PSYCHOLOGICAL SCIENCE

Research Report

Gamma Oscillations Distinguish True From False Memories

Per B. Sederberg,[1] Andreas Schulze-Bonhage,[2] Joseph R. Madsen,[3] Edward B. Bromfield,[4] Brian Litt,[1] Armin Brandt,[2] and Michael J. Kahana[1]

[1]*University of Pennsylvania;* [2]*University Hospital of Freiburg, Freiburg, Germany;* [3]*Children's Hospital Boston, Boston, Massachusetts and* [4]*Brigham & Women's Hospital, Boston, Massachusetts*

ABSTRACT—*To test whether distinct patterns of electrophysiological activity prior to a response can distinguish true from false memories, we analyzed intracranial electroencephalographic recordings while 52 patients undergoing treatment for epilepsy performed a verbal free-recall task. These analyses revealed that the same pattern of gamma-band (28–100 Hz) oscillatory activity that predicts successful memory formation at item encoding—increased gamma power in the hippocampus, prefrontal cortex, and left temporal lobe—reemerges at retrieval to distinguish correct from incorrect responses. The timing of these oscillatory effects suggests that self-cued memory retrieval begins in the hippocampus and then spreads to the cortex. Thus, retrieval of true, as compared with false, memories induces a distinct pattern of gamma oscillations, possibly reflecting recollection of contextual information associated with past experience.*

Both lesion and functional neuroimaging studies of human memory implicate the hippocampus and surrounding medial temporal lobe (MTL) regions in the encoding and recollection of past events (Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Schacter & Wagner, 1999; Squire, 1992). A topic of great interest is whether neural activity in these regions can distinguish true from false memories.

False memories are thought to occur when one incorrectly recollects the source of a memory, failing to retrieve the temporal context in which the event occurred (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001). In laboratory studies, subjects often incorrectly recall nontarget items that are semantically related to a target item, have been studied on a recent prior list, or both (Roediger & McDermott, 1995; Zaromb et al., 2006). Whereas behavioral studies of false memory have largely focused on recall errors (intrusions) in free recall, neuroimaging studies have generally made comparisons between correct and incorrect recognition. These neuroimaging studies suggest that activity in many of the same MTL regions that are critical for episodic memory retrieval in general can also distinguish true from false recognition. Specifically, increases in hippocampal (Dobbins, Rice, Wagner, & Schacter, 2003) and parahippocampal (Cabeza et al., 2001; Dobbins et al., 2003) activation after the presentation of a memory probe have been shown to indicate true recollection of the contextual details of an encoding event.

However, the low temporal resolution and susceptibility to vocalization artifacts inherent in functional neuroimaging methods limit researchers' ability to understand the neurophysiological basis of self-cued recall of past experiences. Implanted multielectrode recordings, widely used in animal studies, can measure the electrical activity within small brain regions, and thereby characterize how the neural assemblies in those regions react to changes in the animal's behavioral or cognitive state. In humans, implanted electrodes are largely unaffected by artifacts due to vocalizations and are the only means of obtaining fine spatial and temporal resolution in recordings of electrical activity in deep structures, such as the hippocampus, whose activity is not discernible by means of magnetoencephalogram (MEG) or scalp electroencephalogram (EEG).

To uncover the electrophysiological correlates of successful memory formation and self-cued recall, we tested 52 patients with drug-resistant epilepsy who had arrays of subdural or depth electrodes (or both) surgically implanted for 1 to 2 weeks to localize the site or sites of seizure onset. The clinical team determined the placement of these electrodes with the goal of localizing suspected epileptogenic foci and identifying functional regions to be avoided in surgery.

Subjects first studied lists of common nouns. Then, after performing a brief distractor task, they attempted to recall the nouns in any order. To assess how brain oscillations differed between successful and unsuccessful memory formation, we

---

Address correspondence to Michael Kahana, Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, e-mail: kahana@psych.upenn.edu.

compared oscillatory power during presentation of words that were subsequently recalled and words that were not subsequently recalled. To uncover differences in brain oscillations between true and false memories, we compared oscillatory power prior to correct and incorrect recalls. These comparisons were made separately at each of the 3,677 electrodes and for six distinct frequency bands: 2 to 4 Hz (delta), 4 to 8 Hz (theta), 10 to 14 Hz (alpha), 16 to 26 Hz (beta), 28 to 42 Hz (low gamma), and 44 to 100 Hz (high gamma).

METHOD

We recorded intracranial EEG from 52 patients (ages 8 to 53; 22 female, 30 male) while they studied and recalled lists of words in a delayed free-recall task. The lists were composed of 15 or 20 common nouns, chosen at random and without replacement from a pool of high-frequency words (either English or German, depending on the subject's native language). Twenty-one subjects received 20-item lists, and the remaining 31 subjects received 15-item lists. Subsets of these patients were included in prior studies of how brain activity during encoding predicts subsequent recall (Sederberg, Kahana, Howard, Donner, & Madsen, 2003; Sederberg et al., 2007).

Our research protocol was approved by the appropriate institutional review boards, and informed consent was obtained from the subjects and their guardians. More detailed methods, along with supplementary results, are available on the Web at http://memory.psych.upenn.edu/publications/.

Data Processing

Vocal responses were digitally recorded and scored for analysis following each session (Sederberg et al., 2003, 2006, 2007). EEG signal during individual events (word presentation or retrieval) was scanned for artifacts (e.g., spikes), and an event was discarded if the kurtosis of the amplitude distribution of the signal exceeded a threshold of 5 (Delorme, Sejnowski, & Makeig, 2007). Data were notch-filtered at 50 or 60 Hz with a Butterworth filter with zero phase distortion, to eliminate electrical-line and equipment noise, and were down-sampled to 200 Hz. For all our EEG signals, we used the Morlet wavelet transform (with a wave number of 6) to compute spectral power at 46 logarithmically spaced intervals between 2 and 100 Hz as a function of time. We then log-transformed the power and down-sampled it to 50 Hz.

For EEG analysis of the retrieval events, any response that occurred within 2 s of the onset of a prior response was discarded, to avoid overlap with the prior vocalization. Furthermore, any subject who did not make at least six incorrect responses was discarded from the EEG retrieval analysis. After these criteria were applied, 32 subjects remained in the retrieval analysis.

Data Analysis

We used a Wilcoxon rank sum test for the two comparisons: between recalled and not recalled encoding events and between correct and incorrect recall events. For encoding, we tested for differences in the mean log-transformed wavelet power during the period from 0 to 2,000 ms after presentation onset. For retrieval, we tested for differences in mean log-transformed wavelet power in the 500 ms prior to a response. These comparisons were made separately for each electrode and at each frequency.

Because the large sample of electrodes provided widespread coverage of most brain regions (see Table 1 in the supplemental material available at http://memory.psych.upenn.edu/publications/), we were able to aggregate data across both subjects and electrodes to determine if there was a statistically significant effect within a given brain region (Sederberg et al., 2007). We thus calculated between-subjects statistics within each Brodmann's area (BA), as determined with the Talairach Daemon (Lancaster et al., 2000), and for the left and right hippocampal areas, as determined by the clinical team.

To calculate the significance of the power differences aggregated across subjects for specific regions, we performed a permutation test at each electrode and combined the resulting empirical distributions across all the electrodes in each region. This permutation procedure allowed us to control for multiple comparisons while maintaining a fixed Type I error rate (Efron, 1979; Gibbons & Shanken, 1987). We considered a region to exhibit a significant aggregate effect across subjects only if at least 5 subjects contributed electrodes to that region. Finally, we determined a single significance threshold by means of the false-discovery-rate method (with $\alpha = .1$), and then applied this threshold to all regions (Genovese, Lazar, & Nichols, 2002).

To determine the correlation between (a) the power differences between successful and unsuccessful encoding and (b) the power differences between correct and incorrect responses, we calculated Pearson's correlations (across all regions with electrodes from greater than 5 subjects) between the Z-transformed p values at encoding and retrieval. To test for differences between correlations of different frequency bands, we performed a Fisher's two-tailed test for the difference between two independent correlations.

RESULTS AND DISCUSSION

The 52 subjects recalled 23.2 ± 1.2% of the words on each list. Of the words recalled, 72.9 ± 2.4% were correct; the incorrect responses (intrusions) were either words from prior lists or words that were not presented (repetitions were excluded).

We found that increased gamma power in the hippocampus, left temporal lobe (BA 20, 21, 22, and 37), inferior prefrontal cortex (BA 47 and 11), and occipital lobe (BA 18 and 19) during encoding predicted whether the studied item would be subsequently recalled (Fig. 1, left column). A similar pattern of increased gamma power prior to a retrieval predicted whether the recalled item was on the just-presented target list (Fig. 1, right column). Gamma power increased in the hippocampus, left temporal lobe (BA 20 and 21), and inferior prefrontal cortex (BA 47 and 11) during the 500 ms prior to correct, relative to incorrect, recalls.

We then asked whether retrieval of a target item represents a recapitulation of the perceptual and cognitive processes that took place during successful encoding (Kahn, Davachi, & Wagner, 2004; Kolers & Roediger, 1984). To answer this question, we assessed the correlation between the oscillations predicting successful encoding and the oscillations predicting correct recall across the 48 brain regions sampled. The 44- to 100-Hz high gamma band exhibited a significant positive correlation (Pearson's $\rho = .68, p < .0001$), which indicates that the gamma power associated with successful encoding of list items reemerges with the retrieval of those items. The delta and low gamma bands also exhibited reliable, but smaller, positive correlations between these two effects (delta: $\rho = .30, p = .04$; low gamma: $\rho = .43, p = .002$), whereas correlations in the theta, alpha, and beta bands were not statistically significant. A direct comparison of the correlations in the different frequency bands (Fisher's test for the difference between two independent correlations) revealed that the correlation in the high gamma band was significantly greater than the correlations in the delta ($p = .0092$), theta ($p < .0001$), alpha ($p = .0048$), and beta ($p = .0001$) bands. Thus, more than activity in any other frequency band, gamma band activity, which was shown to be significant in the hippocampus, prefrontal cortex, and temporal lobe, accompanies both encoding and retrieval processes.

Because the placement of the electrodes was determined for clinical purposes, few subjects had electrodes in all regions of interest (in particular, the hippocampus, temporal lobe, and prefrontal cortex). Even so, we were able to harness the high temporal resolution of intracranial EEG to trace changes in power that differentiated true and false recall within individual regions. The earliest increases in gamma power differentiating correct from incorrect recall occurred in the right and left

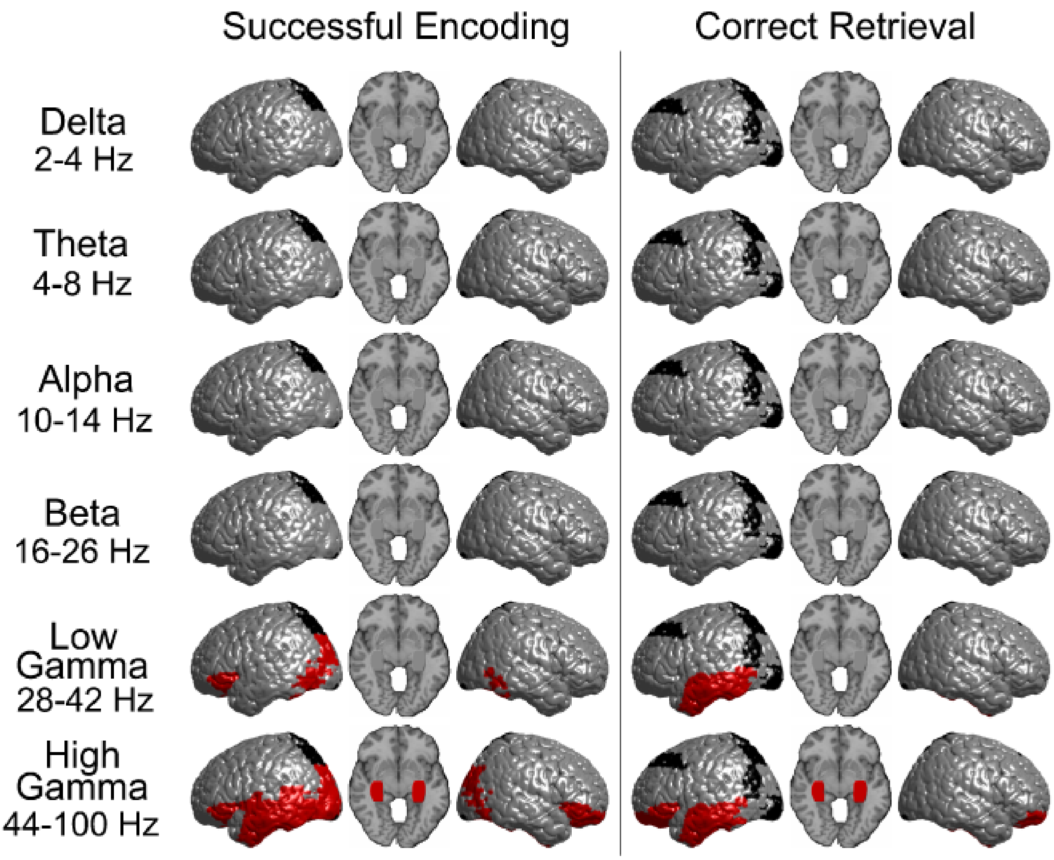

Fig. 1. Regions exhibiting significant increases in mean power during successful (relative to unsuccessful) encoding (left column) and prior to correct (relative to incorrect) recall (right column). Areas with significant increases are shown in red, and nonsignificant regions are shown in gray. Regions containing electrodes from fewer than 5 subjects are shown in black. Each row presents results for a distinct frequency band. For each comparison, the figure shows, from left to right, a left cortical view, a brain slice through the hippocampus, and a right cortical view. The cortical views present results for regions defined in terms of Brodmann's areas; the slices present results for the left and right hippocampus. Note that fewer subjects were included in the retrieval than in the encoding analysis, giving rise to the discrepancy in regions with fewer than 5 subjects.

hippocampus at approximately 500 ms and 400 ms prior to the response (Figs. 2b and 2a, respectively). This increase in hippocampal gamma activity was followed by increased gamma activity in the left temporal lobe and prefrontal cortex in the 250 ms prior to the response (Figs. 2c–2f). Thus, the neural activity uniquely associated with recalling an item that was present in the studied list (as compared with incorrectly recalling a nonpresented item) propagated from the hippocampus to temporal and prefrontal regions. This sequential timing of changes in gamma oscillations lends support to the hypothesis that episodic retrieval begins in the hippocampus and that neuronal activation then spreads to the cortex (Naya, Yoshida, & Miyashita, 2001).

We have shown that increased gamma oscillations immediately preceding a response distinguish true from false memories. Recall of a true memory is preceded by an increase in gamma oscillations in the hippocampus (bilaterally) and in the temporal and prefrontal cortices (primarily in the left hemisphere). The topography of the gamma-power increases predicting true recall matched the topography of the gamma-power increases predicting successful encoding. The similarity of these topographies suggests that neural mechanisms underlying successful encoding of an item into a list context reemerge at retrieval to distinguish items that actually were on the list from those that were not (Daselaar, Fleck, Prince, & Cabeza, 2006; Kahn et al., 2004; Osipova et al., 2006; Polyn, Natu, Cohen, & Norman, 2005; Slotnick & Schacter, 2004). Two recent functional magnetic resonance imaging studies have found a similar correspondence between brain activity at encoding and retrieval. Slotnick and Schacter (2004) proposed a "sensory reactivation hypothesis" based on their finding that visual cortical and hippocampal regions that are activated during sensory experience

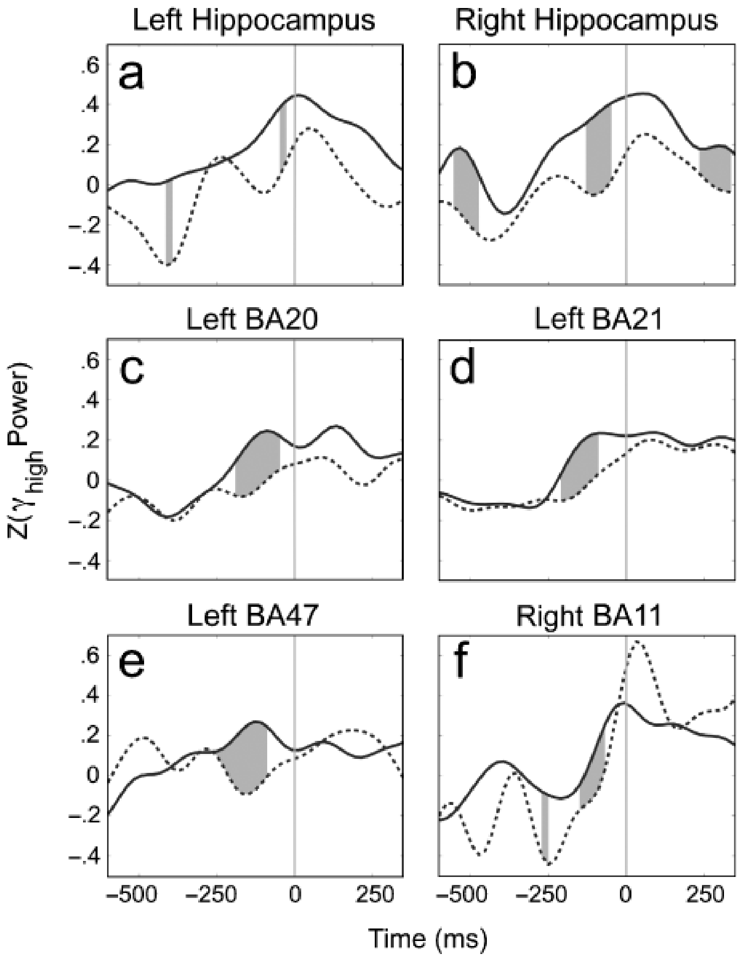

Fig. 2. Time course of mean Z-transformed power for correct (solid lines) and incorrect (dashed lines) recall. Mean power was calculated in the high gamma band for six distinct regions: (a) left hippocampus, (b) right hippocampus, (c) left inferior temporal lobe (Brodmann's area, BA, 20), (d) left lateral temporal lobe (BA 21), (e) left inferior prefrontal cortex (BA 47), and (f) right inferior prefrontal cortex (BA 11). Shaded areas indicate times when the power was significantly different between conditions. Time 0 ms indicates the onset of the response vocalization.

show increased blood-oxygenation-level-dependent (BOLD) activation during true recognition of shapes. Similarly, in a recent free-recall study, the distributed pattern of BOLD activity associated with studying words of a particular semantic category reemerged prior to recall of words from that category (Polyn et al., 2005).

There are at least two possible basic interpretations of the increases in gamma oscillations that predicted successful encoding and correct recall. These increases may have reflected a shift in overall brain state; that is, changes in oscillatory power may accompany brain states that are conducive to memory encoding or retrieval. For example, increased gamma oscillations could be associated with increased attention or arousal, which would support both memory encoding and retrieval processes (Sederberg et al., 2006; Tallon-Baudry, Bertrand, Henaff, Isnard, & Fischer, 2005; Womelsdorf, Fries, Mitra, & Desimone, 2006). Alternatively, the observed changes in brain oscillations may have reflected item-specific memory processes (Fabiani, Stadler, & Wessels, 2000; Gonsalves & Paller, 2000; Urbach, Windmann, Payne, & Kutas, 2005). For example, an increase in gamma oscillations may indicate successful storage and also successful retrieval of a conjunctive representation of the item's semantic information and the temporal context of the list (Howard & Kahana, 2002).

Acknowledgments—We thank J. Jacobs, K. Norman, S. Polyn, and M. van Vugt for helpful comments on the manuscript. We acknowledge support from the Swartz Foundation, National Institutes of Health (Research Grants MH61975, MH062196, and MH072138), and National Science Foundation (Grant SBE0354378).

REFERENCES

Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R., & Schacter, D.L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, USA*, *98*, 4805–4810.

Daselaar, S.M., Fleck, M.S., Prince, S.E., & Cabeza, R. (2006). The medial temporal lobe distinguishes old from new independently of consciousness. *Journal of Neuroscience*, *26*, 5835–5839.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*, 1443–1449.

Dobbins, I.G., Rice, H.J., Wagner, A.D., & Schacter, D.L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia*, *41*, 318–333.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, *7*, 1–26.

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.

Fabiani, M., Stadler, M.A., & Wessels, P.M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience*, *12*, 941–949.

Genovese, C.R., Lazar, N.A., & Nichols, T.E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878.

Gibbons, M.R., & Shanken, J. (1987). Subperiod aggregation and the power of multivariate tests of portfolio efficiency. *Journal of Financial Economics*, *19*, 389–394.

Gonsalves, B., & Paller, K.A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience*, *3*, 1316–1321.

Howard, M.W., & Kahana, M.J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299.

Kahn, I., Davachi, L., & Wagner, A.D. (2004). Functional-neuroanatomic correlates of recollection: Implications for models of recognition memory. *Journal of Neuroscience*, *24*, 4172–4180.

Kolers, P.A., & Roediger, H.L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*, 425–449.

Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*, 120–131.

Naya, Y., Yoshida, M., & Miyashita, Y. (2001). Backward spreading of memory-retrieval signal in the primate temporal cortex. *Science*, *291*, 661–664.

Osipova, D., Takashima, A., Oostenveld, R., Fernandez, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience*, *26*, 7523–7531.

Polyn, S.M., Natu, V.S., Cohen, J.D., & Norman, K.A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, *310*, 1963–1966.

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

Schacter, D.L., & Wagner, A.D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus*, *9*, 7–24.

Sederberg, P.B., Gauthier, L.V., Terushkin, V., Miller, J.F., Barnathan, J.A., & Kahana, M.J. (2006). Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage*, *32*, 1422–1431.

Sederberg, P.B., Kahana, M.J., Howard, M.W., Donner, E.J., & Madsen, J.R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, *23*, 10809–10814.

Sederberg, P.B., Schulze-Bonhage, A., Madsen, J.R., Bromfield, E.B., McCarthy, D.C., Brandt, A., et al. (2007). Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex*, *17*, 1190–1196.

Slotnick, S.D., & Schacter, D.L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*, 664–672.

Squire, L.R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Tallon-Baudry, C., Bertrand, O., Henaff, M.-A., Isnard, J., & Fischer, C. (2005). Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cerebral Cortex*, *15*, 654–662.

Urbach, T.P., Windmann, S.S., Payne, D.G., & Kutas, M. (2005). Mismaking memories: Neural precursors of memory illusions in electrical brain activity. *Psychological Science*, *16*, 19–24.

Womelsdorf, T., Fries, P., Mitra, P.P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, *439*, 733–736.

Zaromb, F.M., Howard, M.W., Dolan, E.D., Sirotin, Y.B., Tully, M., Wingfield, A., & Kahana, M.J. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 792–804.

(RECEIVED 12/22/06; ACCEPTED 2/26/07;
FINAL MATERIALS RECEIVED 3/14/07)