Brain potentials associated with recollective processing of spoken words

BRIAN GONSALVES and KEN A. PALLER
Northwestern University, Evanston, Illinois

The neurocognitive foundations of recollection can be explored by recording event-related potentials (ERPs) from the human brain. In the present study, we monitored brain activity while participants heard a series of words, first in a study phase and again, 1-2 min later, in a test phase, when both priming and recognition were measured. Level of processing at study was manipulated within-subjects via instructions either to visualize the referent of each word (the *image* task) or to detect the presence of target letters (the *letter* task). Priming of lexical decision response time was observed but did not differ across study task, whereas recognition was better for image- than for letter-task words. Brain potentials recorded at test revealed a *task effect*, wherein ERPs were more positive for image- than for letter-task words approximately 600-900 msec after word onset. The task effect was restricted to posterior scalp locations and was interpreted as an indication of visual imagery triggered by spoken words. Given that similar potentials were also elicited at study, we speculate that accurate recognition of words from the image task involved the recapitulation of the visual imagery that was initially engaged during the study phase.

One way an individual can recognize that a stimulus has been perceived before is by retrieving details of the episode in which the initial perception occurred. The subjective experience of remembering, *conscious recollection*, is often cued when details of the spatiotemporal context of an earlier episode are retrieved. Recollection can also be provoked by records of the cognitive processing engaged during the initial experience (M. K. Johnson, Hashtroudi, & Lindsay, 1993). In particular, the use of visual imagery at encoding and again during retrieval enhances memory for verbal items (Paivio, 1992). Moreover, in an experiment in which individuals judged their subjective experience of remembering, using the *remember/ know* procedure (Gardiner, 1988), recollection was more likely when visual details of the initial experience could be retrieved (Dewhurst & Conway, 1994).

To achieve a better understanding of the cognitive processes associated with recollection, it would be helpful to be able to monitor them as they occur. To this end, neural activity associated with cognitive processing can be observed by recording event-related potentials (ERPs) produced by the brain. ERPs recorded from electrodes on the scalp reflect the collective response of groups of neurons,

the activity of which is synchronized to particular cognitive events (Allison, Wood, & McCarthy, 1986; Kutas & Dale, 1997). ERPs provide measures with a temporal resolution that is sufficient for examining such memory functions and that is unsurpassed by functional neuroimaging techniques, such as positron emission tomography and functional magnetic resonance imaging.

ERPs have been used extensively to investigate recognition memory for items previously presented in an experimental context (for reviews, see R. Johnson, 1995; Paller, 1993, in press; Rugg, 1995). In many experiments, ERPs elicited by studied words were more positive than ERPs elicited by unstudied words at a latency of approximately 400–800 msec after word onset. Some investigators have distinguished between *early* and *late* portions of this ERP repetition effect (Smith & Halgren, 1989; Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991). The early portion is thought to be a modulation of the N400 component of the ERP (which has been studied in a variety of linguistic paradigms; see Kutas & Van Petten, 1988, for a review), whereas the late portion is thought to be a potential index of recognition memory. Other investigators have identified additional subcomponents of ERP repetition effects (e.g., R. Johnson, Kreiter, Russo, & Zhu, 1998). Indeed, ERP repetition effects can be conceived of as conglomerations of multiple repetition-sensitive effects that can vary independently across experiments. For example, recent results have suggested that repetition priming of visual word-form produces an early portion of the ERP repetition effect distinct from the N400 component (Paller & Gross, 1998; Paller, Kutas, & McIsaac, 1998).

Accordingly, it is critical to take the contrast between recollection and priming into account in ERP studies of memory. *Priming*, in this context, refers to altered per-

formance for studied items on certain *implicit memory tests*, in which no reference is made to prior study episodes. The idea that priming and recollection depend on different neural processing is supported by neuropsychological evidence showing preserved priming in amnesic patients (Schacter, 1987; Shimamura, 1986; Squire & Knowlton, 1995). The analysis of ERP effects solely on the basis of repetition is problematic, because both recollection and priming differ between studied and unstudied items. Furthermore, contrasting ERPs across implicit and explicit memory tests is also problematic, because of various confounding factors (Badgaiyan & Posner, 1997; Paller, 1993).

Paller and Kutas (1992) attempted to avoid these pitfalls by investigating ERP correlates of incidental recollection during an implicit memory test. A levels-of-processing manipulation at study was used to dissociate priming and recollection, based on prior reports of memory dissociations achieved in this manner (e.g., Jacoby & Dallas, 1981). At study, participants read words and made either a size judgment based on visualizing the referent of each word (*image task*) or a decision concerning the number of *e*s in each word (*letter task*). Subsequently, masked words were presented in a word identification test. Identification was more accurate for studied than for unstudied words but did not differ as a function of study task. In contrast, recognition tested later was more accurate for words from the image task. Likewise, ERPs recorded during the word identification test showed an enhanced positivity for words from the image task relative to those from the letter task. Given that study task had a strong influence on recognition but not on priming, the task effect on test phase ERPs was interpreted as a reflection of incidental recollection that was greater for words from the image task than for words from the letter task. The use of an implicit memory test in this experiment was important because behavioral responses in this test did not differ as a function of study task, thus controlling for priming effects and response-related factors that might otherwise have confounded the critical ERP comparisons.

Paller, Kutas, and McIsaac (1995) subsequently extended these findings, using a different priming measure, lexical decision response time. An ERP correlate of recollection was again found, and furthermore, it was larger for participants who were required to follow each lexical decision with a recognition judgment than for those assigned only the implicit memory test. This experimental approach is thus an effective means for isolating ERP correlates of recollective processing.

The present study extends the findings of Paller and Kutas (1992) and Paller et al. (1995) and uses a similar levels-of-processing manipulation at study to dissociate recollection and priming. Whereas visually presented word stimuli were used previously, we tested the generalizability of these findings by using auditorily presented word stimuli. Priming was measured, using lexical decision response time, and recognition responses were ob-

tained for each word as well. Although this compound test does not constitute an implicit memory test, it has the advantage of providing measures of both types of memory at the same retention delay. Furthermore, robust ERP correlates of recollection were found when the same sort of compound test was used by Paller et al. (1995, Experiment 2). On the basis of behavioral and ERP results, Paller et al. (1995) inferred that the added recognition requirement led participants to allocate attention to remembering whether each word was recently encountered and that this occurred prior to each lexical decision response. In the present experiment, we predicted that recollection would likewise influence ERPs during the lexical decision portion of the compound test. The ERPs of interest were, therefore, those elicited during the lexical decision test in response to two classes of studied items, words from the letter task and the image task. We predicted that priming, as indexed by lexical decision response time, would not differ between the two classes of studied words and that ERP effects related to the response and to priming would thus be matched. On the basis of this reasoning, the differential brain responses that were recorded for the two classes of studied words were attributed to a process associated with the experience of recollection.

METHOD

Participants

Six men and 6 women, 18–28 years of age, were recruited from the Northwestern University community. All were right-handed, native English speakers with no history of epilepsy or other neurological disease and no recent use of psychoactive medications. Two additional participants were excluded from the study for failure to complete the experiment (owing to personal discomfort in one case and equipment problems in the other).

Stimuli

The stimuli were 285 words and 95 pseudowords spoken by one of the experimenters. The words were nouns from an on-line psycholinguistic database (Coltheart, 1981; including imageability norms from Gilhooly & Logie, 1980; Paivio, Yuille, & Madigan, 1968; Toglia & Battig, 1978). Imageability and concreteness ratings for each word were greater than 500 (mean imageability = 585, mean concreteness = 588). The mean written frequency of the words was 54.5 occurrences per million (Kučera & Francis, 1967). The pseudowords were created by rearranging syllables of English words. Words and pseudowords were recorded digitally at a sampling rate of 22 kHz. The mean stimulus duration was 503 msec (*SE* = 5 msec) for words and 524 msec (*SE* = 7 msec) for pseudowords. Another set of 15 words and 5 pseudowords was used in the practice phase of the experiment and was not used again.

Procedure

After electrodes were attached (see below), each participant was led to a sound-attenuating chamber and was seated in a chair about 140 cm from a computer monitor and a pair of speakers. Communication with the experimenter was possible throughout the experiment via an intercom system. To reduce ERP artifacts, the participants were instructed to relax muscles, to blink as infrequently as possible during experimental runs, and to minimize body movement and eye movement. Once acceptable bioelectric signals were observed, the participants received instructions and completed a short practice

version of each task. Responses were registered with two buttons, one held in each hand. The participants were allowed to repeat these practice runs until they felt comfortable with each task.

The experiment consisted of six runs, each of which consisted of two study tasks and a test phase. Brief instructions were presented on the screen prior to each task to review specific response requirements. The study tasks will be referred to as the *image task* and the *letter task*. In the image task, the participants were required to visualize the referent of each word and decide whether the object was larger or smaller than the computer monitor in front of them. In the letter task, they were required to decide whether each word contained the letters *d* or *t*, responding with one hand if one or both were present and with the opposite hand if neither was present. The order of the image and letter tasks was counterbalanced across the six runs of the experiment. In each run, 15 words were presented in each task at a rate of one word every 2 sec.

The test phase began immediately upon completion of the study tasks and consisted of 60 items in random order: the 15 words heard in the image task, the 15 words heard in the letter task, 15 unstudied words, and 15 pseudowords. Two responses were required for each item. First, the participants made a lexical decision, indicating whether the item was a word or a pseudoword. Then, a tone presented 2 sec after the onset of each word signaled the recognition stage. The participants had up to 2.5 sec to make their recognition response, indicating whether a studied or an unstudied word had been heard. No recognition response was required for pseudowords. The next item was presented 500 msec after the recognition response was registered or, for pseudowords, 1,500 msec after the tone.

Assignment of words to conditions was counterbalanced across participants, so that each word appeared equally often in image, letter, and unstudied conditions. The hand associated with each response type in the test phase was also counterbalanced across participants. Upon completion of the experiment, the participants were debriefed and paid for their participation.

Electrophysiological Recording

The electroencephalogram (EEG) was recorded by using tin electrodes embedded in an elastic cap. The following 21 scalp electrodes from the International 10–20 system (Jasper, 1958) were used: Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, and T6. Recordings were referenced initially to the left mastoid and digitally rereferenced off line to the average of the left and right mastoid. The electrooculogram (EOG) was recorded by using an electrode placed beneath the right eye referenced to the left mastoid and electrodes placed lateral to each eye referenced to one another. Signals were amplified with a 0.1–100 Hz bandpass and were digitized at a rate of 250 samples/sec continuously. For averaging, data were resampled at a rate of 125 samples/sec. Each trial consisted of a 100-msec prestimulus baseline and a 1,940-msec period after the onset of the stimulus. Data were examined for EOG and muscle artifacts, and trials with such artifacts were removed using an automated procedure (an average of 13.6% of the trials were rejected). ERPs were initially quantified using mean amplitude measurements at each electrode location over various latency ranges. Measurements were also made for a negative peak between 200 and 1,000 msec at Cz and a positive peak between 500 and 1,300 msec at Oz. These peaks were labeled on the basis of the usual convention of polarity (P or N) and approximate peak latency. Results were consistent across mean amplitude and peak amplitude measurements, so for simplicity the latter are emphasized here. Peak amplitude and latency measurements were made following a low-pass filter at 10 Hz, to avoid measurement of spurious peaks caused by high-frequency noise. Each measurement was subjected to a repeated measures analysis of variance (ANOVA), with condition (unstudied, letter task, and image task) as a factor

and an alpha level of .05. Significance levels were adjusted by using the Greenhouse–Geisser correction, when appropriate (indicated by p_{GG}). For significant main effects of condition, pairwise comparisons were conducted using the Bonferroni correction for multiple comparisons. The *p* values given for these comparisons are the Bonferroni adjusted values.

RESULTS

Behavioral Performance

At study, the participants were 91% accurate in the letter task and 85% accurate in the image task (although the nominally correct response for some words in the image task was debatable). Mean response times were 1,122 msec (SE = 36 msec) for the letter task and 1,076 msec (SE = 27 msec) for the image task.

At test, repetition priming was observed in measures of lexical decision response times for correct trials, as is shown in Figure 1A. The mean difference in response times to studied and unstudied words was 76 msec (SE = 13 msec), a significant priming effect [$t(11)$ = 5.77, p < .001]. Response times for the two types of studied words were also assessed, and as predicted, the mean difference between image-task and letter-task words (12 msec; SE = 10 msec), was nonsignificant [$t(11)$ = 1.29, p = .22].
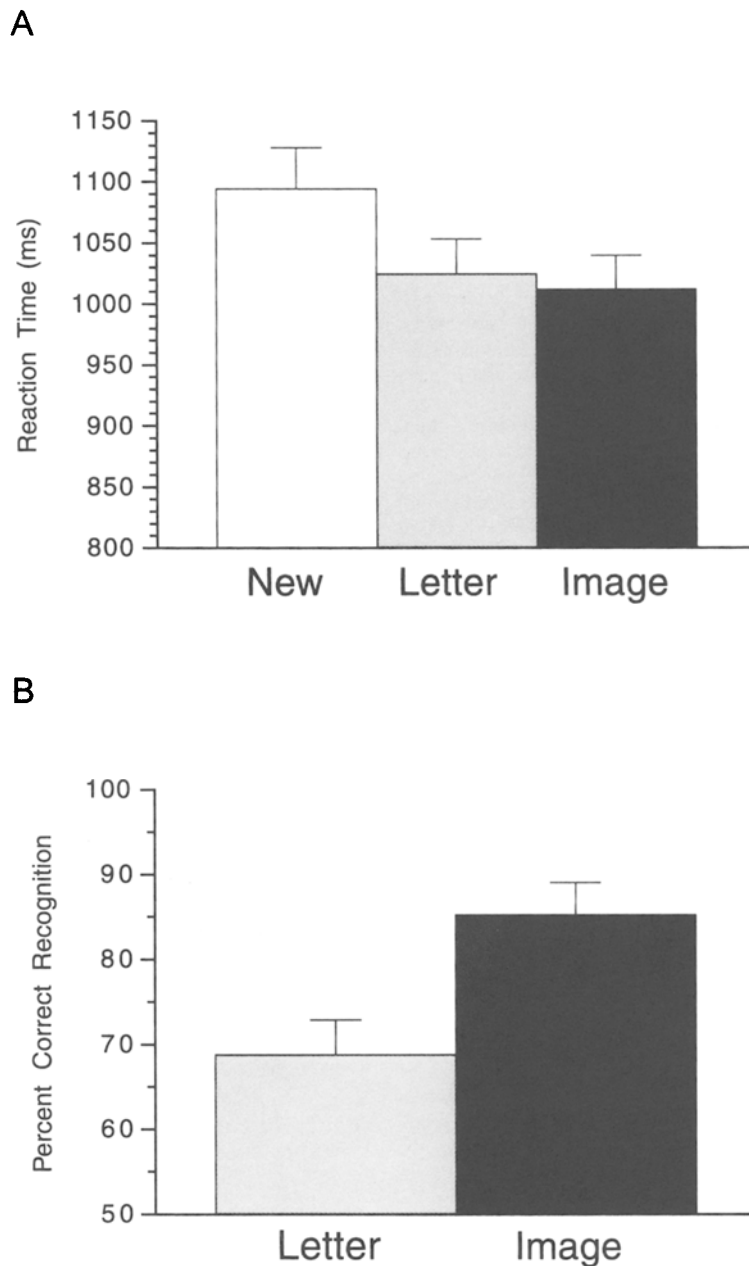
Response accuracy for lexical decisions averaged 94% (SE = 0.8%) for unstudied words and 98% (SE = 0.3%) for studied words, a significant difference [$t(11)$ = 4.78, p = .001]. Response accuracy did not differ significantly for studied words as a function of study task [$t(11)$ < 1]. Responses to pseudowords were correct 95% of the time (SE = 0.8%). For correct trials, mean response time was 1,036 msec (SE = 30 msec) for words and 1,084 msec (SE = 28 msec) for pseudowords.

Mean accuracy in the recognition test is shown in Figure 1B. The difference in recognition accuracy between words from the letter and image tasks was 16.4% (SE = 2.5%), a difference that was significant [$t(11)$ = 4.45, p = .001]. The correct rejection rate for unstudied words was 78% (SE = 4.7%).

Event-Related Potentials

N600 in test phase. ERPs to unstudied, letter-task, and image-task words presented during the lexical decision portion of the test phase are shown in Figure 2. ERP differences between studied and unstudied words were evident at the latency of a large negative deflection labeled N600. Because N600 in the grand average was maximal at central scalp locations, peak amplitudes and latencies were measured at the central midline electrode (Cz). Across subjects, the mean peak latency of N600 was 566 msec for unstudied words (SE = 39 msec), 617 msec for letter-task words (SE = 32 msec), and 608 msec for image-task words (SE = 30 msec), but these differences were not significant [$F(2,22)$ = 1.69, p = .22].

An analysis of N600 peak amplitude yielded a main effect of condition [$F(2,22)$ = 8.97, p_{GG} = .003], which was

A

B

Figure 1. A, mean lexical decision response time for unstudied words, studied words from the letter task, and studied words from the image task. B, mean recognition accuracy for studied words as a function of study task.

followed up by pairwise comparisons. N600 to unstudied words was reliably more negative than N600 to letter-task words [difference = 1.75 μV, SE = 0.57 μV; $t(11)$ = 3.07, p = .03] and image-task words [difference = 1.83 μV, SE = 0.37 μV; $t(11)$ = 4.89, p = .001]. The amplitude of N600 did not differ between image-task and letter-task words [difference = 0.01 μV, SE = 0.50 μV; $t(11)$ < 1].

Occipital P820 in test phase. ERP differences between letter-task and image-task words were apparent at the latency of a positive deflection that was maximal at occipital sites, labeled P820 on the basis of mean latencies of the peak for image-task and letter-task words (Figure 2). Analyses focused on P820 amplitude and latency as measured from the midline occipital electrode (Oz), but comparable results were found at nearby electrodes and for mean amplitudes over various latency ranges.[1] P820 latency at Oz varied by condition [$F(2,22)$ = 7.63, p_{GG} = .004], reflecting a longer latency for unstudied words (928 msec, SE = 28 msec) than for letter-task words (840 msec, SE = 40 msec) and image-task words
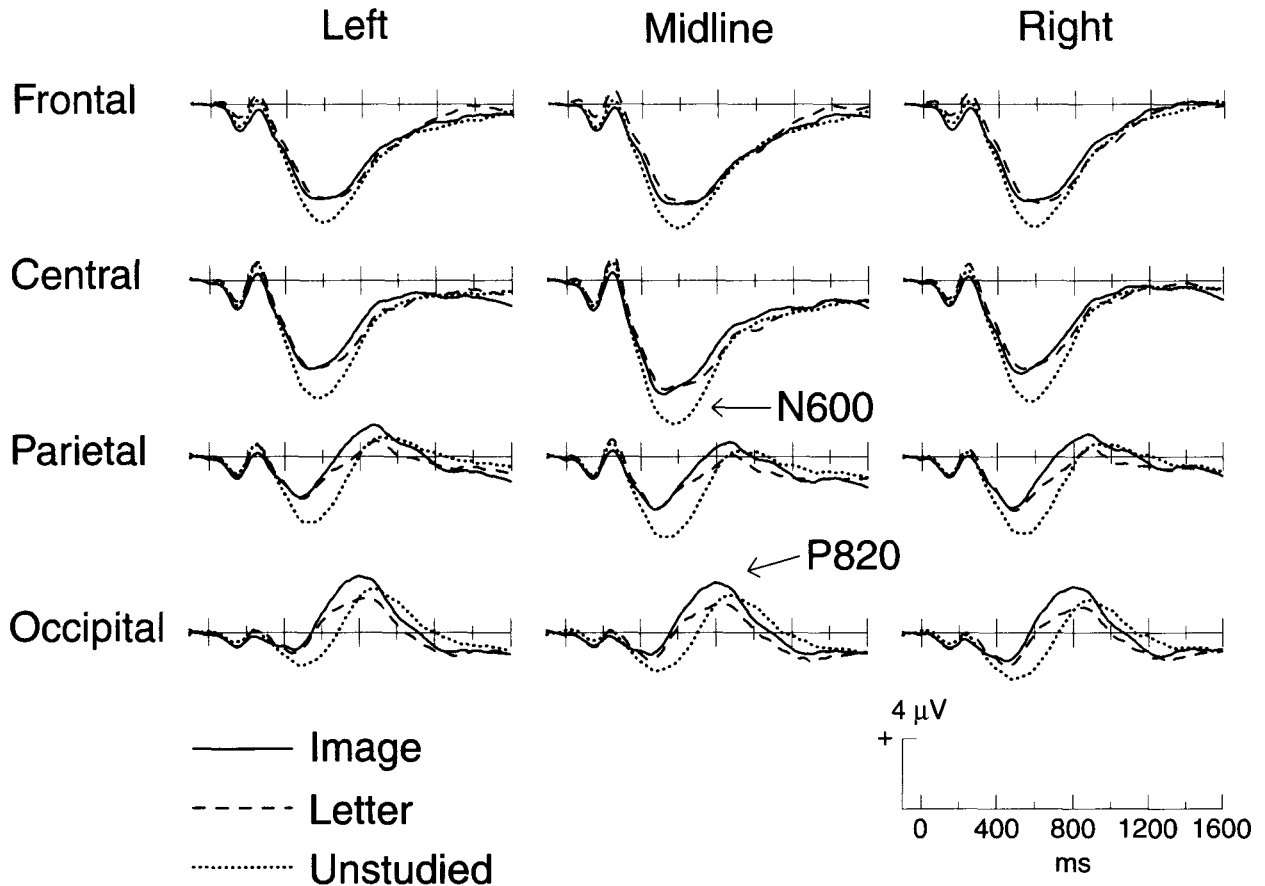
Figure 2. Event-related potentials to image, letter, and unstudied words presented during the lexical decision test. Plots represent voltage in microvolts as a function of time in milliseconds (time 0 = word onset).

(802 msec, SE = 21 msec). Pairwise comparisons revealed that only the difference between unstudied and image-task words (difference = 126 msec, SE = 29 msec) was significant [$t(11)$ = 4.34, p = .004].

Most important, P820 amplitude was found to vary across condition [$F(2,22)$ = 5.52, p_{GG} = .014], in that peak amplitudes to image-task words (3.67 μV, SE = 0.69 μV) were largest, followed by unstudied words (3.05 μV, SE = 0.94 μV) and letter-task words (2.53 μV, SE = 0.71 μV). Pairwise comparisons revealed that only the difference between image-task and letter-task (difference = 1.14 μV, SE = 0.29 μV) was significant [$t(11)$ = 3.99, p = .006].

The occipital task effect at test is shown again in the lower portion of Figure 3, including difference waves computed by subtracting the ERP to words from the letter task from the ERP to words from the image task. To obtain more precise information about the time-course of the task effect, mean amplitude measurements were taken over consecutive 20-msec intervals from 500 to 1,000 msec at the Oz electrode and were submitted to one-way repeated measures ANOVAs. Significant main

effects of condition were found for all 12 intervals between 640 and 880 msec (ps < .05).

Occipital P970 in study phase. ERPs recorded during the image and letter tasks at study are shown in the upper portion of Figure 3. The amplitude and latency of a positive occipital peak, P970, were measured at the Oz electrode. P970 latency for words in the image task (974 msec, SE = 57 msec) and words in the letter task (961 msec, SE = 55 msec) did not differ [$F(1,11)$ < 1]. P970 amplitude varied reliably across condition [$F(1,11)$ = 8.83, p = .013], since peak amplitudes in the image task (4.10 μV, SE = 0.63 μV) were greater than those in the letter task (2.71 μV, SE = 0.67 μV).

Comparisons between occipital peaks at study and test. Comparisons between the study phase P970 and the test phase P820 (Figure 3) at Oz were made via two-way repeated measures ANOVAs, with phase and task as factors. For peak amplitude, there was a main effect of task [$F(1,11)$ = 14.92, p = .003], substantiating the prior findings that amplitudes were greater for image-task than for letter-task words. This effect was similar for study and test phases, in that the main effect of phase and the inter-
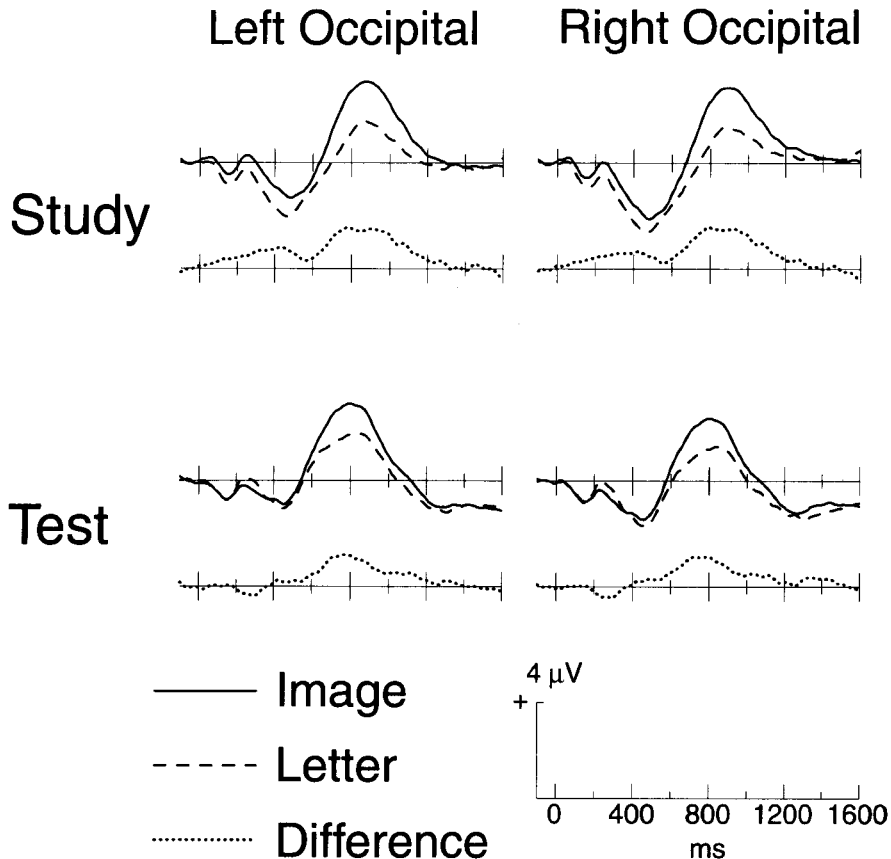
Left Occipital   Right Occipital

Figure 3. Occipital event-related potentials (ERPs) recorded during the study phase and during the lexical decision portion of the test phase to words from the image task and the letter task. Difference waves computed by subtracting letter-task from image-task ERPs are also shown.

action of phase and task were not significant ($Fs < 1$). For peak latency, there was a main effect of phase, indicating that peak latencies were shorter at test (821 msec, $SE = 27$ msec) than at study [968 msec, $SE = 35$ msec; $F(1,11) = 16.23, p = .002$]. There was no main effect of task and no interaction of phase and task for the latency data ($Fs < 1$), indicating that a similar latency shift from study to test was present for both conditions. A comparison between the peak latency of the occipital potential elicited by unstudied words in the test phase (928 msec) and that elicited by words in the study phase, collapsed across both tasks (968 msec), showed that they were not significantly different [$t(11) = 1.04, p = .32$].

**Topographic comparisons between study and test.** To compare the scalp distribution of the ERP difference between image-task and letter-task words at study with that at test, mean amplitude measurements were taken over a 200-msec interval centered on the mean peak latencies in these two conditions (i.e., 870–1,070 msec for study ERPs and 720–920 msec for test ERPs). Mean amplitude differences (image ERP minus letter ERP) were computed for each electrode location and are displayed as topographic maps for the study phase and the test phase

(Figure 4). Difference measurements were also scaled according to the vector-length method (McCarthy & Wood, 1985) and then compared in a two-way ANOVA, with phase (study vs. test) and electrode (all 21 scalp sites) as factors. A phase × electrode interaction in this analysis can be interpreted as a difference in scalp distributions between study and test, independent of any amplitude differences. The interaction was not significant [$F(1,11) < 1$], indicating that the scalp distribution of the ERP differences did not differ across study and test phases.

DISCUSSION

The key electrophysiological result was that brain potentials to words presented for lexical decisions differed according to how those words had been processed earlier in the experiment. Specifically, the amplitude of a P820 potential at occipital scalp regions was larger for the image-task words than for the letter-task words. In addition, earlier ERPs differed between studied and unstudied words, as captured by measurements of N600. However, as was discussed above, interpretations of ERP differences between old and new items are problematic, be-

Study

-1.0μV   1.5μV
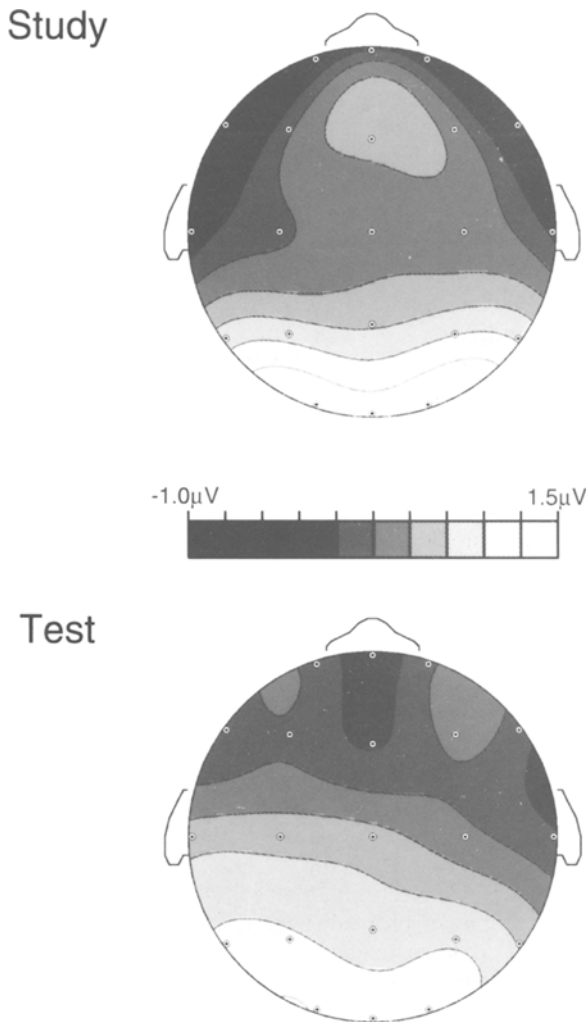
Test

Figure 4. Topographic representation of the event-related potential (ERP) task effect at study and at test. ERP amplitudes were measured between 870 and 1,070 msec for study phase ERPs and between 720 and 920 msec for test phase ERPs for each electrode site. Shaded regions represent the difference in microvolts for ERPs to words from the image task minus ERPs to words from the letter task, as computed by a surface spline interpolation.

cause these differences may reflect processes related to both priming and recollection. The present study was designed to eliminate this confound by employing a levels-of-processing manipulation at study and shifting the focus to comparisons between the two types of studied words.

The key behavioral results fulfilled our predictions: Priming levels, as indexed by lexical decision response time, did not differ as a function of study task, whereas recognition accuracy was greater for words from the image task than for words from the letter task. Accordingly, we conclude that the *ERP task effect*, the larger P820 for image-task words than for letter-task words,

was associated with differential recollective processing between the two conditions.

Put another way, the task manipulation in the study phase reliably influenced both recognition accuracy and occipital potentials recorded during the test phase, as captured by P820 amplitude measurements. Because this manipulation did not differentially influence lexical decision times, we can be confident that the ERP task effect is not an ERP correlate of perceptual priming or of latency-shifted response processes. Thus, a more rigorous association with recollection is possible with the ERP task effect than with ERP differences between old and new items.[2]

Although the ERP task effect in the present experiment was restricted to occipital scalp regions, corresponding ERP task effects in parallel studies in the visual modality were observed at virtually all scalp locations sampled, with the largest amplitudes near anterior regions (Paller & Kutas, 1992; Paller et al., 1995). Paller and colleagues hypothesized that the ERP task effects they recorded indexed brain processing associated with recollecting a recent experience with the same word. Other investigators have arrived at similar conclusions, using the remember/know procedure (Smith, 1993), by requiring participants to make source attributions for studied items (Senkfor & Van Petten, 1998; Wilding, Doyle, & Rugg, 1995; Wilding & Rugg, 1996) or by manipulating levels of processing at study (Rugg et al., 1998). By one account, ERPs related to recollection include a left-parietal effect reflecting retrieval of information about the study episode and a right-frontal effect reflecting processes associated with making source attributions about studied items (Wilding & Rugg, 1996). Other experiments have linked frontal ERPs with strategic processes engaged during memory retrieval (e.g., Ranganath & Paller, 1999). Frontal ERPs in the present experiment were remarkably similar across conditions. Presumably, strategic processing controlled by the prefrontal cortex was engaged to a similar extent for words from the image task, words from the letter task, and new words. However, several additional points must be raised before we can fully interpret the differences between results obtained using auditory versus visual word stimuli.

One potential ramification of investigations of ERP correlates of recollection is to provide a new means to study the component cognitive processes responsible for recollection. Undoubtedly, recollection can emerge as the result of multiple processes that can vary across different situations. Insights about retrieval processes in the present experiment can be gained by considering the topographic results. The ERP task effect was maximal over cortical regions responsible for processing and representing visual information, yet these ERPs were elicited by auditory stimuli presented while the visual environment remained constant. Nonetheless, we interpret the ERP task effect as a reflection of visual information processing associated with the generation of visual images.

This imagery hypothesis receives additional support from the striking similarities between occipital ERPs

recorded at study and at test (Figure 3). Study–test comparisons are complicated by the different sorts of processing engaged by the two tasks. In the letter task, the participants analyzed phonology and/or orthography in order to perform accurately, although it is likely that some semantic processing and visual imagery occurred as well. Performance in the image task principally required visual imagery, including a size comparison, so that there were multiple differences between the tasks. Yet, the most prominent ERP difference between the two study conditions was the discrepancy in P970 amplitude at occipital locations. The same sort of amplitude modulation was found in the test phase, when the task requirements differed from those in the study phase but were the same for the image and the letter conditions. In short, the most likely explanation for occipital peak amplitude differences between image and letter conditions, in both the study phase and the test phase, is that they reflect the differential engagement of visual imagery.

We thus propose that recognition performance was strongly determined by visual imagery during the study phase and that occipital ERPs provided an index of this imagery. When a word in the test phase provoked an individual to generate the same visual image that was generated a few minutes earlier, that word was more readily recognized as old (as in encoding specificity; Tulving & Thompson, 1973). Comparisons between study and test ERPs are also intriguing, because occipital peaks reached maximal amplitudes reliably earlier at study than at test (by 172 msec for image words and 121 msec for letter words). Speeded image generation may be responsible for this change in latency, so that P820 can be conceived of as a quicker version of P970. Even though both peaks may be composites of multiple overlapping components,[3] they are still useful for monitoring visual imagery. The idea that visual images were generated more quickly in the test phase for repeated words is reminiscent of the idea of perceptual fluency (e.g., Jacoby & Dallas, 1981), although fluent processing in this case concerns internally generated information rather than sensory information. This facilitation was presumably not the basis for either priming or familiarity in the present experiment but rather functioned as a principal cue supporting recollection of the study episode.

Another observation that is consistent with these interpretations was made with regard to the occipital ERP peak elicited by unstudied words in the test phase. The latency of this peak was longer than that for studied words. In fact, for words heard for the first time, occipital peak latency was similar in the study phase and the test phase, as if image generation required the same amount of time regardless of the task or phase of the experiment.

Evidence from neuroimaging also supports the idea that the visual cortex is active during visual imagery tasks (D'Esposito et al., 1997; Kosslyn et al., 1993; Le Bihan et al., 1995). In addition, activity in the visual cortex due to visual stimulation may interfere with activity required for visual image generation (see, e.g., Brooks, 1970;

Craver-Lemley & Reeves, 1987; Segal & Fusella, 1970; for reviews, see Craver-Lemley & Reeves, 1992; Kosslyn, 1994). ERP results also suggest that visual imagery and visual processing share, and thus compete for, the same processing resources in the visual cortex (Farah, Peronnet, Gonon, & Giard, 1988). Furthermore, brain potentials were associated with visual imagery in an experiment in which participants generated visual images for words presented in visual and auditory modalities (Farah, Peronnet, Weisberg, & Monheit, 1990). In both modalities, visual imagery, relative to a no-imagery control condition, led to an enhanced positivity, 600–1,100 msec poststimulus. This effect was maximal at occipital electrode sites and was larger for auditory stimuli, consistent with the results of the present experiment.

We can now return to the question of why an occipital task effect was found in the present experiment but not in previous experiments with nearly identical designs in the visual modality. First, ERPs in visual paradigms may have reflected various types of visual processing in such a way that superimposed ERP correlates of visual imagery were obscured. Second, we contend that visual imagery was much more robust in the present experiment than in prior experiments with visual stimulation because more resources for visual imagery were available.

In conclusion, we suggest that the participants in the test phase, upon hearing a word studied earlier in the image task, tended to reexperience the visual image generated at study and to use this experience as a basis for classifying the word as *old*. In fact, several participants reported just such an experience upon debriefing. A similar phenomenon probably occurred for words in the letter task, but to a lesser extent. Thus, retrieval of visual images in response to spoken words originally encoded under imagery instructions may be largely responsible for the enhanced recognition memory for these words. This particular levels-of-processing manipulation may be effective precisely because of the potency of participants' repeated experience of visual imagery, a process that was not apparent in their overt behavior but that was monitored by ERP measures. Behavioral and ERP evidence, in combination, can thus provide information about the nature of the cognitive processing responsible for participants' experience of recollection and about the timing of those processes. One such process is the recapitulation of visual imagery that promoted accurate recognition in the present experiment.

REFERENCES

ALLISON, T., WOOD, C. C., & McCARTHY, G. (1986). The central nervous system. In M. G. H. Coles, E. Donchin, & S. W. Porges (Eds.), *Psychophysiology* (pp. 5-25). New York: Guilford.

BADGAIYAN, R., & POSNER, M. I. (1997). Time course of cortical activations in implicit and explicit recall. *Journal of Neuroscience*, 17, 4904-4913.

BROOKS, L. R. (1970). An extension of the conflict between visualization and reading. *Quarterly Journal of Experimental Psychology*, 22, 91-96.

COLTHEART, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.

CRAVER-LEMLEY, C., & REEVES, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception*, **16**, 599-614.

CRAVER-LEMLEY, C., & REEVES, A. (1992). How visual imagery interferes with vision. *Psychological Review*, **99**, 633-649.

D'ESPOSITO, M., DETRE, J. A., AGUIRRE, G. K., STALLCUP, M., ALSOP, D. C., TIPPET, L. J., & FARAH, M. J. (1997). A functional MRI study of mental image generation. *Neuropsychologia*, **35**, 725-730.

DEWHURST, S. A., & CONWAY, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1088-1098.

DONCHIN, E., & COLES, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral & Brain Sciences*, **11**, 357-427.

FARAH, M., PERONNET, F., GONON, M., & GIARD, M. (1988). Electrophysiological evidence for a shared representational medium for visual images and percepts. *Journal of Experimental Psychology: General*, **117**, 248-257.

FARAH, M., PERONNET, F., WEISBERG, L., & MONHEIT, M. (1990). Brain activity underlying mental imagery: Event-related potentials during image generation. *Journal of Cognitive Neuroscience*, **1**, 302-316.

GARDINER, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, **16**, 309-313.

GILHOOLY, K. J., & LOGIE, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, **12**, 395-427.

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

JASPER, H. H. (1958). Report to the committee on methods of clinical examination in electroencephalography: Appendix. The ten–twenty system of the International Federation. *Electroencephalography & Clinical Neurophysiology*, **10**, 371-375.

JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.

JOHNSON, R., JR. (1995). Event-related potential insights into the neurobiology of memory systems. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 10, pp. 135-163). Amsterdam: Elsevier.

JOHNSON, R., JR., KREITER, K., RUSSO, B., & ZHU, J. (1998). A spatiotemporal analysis of recognition-related event-related brain potentials. *International Journal of Psychophysiology*, **29**, 83-104.

KOSSLYN, S. (1994). *Image and brain*. Cambridge, MA: MIT Press.

KOSSLYN, S., ALPERT, N., THOMPSON, W., MALJKOVIC, V., WEISE, S., CHABRIS, C., HAMILTON, S., RAUCH, S., & BUONNANO, S. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, **5**, 263-287.

KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

KUTAS, M., & DALE, A. (1997). Electrical and magnetic readings of mental functions. In M. D. Rugg (Ed.), *Cognitive neuroscience* (pp. 197-242). Cambridge, MA: MIT Press.

KUTAS, M., & VAN PETTEN, C. (1988). Event-related brain potential studies of language. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 3, pp. 139-187). Greenwich, CT: JAI Press.

LE BIHAN, D., TURNER, R., ZEFFIRO, T. A., CUENOD, C. A., JEZZARD, P., & BONNEROT, V. (1995). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, **90**, 11802-11805.

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

PAIVIO, A. (1992). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, **45**, 255-287.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monographs*, **76**(3, Pt. 2).

PALLER, K. A. (1993). Elektrophysiologische Studien zum Menschlichen Gedächtnis [Electrophysiological studies of human memory]. *Zeitschrift für Elektroenzephalographie, Elektromyographie und verwandte Gebiete*, **24**, 24-33.

PALLER, K. A. (in press). Neurocognitive foundations of human memory. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 40). San Diego: Academic Press.

PALLER, K. A., & GROSS, M. (1998). Brain potentials associated with perceptual priming versus explicit remembering during the repetition of visual word-form. *Neuropsychologia*, **36**, 559-571.

PALLER, K. A., & KUTAS, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between recollection and priming. *Journal of Cognitive Neuroscience*, **4**, 375-391.

PALLER, K. A., KUTAS, M., & MCISAAC, H. K. (1995). Monitoring conscious recollection via the electrical activity of the brain. *Psychological Science*, **6**, 107-111.

PALLER, K. A., KUTAS, M., & MCISAAC, H. K. (1998). An electrophysiological measure of priming of visual word form. *Consciousness & Cognition*, **7**, 54-66.

RANGANATH, C., & PALLER, K. A. (1999). Frontal brain potentials during recognition are modulated by requirements to retrieve perceptual detail. *Neuron*, **22**, 605-613.

RUGG, M. D. (1995). Event-related potential studies of human memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 789-801). Cambridge, MA: MIT Press.

RUGG, M. D., MARK, R. E., WALLA, P., SCHLOERSCHEIDT, A. M., BIRCH, C. S., & ALLAN, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, **392**, 595-598.

SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 501-518.

SCHACTER, D. L., & BUCKNER, R. L. (1998). Priming and the brain. *Neuron*, **20**, 185-195.

SEGAL, S. J., & FUSELLA, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology*, **83**, 458-464.

SENKFOR, A. J., & VAN PETTEN, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1005-1025.

SHIMAMURA, A. P. (1986). Priming effects in amnesia: Evidence for a dissociable memory function. *Quarterly Journal of Experimental Psychology*, **38A**, 619-644.

SMITH, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience*, **5**, 1-13.

SMITH, M. E., & HALGREN, E. (1989). Dissociation of recognition memory components following temporal lobe lesions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 50-60.

SQUIRE, L. R., & KNOWLTON, B. J. (1995). Memory, hippocampus, and brain systems. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 825-837). Cambridge, MA: MIT Press.

TOGLIA, M. P., & BATTIG, W. R. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.

TULVING, E., & THOMPSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.

VAN PETTEN, C., KUTAS, M., KLUENDER, R., MITCHINER, M., & MCISAAC, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience*, **3**, 129-150.

WILDING, E., DOYLE, M. C., & RUGG, M. D. (1995). Recognition memory with and without retrieval of context: An event-related potential study. *Neuropsychologia*, **33**, 743-767.

WILDING, E., & RUGG, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, **119**, 889-905.

NOTES

1. In one analysis, we compared mean amplitudes between image-task and letter-task ERPs over five 300-msec time windows, separately for sets of electrodes grouped by region. All tests were nonsignificant,

except for a main effect of condition at occipital sites (O1, Oz, O2) from 600 to 900 msec [$F(1,11) = 10.26, p = .008$] and a marginal main effect of condition at parietal and posterior temporal sites (T5, P3, Pz, P4, T6) from 600 to 900 msec [$F(1,11) = 4.74, p = .052$], both reflecting greater amplitudes for image-task words relative to letter-task words.

2. Levels-of-processing manipulations often affect conceptual priming (Schacter & Buckner, 1998), which raises the possibility that the ERP task effect could reflect brain processes associated with conceptual priming rather than conscious recollection. Nevertheless, neuroimaging studies have consistently shown that conceptual priming is associated with reductions of activation in the left inferior prefrontal cortex (reviewed by Schacter & Buckner, 1998), whereas ERP effects in the present experiment were restricted to occipital regions. A connection between the ERP task effect and conceptual priming thus seems unlikely.

3. One possibility is that P970 and P820 are related to the P300 potential that has been studied extensively (Donchin & Coles, 1988). However, there are several differences between P970/P820 and the typical characteristics of P300. Differences in scalp topography are obvious, since P300 is generally largest at central and parietal scalp locations. Moreover, pseudowords that were presented on 25% of the trials in the test phase elicited a positive ERP likely akin to P300, and the topography of this positive ERP was not occipital. Also, current accounts of the functional significance of P300 cannot easily explain the present results, whereas our imagery account provides a unitary explanation for occipital amplitude differences between conditions and for latency differences between study and test.