Medial temporal lobe high-frequency activity at feedback signals error and predicts subsequent memory during spatial learning

Rebecca Stevenson  
UC Berkeley

John Janecek  
UC Irvine

Myra Larson  
University of California, Irvine, CA

Lilit Mnatsakanyan  
University of California, Irvine

Sumeet Vadera  
University of California, Irvine

Jack Lin  
University of California, Irvine  
https://orcid.org/0000-0003-1304-227X

Michael Yassa (✉️ myassa@uci.edu)  
Center for the Neurobiology of Learning and Memory and Departments of Neurobiology and Behavior, Psychiatry and Human Behavior, Neurology, Psychological Science, University of California, Irvine, CA  
https://orcid.org/0000-0002-8635-1498

Article

Keywords: medial temporal lobe, high-frequency activity, feedback, learning, intracranial

DOI: https://doi.org/10.21203/rs.3.rs-194736/v1

License: ☕️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Title:
Medial temporal lobe high-frequency activity at feedback signals error and predicts subsequent memory during spatial learning

Authors:
Rebecca F. Stevenson¹, John T. Janacek², Myra S. Larson², Lilit Mnatsakanyan³, Sumeet Vadera⁴, Jack J. Lin³,⁵,⁶, Michael A. Yassa²,³,⁷*

¹Helen Wills Neuroscience Institute, University of California, Berkeley
²Department of Neurobiology and Behavior, University of California, Irvine
³Department of Neurology, University of California, Irvine
⁴Department of Neurological Surgery, University of California, Irvine
⁵Department of Biomedical Engineering, University of California, Irvine
⁶Department of Anatomy and Neurobiology, University of California, Irvine
⁷Center for the Neurobiology of Learning and Memory, University of California, Irvine

*Correspondence: myassa@uci.edu

Michael A. Yassa (Contact): Department of Neurobiology and Behavior, 1418 Biological Sciences 3, Irvine CA, 92697, Tel: (949) 824-1687, Email: myassa@uci.edu

Author contributions:
R.F.S. and M.A.Y designed the research, R.F.S. collected the data, L.M., S.V., and J.J.L. managed patients and surgeries, R.F.S. analyzed the data, J.T.J and M.S.L assisted with preprocessing and data analysis, R.F.S., J.J.L., and M.A.Y. interpreted the data and wrote the paper with input from all authors.

The authors declare no conflict of interest.

Data sharing: Data and code are available upon request from the corresponding author.

Keywords: medial temporal lobe; high-frequency activity; feedback; learning; intracranial
Abstract:

The ability to incorporate information about feedback is critical for associative learning. The medial temporal lobe (MTL) and prefrontal cortex (PFC) are thought to be involved in processing feedback as new associations are learned. However, the relative contributions of these regions to feedback processing and subsequent memory performance in humans are poorly understood. To address this question, we tested pre-surgical epilepsy patients with depth electrodes implanted in the MTL and PFC using a spatial memory task in which subjects learned object-location associations over time. During encoding, subjects were shown objects at random locations along the circumference of an invisible circle. For each training block, the same objects were shown at the top of the circle and subjects used a mouse wheel to rotate the object to where it appeared during encoding. After subjects finished placing each object, the object was shown in the correct location for one second as feedback. We found increased high-frequency activity (HFA; 40-100 Hz), thought to reflect local excitatory activity, in the MTL and dorsolateral PFC (dLIPC) at feedback for high error trials. In the MTL, this HFA error signal predicted greater trial-by-trial decreases in error from one training block to the next indicating that these signals are involved in updating memory representations or modifying incorrect associations during learning. The opposite pattern of activity was observed during retrieval, with greater MTL and dLIPC HFA predicting lower error, replicating previous results from our group. Overall, these data suggest putative mechanisms for the learning of object-location associations.
Introduction:

A wealth of evidence indicates that medial temporal and prefrontal networks play a crucial role in tasks involving learned associations, such as between an object and a location (Blumenfeld & Ranganath, 2007; Eichenbaum, Yonelinas, & Ranganath, 2007). Convergent studies in animals and humans have found that activity within the medial temporal lobe (MTL) and lateral prefrontal cortex during encoding and retrieval predicts performance on associative memory tasks (Brincat & Miller, 2015; Histed, Pasupathy, & Miller, 2009; Igarashi, Lu, Colgin, Moser, & Moser, 2014; Ison, Quian Quiroga, & Fried, 2015; Law et al., 2005; Rutishauser, 2008; Suthana et al., 2015; Wirth et al., 2009). As novel associations are often not learned in a single episode, the ability to process and incorporate information about feedback is an important component of associative learning. Both MTL and PFC activity has been found to signal trial outcome at feedback, with single unit recordings in monkeys showing cells that selectively increase firing for both correct and incorrect trials (Brincat & Miller, 2015; E. L. Hargreaves, A. T. Mattfeld, C. E. L. Stark, & W. a. Suzuki, 2012; Wirth et al., 2009). However, the ways in which activity within these networks reflects trial outcome in humans are poorly understood, as are the ways in which MTL and PFC trial outcome signals contribute to learning.

To address this question, we tested pre-surgical epilepsy patients with depth electrodes implanted in both the MTL and dorsolateral PFC (dIPFC) on a variant of a commonly used object-location memory task (Reagh & Yassa, 2014; Rudoy, Voss, Westerberg, & Paller, 2009). During encoding, 30 objects were presented at random locations along the circumference of an invisible circle. For each of the three training blocks, the same objects were shown one at a time at the top of the circle and subjects used a mouse wheel to rotate the object to where it appeared during encoding. After subjects finished placing each object, it was shown in the correct location for one second as feedback. After the third training block, a final test was given
with no feedback. Performance was measured in angular error, or the distance (in degrees) between where subjects placed the object and the correct location. We predicted that MTL and dlPFC high-frequency activity (40-100 Hz) at feedback would predict error, and that the magnitude of this response for specific object-location associations would predict subsequent performance on that association. Using a variant of this task in which subjects were given only one shot to learn the object-location associations (i.e. no training), we previously found a negative correlation between MTL and dlPFC HFA, thought to reflect local excitatory activity, and error at retrieval, indicating that activity within these regions tracks representational fidelity (Stevenson et al., 2018). Here, we predicted that we would replicate these results, showing increased MTL and dlPFC high-frequency activity (HFA) at retrieval for low error trials.

Results:

Task performance. Subjects (21 sessions from 9 patients) performed a spatial learning task as we recorded intracranial electroencephalogram (EEG). During the encoding phase, 30 objects were presented, one at a time, at random positions around the circumference of an invisible circle (Fig. 1). Prior to the encoding phase, subjects were told that they would be tested on the location of each object. Following encoding there were three training blocks during which each object was shown again, this time at the top of the circle, and subjects were instructed to use the mouse wheel to rotate the object to where it appeared during encoding. Subjects were asked to wait 1 second (until text that read ‘Wait…’ disappeared from the screen) before beginning to rotate the object and to press the space bar when they were finished placing the object. After each trial, the object appeared on the screen in the correct location for 1 second as feedback. During the final test phase, subjects were again asked to rotate each object to the correct location, though this time no feedback was given. The inter-trial interval (ITI) was 1.2 +/- 0.2 seconds and the interstimulus interval (ISI), meaning the time between when the subjects
finished placing the object and when it was shown in the correct location as feedback was fixed at 0.5 seconds for subjects 1-2 and variable (0.4 +/- 0.2 seconds) for subjects 3-9. Performance was measured in angular error, or the distance (in degrees) between where subjects placed the object and the correct location. If subjects performed more than one session, new objects and locations were used. No more than one session was performed on each day.

Figure 1. Task schematic and performance. A. During encoding, 30 objects were presented at random positions along the circumference of an invisible circle. During each of the three training blocks, the same objects were presented at the top of the screen. After a 1 second wait period, subjects rotated the object to where it appeared during encoding. After subjects finished placing the object, it was shown in the correct location for 1 second as feedback. No feedback was given during the final test phase. B. Example of a mixture model fit of all trials for subject 1. The cut off for high error trials across blocks was derived from the cumulative distribution function of the von Mises distribution, i.e. 90% of trials estimated to be remembered fell within +/-38 degrees of the correct location. C-D, Mean guess rate (C), and SDMem across sessions for training blocks 1-3 (T1-T3) and test.

Figure 1B shows the distribution of error across all trials for subject 9 (session 21) (see SI Fig. S1 for error distributions for all sessions). These error distributions can be modeled as a mixture of two distributions: a uniform distribution of errors and a von Mises distribution of errors (Fig. 1B) (Sutterer & Awh, 2016; Zhang & Luck, 2008). The uniform distribution reflects trials on
which the subject guessed randomly. The von Mises distribution reflects trials on which the
subject remembered the location of the object with varying precision. We used the cumulative
distribution function of the von Mises distribution estimated for each session to determine which
trials to place in the high error condition. Trials that had less than a 10% chance of being
remembered with some degree of precision were put into the high error condition. For example,
in subject 1, the middle 90% of the von Mises distribution spans +/-38°, so trials with error
greater/less than +/-38° were designated as high error trials (Fig. 2B). The remaining trials were
sorted by error and split evenly into the medium and low error conditions.

We used the MemFit function of Memtoolbox in MATLAB (Suchow, Brady, Fougnie, & Alvarez,
2013), to obtain an estimate of two parameters describing these distributions: the guess rate (g),
which reflects the area under the uniform distribution, and the standard deviation of the von
Mises distribution (SDMem), which reflects the overall precision of remembered responses.
Figure 1C-D shows the mean value of these parameters for T1-3 and for the final test (see SI
Fig. S1 for parameter values across blocks for all sessions). Guess rate, SDMem, and angular
error all decreased over task blocks, indicating learning of the object-location associations
(linear trend, guess rate: F(1, 60) = 14.6, p = 0.0003, SDMem: F(1, 60) = 23, p < 0.0001; error:
F(1, 60) = 173, p < 0.0001).

**Electrode localization.** Electrodes were localized using co-registered pre-implantation MRIs,
post implantation CTs, and post-implantation MRIs (when available) as well as registration to a
high-resolution anatomical atlas with manual tracings of MTL subregions (Zheng et al., 2017).
Five subjects had electrodes localized to the MTL and 9 subjects had electrodes localized to the
dIPFC (see Table 1 for number of MTL and dIPFC contacts for each subject). Only data from
recordings contralateral to the seizure source or outside of the seizure onset zone were used in subsequent analyses.

Figure 2. MTL and dlPFC activity signals error at feedback. Spectrograms showing mean MTL and dlPFC power across all trials at retrieval (A-B) and feedback (C-D). Time course of MTL and dlPFC HFA, normalized within blocks, for high, medium, and low error trials at retrieval (E-F) and feedback (G-H). Gray horizontal lines indicate time points where there are significant differences between conditions (ANOVA, p < 0.05, cluster-corrected). Colored shaded regions indicate s.e.m. The start and end of the retrieval/feedback window are indicated by vertical dotted lines.

MTL and dlPFC HFA predicts error within blocks at retrieval and feedback. We first examined how HFA changed across the training and test blocks (T1-3 and test) at retrieval and feedback. Spectrograms for MTL and dlPFC power across all trials at retrieval (A-B) and feedback (C-D) are shown in Figure 2A-D. The retrieval window was defined as the one second period after the object appeared on the screen but before subjects were able to start moving it. The feedback window was defined as the one second period during which the object was shown in the correct location. Using a cluster-based permutation approach to correct for multiple comparisons across time points, we found that HFA decreased across blocks in the MTL at both retrieval and feedback (SI Fig. S2A, H) and in the dlPFC at feedback (SI Fig. S2N) (ANOVA p < 0.05, cluster-corrected). At feedback, decreases in HFA were observed in nearly all brain
regions we recorded from, including the lateral temporal, orbitofrontal cortex, and anterior
cingulate cortex (see SI Fig. S2 for results from all regions at retrieval and feedback), though
there were regions (e.g. the insula) and time periods where there was no significant difference
in HFA across blocks, suggesting that these changes were not driven by non-neural sources
such as changes in electrode and impedance over time. However, there are a variety of
cognitive processes that could contribute to these decreases in HFA signal, including increases
in familiarity, decreases in novelty, as well as learning of the object-location associations.

In order to control for these nonspecific changes in HFA across blocks, we first normalized HFA
within each block (T1-3 and test). Replicating previous results from our group, we found greater
HFA for lower error trials in the dlPFC at retrieval (ANOVA, p < 0.05, cluster-corrected)
(Stevenson et al., 2018). We found the opposite pattern of activity during feedback, with greater
HFA for high error trials in both regions (Fig. 3C-F). No other region we recorded from showed a
significant difference in HFA between high, medium, and low error trials during the retrieval
window (SI Fig. S3). However, multiple regions, including the anterior cingulate cortex and
caudal prefrontal cortex, showed increased HFA for high error trials at feedback, indicating that
this effect was relatively widespread (ANOVA, p < 0.05, cluster-corrected) (SI Fig. S3).

| Subject | Hand | Epileptogenic region       | Coverage | MTL Hemisphere analyzed | Number of sessions | MTL contacts | dlPFC contacts |
|---------|------|-----------------------------|----------|-------------------------|--------------------|--------------|---------------|
| 1       | R    | Right TLE                   | SEEG     | L                       | 1                  | 4 (1 HC)    | 4             |
| 2       | R    | Right TLE                   | SEEG     | N/A                     | 3                  | 0            | 3             |
| 3       | L    | Right medial frontal/SMA    | SEEG     | Bilateral               | 4                  | 9 (6 HC)    | 6             |
| 4       | R    | Left TLE                    | SEEG     | R                       | 4                  | 5 (4 HC)    | 7             |
| 5       | A    | Bilateral TLE               | SEEG     | N/A                     | 2                  | 0            | 7             |
| 6       | R    | Right TLE                   | SEEG     | L                       | 3                  | 4 (1 HC)    | 7             |
Since there were fewer high error trials than there were medium or low error trials, we performed control analyses to make sure observed differences between conditions were not due to differing numbers of trials across conditions. To address this issue, we sorted trials within each block by error and split them into thirds, putting the top 1/3 trials with the highest error from each block into the high error condition, the middle 1/3 trials into the medium error condition, and the bottom 1/3 trials with the least error from each block into the low error condition. We observed the same pattern of effects in both regions for both retrieval and feedback when trials were sorted evenly in this way. We additionally performed control analyses to ascertain that the gamma effects in the MTL and dlPFC at retrieval and feedback were not associated with the distance the object was moved on the screen (p > 0.05, cluster corrected).

Testing for correlations between HFA and error within blocks in the MTL and dlPFC we found negative correlations between HFA and logged error in both regions at retrieval (Figure 3A-B) (Pearson, p < 0.05, permutation corrected), consistent with previous results from our group (Stevenson et al., 2018). At feedback, we found the opposite pattern of activity in both regions, with increasing HFA predicting greater error (Fig. 3A-B).
**Figure 3. MTL HFA predicts learning across blocks.** Correlations between MTL (A) and dIPFC (B) HFA and current error for at retrieval (blue), and at feedback (red), and partial correlations between HFA at feedback and the subsequent change in error for high error trials from one task block to the next (green, controlling for current error). The start and end of the retrieval/feedback window are indicated by vertical dotted lines.

**MTL HFA at feedback predicts accuracy on subsequent trials.** We then wanted to determine if the observed error signals improved performance on subsequent trials. We found a significant partial correlation between MTL HFA at feedback for high error trials and the change in error from one block to the next (e.g. T1 to T2) for specific object-location associations, controlling for error on the trial occurring in the earlier block (current error) (partial correlation, Pearson, p < 0.05, permutation corrected) (Fig. 3A-B). This was a negative correlation, indicating that the bigger the HFA error signal in the MTL was, the bigger the decrease in error from one block to the next. Importantly, this partial correlation controls for the amount of current error, indicating that this effect is not due to ceiling/floor effects stemming from the amount of current error (i.e. the larger the current error is, the more room there is for decreases in error).

**Discussion:**
Prior work has shown that MTL and dlPFC activity at feedback signals trial outcome (correct vs. incorrect). However, the contributions of these signals to performance on subsequent trials (i.e. learning) are poorly understood. Here, we found that increased MTL and dlPFC HFA, thought to reflect local excitatory activity, signaled increased error at feedback in an object-location associative learning task. Increased MTL activity at feedback also predicted greater decreases in error from one training block to the next for specific object-location associations, indicating that MTL error signals are involved in updating incorrect or imprecise associations during learning.

A variety of processes could contribute to the observed error signals at feedback. Accumulating evidence suggests that the hippocampus plays a crucial role in developing models, or predictions, about the environment. Numerous studies have shown that hippocampal activity increases when predictions about the environment are violated, resulting in what has been termed a ‘match/mismatch’ signal (Duncan, Ketz, Inati, & Davachi, 2012)). These ‘match/mismatch’ signals might contribute to the observed error signals if subjects’ placement of an object was less accurate than they had expected. Similarly, prediction errors signal discrepancies between actual and expected outcomes and are thought to drive learning by updating expectations to make predictions more accurate (Mattar & Daw, 2018; Stachenfeld, Botvinick, & Gershman, 2017). Given that it is generally preferable to do well on any given task, reward prediction errors, or the difference between an actual and expected reward might contribute to the observed error signals. In this case, the reward would be good performance on the task (low error), and a reward prediction error would occur if subjects’ placement of an object was less accurate than they had expected. The increased activity for high error trials might also reflect encoding (or re-encoding) of the correct location if the location of the object was forgotten, misremembered, or imprecise. Regardless of whether subjects had a specific
prediction about where the object would be when it was shown in the correct location at feedback, there would likely be a measure of increased surprise and novelty for higher error trials since the object would appear at an unexpected location. As such, attentional effects that accompany surprise or novelty might also contribute to these signals (Corbetta & Shulman, 2002). However, the error signal observed in the MTL at feedback predicted subsequent performance on specific object-location associations, indicating that this signal is not just reflecting surprise or novelty but that it is contributing to associative learning.

At retrieval, we found negative correlations between increased MTL and dlPFC HFA and error, replicating results from our previous study (Stevenson et al., 2018). These results provide additional evidence that MTL and dlPFC activity does not just reflect a binary signal of retrieval success versus failure, but rather that activity within these regions tracks representational fidelity. This adds to a growing body of literature implicating the extended hippocampal network in spatial memory precision (Koen, Borders, Petzold, & Yonelinas, 2017; Kolarik et al., 2016; Nilakantan, Bridge, Gagnon, VanHaerents, & Voss, 2017).

Previous studies in monkeys have found MTL cells that increase firing at feedback for both correct and error trials (Brincat & Miller, 2015; Wirth et al., 2009). However, the ways in which error signaling drives population activity as recorded by local field potential (LFP), intracranial EEG, or BOLD fMRI in animals and humans is poorly understood. In a cross-species study that examined feedback signals in both monkeys (using LFP) and humans (using fMRI), increased MTL HFA (30-100 Hz) was found to predict error (incorrect > correct) in monkeys, while the BOLD fMRI signal in humans at feedback showed the opposite polarity, with increased MTL beta values for correct trials (correct > incorrect) (E. L. Hargreaves, A. T. Mattfeld, C. E. Stark, & W. A. Suzuki, 2012). High-frequency activity is thought to be generated by increased spiking,
gamma oscillations, or by a combination of these two processes (John F. Burke, Ramayya, & Kahana, 2015). In each case, this would mean that the HFA error signals reflect increased local excitatory activity, as noted above. Although the BOLD signal has been found to be correlated with the local field potential, this signal is thought to be generated by complex neurovascular coupling, making the polarity of this signal difficult to interpret (Logothetis, 2001). The results from the current study support the idea that, at a population level, MTL excitatory activity in humans is driven by increased error during associative learning at feedback, though other task designs could result in differing effects.

We also observed decreases in MTL and dIPFC HFA across task blocks at retrieval and feedback. Although these decreases in activity paralleled decreases in error across blocks, low error trials within blocks were associated with increased MTL and dIPFC activity at retrieval, indicating that these shifts in power across blocks were not driven solely by object-location associative learning. Interestingly, multiple other studies have shown decreases in gamma activity over the course of experimental sessions, linking these gradual shifts to decreases in novelty and increases in familiarity as well as to decreases in attention (Park et al., 2014; Sederberg et al., 2006; Serruya, Sederberg, & Kahana, 2014). Decreases in HFA have also been linked to repetition suppression, where activity elicited by a stimulus decreases from the first to the second presentation (Rodriguez Merzagora et al., 2014). Additionally, these decreases in activity could be related to increases in processing efficiency or learning of the task environment and/or stimuli. Future work will be needed to determine how shifts in HFA relate to changes in novelty, attention, processing efficiency, or item recognition.

A limitation of the current study is that the research was conducted with patients with epilepsy, whose brains may undergo epilepsy-related changes. However, in line with recommendations
outlined in a review by human and nonhuman primate intracranial researchers, we excluded trials that contained epileptiform discharges and only included recordings from non-epileptic tissues (Parvizi & Kastner, 2018).

The ability to update incorrect or imprecise associations based on negative feedback is critical for the formation of accurate representations of the environment that can be used to guide behavior. Our results suggest an important role for the MTL in signaling error at feedback and in updating representations of object-location associations. Future work should aim to identify the mechanisms by which these signals are integrated into mnemonic representations and to assess whether these mechanisms extend to other types of associations in episodic memory.

**Materials and methods:**

**Participants**

Subjects were 9 patients (5 female, 4 male, age 21-69) who had stereotactically implanted intracranial depth electrodes (Integra or Ad-Tech, 5-mm inter-electrode spacing) placed at the University of California, Irvine Medical Center to localize the seizure onset zone for possible surgical resection. Informed consent was obtained from each subject prior to testing and the research protocol was approved by the IRB of the University of California, Irvine. Electrode placement was exclusively guided by clinical needs.

**Spatial learning task**

The task was presented on a laptop computer screen set at a comfortable distance from the patient. Three hundred and sixty locations were generated along the circumference of a circle centered on the screen with a spacing of 1°. Thirty images of common objects were selected from a set previously used by our group (Stark, Stevenson, Wu, Rutledge, & Stark, 2015). Prior
to encoding, subjects were told that they would be shown objects at different locations on the
computer screen and were asked to try to remember the location of each object. During
encoding, the 30 objects appeared one at a time at pseudorandomly assigned circle locations
(Figure 1; 1.2 +/- 0.2 second ITI). During each of the three training blocks, subjects were shown
the same objects again in pseudorandom order, this time at the top of the screen. Subjects were
instructed to wait 1 second (until text that read ‘Wait…’ disappeared from the screen) before
using a mouse wheel to move the object to where it appeared during encoding. Subjects
pressed the space bar to indicate that they were finished placing the object. The object was
then shown in the correct location for one second as feedback. After the third training block,
there was a final test block in which no feedback was given. There was a short (<1 minute)
break after the encoding phase and between each training/test block during which task
instructions were read. If subjects performed more than one session, new objects and locations
were used. No more than one session was performed on each day.

Behavioral analysis

Error on the spatial learning task was measured as the number of degrees between where
subjects placed the object and the correct location. Histograms were used to examine the
distribution of error values across and within task blocks. We used mixture modeling, as
implemented by the MemFit function of Memtoolbox (Suchow et al., 2013), to obtain an estimate
of two parameters describing these distributions: the guess rate (g), which reflects the area
under the uniform distribution, and the standard deviation of the von Mises distribution
(SDMem). We first fit the model using trials from all training and test blocks. Trials that had less
than a 10% chance of being remembered with some degree of precision based on the von
Mises distribution of this model were placed in the high error condition. The remaining trials
were sorted by error and split evenly into the medium and low error conditions. We then fit the
model using trials within each block (i.e. the model was fit separately for training blocks 1 and 2) to determine how the guess rate and precision (SDMem) changed across task blocks.

Electrode localization

The electrode localization was performed using pre- and post-implantation structural T1-weighted 1mm isotropic MRI scans as well as post-implantation CT scans. For each participant, the post-implantation MRI and CT scans were registered to the pre-implantation scan using a 6-parameter rigid body transformation implemented with Advanced Normalization Tools – ANTs (Avants et al., 2011). Electrodes were localized within MTL subregions using a high-resolution (.55 mm) in-house anatomical template with manual tracings of hippocampal subfields and parahippocampal gyrus subregions (Yassa & Stark, 2009). Hippocampal subfield segmentation followed our previously published protocols (Yassa & Stark, 2009). The labeled template was resampled and aligned to each subject’s pre-implantation scan using ANTs Symmetric Normalization, so that the labels could be used to guide localization. Each electrode location was determined by examining the co-registered pre- and post-implantation MRIs and identifying the ROI that corresponded to the center of the electrode artifact in the post-implantation MRI and CT. Cases in which electrodes were on the border between ROI’s or between gray matter and white matter were noted as such. Outside the MTL, electrode localization was guided by a FreeSurfer cortical parcellation of the pre-implantation MRI (Fischl et al., 2004).

Data collection and preprocessing

Intracranial EEG data were recorded using a Nihon Khoden recording system, analog-filtered above 0.01 Hz and digitally sampled at 5000 Hz. After acquisition, data were demeaned and band-pass filtered from 0.3 Hz to 350 Hz using a two pass zero phase delay Butterworth infinite impulse response (IIRR) filter. Power spectra were examined to identify line noise and a
Butterworth notch filter was used to remove 60 Hz noise and harmonics. All electrodes were re-referenced to a white matter electrode located on the same depth electrode probe. A neurologist (J.L.) with subspecialty training in epilepsy visually inspected continuous recordings from each session to identify all data with interictal epileptiform discharges. Data were also inspected for excessive noise, including broadband electromagnetic noise from hospital equipment. To avoid potentially biasing the results, the neurologist was blinded to trial information (e.g. stimulus onset and behavioral performance) as well as to electrode location. Only data from recordings contralateral to the seizure source or outside of the seizure onset zone were used in subsequent analyses.

**HFA Analyses**

Intracranial recordings were broken into event-related epochs (3 seconds pre-stimulus onset and 3 seconds post-stimulus offset) and convolved with complex Morlet wavelets, implemented using the FieldTrip toolbox, to obtain a measure of instantaneous power (Oostenveld, Fries, Maris, & Schoffelen, 2011). Center frequencies ranged from 1 to 150 Hz, with a spacing of 1 Hz and a variable cycle number of 4-15. Power was baseline corrected to the average pre-stimulus power (0.5 to 0.2 prior to stimulus onset), resulting in a measure of the relative change in power per frequency at each time point. For analyses that examined power across training blocks, power was z-transformed separately within each session to account for differences in power and noise across sessions. For analyses that examined power within training blocks, power was z-transformed within each block. We then averaged normalized power over our a priori gamma frequency range of 40-100 Hz, based on prior literature showing MTL and dIPFC gamma activity in this range (J. F. Burke et al., 2014; Greenberg, Burke, Haque, Kahana, & Zaghoul, 2015; Sederberg et al., 2007; Stevenson et al., 2018). Next, we averaged HFA across electrodes located within the MTL, including electrodes in the hippocampus as well as the entorhinal,
perirhinal, and parahippocampal cortices, and within the dIPFC (Broadman area (BA) 9/10/46).

To examine the specificity of effects, we also ran analyses on regions outside of the MTL and dIPFC including the lateral temporal cortex, the insula, the caudal prefrontal cortex (BA 6/8), orbitofrontal cortex, and anterior cingulate cortex.

We used a cluster-based permutation approach implemented using the FieldTrip toolbox to test for differences in HFA across conditions (high, medium, and low error) at each time point within each region (Oostenveld et al., 2011). In order to test for correlations between HFA and error, we averaged HFA over 100ms sliding windows and calculated the Pearson correlation between HFA and angular error at intervals of 20ms. For this analysis, angular error was logged to account for the non-normal distribution of error. We used permutation testing to ensure that the observed correlations were not driven by outliers or other biases in the data. A null distribution of r values was created by shuffling the trial labels between conditions 1000 times. We derived p-values for the observed r values using the cumulative distribution function of these distributions. We used partial correlations to test for associations between HFA at feedback for high error trials and the subsequent change in error from one block the next, controlling for error on the earlier task block (i.e. the association between HFA at feedback on a high error trial at T1 and the change in error for that object-location association from T1 to T2, controlling for error on that trial at T1). P-values for the observed rho values were obtained via permutation testing as described above.

Acknowledgments. We thank C.E.L. Stark and R.T. Knight for discussion, L.D. Harriger, J. Zheng, H. Zhang, and I. Skelin for data collection, M. Hollearn, A. Chun, and B. Houalla for assistance with data preprocessing, the staff of the Epilepsy Monitoring Unit at the University of
California Irvine Medical Center, and the patients who volunteered in this research. This work
was supported by NIMH R01MH102392 and NIA R21AG053555 (PI: Yassa). R.F.S. was
supported by NINDS T32NS45540 (PI: T.Z. Baram).
REFERENCES

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage, 54*(3), 2033-2044. doi:10.1016/j.neuroimage.2010.09.025

Blumenfeld, R., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging. *The Neuroscientist, 13*, 280-291. doi:10.1177/1073858407299290

Brincat, S. L., & Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nat Neurosci, 18*(4), 576-581. doi:10.1038/nn.3954

Burke, J. F., Long, N. M., Zaghloul, K. A., Sharan, A. D., Sperling, M. R., & Kahana, M. J. (2014). Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage, 85 Pt 2*, 834-843. doi:10.1016/j.neuroimage.2013.06.067

Burke, J. F., Ramayya, A. G., & Kahana, M. J. (2015). Human intracranial high-frequency activity during memory processing: neural oscillations or stochastic volatility? *Current Opinion in Neurobiology, 31*, 104-110. doi:10.1016/j.conb.2014.09.003

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci, 3*(3), 201-215. doi:10.1038/nnr755

Duncan, K., Ketz, N., Inati, S. J., & Davachi, L. (2012). Evidence for area CA1 as a match/mismatch detector: A high-resolution fMRI study of the human hippocampus. *Hippocampus, 22*(3), 389-398. doi:10.1002/hipo.20933

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of, 123-152. doi:10.1146/annurev.neuro.30.051606.094328

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., . . . Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cereb Cortex, 14*(1), 11-22. doi:10.1093/cercor/bhg087

Greenberg, J. A., Burke, J. F., Haque, R., Kahana, M. J., & Zaghloul, K. A. (2015). Decreases in theta and increases in high frequency activity underlie associative memory encoding. *Neuroimage, 114*, 257-263. doi:10.1016/j.neuroimage.2015.03.077

Hargreaves, E. L., Mattfeld, A. T., Stark, C. E., & Suzuki, W. A. (2012). Conserved fMRI and LFP signals during new associative learning in the human and macaque monkey medial temporal lobe. *Neuron, 74*(4), 743-752. doi:10.1016/j.neuron.2012.03.029

Hargreaves, E. L., Mattfeld, A. T., Stark, C. E. L., & Suzuki, W. a. (2012). Conserved fMRI and LFP signals during new associative learning in the human and macaque monkey medial temporal lobe. *Neuron, 74*, 743-752. doi:10.1016/j.neuron.2012.03.029

Histed, M. H., Pasupathy, A., & Miller, E. K. (2009). Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron, 63*(2), 244-253. doi:10.1016/j.neuron.2009.06.019

Igarashi, K. M., Lu, L., Colgin, L. L., Moser, M. B., & Moser, E. I. (2014). Coordination of entorhinal-hippocampal ensemble activity during associative learning. *Nature, 510*(7503), 143-147. doi:10.1038/nature13162

Ison, M. J., Quian Quiroga, R., & Fried, I. (2015). Rapid Encoding of New Memories by Individual Neurons in the Human Brain. *Neuron, 87*(1), 220-230. doi:10.1016/j.neuron.2015.06.016
Koen, J. D., Borders, A. A., Petzold, M. T., & Yonelinas, A. P. (2017). Visual short-term memory for high resolution associations is impaired in patients with medial temporal lobe damage. *Hippocampus, 27*(2), 184-193. doi:10.1002/hipo.22682

Kolarik, B. S., Shahlaike, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., . . . Ekstrom, A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris Water Maze: A case study. *Neuropsychologia, 80*, 90-101. doi:10.1016/j.neuropsychologia.2015.11.013

Law, J. R., Flanery, M. A., Wirth, S., Yanike, M., Smith, A. C., Frank, L. M., . . . Stark, C. E. (2005). Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J Neurosci, 25*(24), 5720-5729. doi:10.1523/JNEUROSCI.4935-04.2005

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nat Neurosci, 21*(11), 1609-1617. doi:10.1038/s41593-018-0232-z

Nilakantan, A. S., Bridge, D. J., Gagnon, E. P., VanHaerents, S. A., & Voss, J. L. (2017). Stimulation of the Posterior Cortical-Hippocampal Network Enhances Precision of Memory Recollection. *Curr Biol, 27*(3), 465-470. doi:10.1016/j.cub.2016.12.042

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci, 2011*, 156869. doi:10.1155/2011/156869

Park, J., Lee, H., Kim, T., Park, G. Y., Lee, E. M., Baek, S., . . . Kang, J. K. (2014). Role of low- and high-frequency oscillations in the human hippocampus for encoding environmental novelty during a spatial navigation task. *Hippocampus, 24*(11), 1341-1352. doi:10.1002/hipo.22315

Parvizi, J., & Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nat Neurosci, 21*(4), 474-483. doi:10.1038/s41593-018-0108-2

Reagh, Z. M., & Yassa, M. A. (2014). Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proc Natl Acad Sci U S A, 111*(40), E4264-4273. doi:10.1073/pnas.1411250111

Rodriguez Merzagora, A., Coffey, T. J., Sperling, M. R., Sharan, A., Litt, B., Baltuch, G., & Jacobs, J. (2014). Repeated stimuli elicit diminished high-gamma electrocorticographic responses. *Neuroimage, 85 Pt 2*, 844-852. doi:10.1016/j.neuroimage.2013.07.006

Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. a. (2009). Strengthening individual memories by reactivating them during sleep. *Science (New York, N.Y.), 326*, 1079. doi:10.1126/science.1179013

Rutishauser, U. (2008). Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *105.

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*(3), 1422-1431. doi:10.1016/j.neuroimage.2006.04.223

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., . . . Kahana, M. J. (2007). Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cereb Cortex, 17*(5), 1190-1196. doi:10.1093/cercor/bhl030

Serruya, M. D., Sederberg, P. B., & Kahana, M. J. (2014). Power shifts track serial position and modulate encoding in human episodic memory. *Cereb Cortex, 24*(2), 403-413. doi:10.1093/cercor/bhs318

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nat Neurosci, 20*(11), 1643-1653. doi:10.1038/nn.4650
Stark, S. M., Stevenson, R., Wu, C., Rutledge, S., & Stark, C. E. (2015). Stability of age-related deficits in the mnemonic similarity task across task variations. *Behav Neurosci, 129*(3), 257-268. doi:10.1037/bne0000055

Stevenson, R. F., Zheng, J., Mnatsakanyan, L., Vadera, S., Knight, R. T., Lin, J. J., & Yassa, M. A. (2018). Hippocampal CA1 gamma power predicts the precision of spatial memory judgments. *Proc Natl Acad Sci U S A, 115*(40), 10148-10153. doi:10.1073/pnas.1805724115

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *J Vis, 13*(10). doi:10.1167/13.10.9

Suthana, N. A., Donix, M., Wozny, D. R., Bazih, A., Jones, M., Heidemann, R. M., . . . Bookheimer, S. Y. (2015). High-resolution 7T fMRI of Human Hippocampal Subfields during Associative Learning. *J Cogn Neurosci, 27*(6), 1194-1206. doi:10.1162/jocn_a_00772

Sutterer, D. W., & Awh, E. (2016). Retrieval practice enhances the accessibility but not the quality of memory. *Psychon Bull Rev, 23*(3), 831-841. doi:10.3758/s13423-015-0937-x

Wirth, S., Avsar, E., Chiu, C. C., Sharma, V., Smith, A. C., Brown, E., & Suzuki, W. A. (2009). Trial outcome and associative learning signals in the monkey hippocampus. *Neuron, 61*(6), 930-940. doi:10.1016/j.neuron.2009.01.012

Yassa, M. A., & Stark, C. E. (2009). A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *Neuroimage, 44*(2), 319-327. doi:10.1016/j.neuroimage.2008.09.016

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-235. doi:10.1038/nature06860

Zheng, J., Anderson, K. L., Leal, S. L., Shestyuk, A., Gulsen, G., Mnatsakanyan, L., . . . Lin, J. J. (2017). Amygdala-hippocampal dynamics during salient information processing. *Nat Commun, 8*, 14413. doi:10.1038/ncomms14413
Medial temporal lobe high-frequency activity at feedback signals error and predicts subsequent memory during spatial learning

Stevenson et al.

Supplementary information

**Fig S1. Mixture model fit for all sessions.** Histograms showing the distribution of error for all trials pooled across training blocks and mixture model fit for each session. S: session, g: guess rate, sd: standard deviation of the von Mises distribution.
Fig S2. High-frequency activity (HFA) across task blocks at retrieval. Time course of HFA (40-100 Hz) signal for each task block (T1-3, and Test for retrieval, T1-3 for feedback) in all regions at retrieval (A-G) and feedback (H-N). Gray horizontal lines indicate time points where there were significant differences between conditions (ANOVA, p < 0.05, cluster-corrected). The start and end of the retrieval/feedback window are indicated by vertical dotted lines. Colored shaded regions indicate s.e.m.
Fig S3. HFA and error at retrieval and feedback. Time course of HFA signal normalized within blocks for high, medium, and low error trials. Gray horizontal lines indicate time points where there were significant differences between conditions (ANOVA, p < 0.05, cluster-corrected). The start and end of the retrieval/feedback window are indicated by vertical dotted lines. Colored shaded regions indicate s.e.m.