaningful associations, in that responses to old M+ words were 77 msec faster than responses to new M+ words, $t(11) = 2.4$, $p = .03$. In contrast, RTs to M− words did not differ significantly according to whether the word had been studied, $t(11) = 1.0$, $p = .32$. The magnitude of the old/new difference in response times was significantly greater for M+ than for M− words, $t(11) = 2.3$, $p = .04$. Classification accuracy was comparably high for all word categories. Although exposure to a nonword might in some circumstances increase the likelihood of that word later being endorsed as a real

Table 1. Conceptual Priming Specific to M+ Words in Experiment 1

	M+ Words		M− Words	
	Old	*New*	*Old*	*New*
Lexical-decision Test				
Accuracy (%)	83.1 (6.1)	85.9 (5.5)	82.1 (2.9)	84.5 (4.8)
RT (msec)	**690.6 (32.4)**	**767.4 (42.5)**	756.0 (39.4)	731.1 (40.3)
Liking-rating Test				
Rating × 10	**26.2 (0.6)**	**24.7 (0.5)**	22.4 (0.5)	22.7 (0.9)
RT (msec)	982.5 (48.6)	945.1 (36.7)	1038.7 (41.2)	978.6 (50.8)

Accuracy in the lexical-decision test is the percentage of uncommon words that were identified as nonwords because subjects were led to believe that these were not real words. RT in the lexical-decision test is based on correct responses. Values in boldface show significant pairwise old/new differences at $p < .05$ indicating conceptual priming. *SE* is indicated in parentheses.

word, trends for such effects were nonsignificant for obscure words in the present design. The high-frequency words in the test probably served to anchor the criterion for responding "word" at a very high level that was seldom reached by any of the repeated obscure words.

On the basis of the premise that fluent processing can enhance a positive effect (e.g., Winkielman, Halberstadt, Fazendeiro, & Catty, 2006), the liking-rating test used a scale that indicated the extent to which stimuli were liked at an intuitive level. Speed was not emphasized during this test, and priming effects on response times were therefore not predicted. Liking ratings were higher for old M+ words than for new M+ words, $t(11) = 2.3$, $p = .04$, whereas liking ratings for old and new M− words did not differ significantly, $t(11) = 0.4$, $p = .69$. The magnitude of the old/new difference in meaningfulness ratings was significantly greater for M+ than for M− words, $t(11) = 2.2$, $p = .05$. Meaningful words were thus liked more if they had been seen earlier than if they were new.

It is important to note that both of these priming tests could conceivably measure perceptual priming effects, conceptual priming effects, or both together. The finding that priming was observed selectively for words in the M+ category, in both tests, suggests that some level of conceptual processing was necessary. As described in the Materials and methods section, specific stimuli were largely counterbalanced in the M+ and M− categories on the basis of the subjects' idiosyncratic meaningfulness ratings. If perceptual priming was operative, it is reasonable to assume that it would have been operative for M− words, but it was not apparent. Also, given that visual feature processing during the study phase was engaged for both M+ and M− words and that there is no reason to expect perceptual priming only for M+ words, we conclude that the two priming tests used here were not sensitive to perceptual priming effects. The most likely variable to produce priming effects was the repeated processing of meaningful associations, which occurred selectively for M+ words in both tasks.

EXPERIMENT 2: RECOGNITION MEMORY

Materials and Methods

The recognition memory experiment included 10 blocks, each with a unique set of words presented during a study phase and during a test phase. All 400 words described for Experiment 1 were used as stimuli. During each study phase, 20 words were presented individually (2000-msec presentation, ISI randomized between 2000 and 3500 msec). Subjects ($n = 16$, ages 18–25 years, 3 males) were misinformed that each stimulus was either an uncommon word or a pseudoword (all were uncommon words). They rated each word using the 5-point meaningfulness scale described in Experiment 1. Subjects were instructed that memory for the words being rated would subsequently be tested. As in Experiment 1, meaningfulness was rated inconsistently across subjects (average $SD = 1.2$ on the 5-point scale). Many words were assigned to the M+ category by some subjects and to the M− category by other subjects. For instance, 74.8% of the words were designated M+ by at least one third of the subjects while also designated M− by at least one third of the subjects.

Recognition memory was tested after an additional delay of approximately 1 min during which subjects were given instructions. The average retention delay was thus approximately 3 min. Each test included the 20 words viewed during the previous study phase and the 20 entirely novel words, presented in randomized order (1000-msec presentation, ISI randomized between 2000 and 3000 msec). Subjects used five buttons to categorize each stimulus as old or new using response categories on the basis of a modified "remember/know" paradigm (Gardiner & Java, 1991; Tulving, 1985). *Remember* responses indicated that recognition coincided with recall of contextual detail, *know* responses indicated that recognition included an acontextual feeling of familiarity, and *new* responses indicated that the word was not presented earlier. Know responses were subdivided into high-, medium-, and low-confidence categories.

We adopted instructions developed by Montaldi, Spencer, Roberts, and Mayes (2006) to de-emphasize strategic processing that could enhance recollection-based recognition. Each subject received detailed information regarding differences between the experiences of recollection and familiarity. Subjects were asked to strategically limit recollection by attempting to refrain from effortful retrieval of details. However, it was stressed that recollection could occur incidentally and that it was important to accurately report recollection whenever it occurred. Recollection was defined for subjects as the retrieval of any contextual details from the study phase accompanying recognition of a stimulus. Recollection was further explained using examples of possible contextual details, such as recalling a previous decision about whether the item was an uncommon word or a pseudoword, recalling a meaning previously associated with the item, or recalling any other detail for the learning episode. Although this procedure is slightly different from typical remember/know procedures, ERP effects qualitatively similar to those reported below (i.e., associations between LPC and familiarity strength and between FN400 and conceptual priming) were obtained in a pilot experiment that used typical remember/know instructions (Lucas, Voss, & Paller, 2007).

After all study-test blocks were completed, all 400 words were rated for meaningfulness in a final phase of the experiment. This step was included so that new words from recognition tests could be assigned to the M+ and M− categories for analysis.

ERPs were extracted from high-density scalp electroencephalographic recordings made during the study and test phases. Recordings were made from 59 evenly distributed tin electrodes embedded in an elastic cap (Woldorff et al., 2002). Voltage was referenced to a right mastoid electrode and rereferenced off-line to average mastoids. The EOG was recorded from four additional channels (below the center of each eye and on each outer canthus). Electrode impedance was kept less than 5 kΩ. EEG signals were recorded with a band-pass filter of 0.05 to 200 Hz and sampled at a rate of 1000 Hz. Each 1100-msec averaging epoch began 100 msec before stimulus onset. Mean prestimulus amplitudes were subtracted to correct for baseline variability. Epochs containing electro-ocular or other artifacts were excluded from ERP analyses (average of 8.6% per subject, $SE = 2.6\%$). For presentation purposes only, waveforms were low-pass filtered with a 40-Hz zero-phase-shift Butterworth filter.

ERPs were averaged selectively as a function of test-phase responses and study-phase meaningfulness category. Statistical comparisons were performed on amplitudes averaged over three midline electrode clusters and were made using repeated measures ANOVA ($\alpha = .05$), with the Greenhouse–Geisser correction for nonsphericity. Post hoc pairwise comparisons used Bonferroni correction. The clusters included the following electrodes: anterior cluster —Fza, F3s, F4s, Fzp, FC1, FC2, and Cza; central cluster—C1a, C2a, Cz, C1p, C2p, and Pzs; and posterior cluster—Pzi, PO1, PO2, Ozs, O1′, O2′, O1i, O2i, and Ozi. The lowercase letters following electrode labels indicate that the given electrode was slightly anterior (a), posterior (p), superior (s), or inferior (i) to the corresponding locations from the International 10–20 system.

Results and Discussion

As in Experiment 1, obscure words were rated for subjective meaningfulness both in the study phase and when the meaningfulness task was administered at the conclusion of the experiment. In the study phase, words were assigned fairly evenly to the two meaningfulness categories (47.8% M+, $SE = 1.8\%$; 41.4% M−, $SE = 2.3\%$). If a subject claimed to know the definition of a word, that word was excluded from all analyses for that subject (10.8%, $SE = 1.9\%$). In the final meaningfulness task, new words were categorized into M+ words (45.5%, $SE = 3.1\%$) and M− words (42.4%, $SE = 3.7\%$), and those words in a subject's lexicon were excluded (12.1%, $SE = 2.5\%$). For those items that were rated twice, ratings were highly consistent with initial ratings (91% assigned to the same M+ or M− category, $SE = 5.1\%$); initial ratings were used to categorize stimuli for analyses.

An initial analysis was conducted to determine whether behavioral and ERP results in the recognition test differed for new M+ words and new M− words. Response rates for correct rejections (new items endorsed as new) were very similar for these two conditions (0.50 and 0.48, respectively), $t(15) = 1.2$, $p = .26$. Likewise, false alarm rates to new items were very similar for M+ versus M− categories for low-confidence know responses (0.30 and 0.32, respectively), $t(15) = 1.1$, $p = .28$; medium-confidence know responses (0.14 and 0.13, respectively), $t(15) = 1.0$, $p = .33$; high-confidence know responses (0.05 and 0.06, respectively), $t(15) = 1.2$, $p = .24$; and remember responses (0.004 and 0.006, respectively), $t(15) = 0.8$, $p = .41$, also with no reliable differences when collapsed across these four responses, $t(15) = 1.2$, $p = .26$. New M+ and new M− items were therefore combined for the behavioral analysis of performance in the recognition test. Furthermore, as shown in Figure 1, ERPs for these two conditions were virtually identical for all electrodes. Accordingly, correct-rejection M+ and M− items were collapsed together to create the new correct-rejection category for the primary ERP analyses.

Accuracy in the recognition test was reasonably high and varied by meta-memory judgment, as shown in Figure 2. The highest proportion of new words were endorsed as new responses, with progressively lower proportions as know confidence increased and very few remember responses. The highest proportion of old words were endorsed as high-confidence know responses, with progressively lower proportions as know confidence decreased and very few endorsed as new.

Remember responses were registered more often for M+ compared with M− words, $t(15) = 4.0$, $p = .001$, and more often than chance as estimated by the corresponding false alarm rate for new words, $t(15) = 6.8$, $p < .001$ for M+, $t(15) = 3.4$, $p = .004$ for M−. In contrast, the proportion of high-confidence know responses did not vary

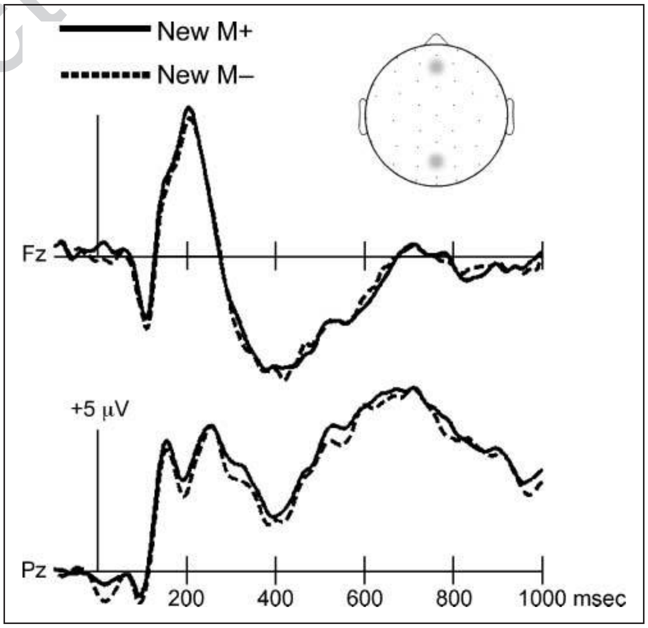

Figure 1. ERP correlates of correctly rejected new words for both meaningfulness levels in Experiment 2. ERP waveforms for new correct rejections during recognition testing are shown separately for items that were subsequently endorsed with M+ and M− ratings. Waveforms are shown for electrode locations approximating Fz and Pz, indicated with gray circles on the schematic diagram of the head. Comparably negligible differences between conditions were evident at all recording sites. Therefore, the new correct-rejection condition (new CR) was computed by collapsing all M+ and M− words together for the primary analyses.

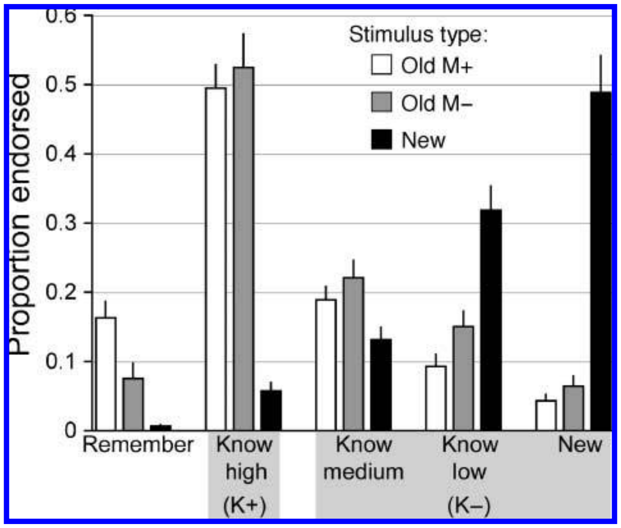

Figure 2. Recognition performance in Experiment 2. Average endorsement rates for old M+ words, old M− words, and new words for each of the five response options during recognition testing. Proportions endorsed for each condition (old M+, old M−, and new) were computed relative to the total number of words in that condition. *Remember* responses were used to indicate that recognition coincided with recall of contextual detail, *know* responses to indicate that recognition included an acontextual feeling of familiarity, and *new* responses to indicate that the word was not presented earlier. Know responses were subdivided into high-, medium-, and low-confidence categories. Items in the new condition were collapsed across M+ and M− categories because performance differences as a function of meaningfulness category were negligible (see text). The assignment of response options to the K+ and K− categories is indicated with shading. Error bars indicate *SE*.

significantly for M+ versus M− words, $t(15) = 0.3$, $p = .74$, and were registered more often than the corresponding false alarm rate for new words for both meaningfulness categories, $t(15) = 12.8$, $p < .001$ for M+, $t(15) = 7.8$, $p < .001$ for M−.[1] Remember and high-confidence know responses indexed veridical memory rather than hits on the basis of pure chance responding, given that, on average, hits outnumbered false alarms by a factor of approximately 18 for remember responses (24.8 for M+ and 11.5 for M−) and by a factor of approximately 9 for high-confidence know responses (8.6 for M+ and 9.0 for M−).

In contrast to the high accuracy achieved with remember and high-confidence know responses, medium- and low-confidence know responses to old items were tantamount to recognition misses. Medium-confidence know responses were registered more often for old items than for new items, $t(15) = 2.9$, $p = .01$ for M+, $t(15) = 3.2$, $p = .006$ for M−, but hits outnumbered false alarms merely by a factor of approximately 1.6 (1.4 for M+ and 1.7 for M−). Therefore, the majority of hits were likely based on uninformed guessing. For low-confidence know responses, hits were registered less often than false alarms, $t(15) = 7.0$, $p < .001$ for M+, $t(15) = 4.6$, $p < .001$ for M−. Therefore, the medium-confidence know, the low-confidence know, and the new response conditions can all be regarded as failing to index recognition at a level of accuracy sufficient for producing memory effects in ERPs.[2] These conditions were combined to define the K− category for some subsequent ERP analyses.

To distinguish neural correlates of conceptual priming from those of explicit memory, our strategy was to contrast ERP responses to two subsets of words (1) that differed systematically in capacity for conceptual priming and (2) that did not differ in explicit memory. To address the first requirement, we relied on results from Experiment 1 indicating that systematic variation in perceived meaning between M+ and M− categories yielded differential levels of conceptual priming, along with the assumption that the relevant processing occurred in the recognition test of Experiment 2. Do M+ and M− words differ in other ways in addition to the capacity for conceptual priming? Greater processing of meaning at encoding certainly leads to better explicit memory. However, subjects were instructed to use the remember response when meaningful associations from the study phase were recalled, and memory results showed that recollection occurred preferentially for M+ words (Figure 2). We instead focused on words recognized using the high-confidence know response, thus avoiding trials with recognition based on the subjective sense of contextual recollection.

Although high-confidence know responses did not yield different endorsement rates for M+ and M− words, meaning processing may have differed between M+ and M− words but not meaning in the sense of explicit recall of meaning from the study phase. Moreover, differential meaning per se was shown to have negligible effects on ERPs in the test phase (Figure 1). Because subjective ratings were used to divide words into the M+ and M− categories, the two conditions could possibly have differed in ways besides the extent to which meaningful associations had been produced in the study phase. For example, words with particular orthographic features may have provoked high meaningfulness ratings. However, concern regarding this possibility is lessened by virtue of the aforementioned finding that stimuli tended to be perceived as high in meaning by some subjects and as low in meaning by others, which provided a degree of stimulus counterbalancing.

To address the second requirement, limiting our M+/M− comparisons to words that were recognized with high-confidence familiarity at test also had the advantage of ensuring a contrast between conditions matched in explicit memory. In another sense, however, familiarity evaluated across the entire set of repeated words would likely be much greater for M+ than for M− words. As already noted, M+ words were disproportionately subject to recollection, and recollection can entail simultaneous familiarity (Yonelinas, 2002). Nevertheless, high-confidence know responses indicated the absence of recollection and reflected a level of familiarity strength (the highest of the three levels) that was equivalent for words in the two meaningfulness conditions. Indeed, accuracy as indicated by discrimination sensitivity did not differ significantly between M+ and M− words that received high-confidence know responses (d' values were 1.6 and 1.7, respectively, computed using the overall false alarm rate for new words), $t(15) = 1.2$, $p = .24$. In addition to this evidence, it is reasonable to assume that subjects used the same criteria for making high-confidence familiarity responses to M+ and M− words, given that the alternative would entail two-stage decisions involving meaningfulness assessment first, followed by familiarity assessment with the adoption of two different criteria sets. We therefore examined ERPs for words endorsed with the high-confidence familiarity response to identify neural correlates of familiarity-based recognition with versus without concomitant conceptual implicit-memory processing (M+ vs. M−).

Figure 3 demonstrates that at a latency of approximately 300–500 msec, positive old/new differences with maximum values at midfrontal electrodes (FN400 effects) were strikingly evident for M+ words endorsed with high-confidence know responses, whereas these differences were not evident for M− words endorsed with high-confidence know responses. Of course, this outcome might be interpreted by claiming that FN400 potentials signal familiarity selectively for M+ words, but that interpretation is not consistent with other results discussed below. Figure 3 also shows later-onset positive old/new effects with posterior distributions (LPC effects) that were evident for both meaningfulness categories with similar magnitudes.

Formal ERP comparisons were made across these three conditions (M+ high-confidence familiarity correct trials or "M+ K+ Hit," M− high-confidence familiarity correct trials or "M− K+ Hit," and new correct rejections or "New CR"; average trial counts for these three conditions were 42, 33, and 89, ranges 22–57, 20–71, and 27–168, respectively). ERP amplitudes were averaged for three midline electrode clusters (frontal, central, and posterior; see Materials and methods) and two latency intervals that captured FN400 potentials (300–500 msec) and LPC potentials (600–900 msec). A three-way Electrode Cluster × Latency Interval × Condition interaction indicated that ERP differences between conditions differed across space and time, $F(2.4, 35.4) = 3.7$, $p = .03$. Assessments were thus made separately for each cluster and interval.

For the interval from 300 to 500 msec, main effects of condition were significant for all three electrode clusters: frontal, $F(1.2, 17.6) = 10.3$, $p = .004$; central, $F(1.5, 22.4) = 12.2$, $p = .001$; and posterior, $F(1.6, 23.4) = 15.7$, $p < .001$. For all three clusters, amplitudes were significantly greater for M+ than for new ($p = .005$, $p = .001$, and $p < .001$, respectively, for the frontal, central, and posterior clusters) and for M+ than for M− ($p = .005$, $p = .004$, and $p = .002$, respectively). In contrast, amplitudes for M− and new did not differ significantly for any cluster ($p = .12$, $p = .22$, and $p = .19$, respectively).[3]

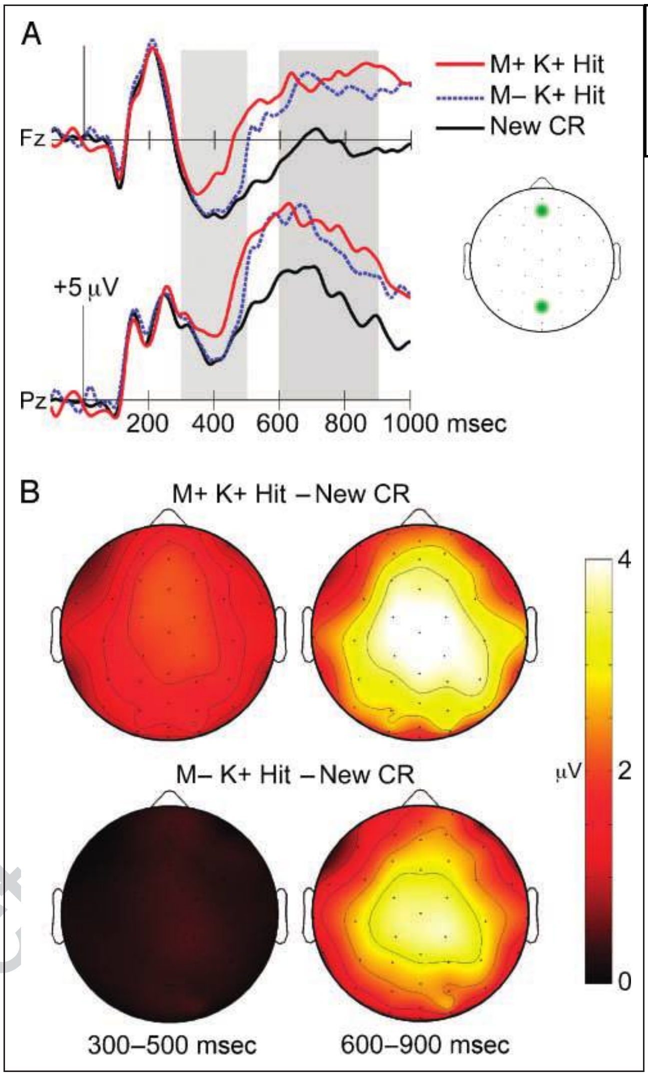

Figure 3. Neural correlates of recognition with and without concomitant conceptual implicit memory. ERP results are compared for M+ and M− words restricted to those recognized with high-confidence know responses (M+ K+ hits and M− K+ hits) and for correctly rejected new words (new CR). ERPs for new words did not differ as a function of meaningfulness category (Figure 1). (A) Time-voltage plots for each condition are shown for electrode locations approximating Fz and Pz, indicated with green circles on the schematic diagram of a head viewed from above. (B) Old-versus-new ERP differences for the 300- to 500-msec interval (left) and for the 600- to 900-msec interval (right) are plotted topographically, as computed separately for the two old–new subtractions. Coloration from dark to light indicates increasing difference amplitude.

For the interval from 600 to 900 msec, main effects of condition were significant for all three electrode clusters: $F(1.8, 27.1) = 11.8$, $p < .001$, $F(1.9, 28.0) = 15.7$, $p < .001$, and $F(1.9, 28.3) = 10.1$, $p = .001$. Amplitudes were significantly more positive for M+ than for new ($p < .001$, $p < .001$, and $p = .001$, respectively). Likewise, amplitudes were significantly more positive for M− than for new ($p = .006$, $p < .001$, and $p = .003$, respectively).[4] Amplitudes did not differ significantly for M+ versus M− for any cluster ($p = .27$, $p = .47$, and $p = .26$, respectively).

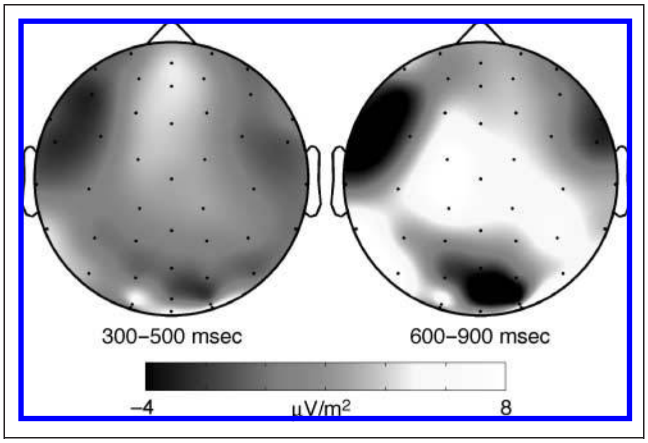

Figure 4. Current-source density (CSD) maps of M+ old/new ERP effects with high-confidence familiarity responses. Three-dimensional surface Laplacian transformations were applied to ERP difference waveforms to create spherical spline CSD maps (Kayser & Tenke, 2006). The distribution of the 300- to 500-msec CSD values differed significantly from the distribution of the 600- to 900-msec values, as indicated by a significant electrode-by-value interaction, $F(6.4, 95.7) = 5.2$, $p < .001$.

An assessment of the topographic distribution of the aforementioned ERP effects used the vector normalization approach (McCarthy & Wood, 1985). Averaged amplitude values from all scalp electrodes were compared for two conditions after overall amplitude differences were removed. The first comparison sought to determine if the M+ old/new effect differed spatially from 300 to 500 versus 600 to 900 msec. A marginal Electrode × Condition interaction, $F(4.5, 67.9) = 2.0$, $p = .09$, substantiated the observation that the effect tended to be more anterior for the earlier compared with the later interval (Figure 3B), and additional evidence for this anterior/posterior distinction was obtained via current-source density maps for the two intervals (Figure 4). Collectively, these distributional analyses indicated that the intervals (300–500 and 600–900 msec) captured contributions from distinct neural populations to old/new effects rather than merely showing activity for the same population at different intervals.

The second topographic comparison sought to determine if the 600- to 900-msec old/new effects differed spatially for M+ versus M− conditions. A nonsignificant Electrode × Condition interaction on vector-normalized amplitude values, $F(3.8, 56.4) = 1.4$, $p = .25$, indicated that the distributions did not differ reliably.

To summarize, statistical analyses substantiated visual impressions of the data. ERP old/new effects for familiar versus new items included LPC effects at 600–900 msec that were highly similar in amplitude and topography for M+ and M− words. FN400 effects at 300–500 msec were evident only for M+ words. Given the reasoning for comparing these conditions, LPC effects can be attributed to familiarity, which was produced by both types of old word, and FN400 effects to conceptual priming, which was shown to be selective for M+ words.

An alternative interpretation is that FN400 effects reflected familiarity rather than conceptual priming and that high-confidence familiarity was stronger for M+ versus M− words. For instance, the neural events responsible for familiarity may have unfolded earlier for M+ words than for M− words, appearing in the 300- to 500-msec interval only for M+ words. We assessed this possibility by identifying ERPs that varied with familiarity strength but for which subjective meaningfulness was matched. ERPs for high-confidence know responses (K+) were compared with ERPs for a combined low-familiarity condition (K−) comprising misses, medium-confidence know responses, and low-confidence know responses. ERPs for M+ and M− words were assessed separately to account for conceptual processing (Figure 5; M+ analyses were conducted without data from four subjects and M− analyses without data from one subject because of insufficient trial counts; average trial counts for included subjects

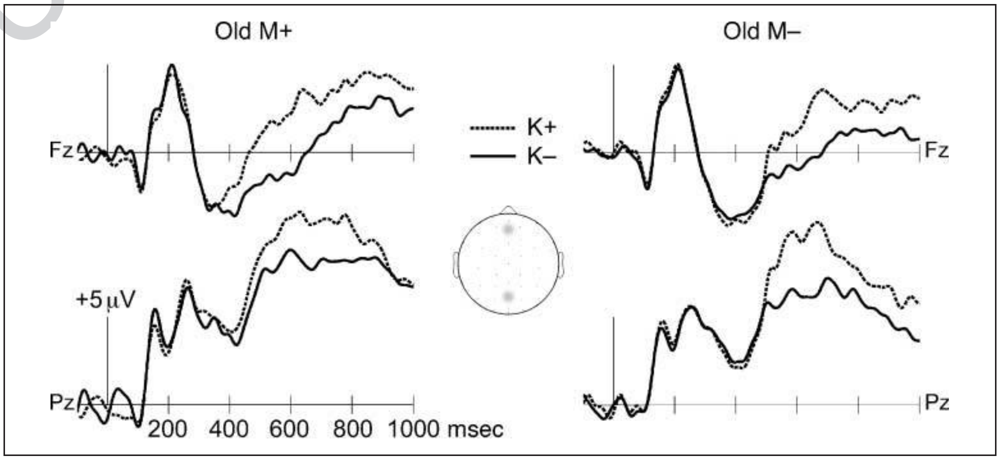

Figure 5. Familiarity strength modulates LPC amplitudes for both meaningfulness categories. ERP results are shown for the strong familiarity condition (K+, high-confidence know responses to old items) and for the weak familiarity condition (K−, collapsing medium-confidence know, low-confidence know, and new responses to old items). K+/K− contrasts were conducted separately for old M+ and old M− words. Waveforms are shown for electrode locations approximating Fz and Pz, indicated with gray circles on the schematic diagram of the head.

for the K+ M+, K− M+, K+ M−, and K− M− conditions were 40, 35, 30, and 39, ranges 22–55, 20–55, 20–56, and 17–68, respectively). For both meaningfulness categories, ERP differences between K+ and K− categories resembled the LPC effects found in old/new comparisons (Figure 3).

Formal assessments of amplitudes over the 600- to 900-msec latency interval for each of the three electrode clusters indicated that ERPs for the K+ category were reliably more positive than ERPs for the K− category; main effect of condition, M+, $F(1, 11) = 7.8$, $p = .02$, M−, $F(1, 14) = 10.8$, $p = .005$; nonsignificant Condition × Cluster interaction, M+, $F(2, 22) = 1.1$, $p = .35$, M−, $F(1.3, 18.1) = 2.6$, $p = .12$. Visual inspection of ERP waveforms indicated that ERPs for the K+ and K− categories diverged earlier for M+ than for M−. Indeed, no differences were apparent for M− words before 500 msec, whereas differences for M+ words were reliable from 400 to 500 msec, $t(11) = 3.3$, $p = .007$, averaged for the three electrode clusters. No K+/K− differences were apparent from 300 to 400 msec, and this latency interval was therefore excluded from statistical assessments. This finding is consistent with the notion that fluency can reduce the latency of ERP recognition effects (e.g., Woollams, Taylor, Karayanidis, & Henson, 2008). In this case, LPC correlates of familiarity occurred earlier for items subject to conceptual fluency effects. The ERP results summarized in Figure 5 are important because they show that the FN400 differences between M+ and M− words emphasized in Figure 3 were not due to subtle differences for these two categories in the level of familiarity captured by know responses; variations in familiarity strength were related to variations in LPC amplitude, not FN400 amplitude. ERPs in Figures 3 and 5 thus show a striking dissociation between familiarity-based recognition and FN400 potentials; when M− words were endorsed with highly accurate familiarity responses, no FN400 effects were exhibited either in a standard old/new contrast (Figure 3) or in a contrast between confident-familiarity and low-familiarity conditions (Figure 5).

To determine if ERP correlates of recollection for obscure words in the present study match ERP correlates of recollection described in the extant literature for other stimulus categories, we computed ERPs for old words endorsed with remember responses (Figure 6). Because of low trial counts, M+ and M− words were considered together, and data from 11 subjects were included (average trial counts for the remember and new conditions for included subjects were 33 and 96, ranges 19–69 and 44–168, respectively). ERPs were more positive for old words than for new words, and the old–new effects can be characterized as comprising FN400 and LPC effects, as are frequently found in association with recollection (reviewed earlier). Moreover, these effects closely resembled old–new effects for M+ words endorsed with K+ responses (Figure 3). These findings show that recollected obscure words elicit electrical signals comparable with those attributed to recollection in prior studies. The current design was not suitable for determining the influence of conceptual priming on ERP correlates of recollection because M+ and M− words could not be considered separately and because recollection was likely much stronger for M+ words than for M− words due to the source-specifying information provided by unique meaningful associations for M+ words.

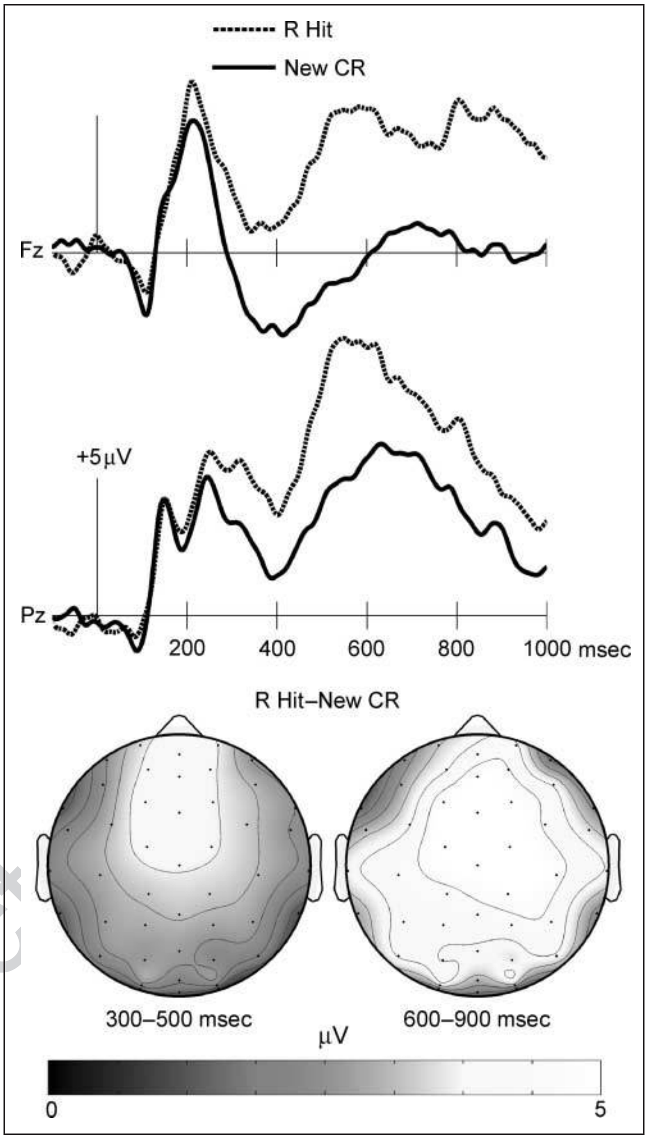

Figure 6. Neural correlates of recollection. ERPs for remember responses (R) were computed by collapsing the M+ and M− old word categories and are displayed along with ERPs for correctly rejected new words (also collapsing M+ and M− categories) at the same anterior and posterior recording sites used in previous figures. Old-versus-new ERP differences for the 300- to 500-msec interval and for the 600- to 900-msec interval are plotted topographically.

GENERAL DISCUSSION

Understanding implicit and explicit memory requires a valid characterization of their distinct neural substrates. The difficulties generally associated with separating neural correlates of familiarity from those of conceptual priming were surmounted in the present research by taking several novel steps. First, we used a large set of obscure words in the English language and, on the basis of individualized assessments of meaningfulness, included only those that were not in subjects' lexicons. Further, from the remaining words, we found that only those words that elicited meaningful associations (M+ words) were capable of supporting conceptual priming. As a whole, these M+ words also led to stronger explicit memory compared with M− words. However, sets of M+ and M− words were endorsed with familiarity-based recognition responses of equally high confidence and were therefore equated in explicit-memory strength. By focusing on these trials, neural correlates of conceptual priming were dissociated from those of familiarity.

ERP old–new effects corresponding to high-confidence familiarity included FN400 potentials only for M+ words, reflecting concurrent conceptual priming for these stimuli. Although a recognition test does not include any behavioral measures to show whether conceptual priming occurs concurrently, results from Experiment 1 showed that conceptual priming occurs selectively for repeated M+ words. The reasoning underlying our approach would be weakened if M+ and M− words in this contrast were not actually matched in explicit memory (e.g., if subjects recollected on M+ trials but then mistakenly registered a high-confidence familiarity response, if different types of familiarity occurred for M+ versus M− items, or if the high-confidence familiarity were systematically inaccurate in other ways). However, these concerns can be dismissed on the basis of results from analyses focused on ERP correlates of familiarity in the present data set. In particular, variations in familiarity strength for M− words were found to be associated with variations in LPC amplitude, not FN400 amplitude (Figure 5). The results from these experiments thus indicate that FN400 potentials constituted electrical correlates of conceptual priming, whereas LPC potentials from 500 to 700 msec were associated with familiarity-based recognition.

Contrasting neural correlates of these two types of memory were obtained by virtue of the fact that accurate explicit recognition can occur with or without stimulus meaning, whereas stimulus meaning is a prerequisite for conceptual priming. Here, behavioral responses revealed subsets of M+ and M− words that were matched in explicit familiarity. These stimuli were recognized one the basis of surface-level features, without contextual retrieval, as recollective trials were categorized separately. For M+ words, explicit retrieval of the perceived meaning initially elicited during the meaningfulness rating task would tend to result in a remember response. Indeed, subjects were explicitly instructed to give "remember" responses to stimuli for which a previously assigned meaning was recalled, and we found that these responses were more prevalent for M+ than for M− words. A reasonable assumption is thus that high-confidence familiarity responses were based on feelings of familiarity (noetic awareness of repetition; cf. Tulving, 1985) rather than on the subjective experience of meaningfulness, which is consistent with our finding that the prevalence and accuracy of high-confidence familiarity responses did not vary with meaningfulness.

The present results demand a reappraisal of prior studies that have garnered support for dual-process models of recognition from associations between FN400 potentials and familiarity. For words or other meaningful stimuli, repetition can induce both familiarity and conceptual priming, such that FN400 potentials cannot be ascribed to either type of memory in the absence of further evidence, preferably direct behavioral evidence for a dissociation between familiarity and conceptual priming (Paller et al., 2007). On the basis of the assumption that only meaningful stimuli can support conceptual priming, several studies attempted to record electrophysiological correlates of familiarity by using minimally meaningful stimuli such as pseudowords (Curran, 1999) and abstract geometrical shapes (Groh-Bordin, Zimmer, & Ecker, 2006; Curran, Tanaka, & Weiskopf, 2002). In these studies, FN400 potentials appeared to reflect familiarity, and results were thus taken as support for dual-process models. However, these investigators assumed that a lack of consensus stimulus meaning was sufficient to eliminate conceptual processing and thus to prevent conceptual priming from contaminating neural measures of familiarity. The results of the present study challenge this assumption. It is very likely that subjects in previous studies found meaning in stimuli that investigators presumed to be meaningless, especially given that subjects could improve their performance in anticipated memory tests by doing so. Subjects in memory experiments are remarkably adept in finding meaning in their own way when viewing to-be-remembered abstract or minimalist stimuli (Voss & Paller, 2007). Therefore, FN400 potentials to such stimuli in recognition tests could have reflected conceptual priming rather than familiarity. Conceptual priming with obscure words was examined in Experiment 1 under circumstances comparable with those used for recognition testing in Experiment 2. Notably, the same study tasks were used in both cases. We thus posit that the facilitated conceptual processing that drives conceptual priming with obscure words is also operative during recognition testing.

In convergence with these views, we recently found that abstract geometric shapes could be perceived as meaningful in idiosyncratic ways and that conceptual priming was indexed by FN400 potentials under these circumstances (Voss, Schendan, & Paller, n.d.; Voss & Paller, 2007). Together with these prior results, the present results using minimally meaningful words indicate that neural correlates of conceptual priming can contaminate those of explicit memory during recognition testing. This conclusion appears to hold both for abstract geometric shapes and for verbal materials akin to those commonly used in memory research.

Dual-process models of recognition have gained substantial support from associating recollection and familiarity with LPC and FN400, respectively (reviewed in Eichenbaum et al., 2007; Rugg & Curran, 2007; Mecklinger, 2006). The present results suggest that in prior ERP studies of memory using verbal stimuli—which constitute the bulk of the existing literature—neural correlates of conceptual priming were misattributed to familiarity. Indeed, a recent study that used both verbal stimuli (words and pictures) and nonverbal stimuli (dot patterns and spatial gratings) during a working memory task found FN400 effects selectively for the verbal stimuli, which could presumably support conceptual priming, whereas both stimulus types produced LPC effects (Danker et al., 2008). Familiarity in the current experiment was associated with LPC potentials, suggesting that familiarity and recollection produce largely similar electrical signals (see also Voss & Paller, 2008a, 2009; Paller et al., 2007).

Many findings in amnesic patients accommodate the dual-process perspective (e.g., Holdstock et al., 2002; Yonelinas et al., 2002; Duzel, Vargha-Khadem, Heinze, & Mishkin, 2001) by suggesting that recollection depends on the integrity of the hippocampus whereas familiarity is supported by the adjacent neocortex of the parahippocampal gyrus (Aggleton & Brown, 2006). Other evidence casts doubt on this simple dichotomy, however, by indicating that the hippocampus is critical for both recollection and familiarity (Wais et al., 2006; Wixted & Squire, 2004). In healthy subjects, fMRI measures have often indicated that recollection recruits greater activity within the hippocampus than does familiarity, whereas familiarity is more directly tied to activity within adjacent neocortex (e.g., Montaldi et al., 2006; Yonelinas, Otten, Shaw, & Rugg, 2005; Ranganath et al., 2004; Davachi, Mitchell, & Wagner, 2003; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000). However, patterns of activity associated with both recollection and familiarity often encompass both the hippocampus and the adjacent neocortex to different degrees for each memory type (Wais, 2008), again casting doubt on any clear dichotomy between corresponding neural substrates. Furthermore, a possibility suggested by the present results and by a critical examination of the ERP literature is that fMRI correlates of recognition in part reflect other co-occurring conceptual memory processes (Voss & Paller, 2008a; Paller et al., 2007). Future fMRI studies of recognition should consider a wider variety of mnemonic phenomena, including conceptual priming.

In conclusion, by dissociating electrophysiological correlates of conceptual priming and familiarity during recognition, we have shown that neural events associated with both types of memory are operative during recognition testing. Moreover, differential findings for conceptual priming and familiarity have critical implications for dual-process models of recognition as well as for determining the extent to which conceptual fluency can function to influence performance in recognition tests. The finding that familiarity can produce an electrical signature similar to that frequently attributed to recollection raises serious doubts about other evidence previously used in many studies of contrasts between recollection and familiarity, including results that have been used to support dual-process models. These findings thus constitute a step forward in constructing valid models of the neural substrates of these distinct forms of memory.

APPENDIX A

List of the 400 Uncommon Words Used as Stimuli.

abacinate
abra
abscind
abstruse
accipiter
accubation
acoria
aculeate
acumen
addle
adenia
adipic
adze
aestival
brogan
buccal
burke
byre
byssus
cache
caecity
cafard
calamus
calver
camber
camorra
canard
cang
drail
drung
dryad
ecbolic
eclat
egress
elapid
eldritch
empasm
enchorial
entasis
epicure
epopt
esker
hoyden
hyoid
hypural
ichnite
ignavia
immix
incanous
incept
incult
indurate
introrse
inumbrate
issles
jamb
monture
morass
morkin
mucin
mulm
nadir
nasard
nascent
naze
nepenthe
nidus
noxal
obelus
ocarina
robur
romage
rosin
ruelle
sagacity
sapid
sarment
scalary
scelerat
sedilia
seine
sellate
sepiment
seric

APPENDIX A *(continued)*

afferent
agapet
agger
ague
alacrity
alar
albedo
algid
alkanet
allodic
almuce
alpaca
alvine
ament
amity
anatine
anguine
anneal
anthelion
antiphon
arefy
argol
arras
artifice
asperity
assibilate
aval
avaunt
avowal
baculus
baft
bagnio
bannock
barton
bavin
bechic
beloid
bema
benthic
betel
canthus
capias
carnifex
catena
cavil
cenoby
cerberic
chantry
chaton
choller
cimex
cippus
cladose
claver
cleg
cloaca
coeval
coffle
coif
colubrine
columella
comate
comestible
conatus
copacetic
corf
cornice
corrody
cortege
crampon
crebrous
crotal
cruet
culet
cupel
curtate
cylix
dalliance
dammar
dapatical
essive
eutexia
exarate
facula
falcate
falderal
fanal
fane
fardage
farrow
farthing
fenestral
ferity
filature
fipple
flacon
flench
foehn
foin
forel
forfex
fossor
frisson
frith
frottage
fulgent
fusain
gadoid
galliard
gavage
gemel
gerent
gingham
glozing
glyptic
godet
goral
gramary
gricer
grilse
jarta
jerid
jib
jorum
jubate
jupon
kale
kalpis
kedge
kermes
kincob
knoll
knout
knubble
krobylos
kurgan
kyrie
lacis
lagan
lambent
laniary
larkspur
larrup
latrant
leal
lepid
leporine
levant
lictor
limn
limous
littoral
livedo
lorgnon
ludic
mabble
mabsoot
macrural
maffick
manal
olamic
onager
operculum
operose
orant
osculation
pabulum
paean
palabra
palfrey
palinola
palisade
pangram
parnel
patulous
pavis
pedicle
pergola
pericope
phaeton
phratry
pinder
pinguid
pizzle
placoid
plangent
podex
pone
praxis
prolate
prosaic
provender
prow
pugilist
pullet
purloin
putcher
quab
quarion
quirt
sessile
shindle
sideral
simity
skelder
slade
solander
storge
strepent
strop
sural
swage
sward
synod
tabor
tain
tapis
tarn
telary
telson
tephra
tewel
thane
thionic
tomium
tonsure
tornote
trammel
trellis
unction
undine
urman
ustion
vapulate
varec
venatic
vestal
visard
vitriol
wanigan

APPENDIX A *(continued)*

bibelot
bifid
bilious
birl
blench
blissom
blunge
blype
bodge
bollard
brayer
bream
bricole
brio
dapifer
darby
debel
deckle
dedans
delf
demotic
desinent
desudation
diadem
didact
dight
dinic
doxastic
gurlet
gybe
habile
halation
hallux
hamartia
haslock
hebamic
henotic
hent
hiant
hircine
hopple
hornito
manus
maremma
marl
mascaron
matinal
maunder
menald
metis
metope
mica
minaret
minatory
mogadore
monad
raceme
rachis
ragmatical
raguly
raiment
ranarian
rebus
recreant
rectrix
remora
reptant
rhumb
ridotto
rillet
wimple
witan
yapness
yarling
zatch
zule

Acknowledgments

The authors thank Ian M. Finn for providing comments on an earlier draft of this manuscript. This material is based on work supported by the National Science Foundation under grant numbers BCS-0518800 and BCS-0818912. J. L. V. was supported by an NIH grant T32AG20506.

Reprint requests should be sent to Joel L. Voss, Beckman Institute, 405 North Mathews Avenue, Urbana, IL 61801, or via e-mail: joelvoss@illinois.edu.

Notes

1. The same old/new behavioral results were identified when new items were subdivided into M+ and M− categories: remember M+ old versus new, $t(15) = 6.7$, $p < .001$; remember M− old versus new, $t(15) = 3.7$, $p = .002$; high-confidence know M+ old versus new, $t(15) = 13.7$, $p < .001$; and high-confidence know M− old versus new, $t(15) = 9.1$, $p < .001$.
2. The same old/new behavioral results were obtained when new items were subdivided into M+ and M− categories: medium-confidence know M+ old versus new, $t(15) = 2.2$, $p = .04$; medium-confidence know M− old versus new, $t(15) = 3.6$, $p = .003$; low-confidence know M+ old versus new, $t(15) = -5.8$, $p < .001$; low-confidence know M− old versus new, $t(15) = -4.7$, $p < .001$; new-response M+ old versus new, $t(15) = -7.8$, $p < .001$; and new-response M− old versus new, $t(15) = -7.7$, $p < .001$.
3. The same old/new ERP results for 300–500 msec were obtained when new items were subdivided into M+ and M− categories. For the frontal, central, and posterior clusters, pairwise old/new comparisons for M+ yielded p values of .001, .001, and .004, respectively, and pairwise old/new comparisons for M− yielded p values of .14, .18, and .09, respectively.
4. The same old/new ERP results for 600–900 msec were obtained when new items were subdivided into M+ and M− categories. For the frontal, central, and posterior clusters, pairwise old/new comparisons for M+ yielded p values of <.001, <.001 and .003, respectively, and pairwise old/new comparisons for M− yielded p values of <.001, <.001, and .001, respectively.

REFERENCES

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences, 10,* 455–463.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia, 37,* 771–785.

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective & Behavioral Neuroscience, 2,* 1–18.

Danker, J. F., Hwang, G. M., Gauthier, L., Geller, A., Kahana, M. J., & Sekuler, R. (2008). Characterizing the ERP old–new effect in a short-term memory task. *Psychophysiology, 45,* 784–793.

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, U.S.A., 100,* 2157–2162.

Duzel, E., Vargha-Khadem, F., Heinze, H. J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 8101–8106.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30,* 123–152.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience, 3,* 1149–1152.

Friedman, D., & Johnson, R., Jr. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51,* 6–28.

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49,* 87–115.

Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition, 19,* 617–623.

Groh-Bordin, C., Zimmer, H. D., & Ecker, U. K. (2006). Has the butcher on the bus dyed his hair? When color changes modulate ERP correlates of familiarity and recollection. *Neuroimage, 32,* 1879–1890.

Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., et al. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus, 12,* 341–351.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Kayser, J., & Tenke, C. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology, 117,* 348–368.

Lucas, H. D., Voss, J. L., & Paller, K. A. (2007). *Characterizing neural correlates of familiarity and conceptual priming*. Paper presented at the annual meeting of the Organization for Human Brain Mapping, Chicago, IL.

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

Mandler, G. (2008). Familiarity breeds attempts. *Perspectives on Psychological Science, 3,* 390–399.

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

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience, 37,* 292–299.

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus, 16,* 504–520.

Paller, K. A. (2004). Electrical signals of memory and the awareness of remembering. *Current Directions in Psychological Science, 13,* 49–55.

Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences, 11,* 243–250.

Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1070–1074.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42,* 2–13.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475–543.

Rugg, M. D., & Curran, T. (2007). ERPs and recognition memory. *Trends in Cognitive Sciences, 11,* 251–257.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience persepective. *Trends in Cognitive Sciences, 7,* 313–319.

Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron, 20,* 185–195.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82,* 171–177.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8,* 872–883.

Tulving, E. (1985). Memory and consciousness. *Canadian Journal of Psychology, 26,* 1–12.

Voss, J. L., & Paller, K. A. (2007). Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. *Learning & Memory, 14,* 259–267.

Voss, J. L., & Paller, K. A. (2008a). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. *Neuropsychologia, 46,* 3021–3029.

Voss, J. L., & Paller, K. A. (2008b). Neural substrates of remembering: Electroencephalographic studies. In J. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vol. 3, pp. 79–97). Oxford: Elsevier.

Voss, J. L., & Paller, K. A. (2009). Remembering and knowing: Electrophysiological distinctions at encoding but not retrieval. *Neuroimage, 46,* 280–289.

Voss, J. L., Schendan, H. E., & Paller, K. A. (n.d.). *Brain potentials that differentiate perceptual and conceptual priming*. Manuscript submitted for publication.

Wagner, A. D., Gabrieli, J. D., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 305–323.

Wais, P. E. (2008). fMRI signals associated with memory strength in the medial temporal lobes: A meta-analysis. *Neuropsychologia, 46,* 3185–3196.

Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron, 49,* 459–466.

Whittlesea, B. W., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica (Amsterdam), 98,* 141–165.

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science, 17,* 799–806.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114,* 152–176.

Wixted, J. T., & Squire, L. R. (2004). Recall and recognition are equally impaired in patients with selective hippocampal damage. *Cognitive, Affective, and Behavioral Neuroscience, 4,* 58–66.

Woldorff, M. G., Liotti, M., Seabolt, M., Busse, L., Lancaster, J. L., & Fox, P. T. (2002). The temporal dynamics of the effects in occipital cortex of visual–spatial selective attention. *Brain Research: Cognitive Brain Research, 15,* 1–15.

Woollams, A. M., Taylor, J. R., Karayanidis, F., & Henson, R. N. (2008). Event-related potentials associated with masked priming of test cues reveal multiple potential contributions to recognition memory. *Journal of Cognitive Neuroscience, 20,* 1114–1129.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language, 46,* 441–517.

Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauve, M. J., Widaman, K. F., et al. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience, 5,* 1236–1241.

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,* 3002–3008.