SUPPLEMENTARY INFORMATION

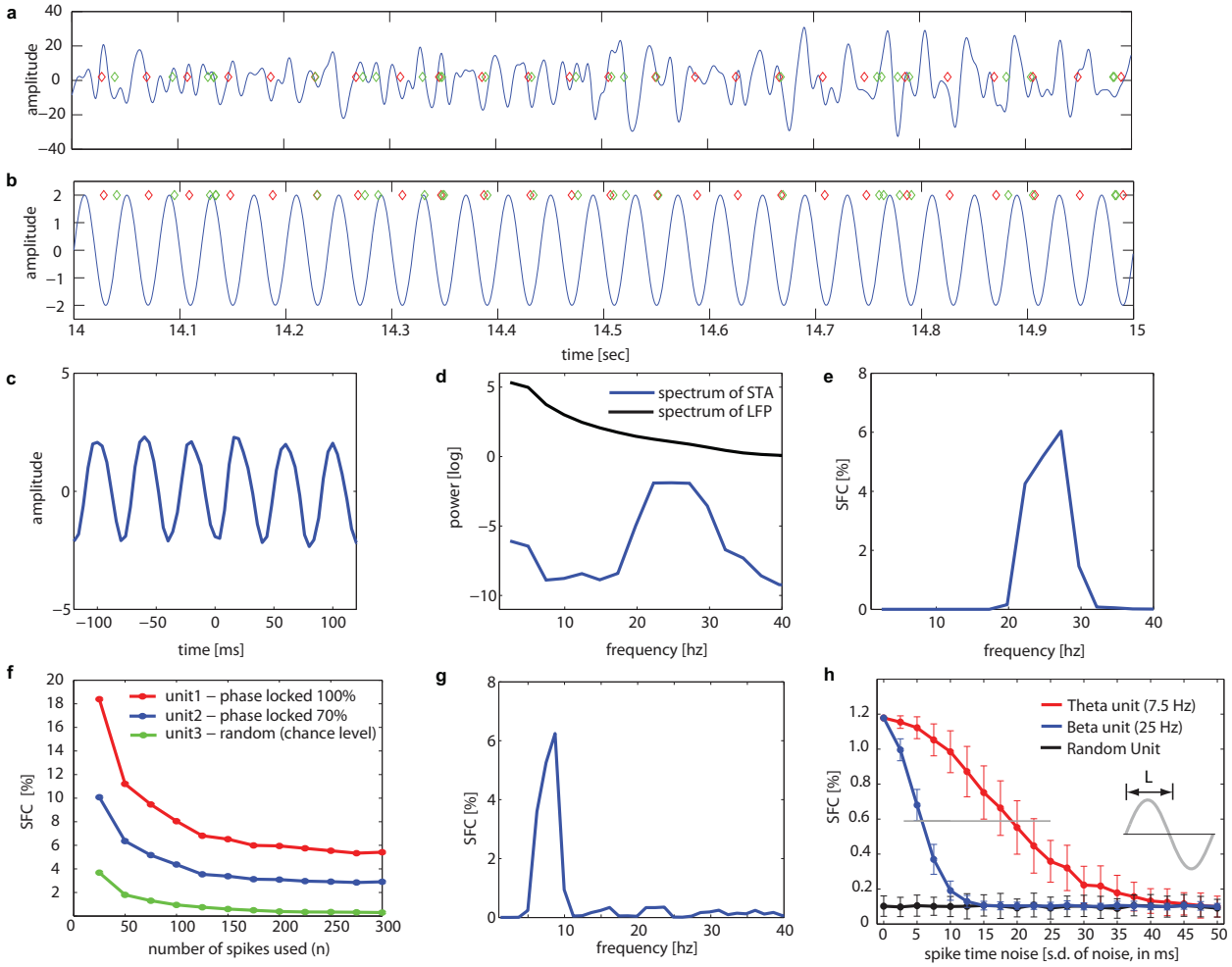

Figure S1: Calculation of spike-field coherence (SFC) with simulated data**. (A)** Bandpass filtered version of the simulated raw signal (10-300 Hz). **(B)** One part of (A) is a 25 Hz oscillation of amplitude 2. The red and green dots indicate the spikes of two simulated neurons. The red neuron follows the 25Hz oscillation whereas the green neuron fires randomly (but with the same rate). **(C-E)** Procedure to calculate the SFC. **(C)** The spike triggered average (STA) for the red simulated neuron shown in (B), calculated using 1400 spikes (spikes are at t = 0). **(D)** The power spectrum of the STA (blue) shown in (C). Notice the distinct peak around 25Hz. The spike-triggered power (STP) is the average spectrum of all LFP segments used to construct the STA (black). **(E)** The SFC as a function of frequency. Note the peak of the SFC at ~25 Hz. Also note the numerical value of the peak SFC (~ 6%, despite almost perfect phase locking). The spectra were calculated using multi-taper estimates with a half-width of 10 Hz. This is visible as the width of the bump in (E). A smaller half width was used for the real data. **(F)** The mean peak height of the SFC at 25 Hz as shown in (E) as a function of the number of spikes used for its calculation. Unit 1 (red) and 3 (green) are as described in (B). The SFC shown in (C-E) is for unit 1. The third unit shown (blue) fires the exact same spikes as does unit 1, except that 30% of its spikes were randomly shuffled in time. Thus, this unit fires the same number of spikes but only 70% of them are phase-locked. Note the influence of this manipulation on the SFC (about a 50% reduction in SFC,

regardless of the number of spikes used). Also note that it is crucial that the same numbers of spikes are considered for all experimental conditions. This is always the case in our comparisons. The reason for this bias is that, under the null hypothesis of no phase locking, the amplitude of the STA tends to zero for large samples (the green line). Due to imperfect phase locking and noise of real neurons, the amplitude of the STA gets smaller as a function of the number of spikes considered. As a result, the relative difference in SFC between conditions, rather than the SFC magnitude, is important. **(G)** SFC as a function of frequency (same as in panel E), but for a frequency of 8 Hz and a half-width of 5 Hz. n = 250 spikes were used. **(H)** The SFC as a function of spike timing jitter (in terms of s.d. of a normal distribution). s.d.=0 is perfect phase locking. Note how the SFC decays as spike timing accuracy decreases. The SFC is reduced by 50% (gray line) for spike-timing jitter equal to 30% of the half-cycle length (the time it takes the oscillation from peak to trough, as illustrated by L in the inset) of the underlying oscillation (red: 7.5Hz, half-cycle length 66.7ms; blue, half-cycle length of 20ms). Thus, for the theta oscillation (red) the 50% point is reached for jitter of s.d.= 20 ms and for the beta unit (blue) for s.d. = 6 ms. Phase-locking becomes indistiguishable from no phase locking if the noise exceeds ~75% of the half-cycle length (red: 44ms, blue: 13ms). Each datapoint represents the SFC calculated for 1000 spikes. Errorbars shown are ±s.d. for 100 simulations.

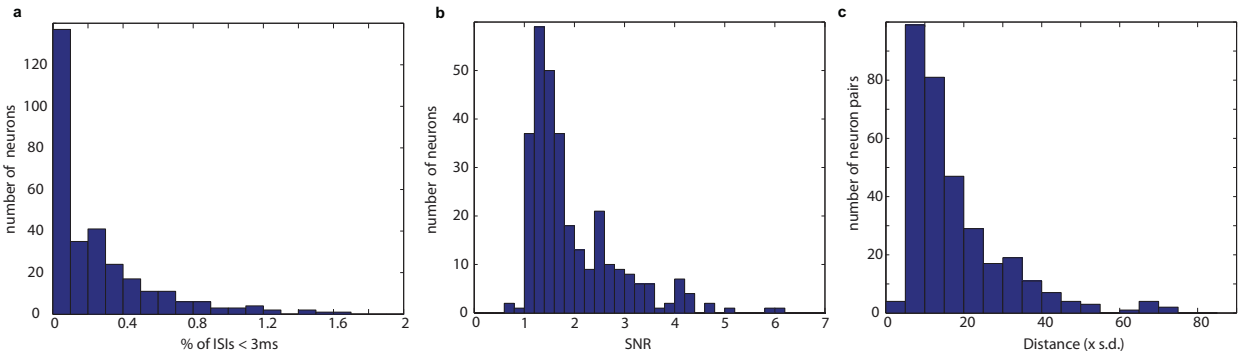

Figure S2: Spike sorting quantification. There were a total of 328 neurons ("putative single units"). **(A)** Summary of the inter-spike intervals (ISIs) of all neurons. Shown is a histogram of the proportion of ISIs that were smaller than 3ms. A large majority of clusters lie below 0.5%, indicating good isolation of units. The mean percentage of ISIs below 3% was 0.26 ± 0.02. **(B)** Histogram of the signal-to-noise ratio (SNR) of the mean waveform for each neuron. The mean SNR was 2.0 ± 0.05. **(C)** Pairwise distance of neurons isolated on the same wire. For this plot, only units on wires (n=100) with multiple neurons are considered. There were a total of n = 328 pairs. On a wire with more than one neuron, there were on average 2.8 ± 0.1 neurons. The distance was calculated using the projection test (see methods). It is quantified in terms of s.d. of the variance around the mean waveform of the neuron. The mean distance was 18.0 ± 0.7 s.d. All errors are ± s.e.m.

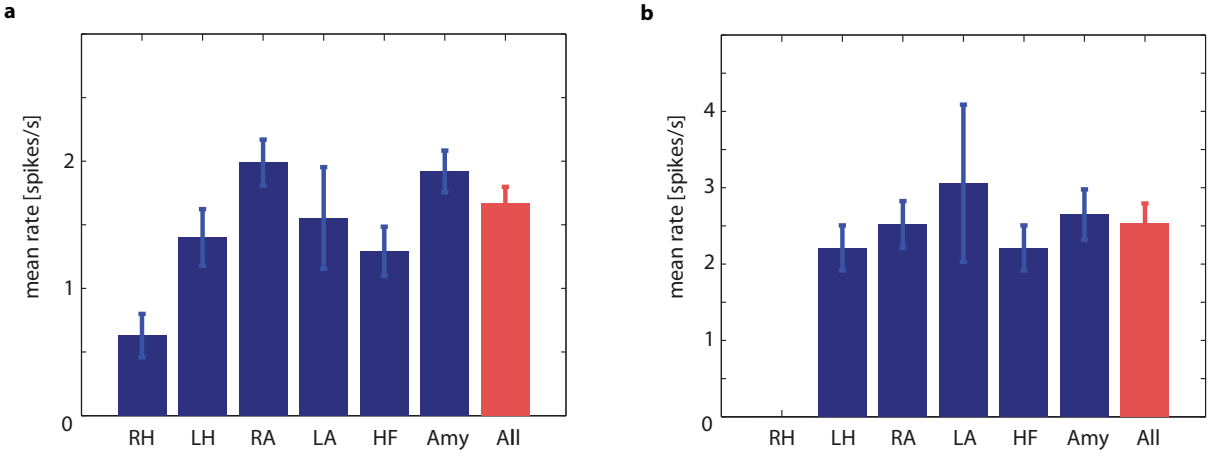

Figure S3: Mean firing rates and numbers of neurons from different brain areas. **(A)** Mean firing rates of neurons for each area. All neurons are included, regardless of firing rate. The number of units for each area was as follows, from left to right: 17,103, 155, and 30. **(B)** Mean firing rates of neurons that were significantly phase-locked to the theta oscillation (compare to Fig 2). Number of neurons for each area was as follows, from left to right: 0, 13, 29 and 9 (total 51). Errorbars are ± s.e.m. Abbreviations: right hippocampus (RH), left hippocampus (LH), right amygdala (RA), left amygdala (LA), Hippocampus (HF), Amygdala (Amy).

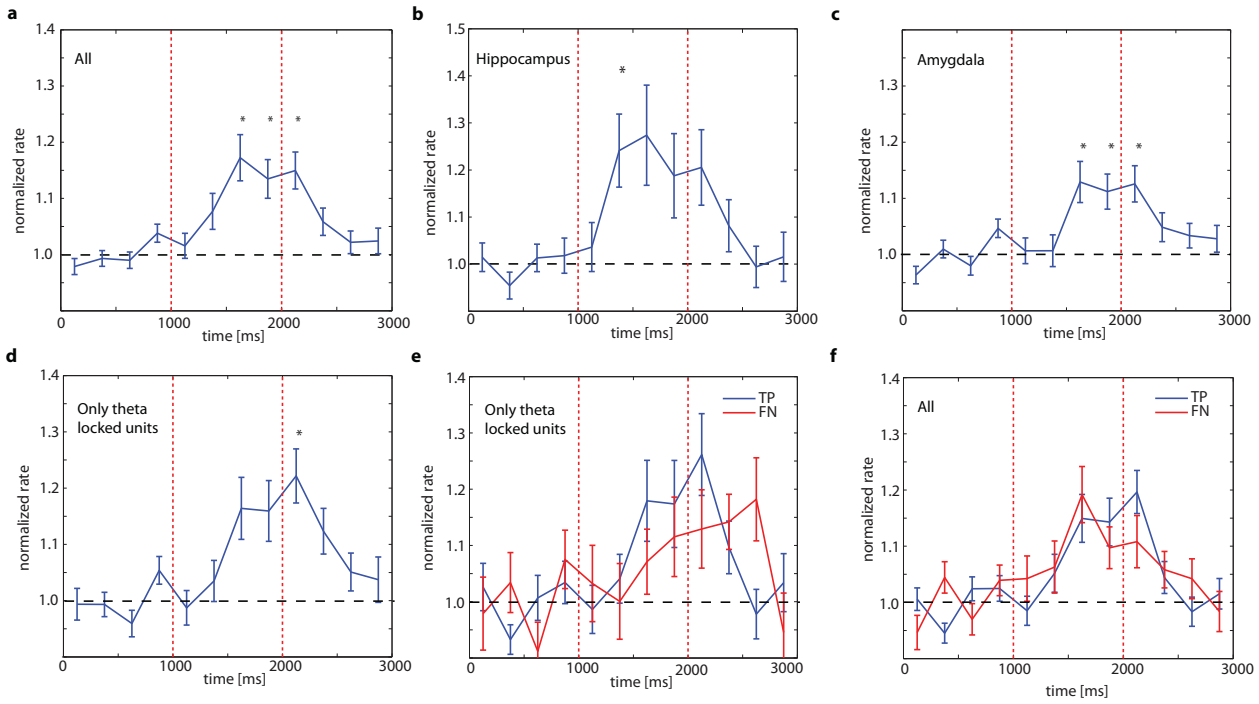

Figure S4: Mean firing rate as a function of time, relative to stimulus onset. On average, neurons increased their firing rate after stimulus onset. This was true for hippocampal as well as amygdala units. The firing rate change did not distinguish between stimuli that were later remembered or forgotten. **(A)** All recorded neurons (during learning) that fired at least 50 spikes during the learning trials (n = 107). The first significant response to the stimulus was in the 250-500ms period. **(B)** Same as A, but only for hippocampal units (n = 32). **(C)** Same as (A), but only for units recorded from the amygdala (n =75). **(D)** Same as (A), but only for units that were used for the SFC analysis (n = 33, see Fig 3 and 4). The firing rate response of the neurons that were theta-phase locked was very similar to the overall population. **(E)** The mean firing rate of the same neurons as in (D), but separately for trials that were later remembered (TP, blue) and forgotten (FN, red). There was no significant difference. **(F)** The mean firing rate separately for TP and FN trials for all neurons (n = 107, as in (A)). There was no significant difference. All comparisons were made using paired t-tests and multiple-comparison corrected for 12 comparisons ($p<=0.0042$, Bonferroni corrected $p<0.05$). For (A-D), the statistical comparisons were made against the baseline pre-stimulus firing rate (1s baseline). For (D-E), the statistical comparison was between the two behavioral conditions, separately for each bin. The firing-rate was normalized to the 1 sec pre-stimulus baseline. Units that fired less than 0.25 Hz during the baseline were not included to avoid normalization artifacts. The bin size was 250 ms. Errorbars are ± s.e.m. of number of spikes in this bin, across neurons. The dashed vertical lines indicate the stimulus onset/offset.

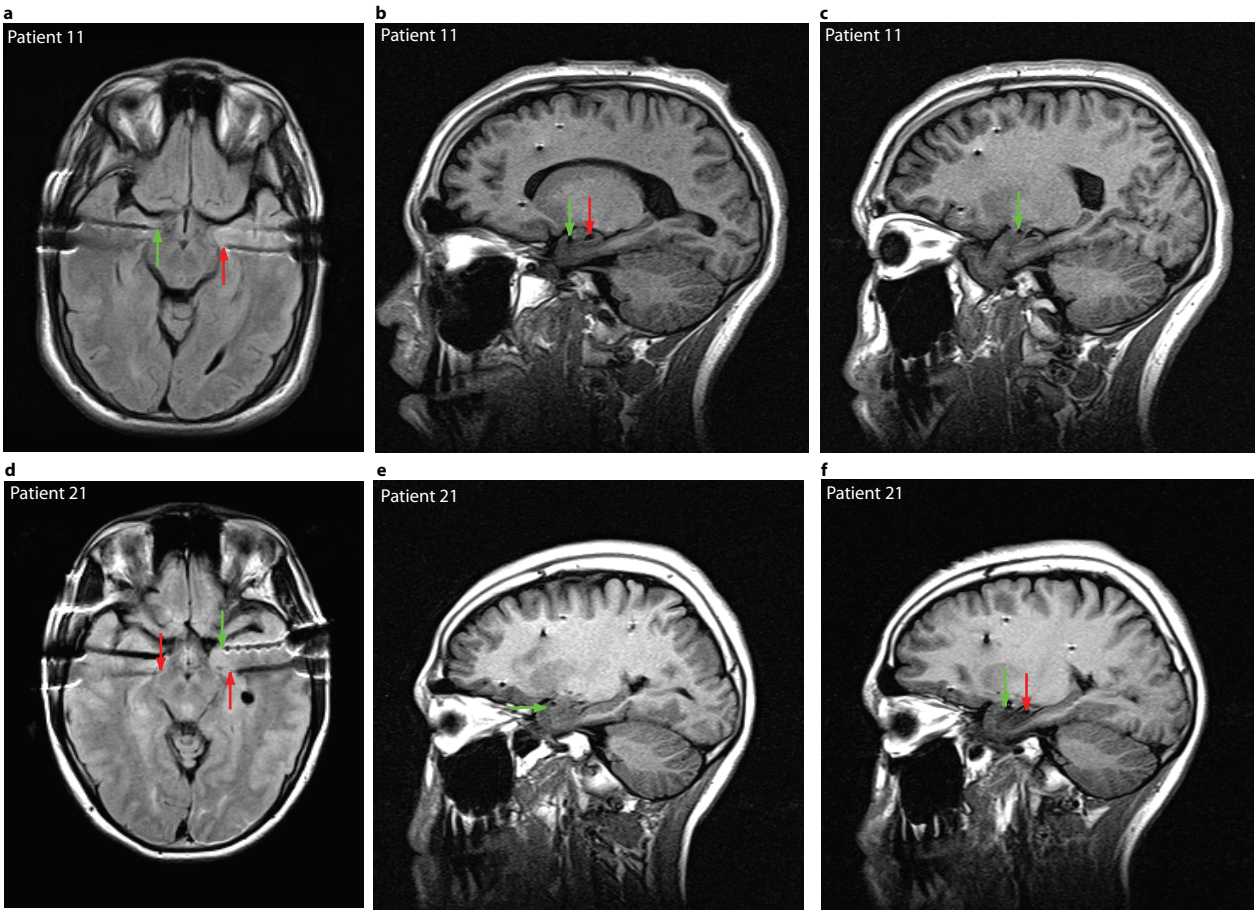

Figure S5: Electrode localization with structural MRIs. Shown are examples from two patients (see table 1 for demographic information). In the axial plane, the patient's left brain is on the right side. All images are T1 FLAIR images, acquired with a 1.5 Tesla clinical MRI (Toshiba) Slices shown are scans with high resolution only in one plane (0.75x0.75mm, slice-spacing 5-6mm). Red and green arrows always indicate an electrode in the hippocampus, green in the amygdala. **(A-C)** First example. **(A)** Axial T1 Flair image. A **(B)** Saggital image of the left side (T1 Flair). **(C)** Saggital image of the right side. **(D-E)** Second example patient. **(D)** Axial T1 image. **(E)** Saggital image of the left side. The hippocampal electrode is more lateral and not visible on this slice. **(F)** Saggital image of the right side.

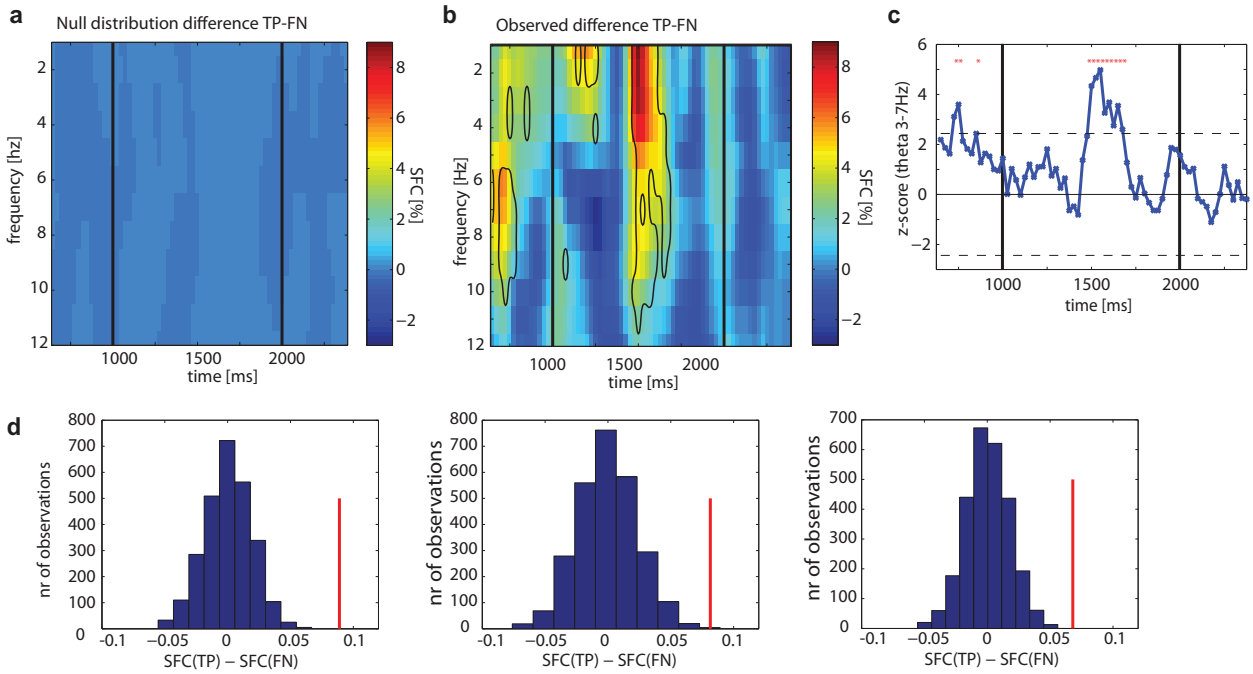

Figure S6: Null distribution for statistical comparison of SFCs. The distribution of the difference between the SFC for TP and FN trials (time course) was estimated by creating 2000 random bootstrap samples at every frequency and time. **(A)** Colorplot of the mean of the estimated null distribution. This is the expected mean difference due to chance. **(B)** The observed difference, as shown in Fig 4 for comparison. Contours indicate regions of significantly higher SFC for TP trials, relative to FN trials (FDR $q<0.05$, corresponding to $Z>2.57$). A similar test for significantly smaller SFC for TP trials shows no significant regions. (**C**) Timecourse of average theta-band (3-7Hz) SFC, quantified as z-scored difference between TP and FN trials. z-scores are estimated empirically against the null distribution (see A). Stars indicate significance at $q<0.05$ (FDR; corresponding to $Z>2.4$; dashed line). **(D)** Three examples of the distribution of chance values and the observed value (red line). Each plot shows the values from one point in the 2D plane of frequency and time as shown in A-B. From left to right: f = 4 Hz, t = 1500 ms; f = 2 Hz, t = 1200 ms; f = 7 Hz, t = 650 ms. Notice the very clear separation of the observed value (red) from the values expected due to chance. The p-values, from left to right, were: 0.0, 0.00075, 0.0 (empirically estimated). The z-values are, from left to right: 4.8, 3.4, 3.9.

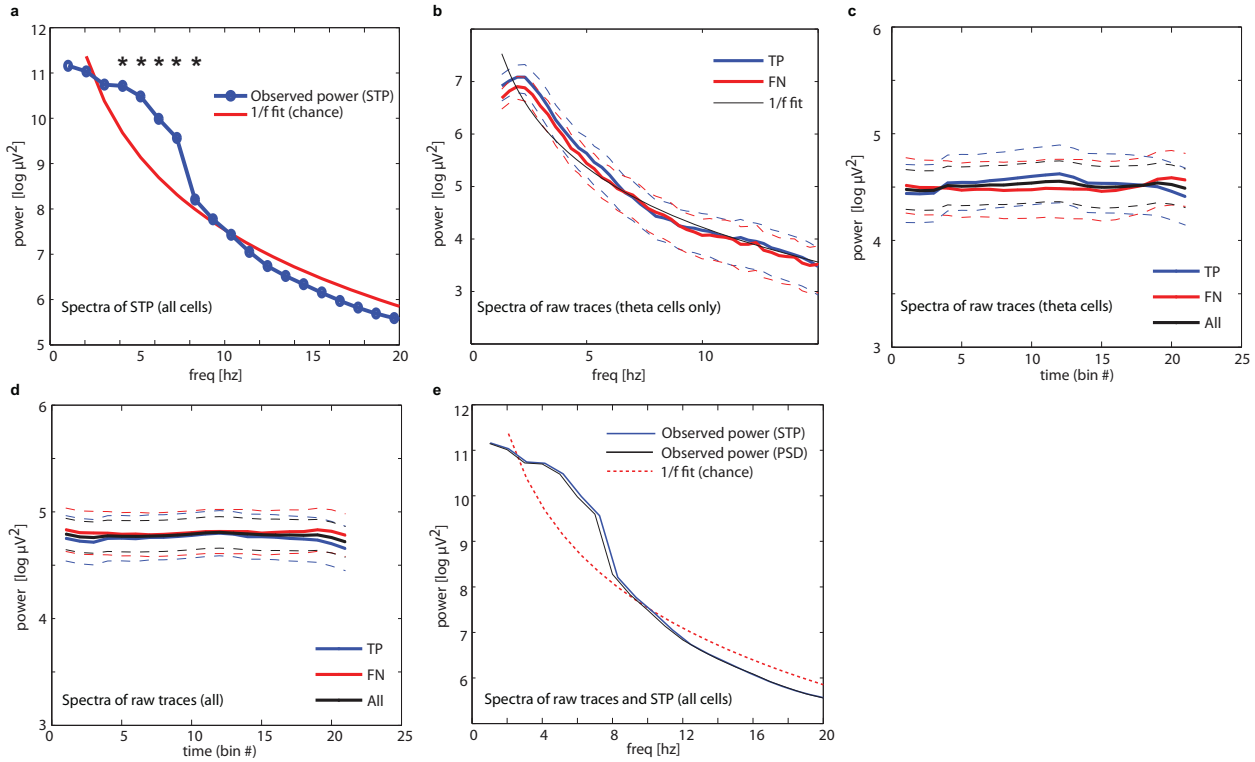

Figure S7: Theta oscillations are present in the power spectra of the LFP and the STP. **(A)** Average spectra of spike-triggered segments of LFP (the STP). This plot shows the average STP for all cells that fired at least 50 spikes to both TP and FN trials (n=112). Cells were not pre-selected for the presence of phase locking. Theta frequency oscillations were clearly present, as indicated in the deviation from the 1/f fit from 4-8 Hz (red, f^{α}, α=--2.10 ± 0.04). Significance was assessed using a Kolmogorov-Smirnov test against the power expected by chance (red), Bonferroni-corrected for multiple comparisons. **(B)** The average power spectrum of the entire 3 s trace of all trials for theta phase-locked neurons (n = 33). Increased theta power was present, however there was no significant difference between the behavioral conditions (TP vs. FN). The dashed lines show ± s.e.m. **(C)** Timecourse of average theta power (log) for all channels (n = 33) that were used for calculating the SFC difference (Fig 3). The spectra were calculated using a moving window of size 1000 ms, which was advanced by a stepsize of 100 ms. The total trial duration shown is 3 s. Thus, the first datapoint is entirely during the baseline whereas data-point 11 is entirely during the stimulus (which is on the screen for 1s). Dashed lines indicate ± s.e.m. The log power was averaged, to reduce the influence of different absolute amplitudes of different channels (the result was unchanged if the raw power was averaged). There was no significant difference between behavioral groups. Thus, theta power was not (on average) strongly modulated by the onset or offset of the stimulus (also see (F)). **(D)** Same as in (D), but for all channels (n = 112) with cells that had at least 50 spikes for all TP and FN trials. **(E)** Comparison of the power spectra of the STP (as shown in A) and the raw traces, for all n = 112 channels with cells with at least 50 spikes. The raw trace spectra (black) are based on the same data as (D), but collapsed over time. Note the close overlap of the two lines, indicating that the STP power and the raw trace power are the same. Thus, theta power triggered by spikes (the STP) is not additionally modulated on a short time scale. The same is true for the spectra of the n = 33 channels used for the SFC (not shown). Note that the differences in amplitude between (A) and (B) are due to different numbers of channels (of different amplitudes). As shown in this panel, there is no amplitude difference between the two (for the same channels). All spectra shown in this figure were calculated using multi-taper analysis (same parameters as for the SFC). The frequency resolution (half-width) for (A) is 4.2 Hz, 1.3 Hz for (B), and 4.0 Hz for (C-D) (due to different lengths).

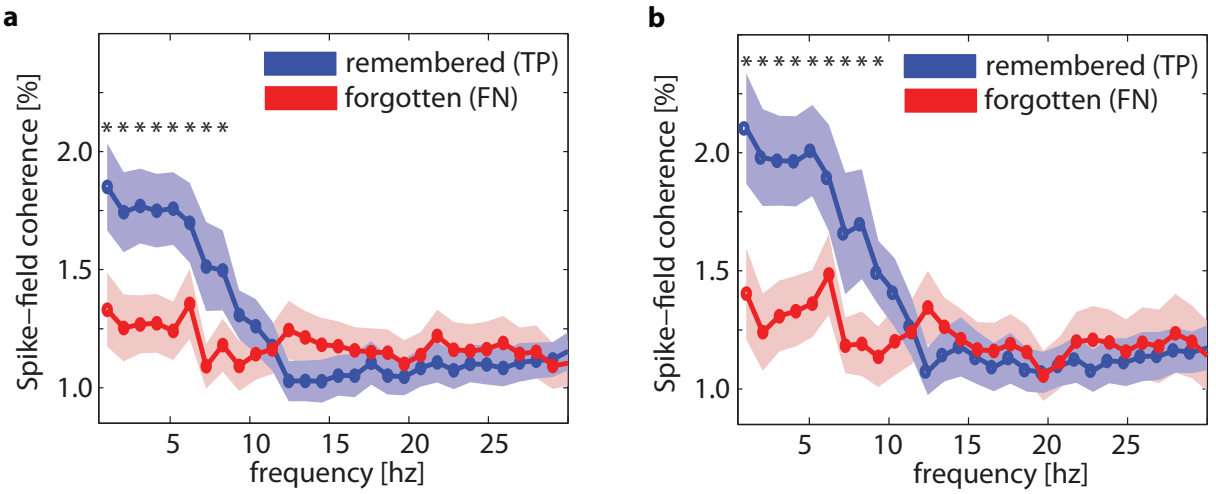

Figure S8: Different methods of selecting theta-frequency phase-locked neurons. In the primary analysis, we selected cells based on all spikes during the entire learning block (including pre-and post stimulus baseline as well as the response period). We also selected cells using two other methods: using only spikes during the 3s learning trials themselves (A) or using only spikes which occurred during significant theta power (greater than the median power, (B)). While different numbers of cells were selected using these selection criteria, the results remained very similar. **(A)** The SFC for remembered compared to forgotten trials (compare to Fig 3). Here, phase-locked units were selected based on only the spikes during the trials themselves instead of the entire task. 65 neurons were selected and had enough spikes to be included in the SFC comparison. The average theta-band SFC for TP/FN trials was 1.70 ± 0.14% and 1.25 ± 0.11%, respectively and the difference was statistically significant ($p<0.00001$). The difference (36%) was comparable to the one shown in (B) and Fig 3. **(B)** Result for selecting using all spikes during the task that occurred during times of significant theta power. N = 34 neurons were selected. The average theta-band SFC for TP/FN trials was 1.90 ± 0.18% and 1.33 ± 0.13%, respectively and the difference was statistically significant ($p<0.00001$). The difference (43%) between the SFC was comparable to the one shown in (A) and Fig 3. We also compared the average SFC for hippocampal and amygdala neurons separately for both conditions (A and B). In both cases, both classes were individually significant (for A: p=0.014 (hippocampus) and p=0.00043 (amygdala), B: p=0.0012 and p=0.0030).

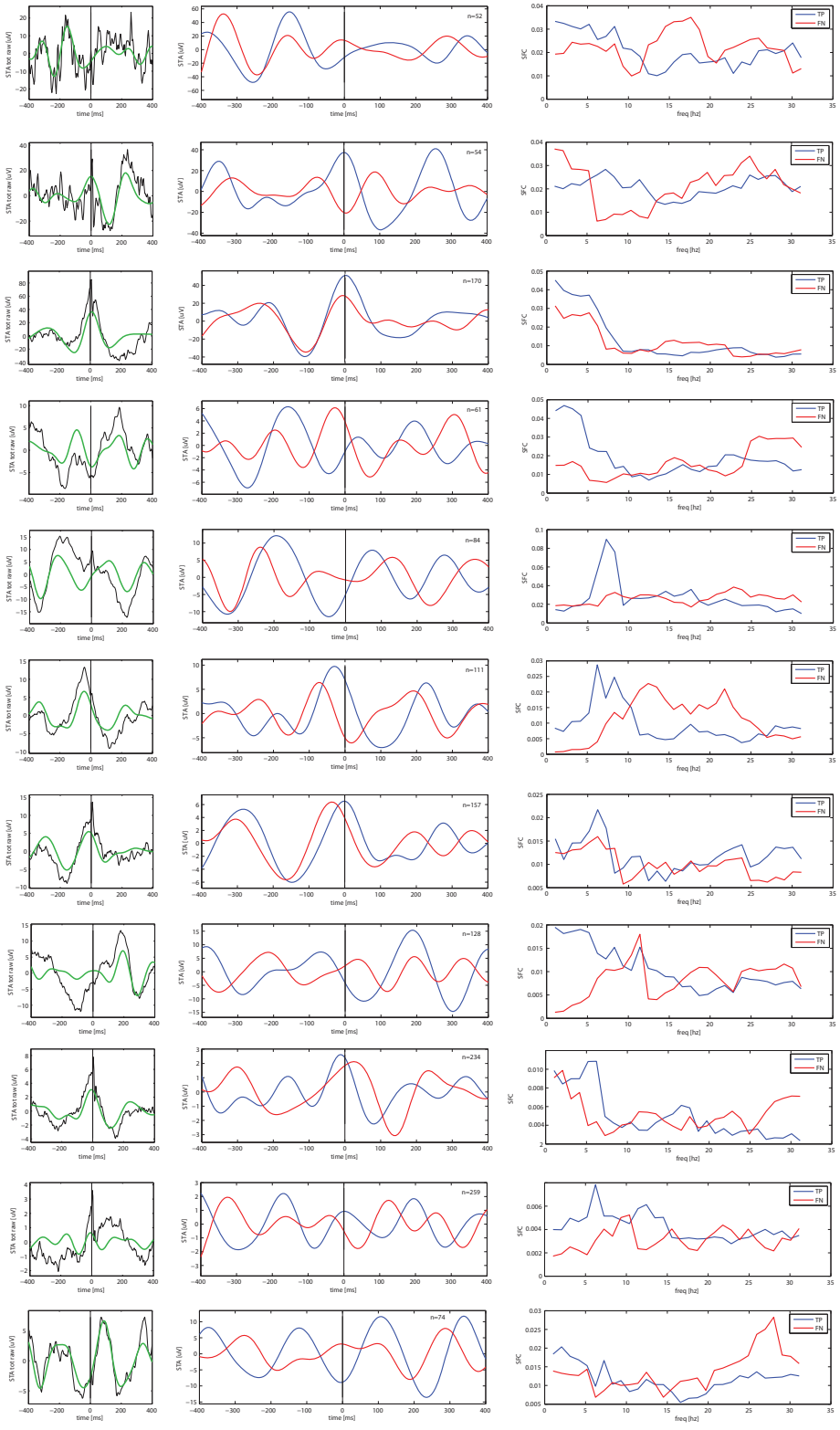

Figure S9: Selection of single-cell examples of spike-triggered averages (STAs). Left column: STA constructed using all spikes, raw (black) and filtered (green, 3-8Hz). Middle column: corresponding STAs (filtered 3-8Hz) for TP and FN. An equal number of spikes (n indicated in each plot) are used for each. Right column: corresponding SFCs. Note that here, the y axis depicts values between 0 and 1. The top 3 traces are from the hippocampus, the remainder from the amygdala. Traces of 6 different patients are shown. Throughout, red represents forgotten trials (FN) and blue remembered trials (TP). Note the different y-axis scale for each unit.

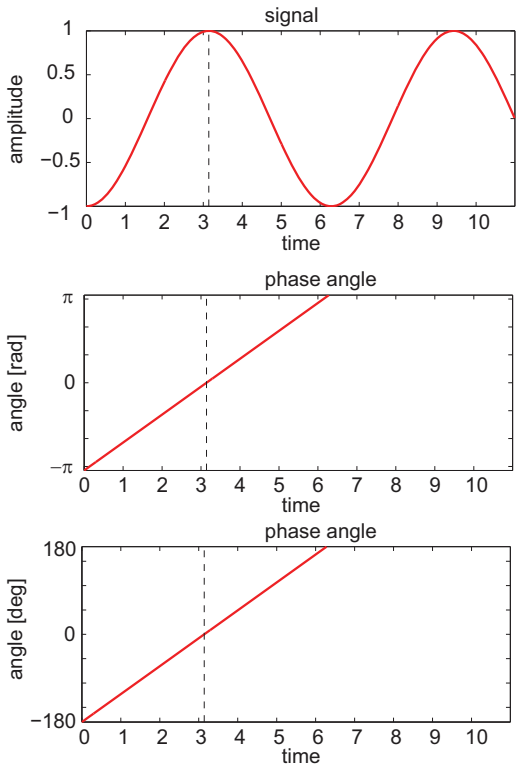

Figure S10: Illustration of the notation used for the phase (angle) of a spike. Shown are 1.5 cycles of an oscillation of cycle (arbitrary units). The corresponding angle is shown in the bottom two rows both in radians (middle) and degrees (bottom). In this notation, the peak of the oscillation corresponds to 0° and the trough to ±180. Note the discontinuity at 180°, when the phase resets.

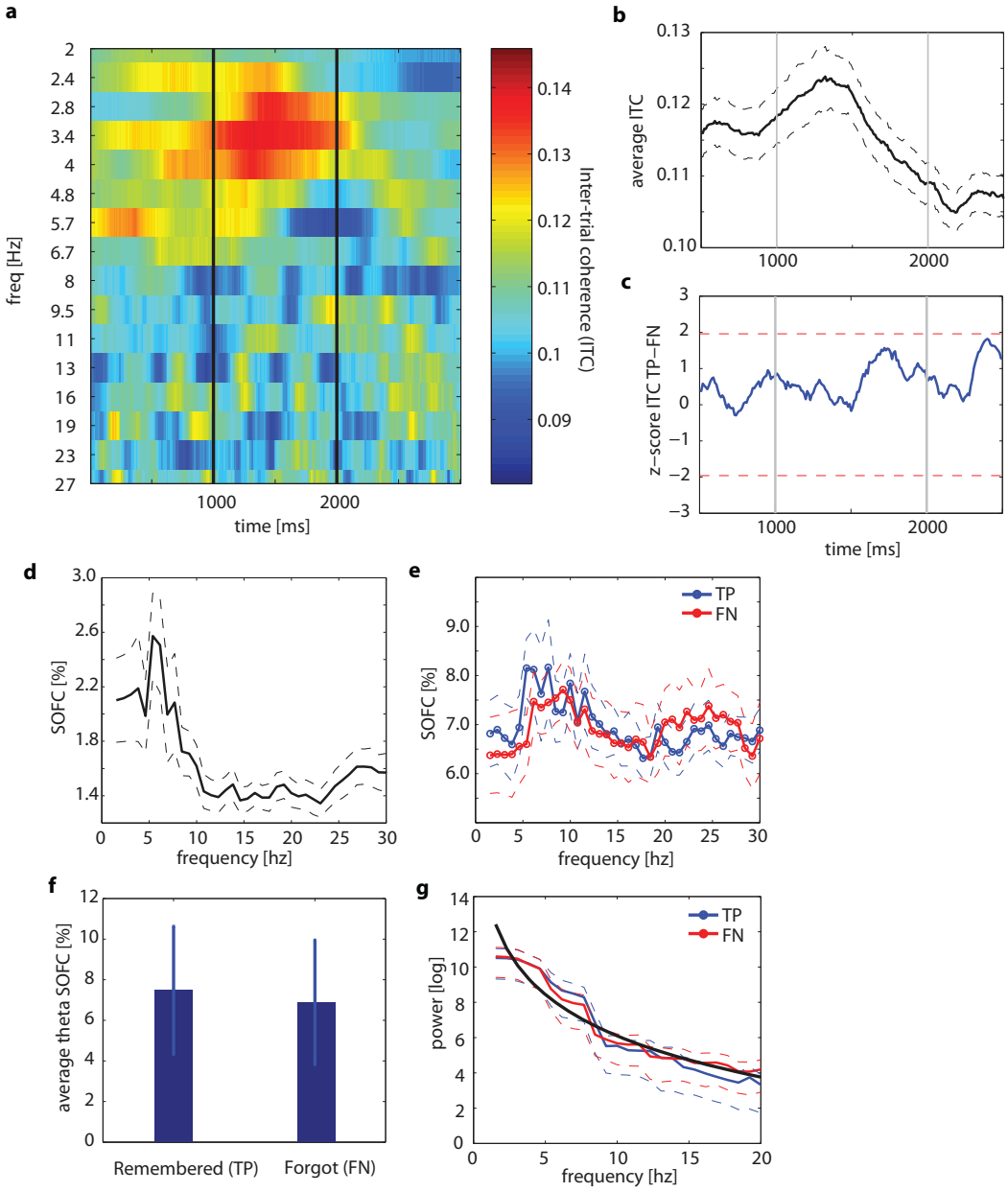

Figure S11: Stimulus-onset triggered phase-resets of ongoing theta oscillations are not different for remembered (TP) and forgotten (FN) trials. Shown are estimates derived using two different techniques: inter-trial coherence (ITC) and stimulus-onset triggered field coherence (SOFC). SOFC is equivalent to SFC, except that the triggering event is the stimulus onset rather than a spike (see methods). **(A-C)** Average ITC for all learning trials and channels (n = 118 unique channels, on which the 305 neurons were detected; channels with units but no LFP were excluded). Stimulus onset/offset are marked by horizontal lines. **(A)** Stimulus onset caused a phase reset of ongoing oscillations (primarily in the 2.4 – 6 Hz range). **(B)** Average ITC in the theta band (3-8 Hz, ±s.e.m.). **(C)** Difference between the average theta ITC (as shown in B) for TP and FN trials (z-scored). Z-scores were calculated based on a bootstrapped null distribution. Red lines indicate the uncorrected p=0.05 threshold. No significant difference was found. **(D-G)** Quantification of phase-resets using the same method as used for assessing SFC. We computed the

SOFC by averaging LFP segments aligned at stimulus onset (-500…800ms relative to stimulus onset, other time-windows revealed similar results) and quantified the power of the resulting average trace using the same procedures used for the SFC (see methods). Power spectra were assessed using multi-tapers with a half-width of 4.6 Hz. This method is more sensitive than the ITC (which we could only reliably estimate for all channels considered together) and thus suitable for smaller number of channels/trials. **(D)** Average SOFC as a function of frequency for all channels on which significantly theta phase-locked cells were detected. If there were no phase-resetting by the stimulus onset, the SOFC would be flat. Phase-resetting is evident in <8Hz, particularly at ~5Hz**.** **(E)** SOFC for TP and FN trials. There was no significant difference (pairwise t-test, FDR-corrected $q<0.05$; uncorrected tests also showed no difference). **(F)** The average SOFC in the theta band (3-8 Hz) was not significantly different ($p=0.33$, paired t-test). **(G)** Power spectrum of the stimulus-onset aligned average LFP segments. There was no significant difference in stimulus-onset triggered average LFP power (paired t-test, uncorrected, $p<0.05$). Notice the presence of excess theta power, indicating phase resets. All errorbars shown are ±s.e. over channels with n as indicated. All pairwise comparisons between TP/FN are corrected for bias by using the same nr of trials for each group. Both ITC as well as the SOFC values depend on the number of trials considered and thus absolute values are not comparable between conditions. Other choices of parameters (window size, subsets of channels) revealed qualitatively similar results.

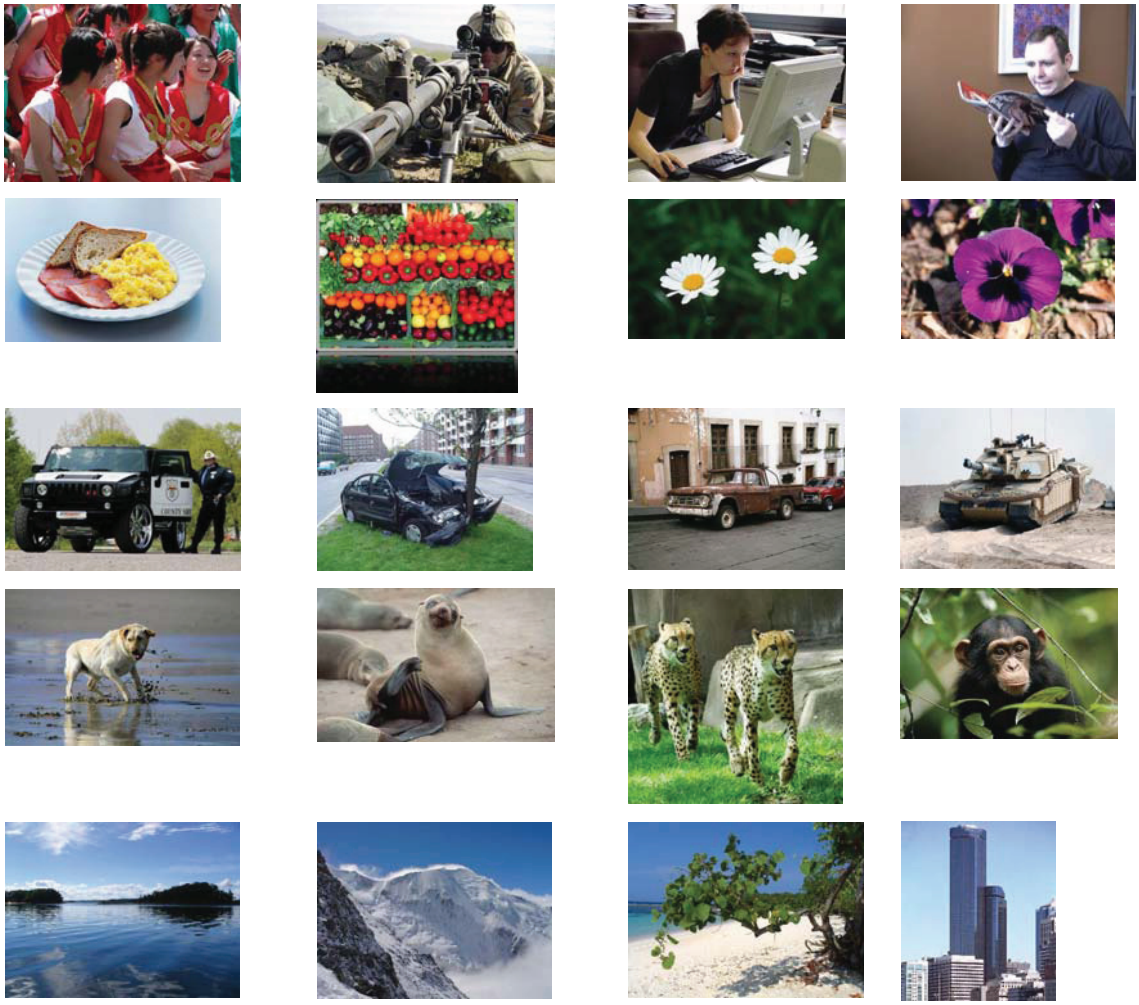

Figure S12: Examples of the stimuli used. Stimuli were chosen from 5 different visual categories (from top to bottom): people, food/flowers, cars/vehicles, animals and outdoor scenes/houses. While we did not explicitly select for emotional content, many pictures did have emotional meaning (i.e. dangerous person, crashed car, cute animal). Images are not shown to scale.

Supplementary Table 1: List of patient demographics, pathology and neuropsychological evaluation. Intelligence was measured with the Wechsler Intelligence Scale (WAIS-III) and memory with the Wechsler Memory Scale Revised (WMS-R). The WMS-R scores are not normalized for age and education. For comparison, the mean and standard deviation of the normal population (from WMS-R) is shown. Abbreviations: Verbal paired associates (VPA), Logical Memory (LM), Visual Reproduction (Vis Rep), performance IQ (PIQ), verbal IQ (VIQ), full scale IQ (FSIQ), perceptual organization index (POI), verbal comprehension index (VCI). Tests indicated with n/a were not performed.

ID	Age	Sex	Hand Dom	Lang Dom	Epilepsy diagnosis	WAIS-III					WMS-R					
						PIQ	VIQ	VCI	POI	FSIQ	VPA 1	VPA 2	LM 1	LM 2	Vis Rep 1	Vis Rep 2
P9	55	M	R	L	Right temporal	97	98	100	103	98	18	7	22	14	32	25
P11	16	M	R	L	Right lateral frontal	84	91	88	84	88	17	8	n/a	n/a	31	29
P14	31	M	L	Bil L>R	Bilateral temporal independent onsets	n/a	n/a	112	111	n/a	11	6	18	11	40	16
P16	34	F	R	n/a	Right frontal	84	68	68	89	74	8	5	16	10	n/a	n/a
P17	19	M	R	L	Left inferior frontal	128	131	122	133	134	24	6	34	37	40	39
P18	40	M	R	L	Right temporal	69	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	25	8
P19	34	M	R	n/a	Left frontal	81	74	76	80	86	n/a	n/a	20	19	35	32
P21	20	M	R	n/a	Not localized	n/a	n/a	93	89	n/a	23	8	34	33	34	32
P23	35	M	R	L	Left temporal	n/a	n/a	74	86	n/a	11	3	13	4	34	32
Average						**90.5**	**92.4**	**91.6**	**96.9**	**96.0**	**16.0**	**6.1**	**22.4**	**18.3**	**33.9**	**26.6**
Normal population (standardization sample, age group 35-44)						*100*	*100*	*100*	*100*	*100*	*20.6*	*7.6*	*26.0*	*21.9*	*32.5*	*29.5*

Supplementary Discussion

Confidence ratings are subjective, i.e. the strength of memory that triggers a "confident" response varies from person to person. We thus adjusted the numerical level of confidence for a trial to be classified as "remembered" individually for each session (Fig 1D) to compare items of equal memory strength as expressed by performance (see methods). Using this dynamic threshold (T=3.9±0.3), the mean recognition rate was 66 ± 3% and 34 ± 3% of trials were categorized as "forgotten".

In our analysis, we used all spikes fired during the learning part of the task to identify whether a neuron is phase-locked to an oscillation in the theta range. We also selected phase-locked neurons using alternative methods (which yielded more neurons in both areas) and compared their SFC and found similar results (Fig S8).

We found that spike-timing accuracy both before as well as after stimulus onset was predictive of memory encoding and strength. While post-learning processing at both the cellular and circuit level are required for later memory retrieval [1,2], our data highlight the importance of neural synchrony immediately preceding and during learning. Our data build on prior studies showing that electrical stimulation of the MTL disrupts memory performance if stimulation occurred within 300ms of the stimulus onset [3], whereas local anesthesia administered 1 min after memory acquisition does not disrupt subsequent memory [4]. To further differentiate between these two possibilities, we compared the theta-SFC (3-7Hz) before and after stimulus onset for each cell. We found that there is no significant correlation between the two (n=33 cells, all trials. Pearson correlation, c=-0.03, p=0.88). This indicates that the two effects were independent and additive, suggesting that both stimulus driven transient activity and state-related activity are important (cf [5]).

Are our findings influenced by the brain pathology (epileptic activity) of the recorded neurons? Several factors lead us to believe that this is unlikely. First, our results concern differences in theta-phase locking between behavioral conditions across the population of neurons recorded. While epilepsy can lead to pathological forms of oscillatory activity, it is unlikely that the timing of spikes relative to these oscillations would differ on a trial-by-trial basis. Also, many of the electrodes were implanted in areas that were later determined to be non-epileptic and half of our patients had epilepsies of origin outside the MTL (see Table 1). Nevertheless, we repeated the analysis excluding all cells from the hemisphere from which tissue was later resected (this was the case for 3 patients). All effects remained significant (n=28 cells of the original 33, average SFC in 3-7Hz band significantly different p=0.00025, paired t-test; compare to Fig 3A).

Our data indicate that the timing of individual spikes during learning trials relative to an ongoing theta frequency oscillation can predict subsequent memory strength. In contrast, neither the firing rate nor the amplitude/phase of the theta oscillations had any apparent predictive value. Increases or decreases of oscillatory power in the intracranial EEG that are predictive of later memory have been reported from a number of brain areas, including the hippocampus [6] . Here, we show that spike-timing accuracy is predictive of encoding success in the absence of such power differences. Between-area synchronization can similarly predict memory success in the absence of any power differences[7]. However, these observations do not rule out the possibility that oscillatory power is different in areas upstream of the MTL neurons we recorded from (thus influencing the timing of propagated spikes).

Supplementary References

1 Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nature reviews* **6**, 119-130 (2005).

2 Squire, L. R., Stark, C. E. & Clark, R. E. The medial temporal lobe. *Annual review of neuroscience* **27**, 279-306 (2004).

3 Ringo, J. L. Brevity of processing in a mnemonic task. *Journal of neurophysiology* **73**, 1712-1715 (1995).

4 Gleissner, U., Helmstaedter, C., Kurthen, M. & Elger, C. E. Evidence of very fast memory consolidation: an intracarotid amytal study. *Neuroreport* **8**, 2893-2896 (1997).

5 Otten, L. J., Quayle, A. H., Akram, S., Ditewig, T. A. & Rugg, M. D. Brain activity before an event predicts later recollection. *Nature neuroscience* **9**, 489-491 (2006).

6 Sederberg, P. B. *et al.* Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex* **17**, 1190-1196 (2007).

7 Fell, J. *et al.* Rhinal-hippocampal theta coherence during declarative memory formation: interaction with gamma synchronization? *The European journal of neuroscience* **17**, 1082-1088 (2003).