Published in final edited form as:

Neuropsychologia. 2013 October ; 51(12): 2344–2351. doi:10.1016/j.neuropsychologia.2013.03.019.

Human Hippocampal Increases in Low-Frequency Power during Associative Prediction Violations

Janice Chen^{a,b,g}, Mohammad Dastjerdi^{a,c}, Brett L. Foster^{a,c}, Karen F. LaRocque^{a,b,f},
Andreas M. Rauschecker^{a,b,d,e}, Josef Parvizi^{a,c,d}, and Anthony D. Wagner^{a,b,d}

^aStanford Human Intracranial Cognitive Electrophysiology Program (SHICEP), Stanford University

^bDepartment of Psychology, Stanford University

^cDepartment of Neurology and Neurological Sciences, Stanford University

^dNeurosciences Program, Stanford University

^eMedical Scientist Training Program, Stanford University

^fCenter for Mind, Brain, & Computation Training Program, Stanford University

^gPrinceton Neuroscience Institute, Princeton University

Abstract

Environmental cues often trigger memories of past events (associative retrieval), and these memories are a form of prediction about imminent experience. Learning is driven by the detection of prediction violations, when the past and present diverge. Using intracranial electroencephalography (iEEG), we show that associative prediction violations elicit increased low-frequency power (in the slow-theta range) in human hippocampus, that this low-frequency power increase is modulated by whether conditions allow predictions to be generated, that the increase rapidly onsets after the moment of violation, and that changes in low-frequency power are not present in adjacent perirhinal cortex. These data suggest that associative mismatch is computed within hippocampus when cues trigger predictions that are violated by imminent experience.

Keywords

hippocampus; associative memory; prediction violation; intracranial EEG; novelty; mismatch

1. Introduction

Hippocampus has long been known to play a central role in episodic memory (e.g., Squire, 1992; Eichenbaum & Cohen, 2001), being particularly important for learning and remembering novel events (e.g., Knight, 1996; Kirchoff, Wagner, Maril, & Stern, 2000).

© 2013 Elsevier Ltd. All rights reserved.

To whom correspondence should be addressed: Janice Chen, Princeton Neuroscience Institute, Green Hall, Princeton, NJ 08542, 650-814-3406, janice@princeton.edu. Anthony D. Wagner, Department of Psychology, Jordan Hall, Bldg 420, Stanford, CA 94305-2130, 650-723-4048, awagner@stanford.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

An emerging perspective posits that hippocampus retrieves episodic memories and compares past events to current experiences; hippocampal detection of a mismatch between past and present is thought to initiate responses in salience and reward processing regions that in turn project back to hippocampus, enhancing memory encoding of the present ‘unexpected’ event (Hasselmo & Wyble, 1997; Vinogradova, 2001; Lisman & Grace, 2005; Kumaran & Maguire, 2009).

The computation of mismatch between past and present requires (a) that cues trigger memory for related past event details (i.e., associative retrieval), where the retrieved memory is a form of prediction about imminent experience (e.g., Buckner, 2010; Chen, Olsen, Preston, Glover, & Wagner, 2011); and (b) a mechanism that detects and signals when associative predictions are violated by subsequent experience (e.g., Hasselmo & Wyble, 1997; Lisman & Grace, 2005). For example, if while watching television one finds that the evening local news program is followed by a particular sitcom, the next encounter with that news program might bring thoughts of the sitcom to mind, as the two were temporally related in the past. Learning about such co-occurrences is advantageous, because when events repeat (as they often do), an accurate memory of what happened in the past enables preparation for upcoming events. At other times, however, such predictions are violated, e.g., the news program is followed by a weather warning. Violated expectations indicate the presence of novelty—be it an unexpected event/item or an unexpected association between two events/items (i.e., both were previously encountered, but not together). In humans, functional MRI (fMRI) data indicate that hippocampal blood-oxygenation-level-dependent (BOLD) activity increases in the presence of associative novelty (Köhler, Danckert, Gati, & Menon, 2005; Kumaran & Maguire, 2006; Chen et al., 2011; Duncan, Ketz, Inati, & Davachi, 2012), and some data suggest that this effect depends on the generation of predictions (Kumaran & Maguire, 2006; Kumaran & Maguire, 2007).

A potentially confusing issue is the terminology surrounding the concept of prediction. The word ‘prediction’ is used here to describe a ‘ready’ state of a neural system (e.g., the hippocampus) that anticipates incoming stimuli (the ‘predicted’ input) based on the history of past inputs; the word ‘prediction’ is not intended to describe an overt, effortful, or explicit expectation or visualization. From our perspective, a subject’s experience of active/explicit expecting may accompany or result from the hippocampal predictive processes explored here, but those experiences are not central to the current investigation.

While BOLD data provide tentative support for a ‘hippocampal mismatch computation’ in humans, a critical test of the hypothesis requires increased precision about the timing of hippocampal responses to associative prediction violations. Prior fMRI findings are open to two interpretations—that mismatch computations originate in the hippocampus or, alternatively, that they originate in extra-hippocampal regions (e.g., prefrontal or parietal cortex) that modulate hippocampal responses. Moreover, because putative hippocampal mismatch computations may trigger a cascade of neural responses that in turn upregulate hippocampal encoding, BOLD measures could reflect a mixture of early violation-related and later encoding-related effects.

Human intracranial electroencephalographic (iEEG) studies have revealed that certain kinds of unexpected events – target stimuli in ‘oddball’ paradigms – evoke a negative potential in human hippocampus that onsets around 300 ms (e.g., Halgren et al., 1980; Halgren et al., 1995; Axmacher et al., 2010). In oddball tasks, subjects are presented with a continuous stream of stimuli (images or sounds); infrequently, target stimuli occur that differ from the majority in a salient dimension (such as color or category). Such targets are *contextually novel* in that the stream of stimuli, all matching along a given dimension (e.g., color), establish an ongoing state of expectancy that is violated when the target stimuli appear. By

contrast, *associative novelty* describes an unexpected association between two specific events/items; a prior association between two events/items enables one to serve as a predictor for the other, and a new association (e.g., the appearance of a different event/item) violates this prediction. Thus, associative novelty also may be described as associative prediction violation. It is not yet clear how hippocampal responses to contextual novelty and associative novelty are related, though it has been suggested that a comparator framework of hippocampal function may be able to encompass both (Kumaran & Maguire, 2007a). While hippocampal responses to contextual novelty have been explored via iEEG, relatively little is known about human hippocampal electrophysiological responses to associative novelty (associative prediction violation).

To directly examine human hippocampal neural activity during associative prediction violation, we used iEEG to record local field potentials (LFPs) during an associative memory sequence task that probed whether differential hippocampal activity: (a) appears rapidly post-violation, (b) is selective to the stimulus that violates the prediction, and (c) depends on the generation of associative predictions.

2. Materials and Methods

2.1 Participants

Seven patients with temporal lobe epilepsy participated in the experiment (Table 1). All patients gave informed consent under a protocol approved by the Stanford University Institutional Research Board. Patients were implanted with depth electrodes as part of a presurgical evaluation to localize their seizure onset zone; locations of implantations were determined based solely on clinical considerations. Participation in this experiment did not cause any additional risk to the patients.

2.2 Task Protocol

Participants viewed sequences of images on a laptop computer while performing an incidental 1-back task (modified from Kumaran & Maguire, 2006). The computer was placed on a table in front of the participant and the distance adjusted to their comfort. Button responses were made using an external ten-key pad.

2.3 Task Parameters

The image stimuli were color photos of objects centered on white backgrounds (Figure 1). Objects were drawn from a set of 752 unique images; participants viewed between 358 and 752 objects, depending on available time and tolerance. The task was run in 8-min blocks, with breaks between blocks at the participant's discretion. The order of images, as well as which images appeared in which experimental conditions, were randomized for each participant. This random assignment of stimuli to conditions and positions ensures that the reported results do not reflect item effects.

Objects appeared one at a time for 1000 ms each, with a 200-ms centered black fixation cross between stimuli. Stimuli were structured into trial-unique sequences of four-object quartets, and a 2000-ms centered red fixation cross followed each quartet. Experimental trials consisted of two quartets, with the first and second quartets containing the same four objects; an initial sequence ('quartet') of four objects (ABCD; mean # of trials = 82.6, s.d. 24.7) was followed by a second quartet, either: (1) in the same order ('Repeat': ABCD; mean # = 22.6, s.d. 10.2), (2) with the last two objects in reverse order ('Mismatch': ABDC; mean # = 25.1, s.d. 7.7), or (3) in a completely rearranged order ('Shuffle': CADB or BDAC; mean # = 21.4, s.d. 10.0). An additional condition consisted of a single quartet shown repeatedly (the same four images in the same order) throughout the experiment, with

the number of repeats dependent on the number of task blocks tolerated by the participant (mean # = 23.7, s.d. 9.3). All trial counts reflect the number of trials retained after artifact rejection (see 2.5 *iEEG Data Analysis*).

During the task, participants made 1-back judgments. Specifically, they were instructed to press a key when two objects in a row were identical. These ‘Target’ events occurred on 12% of quartets (mean # = 24.1, s.d. 7.2) and were evenly distributed across experimental conditions, across first and second quartets, and across object positions within quartets (i.e., positions 2–4). If a Target occurred in a first quartet, no second quartet was shown. Trials in which Targets occurred were coded as a separate condition during analyses.

2.4 iEEG Data Acquisition

LFP data were recorded using laterally penetrating depth electrodes (AdTech Medical Instrument Corporation; <http://www.adtechmedical.com>). These electrodes were cylindrical, 1.3 mm in diameter with $2\pi r(h + r) = 0.12\text{cm}^2$ of exposed recording area and inter-electrode distance of 5 mm. Data were collected at a sampling rate of 3051.8 Hz (high-pass 0.5 Hz, low-pass 300 Hz) through a 128-channel recording system (Tucker Davis Technologies; <http://www.tdt.com>) and analyzed offline using custom software in MATLAB (MathWorks). Trial onsets were recorded using photodiode event markers from the presentation laptop for each experimental trial time-locked with the iEEG recording.

2.5 iEEG Data Analysis

Raw data were notch-filtered at 60 Hz and harmonics to remove power line noise and re-referenced to a common average. The electrodes used for the common average in each patient were all electrodes inside or on the brain that were not deemed contaminated by machine noise, as judged by the recording team (depth and cortical electrodes, no scalp electrodes). The total number of electrodes for subjects 1–7 was 42, 55, 54, 8, 47, and 63, respectively. Artifact detection was implemented by an automatic procedure wherein timepoints were flagged if the notch-filtered data (1) exceeded 6 standard deviations from the signal mean, and/or (2) exceeded 5 standard deviations from the gradient mean, the change from one timepoint to the next (Fell, Ludwig, Rosburg, Axmacher, & Elger, 2008). Any experimental trials that contained flagged timepoints were removed before data were submitted to statistical tests (mean 36.8% trials, s.d. 16.1%). Data were inspected after automatic artifact detection in order to remove remaining noisy trials (mean 2.3%, s.d. 4.1%).

Event-related potentials were constructed by averaging voltage traces across all trials in a condition (that remained after artifact rejection), with the mean of the 100-ms period before stimulus onset subtracted from the voltage trace for each trial.

To assess signal power in specific frequency bands, 42 center frequencies were selected in a logarithmic distribution between 1 to 230 Hz and re-referenced data were bandpass filtered (with a flat-top gaussian) at each frequency. A Hilbert transform was then applied to each of the bandpass filtered time series and the analytic amplitudes squared to obtain instantaneous power (Canolty, Edwards, & Dalal, 2006), allowing event-related analyses of power in desired frequency bands; the resulting power timecourses were resampled to 1/7 the original rate (436 Hz) in order to reduce computing demands. Power values were transformed to percent signal change for each frequency band (% signal change = $((\text{power}/m) - 1) * 100$, where m = mean power over all timepoints in all trials in the recording block that were not removed during artifact detection).

Spectral power analyses were concentrated on low frequencies (< 10 Hz) and frequencies in the gamma range (~30–70 Hz), as prior human medial temporal lobe iEEG studies have

reported that modulations in these ranges are related to cognition (for review, see Jacobs & Kahana, 2010). In humans, signal power in the ‘slow-theta’ range (2.5–5 Hz) has been posited to be particularly relevant to memory-related computations as opposed to the ‘fast-theta’ range (5.5–10 Hz), analogous to the slow- vs. fast-theta distinction in non-human animals (Lega, Jacobs, & Kahana, 2012). In the current study, center frequencies for ‘slow-theta’ were 2.75, 3.63, and 4.79 Hz, selected from the predefined center frequencies to be within the 2.5–5 Hz range; for ‘fast-theta’, center frequencies were 5.75, 6.92, 8.32, and 9.12 Hz. Percent signal change values in each theta range were averaged across these center frequencies before being submitted to statistical analyses.

The distribution of mean power over subjects is theoretically Gaussian, following the central limit theorem. Each power value submitted to analyses was the average over a 500-ms window post-stimulus onset (for rationale see *3.1 Power Modulations in Hippocampus and PrC*), as well as an average over several center frequencies (see prior paragraph), as well as an average over multiple electrodes (e.g., all hippocampal electrodes for a given patient). Histograms of these values across subjects confirmed that the distributions were normal, justifying the use of parametric statistics.

A recording block was defined as a continuous recording session that lasted ~8 min; 3–7 blocks of task were collected from each participant. For all participants, all recording blocks were on the same day, except for one participant (S1) who was recorded on two different days (with one intervening day).

2.6 Electrode Selection

Electrodes of interest were selected conservatively based solely on their anatomical location – falling within hippocampus or peririnal cortex (PrC), as visible on post-surgical computed tomography images of electrodes aligned with structural magnetic resonance images of the brain for each participant (mean # of electrodes = 6.7, s.d. 2.1; see Table 1). Alignment was performed using a mutual-information algorithm implemented in SPM (<http://www.fil.ion.ucl.ac.uk/spm/>). PrC was defined as the gray matter in the fundus and lateral bank of the collateral sulcus, bounded anteriorly by the tip of hippocampus and posteriorly by the end of gyrus intralimbicus (Pruessner et al., 2002; Insausati & Amaral, 2004). The end of gyrus intralimbicus was also used as the boundary segregating anterior and posterior hippocampus. Any electrode for which the anatomical location was ambiguous, i.e., between hippocampus and PrC or at the border of a region such that it may have been in white matter, was excluded. Electrodes in subiculum and entorhinal cortex were excluded because data were not available from all participants in these locations. Data from each electrode were preprocessed separately, and the resulting measures (ERPs and frequency-specific power) from all electrodes in each region (hippocampus or PrC, regardless of hemisphere) were averaged within-participant before being submitted to statistical tests.

All channels from hippocampus and PrC were included regardless of the site of the patient’s seizures. Given clinical evidence that patients with presumably unilateral medial temporal lobe seizures often have significant epileptic activity in the contralateral side (Engel, Williamson, & Wiese, 2008), we decided against designating ‘healthy’ versus ‘pathological’ hemispheres. Instead, we allowed electrodes from both hemispheres into the analyses and made significant effort to eliminate abnormal electrophysiological (i.e., epileptic) activity from our data via the automatic artifact detection method described above (see iEEG Data Analysis). A brief description of each patient’s clinical assessment is listed in Table 1.

3. Results

3.1 Power Modulations in Hippocampus and PrC

The critical associative memory trials consisted of viewing two sequences of four photographs of common objects. Specifically, an initial sequence of four images (ABCD) was followed by a second quartet of the same images, either: (1) in the same order ('Repeat': ABCD), (2) with the last two images in reverse order ('Mismatch': ABDC), or (3) in a completely rearranged order ('Shuffle': CADB or BDAC; Figure 1). On Repeat and Mismatch trials, the re-presentation of the first two images in the sequence (AB) was intended to cue retrieval of the associated image C, establishing a prediction of C's imminent appearance. This prediction was then either confirmed (Repeat: C appeared) or violated (Mismatch: D appeared). On Shuffle trials, the first two images in the second quartet were re-ordered (CA or BD), which is argued to decrease the probability that associative predictions are generated (Kumaran & Maguire, 2006, 2007).

Initial analyses focused on low-frequency power during the second quartets, computed post-onset of the 2nd stimulus (the repeated B on Repeat and Mismatch trials) or post-onset of the critical 3rd stimulus (the expected C, Repeat trials; the prediction-violating D, Mismatch trials). Visual inspection of hippocampus and PrC spectral power plots suggested that there were event-related signal modulations in the ~2–10 Hz range (Figure 2a). We split this range into two bins, 'slow-theta' (2.5–5 Hz) and 'fast-theta' (5.5–10 Hz), following the distinction suggested by Lega et al. (2012), who argued that in humans slow-theta activity is more relevant to memory-related computations than fast-theta. For analyses comparing hippocampus and PrC, only the six participants who had electrodes in PrC were included.

The theory that motivated this study (e.g., Lisman & Grace, 2005) proposes that if hippocampus serves as a novelty detector, two signals should be observed: an earlier novelty-detection signal, and a later encoding-related signal. Axmacher et al. (2010) explored this idea using an oddball task and found two ERP components that seemed to conform to this pattern: an earlier positive deflection (mean 187 ms) modulated by 'expectancy', and a later negative deflection (mean 481 ms, with differences lasting to ~800 ms) modulated by later memory. In the same paper, nucleus accumbens (which is thought to receive novelty signals from hippocampus) was found to differentiate between expected and unexpected stimuli later than hippocampus, at a later ERP component (mean 475 ms). Thus, we surmised a reasonable window in which to seek a mismatch-detection signal in hippocampus would be 0–500 ms.

3.1.1 Prediction Violation Evokes a Low-Frequency Power Increase in Hippocampus

Hippocampal power in the slow-theta range was greater during the prediction-violating D image (Mismatch) relative to the expected C image (Repeat) during the 0–500 ms post-stimulus onset window (Figure 2c and 2e; $t_{(6)} = 4.8$, $p < 0.005$). This hippocampal difference ('prediction violation effect') between Mismatch and Repeat trials was specific to the critical 3rd position items [sequence Position (2nd or 3rd) $\times$ Novelty (Mismatch or Repeat): $F_{(1,6)} = 24.3$, $p < 0.005$]. (The same 0–500 ms window was used for positions 2 and 3, post-onset of the 2nd or 3rd stimulus, respectively.) The mean Mismatch peak occurred at 232 ms (s.d. 89 ms), congruent with single-unit studies in human hippocampus that report stimulus-selective firing starting at around 300 ms after stimulus onset (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). By contrast, a significant violation effect (difference between Repeat and Mismatch) was not observed in PrC ($t < 1$; Figure 2d and 2f), and the effect of prediction violation differed between hippocampus and PrC [Region (hippocampus or PrC) $\times$ Position $\times$ Novelty: $F_{(1,5)} = 14.8$, $p = 0.01$; Figure 2c–f]. Restricting analyses to the 3rd position also yielded a significant Region $\times$ Novelty

interaction ($F_{(1,5)} = 9.7$, $p < 0.05$), demonstrating a greater associative prediction violation response in hippocampus than in PrC.

3.1.2 No Hippocampal Response to Prediction Violation in other Frequency Ranges—The hippocampal associative prediction violation effect seemed to be dominant in the slow-theta range (2.5–5 Hz). All analyses reported above for slow-theta were also conducted for fast-theta (5.5–10 Hz) and gamma (32–64 Hz). For the fast-theta range, a main effect of Region was found, wherein hippocampal power was, on average, greater than in PrC ($F_{(1,5)} = 8.1$, $p < 0.05$). A main effect of Position ($F_{(1,5)} = 12.1$, $p < 0.05$) demonstrated that power was, on average, greater for 3rd position than for 2nd position trials. No other significant effects were found ($ps > 0.1$). For the gamma range, no significant effects were found ($ps > 0.1$).

3.2 Does the Prediction Violation Response Depend on Having Established a Prediction?

The preceding analyses indicate that increased hippocampal power in the slow-theta range had a short latency (within 500 ms of stimulus onset) and was selective to the unexpected stimulus (i.e., Mismatch > Repeat was specific to the 3rd position). We next explored how this hippocampal response was related to the strength of prediction. On Mismatch trials, recapitulation of AB was intended to trigger retrieval of C, establishing a prediction of C's imminent appearance; on Shuffle trials, the first two items were reordered (either CA or BD), which was intended to decrease prediction likelihood of the 3rd position item (D or A; see Figure 1). Thus, we hypothesized that the hippocampal response to prediction violation on Shuffle trials would be intermediate, falling between the response level of Repeat and Mismatch trials.

Numerically, hippocampal power in the slow-theta range on Shuffle trials did fall between the levels observed for Repeat and Mismatch trials at the 3rd position item (Figure 3). However, the response on Shuffle trials did not significantly differ from either Repeat or Mismatch (Figure 3e).

3.3 Anterior vs. Posterior Hippocampus

All seven patients had electrodes in anterior hippocampus, with four also having electrodes in posterior hippocampus (see 2.6 *Electrode Selection* for definition). When analyses were restricted to anterior hippocampal electrodes in all patients, the primary results did not change: hippocampal power in the slow-theta range was greater for Mismatch than Repeat 3rd-position items during the 0–500 ms post-stimulus onset window ($t_{(1,6)} = 5.8$, $p = 0.001$), with the effect being specific to the critical 3rd position items [Position (2nd or 3rd) $\times$ Novelty (Mismatch or Repeat): $F_{(1,6)} = 63.5$, $p < 0.005$]. These comparisons were not significant when analyses were restricted to posterior hippocampal electrodes, though given that only four participants were retained, it should not be inferred that the prediction violation effects were selective to anterior electrodes.

3.4 Event-Related Potentials

As hippocampal power in the slow-theta range differentiated between Repeat and Mismatch trials – i.e., confirmed vs. violated predictions – we examined whether the event-related potential (ERP) also distinguished these two conditions. The same trial window was used as in the spectral analyses (3.1.1 *Prediction Violation Evokes a Low-Frequency Power Increase in Hippocampus*), encompassing positions 2 and 3 of each quartet, and thus the baseline period for this analysis was 100 ms immediately prior to the 2nd position. Hippocampal ERPs at the 3rd position item did not significantly differ between Repeat and Mismatch trials, whether compared over the 0–500 ms window or over any of the five 100-ms windows within that period ($ps > 0.4$; Figure 4a).

Target events elicited a large negative-going ERP in hippocampus (Figure 5a; for display purposes, target items are plotted regardless of whether they appeared in Positions 2, 3, or 4). Target ERPs exhibited a peak average negative deflection at 396 ms (s.d. 35 ms), with Target differing significantly from Repeat and Mismatch Position 3 items over the 200–500 ms window post-onset ($p < 0.05$; for 0–200 ms window, $p > 0.2$); this difference was specific to items that were Targets (Position [N or N-1] x Condition interaction, $F_{(1,6)} = 10.3$, $p < 0.05$). By contrast, PRc ERPs did not exhibit any sensitivity to item status (Mismatch vs. Repeat, $p > 0.4$; Target vs. Repeat and Mismatch, $p > 0.5$ over the 200–500 ms window; Figure 4b and 5b).

Trials from the initial quartets ('First' quartet) that preceded each of the Repeat, Mismatch, and Shuffle quartets (see 2.3 *Task Parameters*) were qualitatively similar to the other conditions in terms of ERPs (Figure 5a–d). Direct comparisons were not made because First items were not matched with Repeat, Mismatch, and Shuffle in terms of the number of times items were viewed (by design, always once for First trials, always twice for Repeat, Mismatch, and Shuffle trials), but the data are displayed for completeness.

The astute reader will notice that Repeat and Mismatch trials diverge at the end of the 2nd position in hippocampus (Figure 4a). If ERPs are calculated using a 100-ms baseline period immediately before the 3rd position (rather than before the 2nd position, as plotted in Figure 4a), Repeat and Mismatch differ significantly ($p < 0.05$) in the 0–300 ms window after the onset of the 3rd position. However, according to the current experiment design, Repeat and Mismatch trials should not differ before the 3rd position, and thus we note this effect with extreme caution. Future studies with more measurements per condition are needed to evaluate the validity of a putative associative mismatch ERP effect in the 0–300 ms range.

4. Discussion

The present electrophysiological data indicate that human hippocampus is sensitive to associative prediction violations, when stimuli predicted based on learned associations do not match encountered stimuli, extending fMRI studies of associative novelty (Köhler et al., 2005; Kumaran & Maguire, 2006; Chen et al., 2011; Duncan et al., 2012). Due to the coarse temporal resolution of fMRI (several seconds), these prior observations left open the possibility that mismatch computations occurred elsewhere in the brain (e.g., prefrontal or parietal cortex) and subsequently modulated hippocampal responses. Additionally, hippocampus is thought to be engaged both by mismatch detection and by the consequent encoding enhancement (Lisman & Grace, 2005), and thus hippocampal BOLD responses to associative mismatch could reflect either or both of these processes. Our data provide strong support for the idea that this hippocampal mismatch effect reflects direct hippocampal computations rather than re-entrant processes, as differences emerged within a few hundred milliseconds after stimulus onset—as early as stimulus-selective firing has been observed in individual human hippocampal neurons (Quiroga et al., 2005), and comparable to stimulus novelty effects observed using scalp EEG (Rugg & Curran, 2007). At the group level, a concomitant sensitivity was absent in PrC, a region known to project visual-object inputs to hippocampus via entorhinal cortex, further supporting the idea that mismatch signals originate in hippocampus (though a subset of participants demonstrated a Mismatch > Repeat PrC response; Figure 2f). The observation of an associative memory-dependent effect in the slow-theta range is broadly consistent with the findings of Lega et al. (2012), who reported that in human hippocampus increased power in slow-theta oscillations is linked to memory processes to a greater degree than fast-theta oscillations.

Based on prior studies, we hypothesized that a hippocampal mismatch response would be maximal in associative cueing conditions that favor the generation of predictions – that is, a

prediction violation response should depend on a prediction being generated in the first place (Fyhn, Molden, Hollup, Moser, & Moser, 2002; Kumaran & Maguire, 2006). The experimental conditions were designed to differ in their effectiveness for evoking predictions, as sequences were initiated in identical order on Repeat and Mismatch trials (high probability of a prediction), but in rearranged order on Shuffle trials (lower probability of a prediction). Indeed, in a version of the task where subjects were asked to make a category judgment for every item, responses were significantly faster to the third item in A–B–C (Repeat) quartets as opposed to the A–B–D (Mismatch) quartets, suggesting that the prediction of C facilitated processing of C (Kumaran & Maguire, 2006). Additionally, responses were slowest overall for Shuffle trials (when considering all four items), suggesting that the rearranged item order negatively affected subjects' ability to predict upcoming items across the quartet (not only at the 3rd position).

Numerically, hippocampal power in the slow-theta range on Shuffle trials fell between the levels observed for Repeat and Mismatch trials at the 3rd position item (Section 3.2, Figure 3e); however, interpretative caution is warranted because the response on Shuffle trials did not significantly differ from either Repeat or Mismatch, and examination at the individual subject level suggests high variability in the Shuffle condition (Figure 3d). Moreover, this pattern is not entirely in accord with prior fMRI work using the same paradigm (Kumaran & Maguire, 2006) wherein hippocampal responses to Shuffle trials were differentiated from Mismatch but not from Repeat trials at the 3rd position, and in the aforementioned category-judgment behavioral version of the task, reaction times for Shuffle trials were significantly slower than Mismatch but not differentiable from Repeat trials at the 3rd position.

One possible interpretation of the current finding – that Shuffle falls in between Repeat and Mismatch in terms of the hippocampal prediction violation response – is that Shuffle trials did give rise to weak or varied predictions. The repetition of items from the First quartet, even if not in the same order, might weakly or infrequently activate predictions of other items in the quartet; such weak/infrequent predictions might give rise to weak/infrequent prediction violations. Perhaps this pattern was not detectable using fMRI but is now hinted at in our intracranial recordings. Future studies should explore the relationship between prediction strength or prediction probability and hippocampal prediction violation responses.

The image-sequence paradigm was designed to elicit predictions whenever image sequences repeat: when A and then B appear, C is predicted to follow simply because the sequence A–B–C occurred in the past (Kumaran & Maguire, 2006). This fundamental feature of memory underlies our claim that when B is followed by D (instead of by C), a prediction is violated. That is, we believe the 'default' behavior of human memory is to predict that events will repeat – it does not require that active or explicit expectations of repetition be established by prior trials within the experiment. Nonetheless, it could be argued that a second level of prediction was accessible to subjects given that C and D were reversed on some trials in a regular manner. If only Repeat and Mismatch Second quartets are considered, then C and D were reversed 50% of the time, ostensibly yielding high uncertainty about whether C will follow B. However, Mismatch Second (ABDC) quartets comprised only 13% of all quartets in the experiment. If one considers only quartets where a previously seen A–B was repeated (Repeat Second, Mismatch Second, and Single), A–B–C occurred 66% and A–B–D occurred 33% of the time. Even if one considers only quartets where B was not novel (Repeat Second, Mismatch Second, Mismatch Shuffle, and Single), C followed B 50% of the time. It is not obvious which is the best method of calculating the probability that C would follow B, but at worst it was 50/50. Thus, we believe that the presence of any second-level prediction may have diluted (the default) A–B–C prediction strength, but did not negate it. That is, the observed effects of mismatch may be underestimated.

Beyond the effects of associative mismatch, the present data also revealed that Target events elicited an increased ERP negativity in hippocampus peaking at 396 ms post-Target onset (Figure 5a). Target events were (a) the only stimuli for which participants were asked to make a manual response, (b) the explicitly sought goal state of the participants, and (c) an infrequent type of event in the experiment. Any one (or a combination) of these variables may have given rise to the observed hippocampal response. The present Target hippocampal ERP is similar to what has been previously observed in 'oddball' paradigms during the presentation of 'unexpected' stimuli – i.e., infrequent stimuli that differed from the majority of stimuli in a salient dimension such as color or category (e.g., Halgren et al., 1980; Axmacher et al., 2010) – as well as in response to goal-relevant stimuli (McCarthy, Wood, Williamson, & Spencer, 1989). It is worth noting that this negative deflection, commonly associated with contextual novelty, was not present in the Repeat, Shuffle, or Mismatch ERPs (compare Figures 4a and 5a), congruent with the idea that contextual and associative novelty are distinct. However, as Target trial responses may reflect a combination of factors rather than being a pure measure of contextual novelty, future studies directly comparing associative and contextual novelty would be informative.

The current data demonstrate that electrophysiological activity in human hippocampus is sensitive to the occurrence of unexpected stimuli, when associative memory-based predictions are violated. Hippocampal detection of when memory-based predictions deviate from current sensory reality positions the hippocampus to trigger a cascade of neural events that facilitates the encoding of novel experiences into memory (Lisman & Grace, 2005).

Acknowledgments

Supported by grants from the National Institutes of Health (R01-MH076932 and R01-NS078396) and Stanford NeuroVentures Program. We thank Avgusta Shestiyuk and Vinitha Rangarajan for assistance with data collection, the patients for donating their time during clinical evaluation at the Stanford Medical Center, and John Lisman for insightful comments on an earlier draft of the manuscript.

References

- Axmacher N, Cohen MX, Fell J, Haupt S, Dümpelmann M, Elger CE, Schlaepfer TE, et al. Intracranial EEG correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. *Neuron*. 2010; 65(4):541–9. [PubMed: 20188658]
- Buckner RL. The role of the hippocampus in prediction and imagination. *Annual Review of Psychology*. 2010; 61:27–48. C1–8.
- Canolty RT, Edwards E, Dalal SS. High gamma power is phase-locked to theta oscillations in human neocortex. *Science*. 2006; 313(5793):1626–1628. [PubMed: 16973878]
- Chen J, Olsen RK, Preston AR, Glover GH, Wagner AD. Associative retrieval processes in the human medial temporal lobe: hippocampal retrieval success and CA1 mismatch detection. *Learning & Memory*. 2011; 18(8):523–8. [PubMed: 21775513]
- Duncan K, Ketz N, Inati SJ, Davachi L. Evidence for area CA1 as a match/mismatch detector: a high-resolution fMRI study of the human hippocampus. *Hippocampus*. 2012; 22(3):389–98. [PubMed: 21484934]
- Eichenbaum, H.; Cohen, NJ. From conditioning to conscious recollection. New York, NY: Oxford UP; 2001.
- Engel, J.; Williamson, P.; Wiese, H. Epilepsy: a comprehensive textbook. 2. Philadelphia, PA: Lippincott-Raven; 2008.
- Fell J, Ludowig E, Rosburg T, Axmacher N, Elger CE. Phase-locking within human mediotemporal lobe predicts memory formation. *NeuroImage*. 2008; 43(2):410–9. [PubMed: 18703147]
- Fyhn M, Molden S, Hollup S, Moser MB, Moser E. Hippocampal neurons responding to first-time dislocation of a target object. *Neuron*. 2002; 35(3):555–66. [PubMed: 12165476]

- Halgren E, Baudena P, Clarke JM, Heit G, Marinkovic K, Devaux B, Vignal JP, et al. Intracerebral potentials to rare target and distractor auditory and visual stimuli. II. Medial, lateral and posterior temporal lobe. *Electroencephalography and Clinical Neurophysiology*. 1995; 94(4):229–50. [PubMed: 7537196]
- Halgren E, Squires N, Wilson C, Rohrbaugh J, Babb T, Crandall P. Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science*. 1980; 210(4471):803–805. [PubMed: 7434000]
- Hasselmo ME, Wyble BP. Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behavioural Brain Research*. 1997; 89(1–2):1–34. [PubMed: 9475612]
- Insausti, R.; Amaral, DG. The human nervous system. 2. Amsterdam: Elsevier; 2004. Hippocampal formation; p. 871–914.
- Jacobs J, Kahana MJ. Direct brain recordings fuel advances in cognitive electrophysiology. *Trends in Cognitive Sciences*. 2010; 14(4):162–71. [PubMed: 20189441]
- Kirchhoff BA, Wagner AD, Maril A, Stern CE. Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *The Journal of Neuroscience*. 2000; 20(16):6173–80. [PubMed: 10934267]
- Knight RT. Contribution of human hippocampal region to novelty detection. *Nature*. 1996; 383(6597):256–259. [PubMed: 8805701]
- Kumaran D, Maguire E. Which computational mechanisms operate in the hippocampus during novelty detection? *Hippocampus*. 2007a; 17(8):735–748. [PubMed: 17598148]
- Kumaran D, Maguire EA. An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biology*. 2006; 4(12):e424. [PubMed: 17132050]
- Kumaran D, Maguire EA. Match mismatch processes underlie human hippocampal responses to associative novelty. *The Journal of Neuroscience*. 2007b; 27(32):8517–24. [PubMed: 17687029]
- Kumaran D, Maguire EA. Novelty signals: a window into hippocampal information processing. *Trends in Cognitive Sciences*. 2009; 13(2):47–54. [PubMed: 19135404]
- Köhler S, Danckert S, Gati JS, Menon RS. Novelty responses to relational and non-relational information in the hippocampus and the parahippocampal region: a comparison based on event-related fMRI. *Hippocampus*. 2005; 15(6):763–74. [PubMed: 15999342]
- Lega BC, Jacobs J, Kahana M. Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*. 2012; 22(4):748–61. [PubMed: 21538660]
- Lisman JE, Grace AA. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*. 2005; 46(5):703–13. [PubMed: 15924857]
- McCarthy G, Wood CC, Williamson PD, Spencer DD. Task-dependent field potentials in human hippocampal formation. *The Journal of Neuroscience*. 1989; 9(12):4253–68. [PubMed: 2593001]
- Pruessner JC, Köhler S, Crane J, Lord C, Byrne A, Kabani N, Collins DL, et al. Entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cerebral Cortex*. 2002; 12(12):1342–1353. [PubMed: 12427684]
- Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. Invariant visual representation by single neurons in the human brain. *Nature*. 2005; 435(7045):1102–7. [PubMed: 15973409]
- Rugg M, Curran T. Event-related potentials and recognition memory. *Trends in Cognitive Sciences*. 2007; 11(6):251–257. [PubMed: 17481940]
- Squire LR. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*. 1992; 99(2):195–231. [PubMed: 1594723]
- Vinogradova OS. Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*. 2001; 11(5):578–598. [PubMed: 11732710]

Highlights

- Associative predictions are generated when previously paired stimuli recur.
- Hippocampus is thought to detect prediction violations, when pairs are mismatched.
- Hippocampal signals were recorded as epilepsy patients viewed image sequences.
- Images appeared in repeated (prediction confirmed) or rearranged (violated) order.
- Associative prediction violations elicit increased hippocampal low-frequency power.

Figure 1.

Task diagram. Each quartet of four images ('First') was followed by a second quartet of the same images in the same order ('Repeat'), a partially recapitulated order ('Mismatch'), or with all images in the second quarter re-ordered (Shuffle). Letters (ABCD) are for illustration purposes and did not appear during stimulus presentation.

Figure 2.

Prediction violation vs. confirmation in hippocampus and perirhinal cortex (PrC). **(a–b)** Event-related spectral power of Mismatch and Repeat trials in hippocampus and PrC at quartet positions 2 and 3. **(c–d)** Percent signal change of power in the slow-theta range in hippocampus and PrC at quartet positions 2 and 3 when images were presented in repeat order (Repeat) or with image D at Position 3, violating predictions (Mismatch). Shaded area represents standard error of subjects. **(e–f)** Individual subject means of power in the slow-theta range during the 0–500 ms window after position 3 onset in hippocampus and PrC. Error bars represent the standard error of trials.

Figure 3.

Comparison of hippocampal power in the slow-theta range during prediction violation (Mismatch) vs. when prediction generation was unlikely (Shuffle). **(a–c)** Percent signal change of power in the slow-theta range at second-quartet positions 1, 2 and 3. Shaded area represents standard error of subjects. **(d)** Individual participant means of power in the slow-theta range 0–500 ms post-stimulus onset. Error bars represent the standard error of trials. **(e)** Mean percent signal change of power in the slow-theta range 0–500 ms post-stimulus onset at position 3 in Repeat, Shuffle, and Mismatch conditions. Error bars represent the standard error of subjects.

Figure 4. Event-related potentials (ERPs) in Hippocampus and PrC. **(a–b)** ERPs for Repeat, Shuffle, and Mismatch Shaded area represents standard error of subjects.

Figure 5.

Event-related potentials (ERPs) in Hippocampus and PrC. **(a–b)** ERPs for First quartet and 1-back Target conditions. All items appearing in First quartets were combined regardless of their position in the quartet (1–4); Target items may have appeared in positions 2, 3, or 4. **(c–d)** ERPs for First quartet and Second quartet items, regardless of their position in the quartet (1–4). Shaded area represents standard error of subjects.

Patient Information

Table 1

Gender, age, clinical information, and number of electrodes in left/right hippocampus/PrC for each patient. Note that Left and Right Hipp are mutually exclusive, but not exclusive from Anterior and Posterior Hipp (i.e., the total number of hippocampal electrodes is found by adding Left and Right Hipp).

	Gender	Age	Left Hipp	Right Hipp	Anterior Hipp	Posterior Hipp	Left PrC	Right PrC	Clinical Description of
S1	F	40	2	1	2	1	2	2	Pathology Primarily right hemisphere
S2	F	25	3	0	1	2	2	2	Primarily right hemisphere
S3	F	21	3	1	4	0	1	0	Bilateral
S4	F	22	2	0	2	0	2	0	Primarily left hemisphere
S5	F	52	5	3	5	3	1	1	Bilateral
S6	M	19	3	3	3	3	0	1	Bilateral
S7	M	50	1	1	2	0	0	0	Bilateral