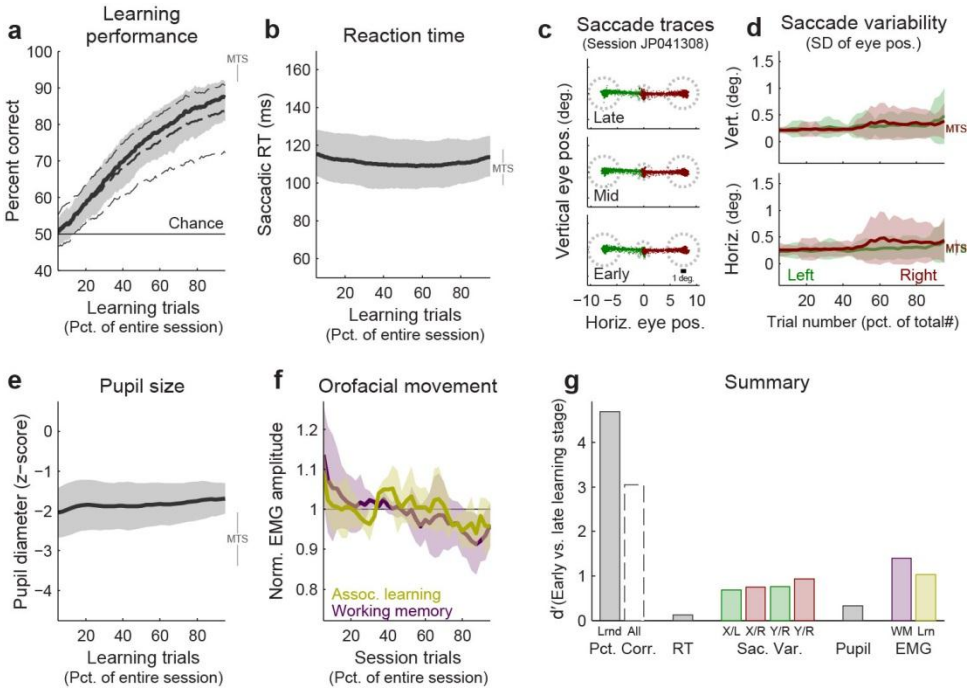
Supplementary Information

Supplementary Figure 1

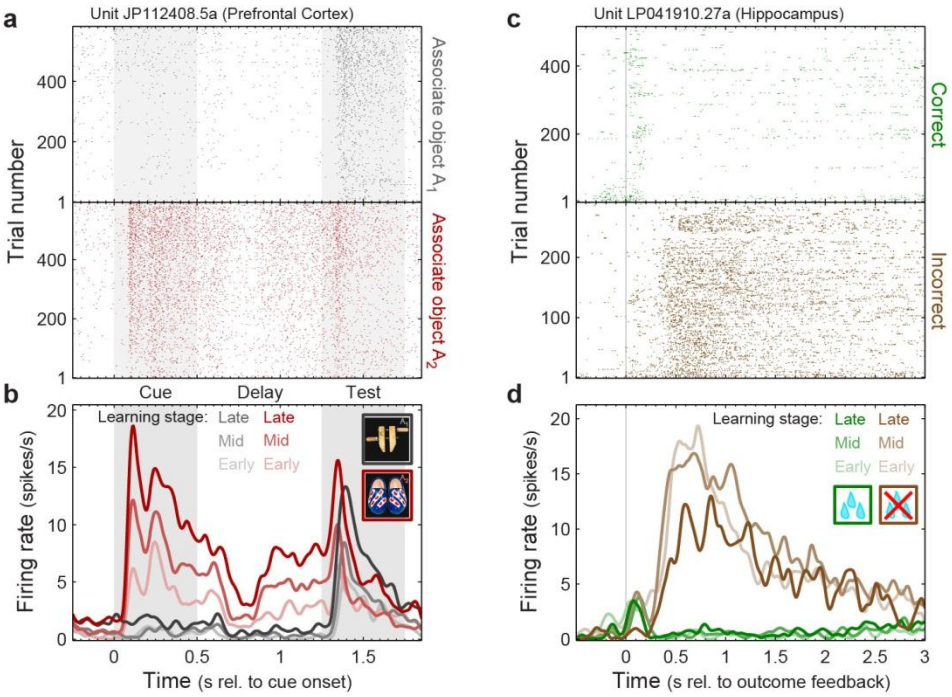


Supplementary Figure 1. Behavioral measures of motor function, motivation, and arousal do not change appreciably with associative learning

In all panels in this figure, behavioral metrics were calculated in identical sliding trial windows (width = 10%, step = 0.5% of each total session length) on the same set of sessions (except for **e**), are plotted in a similar relative scale (~10× average SD across trial windows), and are assayed for changes with learning using the same statistical test (2-sided permutation test on means of early vs. late learning stage [first vs. last third of trials]). (**a**) Learning performance. Across-session (all sessions meeting learning criteria; n = 61) mean ± SD of logit-transformed percent of correct trials, plotted as a function of the percentile of each session's trials. Performance robustly increases across trials ($p \leq 10^{-4}$). This difference is not due to restricting analysis to sessions with successful learning, as it remains significant when all sessions (n = 87) are included ($p \leq 10^{-4}$; dashed curves). Note that this is the same data plotted in main text **Fig. 1c**, but with learning curves pooled (averaged) across all four associations in each session, to match the number of observations for other data in this figure. Also plotted is the across-session mean ± SD ("MTS" to right of main plot) performance for an identity match-to-sample control task (i.e., matching an object to itself, rather than to a learned associate), which was significantly better than for the associative learning task ($p \leq 10^{-4}$). (**b**) Reaction time. Across-session (n = 61) mean ± SD of log-transformed reaction times to response targets. This metric—which may reflect both motor preparatory and motivational factors—does not change with learning (p = 0.49). Note this null result is likely due in part to the enforced delay in our

task between choice object onset and response (cf. **Fig. 1b**), though the fact that match-to-sample task reaction times are significantly faster ($p = 0.008$) indicates that reliable reaction time modulations are possible with this task structure. (**c**) Saccade traces from a typical session. Eye position is plotted for each trial in the early, middle, and late learning stages (top to bottom) to left (green) and right (red) targets, for -130–130 ms relative to saccade onset. Dashed circles: fixation and saccade windows. Scale bar indicates 1 degree of visual angle. (**d**) Saccadic endpoint variability. Across-session ($n = 61$) mean ± SD of the variability of saccade endpoints (across-trial standard deviation of position 30–130 ms post-saccade, when the eyes were typically stable on the target), for vertical (top) and horizontal (bottom) dimensions and saccades to left (green) and right (red) targets. Though saccades do become significantly more variable with learning (all $p < 2\times10^{-4}$), the magnitude of this change is rather small (compare maximum increase of ~0.1 deg. to 1 deg. scale bar in panel **c**; see also panel **g**). (**e**) Pupil size. Across-session ($n = 61$) mean ± SD of pupil diameter during delay period (100–850 ms after start of delay), when pupil size is least influenced by external factors. Within each session, pupil size is expressed as a z-score relative to the fixation period mean and SD. This metric—which is strongly linked to global arousal—does not significantly change with learning ($p = 0.07$), but is significantly decreased for the match-to-sample task ($p \leq 10^{-4}$). (**f**) Lip EMG. Across-session mean ± SD of lip EMG during outcome feedback period (100–1350 ms after outcome feedback), normalized by its mean value for each session. EMG was obtained from two animals performing a working memory–guided saccade task (purple; 4 sessions) or a visuomotor associative learning task (yellow; 6 sessions). Lip EMG—a proxy for reward-related orofacial movements—also shows little change with learning for either the working memory ($p = 0.1$) or learning ($p = 0.43$) tasks. (**g**) Summary of behavioral results. To compare behavioral changes across all reported metrics, relatively independent of the number of observations, we calculated a d′ statistic between the early and late learning stages ($\left|\mathrm{mean}_{\mathrm{early}} - \mathrm{mean}_{\mathrm{late}}\right| / \mathrm{SD}_{\mathrm{pooled}}$). These results reiterate that across-trial changes in motor behavior, motivation, and arousal are relatively minor compared with learning-related changes in performance.

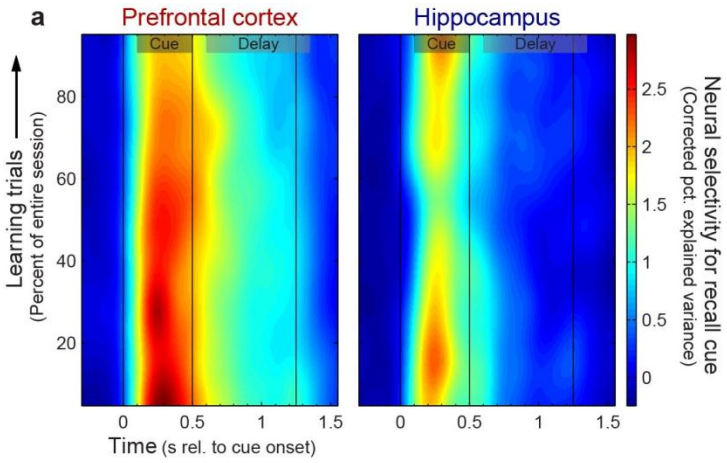
Supplementary Figure 2



Supplementary Figure 2. Population effects are also observed in single neurons

(**a**) Rasters showing spike times for an exemplary PFC neuron on trials in which the monkey is cued to recall associate object A_1 (top panel, gray) or A_2 (bottom, red). Within each panel, learning trials progress from bottom to top. (**b**) Spike density functions (computed with 75 ms Hann window) summarizing the PFC neuron's firing rate when recalling associate A_1 (gray) or A_2 (red) within the early, middle, and late learning stages (light-to-dark colors). Its activity is stronger when associate A_2 is recalled, and this preference increases with learning, reflecting population-level PFC signals for learned associations. (**c**) Rasters showing spike times for an exemplary hippocampal neuron following correct (top panel, green) and incorrect (bottom, brown) trials. Within each panel, learning trials progress from bottom to top. (**d**) Spike density functions showing the HPC neuron's firing rate for correct (green) and incorrect (brown) trials within the early, middle, and late learning stages (light-to-dark colors). Its activity is stronger following incorrect trials, but this preference diminishes with learning, reflecting the robust population-level HPC outcome signals and their shift from error-preferring toward correct-preferring bias with learning.

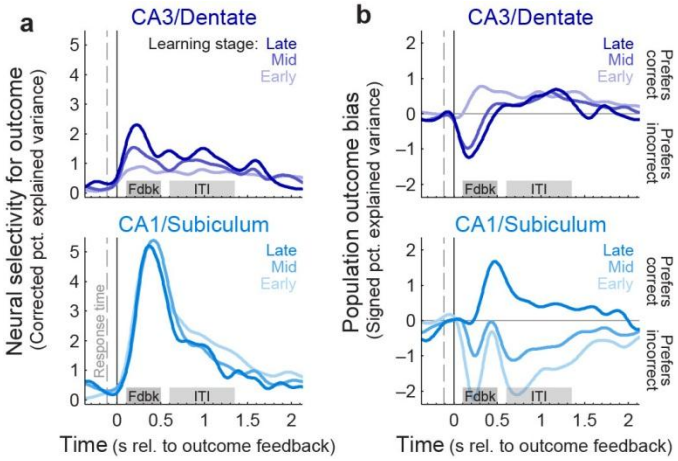
Supplementary Figure 3



Supplementary Figure 3. Hippocampus and PFC carry neural information about the retrieval cue

(**a**) Mean percent of variance in PFC (left) and HPC (right) spiking activity explained by retrieval cue, plotted across time after cue onset and learning trials. Activity reflecting the cues is present in both areas, in contrast to activity reflecting the learned associates (**Fig. 3** in main text), which is only found in PFC. This indicates the lack of associate signals in HPC is not due to a lack of selective visual responses to the stimuli used.

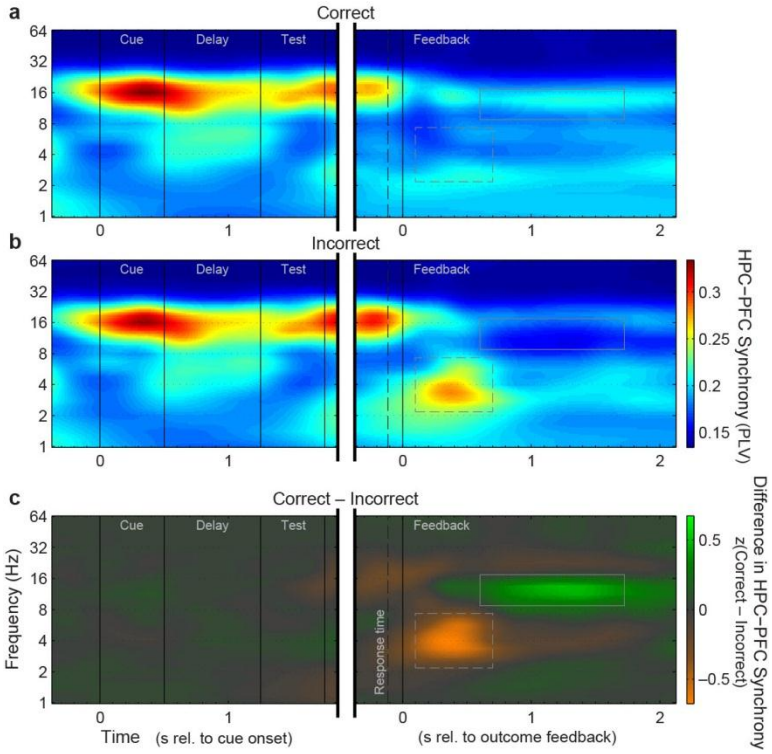
Supplementary Figure 4



Supplementary Figure 4. Trial outcome signals in hippocampal subregions

(**a**) Mean selectivity (PEV) for trial outcome in hippocampal local-projection subregions (top; Dentate gyrus/CA3, n = 104 neurons) and output subregions (bottom; CA1/Subiculum, n = 93) neurons, plotted across learning stages (light-to-dark colors). Outcome signals are significantly stronger overall in the HPC output subregions ($p \leq 10^{-4}$; 2-way subregion × learning-stage ANOVA in outcome feedback epoch). (**b**) Mean bias (signed PEV) in HPC local-projection (top) and output subregion (bottom) neurons for correct (positive values) vs. incorrect (negative values) outcomes. With learning, there was a significant shift from stronger signals for incorrect to correct trials in HPC output subregions (p = 0.005; 2-sided permutation test on ITI epoch signals in early vs. late learning stages), but not local-projection subregions (p = 0.83; interaction in 2-way subregion × learning-stage ANOVA: p = 0.04).

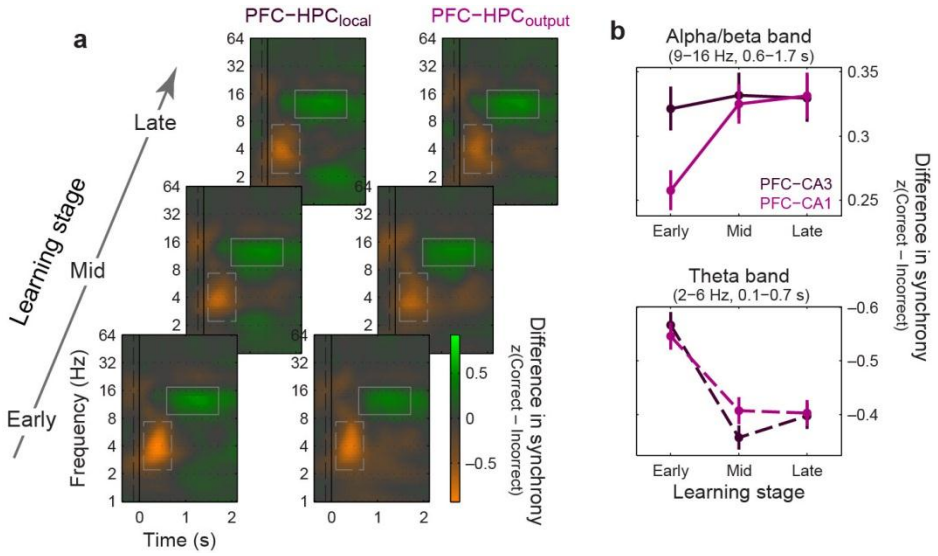
Supplementary Figure 5



Supplementary Figure 5. HPC-PFC synchrony across all trial periods

(**a**) Mean synchrony (PLV) between HPC and PFC LFPs on correct trials, plotted as a spectrogram across time and frequency. Separate plots show time periods during the trial (left; time referenced to retrieval cue onset), and after outcome feedback is given (right; referenced to outcome feedback onset). (**b**) Mean HPC-PFC synchrony on incorrect trials (same conventions and color scale as **a**). (**c**) Mean z-scored difference in HPC-PFC synchrony between correct and incorrect trials, for same time periods as above (right: replot of main text **Fig. 4b**). Though there are clear periods of band-specific synchrony during trial performance (panels **a** and **b**), they are nearly identical for correct and incorrect trials, and thus convey little information about trial outcome.

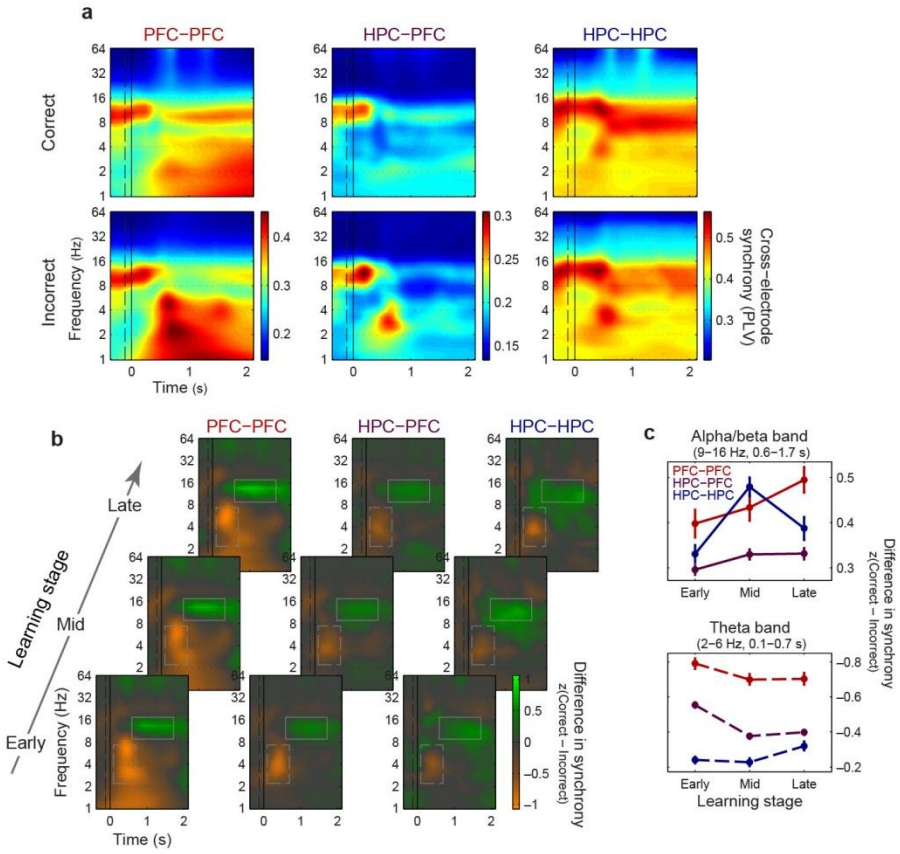
Supplementary Figure 6



Supplementary Figure 6. Learning-related synchrony in hippocampal subregions

(**a**) Mean z-scored difference in synchrony (dPLV) between correct and incorrect trials, calculated between electrodes in PFC and distinct hippocampal subregions (left: PFC & HPC local-projection subregions [dentate gyrus and CA3], n = 558 electrode pairs; right: PFC & HPC output subregions [CA1 and subiculum], n = 407), plotted across learning stages. (**b**) Summary of synchrony learning effects—mean (± SEM) dPLV pooled within the alpha/beta-band (top) and theta-band (bottom) regions of interest, as a function of learning stage. While the theta-band decrease with learning is similar for synchrony between PFC and all HPC subregions ($p \leq 10^{-4}$ for both, 2-sided permutation test on early vs. late learning), the alpha/beta-band increase with learning is only present for synchrony between PFC and HPC output subregions (CA1/Sub.; $p \leq 10^{-4}$), but not for synchrony between PFC and HPC local-projection subregions (dentate/CA3; $p = 0.52$) despite their greater numbers of observations. Synchrony between hippocampal subregions (not shown) is nearly identical to synchrony averaged across all pairs of hippocampal electrodes (**Supplementary Fig. 7**); small numbers of observations precluded meaningful analysis of synchrony between sites within each subregion.

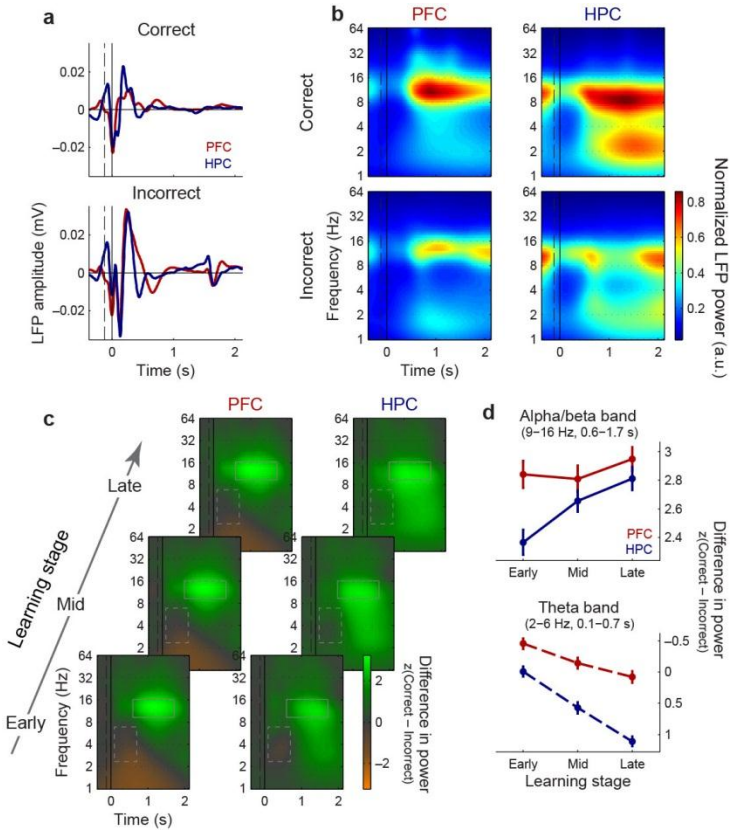
Supplementary Figure 7



Supplementary Figure 7. Learning-related information about trial outcome in oscillatory synchrony between all area pairs

(**a**) Mean synchrony (PLV) spectrograms between pairs of electrodes in PFC (left; n = 648), in HPC (right; n = 694), and between HPC and PFC (center; n = 970), following correct (top) and incorrect (bottom) trials. (**b**) Mean z-scored difference in synchrony (dPLV) between correct and incorrect trials, plotted across learning stages, for all pairs of studied areas (HPC-PFC data replotted from main text **Fig. 4c**). (**c**) Summary of synchrony learning effects—mean (± SEM) synchrony difference pooled within the alpha/beta-band (top) and theta-band (bottom) regions of interest, as a function of learning stage. Synchrony between distinct sites within PFC (red) follows a similar pattern to the cross-area synchrony (purple)—theta decreases ($p = 3\times10^{-4}$), while alpha/beta increases with learning ($p \leq 10^{-4}$, 2-sided permutation test on early vs. late learning). In contrast, intra-hippocampal synchrony increases with learning for both the theta ($p = 2\times10^{-4}$) and alpha/beta bands ($p \leq 10^{-4}$), indicating the observed learning effects do not reflect global state changes that are invariant across all brain areas.

Supplementary Figure 8

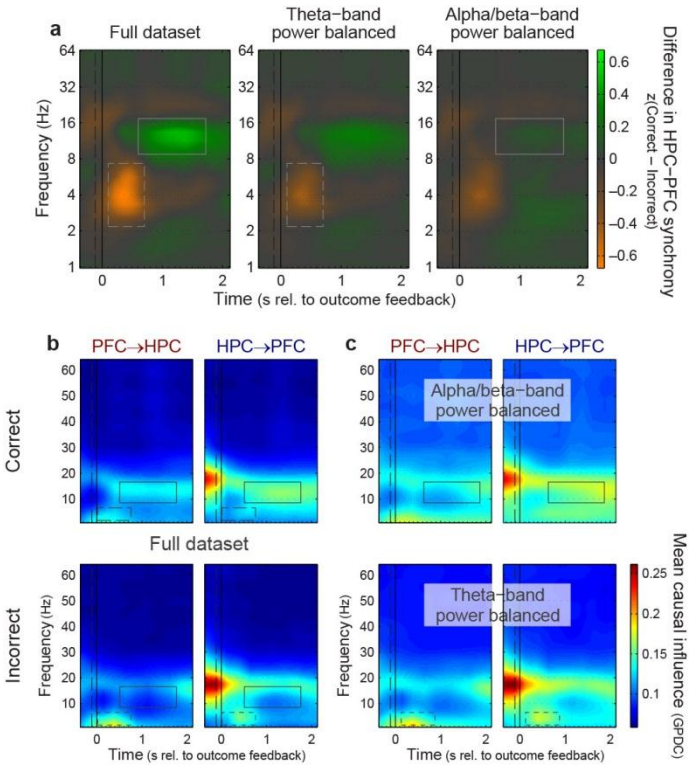


Supplementary Figure 8. Learning-related information about trial outcome in oscillatory power in prefrontal cortex and hippocampus

(**a**) Mean raw LFP signals (evoked potentials) in PFC (red; n = 250 electrodes) and HPC (blue; n = 166), following correct (top) and incorrect (bottom) trials. LFPs are time-locked to the onset of trial-outcome feedback (solid line; dashed line: behavioral response time). (**b**) Mean normalized LFP power spectrograms in PFC (left) and HPC (right), following correct (top) and incorrect (bottom) trials. To highlight components not time-locked to trial events (i.e., induced, rather than evoked, signals), mean raw LFPs were subtracted off each electrode prior to calculation of spectral power. To enhance visualization of band-specific signals relative to the well-known 1/frequency distribution of LFP power, power at each frequency was normalized by 1/frequency for display purposes only. (**c**) Mean z-scored difference in induced power between correct and incorrect trials across learning stages, for PFC (left) and HPC (right). While there is a strong alpha/beta-band signal for correct trials, the theta-band signal for incorrect trials observed in the cross-electrode synchrony results is not as robust in local power. (**d**) Summary of power learning effects—mean (± SEM) induced power difference pooled within the alpha/beta-band (top) and theta-band (bottom) regions of interest, as a function of learning stage. Theta power exhibits a significant positive shift (from incorrect toward correct bias) with learning ($p \leq 10^{-4}$ for both areas), and alpha/beta power also shows a positive trend (significant

only for HPC: $p \leq 10^{-4}$; PFC: $p = 0.06$; 2-sided permutation test on early vs. late learning). These results indicate a similar change with learning for both cross-area synchrony and within-area power.

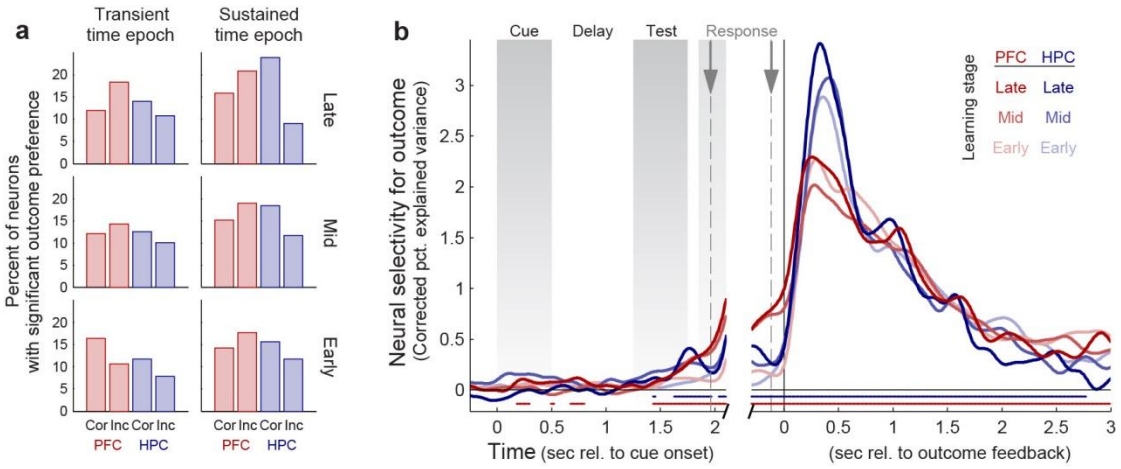
Supplementary Figure 9



Supplementary Figure 9. Balancing band-limited power does not eliminate observed HPC-PFC synchrony and causality effects

(**a**) Mean z-scored difference in HPC-PFC synchrony (dPLV) between correct and incorrect trials for the full dataset (left), and for data subsets where power pooled within the theta-band (middle) or alpha/beta-band (right) regions of interest was balanced across trial outcomes. Though power balancing reduced the effect of trial outcome on neural synchrony, both frequency bands remained significantly different from zero ($p \leq 10^{-4}$ for both; 1-sample bootstrap test). This confirms there is a specific effect of outcome on HPC-PFC synchrony, beyond any possible artifactual effects due to differences in power. (**b**) Frequency-domain directional influences (GPDC) from PFC to HPC (left) and from HPC to PFC (right), following correct (top) and incorrect (bottom) trials for full dataset. (**c**) GPDC for data subsets where power pooled within the theta-band (bottom) or alpha/beta-band (top) regions of interest was balanced across trial outcomes. For the power-balanced controls, alpha/beta-band influences remain stronger from HPC to PFC, and theta-band influences remain stronger from PFC to HPC ($p \leq 10^{-4}$ for both; direction factor in 2-way causal direction $\times$ trial outcome permutation ANOVA). This confirms that the observed directionality effects are not due to any differences in local power within these areas.

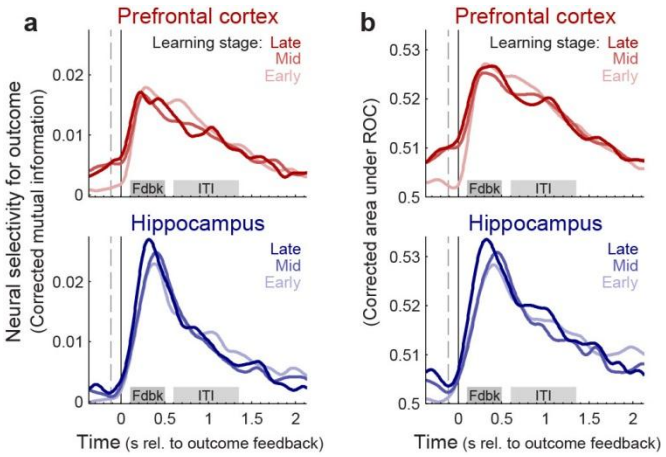
Supplementary Figure 10



Supplementary Figure 10. PFC and HPC neuronal trial outcome signals are distinct from subcortical reward prediction error signals

(**a**) Roughly equal numbers of neurons show stronger activity for correct and incorrect outcomes. Plots show the percent of neurons in PFC (red; n = 319) and HPC (blue; n = 199) with a significant preference for correct (Cor) or incorrect (Inc) outcomes ($p < 0.05$, 2-sided permutation test). Values are plotted separately for early, mid, and late learning stages (bottom to top), and time epochs capturing transient (left; 100–500 ms) and sustained (right; 600–1350 ms) response components. These results are in contrast to previous results from the ventral tegmental area and lateral habenula, which show strong biases toward positive and negative reward prediction errors (roughly, uncertain correct and incorrect outcomes), respectively. (**b**) No robust transfer of outcome signals to earlier trial events. Population mean percent variance explained by trial outcome (correct vs. incorrect) is plotted as a function of time during (left) and after (right) the trial, separately for PFC and HPC and learning stages (see legend). Tick marks at bottom: time points with significant explained variance during the late learning stage ($p <$ 0.05, uncorrected, 1-sample bootstrap test). In the ventral tegmental area and lateral habenula, as reward becomes more predictable during learning, activation shifts from the post-response outcome feedback epoch to earlier trial events predictive of reward. In contrast, post-response trial outcome information in PFC and HPC is present throughout learning, with little shift to earlier time points. These properties, and the learning-related shift from bias toward encoding incorrect to correct outcomes in HPC (main text **Fig. 3c**), distinguish the outcome signals we report from the static reward prediction error signals found in areas such as the ventral tegmental area and lateral habenula.

Supplementary Figure 11



Supplementary Figure 11. Results are similar for neural selectivity measured using mutual information, area under ROC curve, and percent explained variance

(**a**) Mean mutual information between spike counts and trial outcome (correct vs. incorrect) in PFC (top; *n* = 319) and HPC (bottom; *n* = 199) neurons, bias-corrected by subtracting trial-shuffled information. (**b**) Mean area under receiver operating characteristic (ROC) curve for discrimination of spike counts between correct and incorrect trial outcomes in PFC (top) and HPC (bottom) neurons, rectified around 0.5 and bias-corrected by subtracting trial-shuffled area-under-ROC values. Format for both plots is the same as for plots of neural percent explained variance in main text **Fig. 3b**. All three metrics show highly similar results—consistently showing greater selectivity in HPC than PFC—as we have observed for numerous neural activity contrasts.