

# Multimodal Encoding of Novelty, Reward, and Learning in the Primate Nucleus Basalis of Meynert

Clarissa Martinez-Rubio,<sup>1\*</sup>  Angelique C. Paulk,<sup>1\*</sup>  Eric J. McDonald,<sup>1</sup>  Alik S. Widge,<sup>2,3,†</sup> and  Emad N. Eskandar<sup>1,4,†</sup>

<sup>1</sup>Nayef Al-Rodhan Laboratories, Department of Neurosurgery, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114,

<sup>2</sup>Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts 02129, <sup>3</sup>Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02124, and <sup>4</sup>Department of Neurological Surgery, Montefiore Medical Center, Albert Einstein College of Medicine of Yeshiva University, 3316 Rochambeau Avenue, Bronx, NY, 10467

Associative learning is crucial for daily function, involving a complex network of brain regions. One region, the nucleus basalis of Meynert (NBM), is a highly interconnected, largely cholinergic structure implicated in multiple aspects of learning. We show that single neurons in the NBM of nonhuman primates (NHPs;  $n = 2$  males; *Macaca mulatta*) encode learning a new association through spike rate modulation. However, the power of low-frequency local field potential (LFP) oscillations decreases in response to novel, not-yet-learned stimuli but then increase as learning progresses. Both NBM and the dorsolateral prefrontal cortex encode confidence in novel associations by increasing low- and high-frequency LFP power in anticipation of expected rewards. Finally, NBM high-frequency power dynamics are anticorrelated with spike rate modulations. Therefore, novelty, learning, and reward anticipation are separately encoded through differentiable NBM signals. By signaling both the need to learn and confidence in newly acquired associations, NBM may play a key role in coordinating cortical activity throughout the learning process.

**Key words:** encoding cognitive state; learning; local field potential; nucleus basalis of Meynert; single neuron activity

## Significance Statement

Degradation of cells in a key brain region, the nucleus basalis of Meynert (NBM), correlates with Alzheimer's disease and Parkinson's disease progression. To better understand the role of this brain structure in learning and memory, we examined neural activity in the NBM in behaving nonhuman primates while they performed a learning and memory task. We found that single neurons in NBM encoded both salience and an early learning, or cognitive state, whereas populations of neurons in the NBM and prefrontal cortex encode learned state and reward anticipation. The NBM may thus encode multiple stages of learning. These multimodal signals might be leveraged in future studies to develop neural stimulation to facilitate different stages of learning and memory.

## Introduction

In any complex environment, animals must both learn new associations and continue to recall and consider previously formed associations between stimuli, behavior, and outcomes to func-

tion adaptively (Poldrack and Packard, 2003). Both recall and learning depend on the coordinated action of multiple neuronal systems. Converging lines of evidence suggest that brain regions such as the basal forebrain and prefrontal cortex drive these processes across species. A group of neurons in the basal forebrain, the nucleus basalis of Meynert (NBM), is the primary source of cholinergic innervation to the cortex in primates and humans (Mesulam et al., 1983; Mesulam and Mufson, 1984; Struble et al., 1986; Liu et al., 2015). This could mean that the NBM acts as a possible modulator of cortical function, such as in learning (Bakin and Weinberger, 1996; Miasnikov et al., 2008, 2009). Fur-

Received July 17, 2017; revised Nov. 10, 2017; accepted Nov. 27, 2017.

Author contributions: C.M.-R. and E.N.E. designed research; C.M.-R. performed research; A.C.P., E.J.M., and A.S.W. contributed unpublished reagents/analytic tools; A.C.P. analyzed data; C.M.-R., A.C.P., E.J.M., A.S.W., and E.N.E. wrote the paper.

This work was supported by the National Institutes of Health (Grant EY017658 to C.M.-R. and postdoctoral fellowship NS083208 and grants MH086400, DA026297, and EY017658 to E.N.E.), the Epilepsy Foundation (E.N.E.). Analysis of the data was supported in part by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement Number W911NF-14-2-0045 issued by ARO contracting office in support of DARPA's SUBNETS Program. The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

The authors declare no competing financial interests.

\*C.M.-R. and A.C.P. contributed equally to this work.

†A.S.W. and E.N.E. are cosenior authors.

Correspondence should be addressed to Dr. Emad N. Eskandar, Jeffrey P. Bergstein Professor and Chairman, Department of Neurological Surgery, Montefiore Medical Center, Albert Einstein College of Medicine of Yeshiva University, 3316 Rochambeau Avenue, Bronx, NY, 10467. E-mail: eeskandar@mgm.harvard.edu.

DOI:10.1523/JNEUROSCI.2021-17.2017

Copyright © 2018 the authors 0270-6474/18/381942-17\$15.00/0

thermore, NBM neurons have been shown to encode salience, attention, and novelty (Richardson and DeLong, 1986, 1990; Voytko et al., 1994; Voytko, 1996; Masuda et al., 1997; Weinberger, 2003).

In addition, structural dysfunction in the NBM has been correlated with a variety of mental health and neurological disorders (Mesulam, 2013; Grothe et al., 2014; Kilimann et al., 2014; Gratwicke et al., 2015; Liu et al., 2015). Degeneration of the NBM occurs in patients with dementia and Alzheimer's disease, correlating with impaired learning and memory (Mesulam, 2013). Despite its centrality, network connections to multiple regions, and seeming importance in neurodegenerative diseases, the NBM's role in learning and memory is poorly understood. Debate on the role of the NBM in learning and memory is largely due to the highly variable responses observed in the NBM during a wide variety of tasks (Richardson and DeLong, 1986, 1990; Wilson and Rolls, 1990; Voytko et al., 1994; Masuda et al., 1997; Weinberger, 2003). Previous studies examining the role of the NBM in learning and memory using lesions, stimulation, behavioral paradigms, and/or physiological recordings in humans and animal models have not been able to elucidate a unitary role (Wenk, 1997; Ridley et al., 1999; Barefoot et al., 2002; Gibbs and Johnson, 2007; Miasnikov et al., 2008, 2009; Rabiei et al., 2014).

We compared the relative roles of the NBM and dorsolateral prefrontal cortex (dlPFC) in learning by recording neural activity in nonhuman primates (NHPs) performing an associative learning task. We chose the dlPFC because this region is strongly implicated in integrating cognitive functions such as learning, attention, error prediction, and decision making (Wallis and Miller, 2003; Tsujimoto and Sawaguchi, 2005; Ichihara-Takeda and Funahashi, 2008; Asaad and Eskandar, 2011; Kahnt et al., 2011). In addition, neurons in the NBM have been shown to project to the macaque dlPFC, specifically the principal sulcus (Mesulam et al., 1983). Interestingly, there are no observed projections from the dlPFC back to the NBM (Mesulam and Mufson, 1984). Because the dlPFC has been shown to play a significant role in learning and memory (Asaad and Eskandar, 2011), we hypothesized that the NBM could play a role as either a precursor or modulatory structure relative to the learning-related activity seen in the dlPFC.

We found differentiable response profiles and physiological properties, confirming a multifunctional role of the NBM. Single NBM neurons responded primarily to a combination of novelty and early learning through spike rate modulation not seen in dlPFC neurons. Conversely, in the NBM and, to a lesser extent, in the dlPFC, low-frequency population activity encoded learned states, as represented by theta band (4–8 Hz) power in the local field potential (LFP). In contrast, high-frequency LFP power (65–200 Hz) encoded reward anticipation in both the NBM and dlPFC. Strikingly, the spike rates in the NBM were anticorrelated with the simultaneously recorded LFP activity relative to learning state, contrary to spike–LFP relationships demonstrated in many previously reported brain regions (Ojemann et al., 2013; Yazdan-Shahmorad et al., 2013). These results suggest that NBM encodes multiple aspects of the learning process and potentially signals these features to cortex via different types of network activity.

## Materials and Methods

### Animals

Two adult male NHPs (*Macaca mulatta*), “R” (12 kg, 10 years old) and “P” (12 kg, 14 years old) were provided with a balanced diet supplemented with fruits and treats. Subjects were housed in a climate-controlled environment with a 12 h/12 h light/dark cycle and veterinarian-supervised behav-

ioral and social enrichment. Fluid was restricted (70 ml/kg) such that the subjects received the majority of their daily fluid during task performance. All animal care and experimentation was overseen and approved by the Institutional Animal and Care Use Committee at the Massachusetts General Hospital.

### Electrophysiology

A titanium head post and recording chamber (Crist Instruments) were surgically implanted in accordance with applicable Department of Agriculture guidelines. A magnetic resonance image (MRI) scan was acquired and used to plan the recording chamber placement coordinates. The chamber was stereotactically mounted and positioned to provide optimal access to the structures of interest. A second postoperative T1-MRI was performed with fiducial markers to enable mapping of electrode trajectories and to estimate the distance to reach each target brain region. A custom microdrive (Patel et al., 2014) with a 1 mm spaced grid was used to acutely lower two FHC tungsten microelectrodes (600–800 k $\Omega$ ) daily, one to each structure, dlPFC and NBM. A cannula guide tube facilitated access through piercing of the dura mater. Cannula lengths were estimated to penetrate 2 mm into the cortex to minimize damage to the brain.

Using custom MATLAB programs (The MathWorks, RRID:SCR\_001622), a 3D image of each animal's brain was reconstructed, which allowed for the anatomical visualization of the electrode trajectory (Bakker et al., 2015). This trajectory was further confirmed by physiologically mapping the electrode path (Williams et al., 2005). The mapping of the NBM trajectory was performed by sampling the neuronal activity in the cortex, caudate, and putamen, providing control data to compare neuronal firing and to verify cessation of neuronal activity upon traversal of the internal capsule and anterior commissure. For subcortical structures, neuronal activity began at ~18 mm above target, depending on cannula length. A physiological trajectory map was not necessary for the dlPFC. Due to the anatomical positioning of this structure, recordings were performed immediately upon exiting the cannula into cortex, within the confirmed coordinate calculations and depth.

After recordings were complete, a second 3D reconstruction was made to include each anatomically mapped recording site relative to the target brain structures. This process was done by reconstructing, on a slice-by-slice basis, the ventral pallidum (VP), NBM, and dlPFC (area along the principal sulcus) in P's and R's MRIs. Reconstruction was performed relative to three atlases showing the overlap of histology with MRI structures in model macaque brains, as in the Scalable Brain Atlas (Bakker et al., 2015, RRID:SCR\_006934). The three atlases were the Calabrese atlas (Calabrese et al., 2015), the Paxinos atlas (Paxinos et al., 2000), and the Neuromaps Macaque atlas (Dubach and Bowden, 2009; Rohlfing et al., 2012). We classified the recording as VP or NBM based on tip of the electrode overlapped the 3D-reconstructed VP and NBM regions in the brain. Daily coordinates, cannula length, and recording depths were used to identify the individual recording sites for each structure. Only neurons and LFP recordings within the anatomic boundaries of the NBM were physiologically screened for NBM-like characteristics (see below).

Extracellular recordings were digitized at 40 kHz using an OmniPlex system (Plexon, RRID:SCR\_014803) and stored for subsequent spike activity (filtered to 300–5000 Hz) and LFP (filtered to 0.5–500 Hz, down-sampled to 1000 Hz) analyses. Neurons encountered at the calculated target depth range were recorded regardless of characteristic firing rate. We did not move electrodes between blocks in a single session. Experimental blocks were either Novel/Familiar/Recall or Novel/Familiar/Reversal (see below regarding block design), resulting in differing numbers of units per experiment type. Using Offline Sorter (Plexon, RRID:SCR\_000012), data were thresholded offline to identify possible action potentials. Spikes were then sorted manually and clustered in feature space using peak, valley, energy, and both first and second principal components. We identified a total of 322 neurons, with 112 NBM units and 210 dlPFC units.

### Behavior

Two adult rhesus monkeys were trained to perform a visual–motor association task. The subjects learned, by trial and error, to associate spe-

cific novel visual images with a unique saccade direction to one of the four target locations. Eye position was monitored with an infrared video eye-tracking system (ISCAN) that provides eye coordinates to the behavioral control software (MonkeyLogic; Asaad and Eskandar, 2008). Each trial began with a central fixation point (1250–1500 ms, with 250 ms randomized jitter). If fixation was held, a stimulus image was presented. Animals were required to keep holding fixation (1000–1250 ms, with 250 ms randomized jitter) until the stimulus image was cleared and four target objects appeared, allowing the animal to make a choice (1000 ms). Once the animal indicated a choice, the target changed color to either green or red, indicating a correct or incorrect choice (1000 ms), respectively. For every correct choice, the subject received liquid reward. Each block terminated once 18 correct trials had been performed for each of the four images (Williams and Eskandar, 2006a). If an animal broke fixation or failed to meet task criteria, the trial was aborted with no reward. Animals generally completed two or three sessions daily consisting of three blocks each. Each block was modular, in that the experiments were comprised of three blocks: Novel/Familiar/Recall or Novel/Familiar/Reversal. No cues were presented to the animal to signify block change other than changing the four images in use for that block.

During Novel blocks, animals were expected to learn, by trial and error, to make the correct associations of four novel images with their correct target locations. For each session, four new images were randomly chosen from a pool of 1500 pictures to be used in the Novel block. By using new images each day, we could assess both novelty of the stimuli and learning as the block progressed, on a trial-by-trial basis. For Familiar blocks, the animals were required to recall associations between four images that were repeated over the entirety of training and data collection. Animals were well trained in the correct target associations and usually completed the block quickly. Little learning occurred in this block. In addition, this block serves as a comparison with the Novel block in that the animal performed long-term, well known associations compared with learned but newly made associations at the end of the Novel block. The Familiar block also served as a break between Novel and Recall blocks or Novel and Reversal blocks.

During Recall blocks, animals were re tested on the associations learned in that session's Novel block. Little new learning occurs during Recall blocks, but this block allows us to compare the recollection of newly learned associations and contrasts with the long-term memory assessed in the Familiar block. Conversely, in the Reversal blocks, animals were presented with the same four images used during the Novel block, but the associated targets were changed, requiring animals to learn new associations. By reversing associations during this block, we could assess relearning of now-recognizable stimuli and dissociate between learning and novelty because these stimuli were already presented during the Novel block. Overall, the modular design allowed us to examine NBM and dlPFC activity during: (1) learning through operant conditioning, (2) reinforcement through reward, (3) decision making during learning, and (4) the differentiation of newly acquired associations versus well learned associations. The average number of trials per block type across both NHPs were  $140.3 \pm 51.08$  trials (Novel),  $95.4 \pm 15.98$  trials (Familiar),  $142.4 \pm 47.50$  trials (Reversal), and  $103.7 \pm 27.56$  trials (Recall). Familiar blocks contained very few errors, but animals still broke fixation or did not respond on ~25% of trials on any given day.

We quantified learning on a trial-to-trial basis for each image and separately for each block and session. For each image in a block, we estimated a "learning state," also termed "cognitive state," as the latent variable of a linear time-invariant state-space process that gave rise to the observed reaction times and correct/incorrect choices (Smith et al., 2004, 2007, Prerau et al., 2008, 2009), using the expectation-maximization algorithm. We considered the learning of each image as an independent realization of the learning process, yielding four "learning curves" and their confidence bounds per block. These state estimates were standardized to the [0,1] interval for comparison across macaques, blocks, and sessions. For illustrative comparison of the state-space approach to other learning metrics, we also calculated the ratio correct in a sliding five-trial window and then smoothed this with a 10-trial-wide noncausal Gaussian kernel.

### Analysis: single-unit activity

After individual units were identified through spike sorting, we examined the spike waveforms to identify whether they likely originated in the NBM. To be included in the analysis as NBM units, we used predetermined criteria from the literature in that 2 of the following 3 criteria had to be met: (1) a spontaneous firing rate of 5–40 Hz, (2) a coefficient of variation of the interspike interval  $<1$ , and (3) a spike duration  $>180 \mu\text{s}$  (initial negative phase, 200–10k Hz filtering) (DeLong, 1971; Richardson and DeLong, 1986, 1990, 1991). As mentioned above, additionally, 3D mapping had to localize the electrode tip to the NBM region. We did not perform the same check for dlPFC units in that recordings were performed at the exit of the cannula.

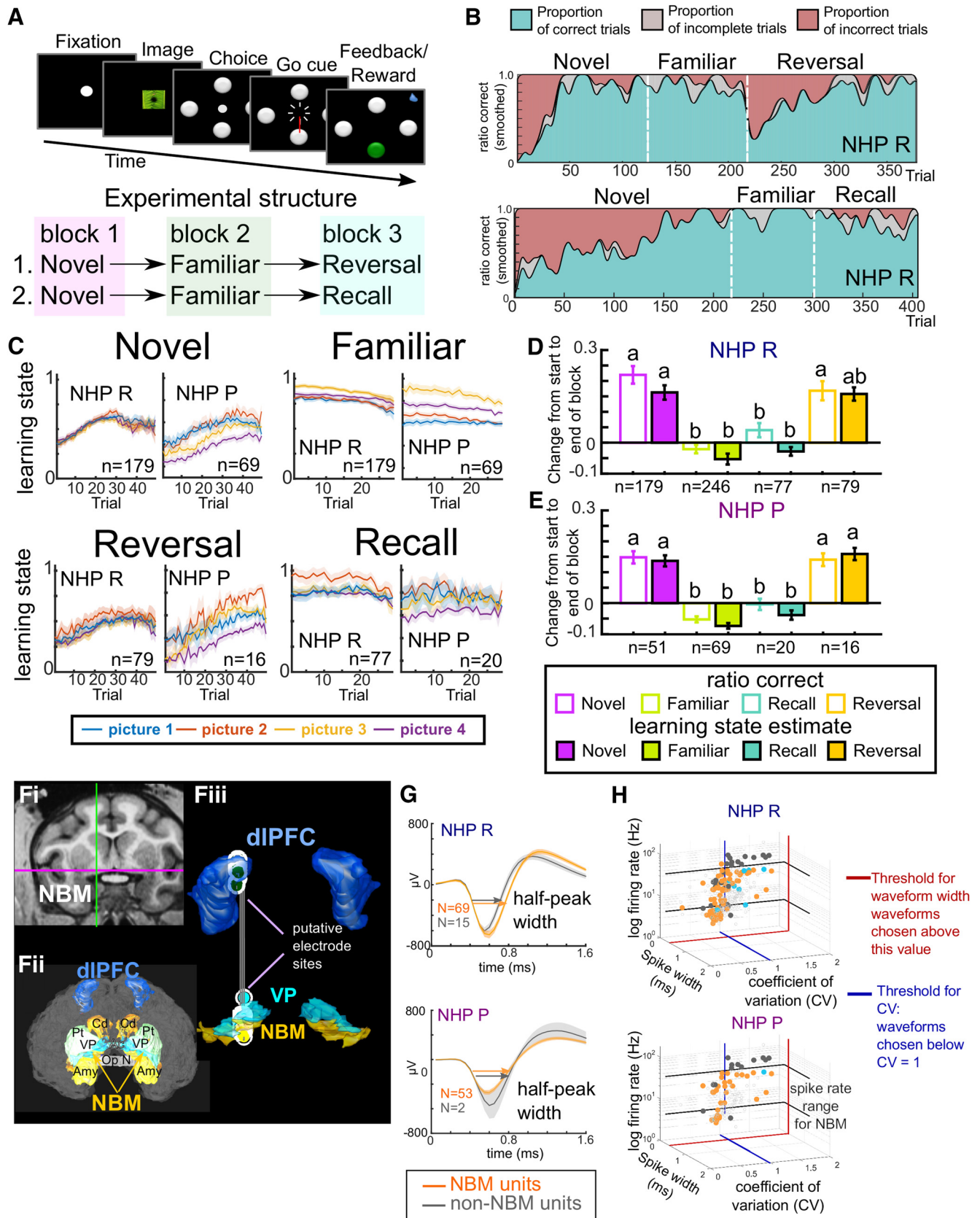
Individual spike times were converted to rates in 5 ms bins and then smoothed by convolution with a Gaussian kernel (50 ms wide, non-causal). Spike rates were then aligned to trial events: fixation, image onset, go cue, and feedback/reward. Then, on a per-trial basis, we defined a "baseline" period as 1 s before the onset of the fixation point. Firing rates during each trial were converted to z-scores based on the mean and SD of that trial's baseline. We then took the absolute value of the z-scored spike rate fluctuations to account for both the negative- and positive-going changes in spike rate that we observed in NBM units (see Results). The variable analyzed over the course of the block is single-trial changes in spike rate driven by task cues (modulation), as opposed to a direct encoding of learning or any other variable in the spike rate (rate coding). We classified units as modulating their firing relative to a task event (e.g., Go cue) when the spike rate exceeded 4 SDs above or below the mean pretrial spike rate for at least 10 5 ms bins. Modulation could be present for only part of a block as long as it was detectable on at least 15 trials.

### Analysis: LFP

On the same channels in which the single-unit activity was recorded, we took the recorded and already filtered and down-sampled LFP and removed line noise (60 Hz). We did this by band-pass filtering the LFP to 55–65 Hz and its harmonics up to 180 Hz and then subtracting these filtered signals from each individual channel. To remove the confound of the average evoked potential and its effect in lower frequency bands, in each block, we subtracted the average LFP evoked potential for correct or incorrect trials from the time courses of the corresponding individual trials. We then performed continuous Morlet wavelet transforms for the frequency range from 1–200 Hz, in 2 Hz steps, to get the Morlet wavelet coefficient amplitude (MWCA, an equivalent of power) using the FieldTrip MATLAB toolbox (Oostenveld et al., 2011) (<http://www.fieldtriptoolbox.org/>, RRID:SCR\_004849). We normalized the MWCA by dividing each time-frequency point by the mean value of the pretrial baseline across all trials in all blocks for each recording session. Normalized power was then averaged to yield power values in the theta (4–8 Hz), alpha (8–15 Hz), beta (15–30 Hz), gamma (30–55 Hz), and high-gamma (65–200 Hz) bands per time point. Recording days for LFP were only excluded for channels localized to outside of the NBM based on anatomical 3D mapping.

### Statistical analyses

Across LFP and spike rate modulation comparisons, we used the Kruskal–Wallis test for nonequivalence of multiple medians and the Wilcoxon rank-sum test (two-sided) for comparisons between individual medians. We additionally used the Wilcoxon signed-rank test (two-sided) for determining whether a distribution's mean was significantly different from zero. *Post hoc* testing of Kruskal–Wallis groups used the Tukey–Kramer method. We corrected for multiple comparisons across frequency bands (for LFP power) and epochs (for spikes and LFP power) by adjusting the target *p*-value with a Bonferroni correction. In addition to the analyses reported here, initial exploratory analyses considered the time of the monkey's response (choice) as an additional epoch and analyzed the delta (1–4 Hz) LFP frequency band. We therefore Bonferroni corrected for comparisons across six bands and five intratrial epochs, yielding a threshold of 0.00167. For time-resolved analysis, we further applied a false discovery rate (FDR) correction to the *p*-value at each time point and only declared a significant result at that time point if the



**Figure 1.** NHPs learn and reverse associations and can do so while recalling already established associations. **A**, Experimental paradigm. During an associative learning task, adult rhesus macaques associated a visual image with one of four target locations. Once the animals indicated a choice by an eye movement, the target changed color to either green or red corresponding with a correct or incorrect choice, respectively (see Materials and Methods). Each experimental session involved three blocks of  $140.3 \pm 51.08$  trials (Novel),  $95.4 \pm 15.98$  trials (Familiar), and either Reversal ( $142.4 \pm 47.50$  trials) or Recall ( $103.7 \pm 27.56$  trials) of associations learned in the same sessions' Novel blocks (average trial numbers across both NHPs). **B**, Representative example of learning behavior for a Novel/Familiar/Reversal experiment (top) and a Novel/Familiar/Recall experiment (bottom) in R. Curves are the ratio correct (Figure legend continues.)

FDR-corrected value was beneath the Bonferroni-corrected target. We did not test for normality because all statistical comparisons were performed using nonparametric tests.

For spike rate and power modulations, we tested correlation between firing rate modulation and behavioral variables or power modulation and behavioral variables using Spearman's rank correlation. For neurobehavioral correlations with two continuous variables (e.g., correlations between spike rate or LFP power and the learning state), we first computed a Spearman correlation coefficient between the two variables for each time point within the trial schema (5 ms nonoverlapping sliding window for spike rates, 50 ms nonoverlapping sliding window for LFP power, aligned to an event of interest such as the Go or Feedback cue). We did this for each experiment block. At each time point, we then collected the correlation coefficients for that time point across all blocks and then tested that distribution against a null hypothesis of zero mean using the Wilcoxon signed-rank test. When displaying time series of such correlations or masking by them (e.g., Fig. 4), we applied an FDR correction to control for multiple comparisons along the time axis and across comparisons, requiring the FDR-corrected  $p$ -values to be under the Bonferroni-corrected threshold described above. For both spike rates and LFP, we did not accept a time point as having a significant correlation unless it was part of a cluster  $>50$  ms in length (10 time bins) for spike rate signals and  $>150$  ms in length (3 time bins) for LFP power comparisons.

To analyze contrasting evolutions of spike and LFP modulation within the Novel block, we realigned these to the standardized learning state. Spike rates were down-sampled to 50 ms bins using MATLAB's "decimate" function to arrive at the same time scale as the LFP power calculations. On each trial and in a 500 ms window before or after each intratrial event, we calculated the mean of either the absolute  $z$ -scored spike rate or the  $z$ -scored LFP power change from baseline (modulation) and then sorted each block's trials by their learning state. Finally, we interpolated the modulation time course for each neural variable to state values between 0.2 and 0.8 in 0.01 standardized unit steps. Few Novel trials had state values outside of the 0.2–0.8 range, making data estimates less stable outside of this range. We did not perform this analysis for Familiar or Recall blocks because the learning state variable did not change sufficiently during those blocks. For illustration, we fitted cubic spline curves to these data (e.g., Fig. 9), but performed analyses directly on the learning-aligned time courses.

To test the correlations per recording between spike rate and LFP power, we performed Spearman correlations between spike rate and LFP power during the Go cue epoch on a per-trial basis and averaged these correlations per combination of single units and the simultaneously re-

corded LFP (e.g., Fig. 9). These analyses were for each experiment type (Novel, Familiar, Recall, and Reversal). The same multiple-comparisons statistical corrections were applied to the dataset as listed above.

Our calculations allowed comparisons and variation between groups and independent samples to be within similar ranges. The groups that we compared were across independent samples (whether neurons or LFP channels), across cognitive tasks (whether learning or short or long-term recall), and across brain regions (NBM and dlPFC). In addition to normalizing data per sample or trial before performing statistical comparisons, the majority of analyses in this study involved correlating a behavioral measure (accuracy, learning, etc.) with a neural measure (such as LFP power or spike rate modulation). The resulting Spearman correlation values resulted in similar bounded ranges between  $-1$  to  $1$  across conditions and groups. These values demonstrated similar variances in both the SE, as shown in the figures, and the statistical comparisons, with only a few groups or comparisons demonstrating significant differences (e.g., Figs. 3, 4, 5, 6, 7, 8).

### Statistical rigor

There was no randomization of the data and no investigator was blinded to the group allocation in the analyses or experimental design. In both animals, we collected recording sessions until we were no longer able to obtain new units on fresh electrode penetrations or when unrelated health issues in R required termination of recordings. We verified that the resulting sample size was adequately powered for the analyses performed. As an example, we found (see Fig. 8) that NBM high-gamma power (HGP) significantly correlated with learning after the Go cue. In the 0.5 s after the Go cue, the correlation (Spearman rho) between learning state and HGP had a mean of 0.368 and SD of 0.15 across 73 independent recording sessions. We generated surrogate data that followed that distribution, for 500 replicates of each putative  $n$  (number of recording sessions). We then tested each replicate for a nonzero median with the Wilcoxon signed-rank test. With the Bonferroni-corrected  $\alpha = 0.00167$  that we used in all LFP analyses, we needed only  $n = 11$  to have 80% power to reject the null hypothesis. Our sample size is nearly seven times larger than the minimum necessary and similar calculations apply to the data of other figures.

### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Results

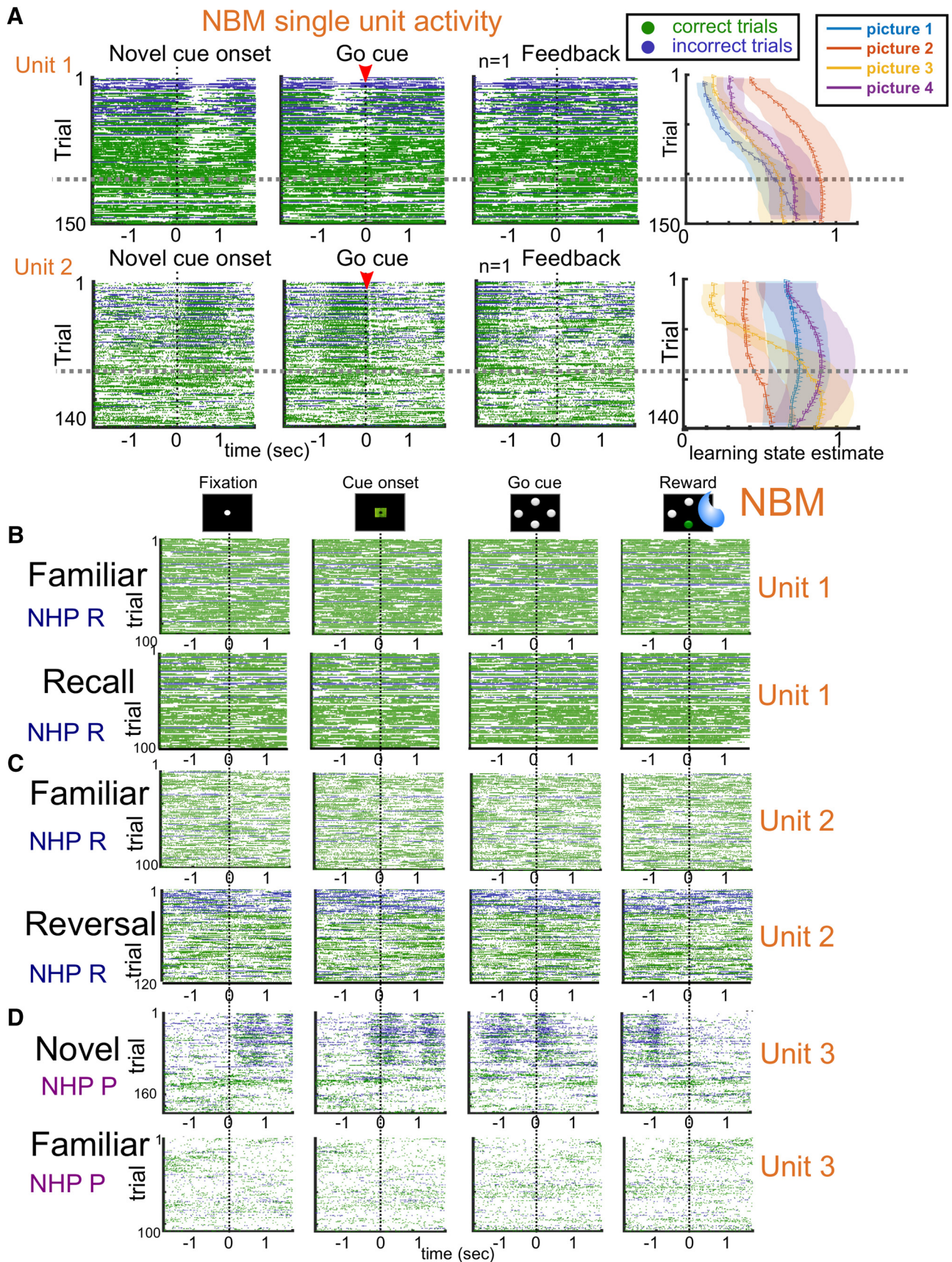
### Learning behavior

We recorded neural activity as two behaving NHPs performed an associative learning task (Williams and Eskandar, 2006). The animals learned, by trial and error, to associate each of four images with one of four target locations (Fig. 1A; see Materials and Methods). The task was subdivided into epochs, beginning with a centralized fixation point (Fixation), followed by a stimulus presentation (Image onset). Next, four target objects appeared and the stimulus image was cleared (Go cue), which signaled to the animals to make a choice by looking at the target. After this eye movement or saccade to the object, the target changed color to provide visual correct/incorrect feedback (Feedback). Correct choices rewarded animals with juice (Reward).

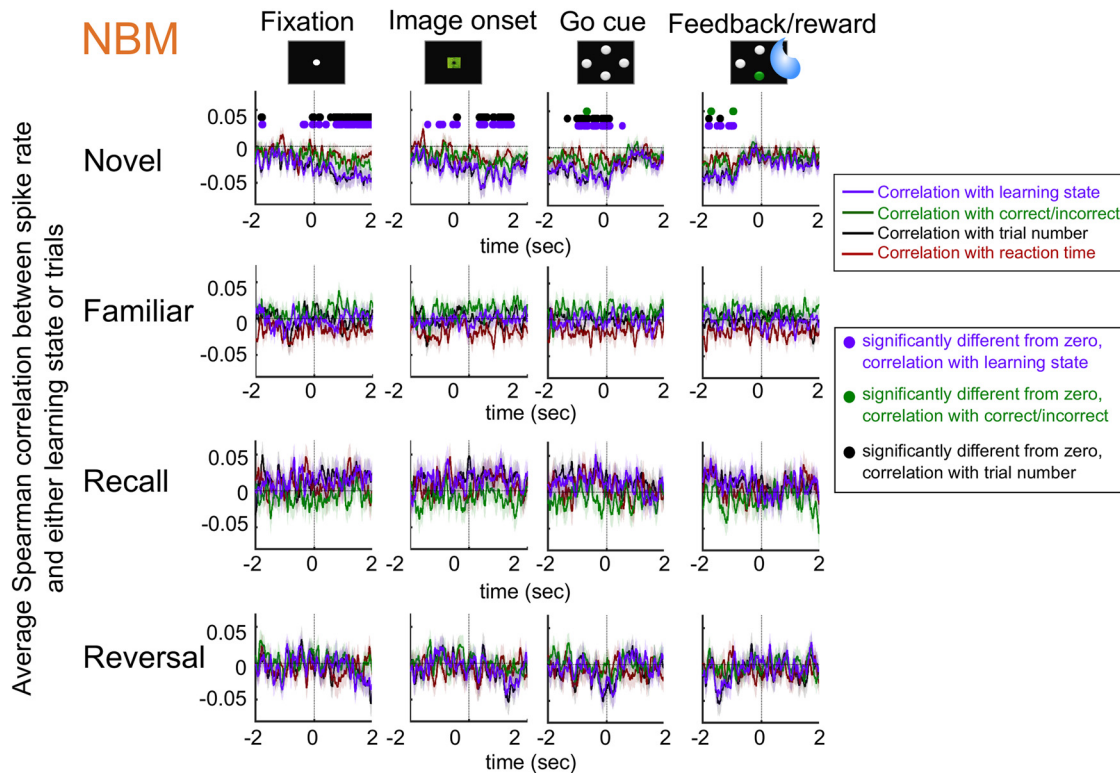
We used a modular design consisting of multiple combinations of four block types: (1) Novel block, in which animals were expected to learn, by trial and error, to make the correct associations of four novel images with their correct target locations; (2) the Familiar block, in which the animals were required to recall associations between four images that were repeated over the entirety of training and data collection (animals were well trained in the correct target associations and usually completed the block quickly); (3) the Recall block, in which animals were retested on the associations learned in the Novel block (this block allowed us

←

(Figure legend continued.) in a sliding five-trial window smoothed by a 10-trial noncausal Gaussian kernel. Performance increases gradually from low levels when learning new associations (Novel, Reversal) and is consistently high when performing learned associations (Familiar, Recall). **C**, Average learning state curves for both NHPs across Novel, Familiar, Reversal, and Recall experiments.  $n$  indicates the number of blocks run per NHP and experimental condition. **D**, **E**, evidence of learning in R and P. We compared ratio-correct (**D**) and learning state (**E**), averaged across each block's four target images, between the first 50 and final 50 trials of each block. In both animals and using either metric, there is significantly greater learning (increase from block start to block end) during Novel and Reversal blocks compared with the Recall and Familiar blocks, in which animals perform already-learned associations. Letters "a" and "b" above bars indicate statistically separable groups by Tukey *post hoc* testing after Kruskal–Wallis test. **D**, R:  $\chi^2 = 161.71, p < 0.000001$ . **E**, P:  $\chi^2 = 215.7, p < 0.000001$ . Error bars indicate SEM. **Fi–Fiii**, T1-weighted MRI with fiducial markers showing electrode trajectories for the NBM and the dlPFC with 3D reconstructions of neighboring brain structures. We use anatomical mapping to additionally separate VP from NBM recordings by reconstructing the areas of the VP and NBM on the MRI and dividing recordings based on the amount of overlap between the electrode tip and the 3D reconstructed structures (see Materials and Methods). Cd, Caudate; Pt, putamen; Ac, anterior commissure; Op N, optic nerve; Amy, amygdala. **G**, Average waveform shape of units classified as NBM or non-NBM (only including recordings which mapped in 3D to the NBM region) illustrating differences in peak width. Shaded regions indicate SEM. **H**, The three criteria per unit and the classification of the units into NBM and non-NBM. The latter were not used for further analysis.



**Figure 2.** NBM neurons modulate firing specifically in response to novel cues. **A**, Representative example of two NBM neurons collected on different days from R. Unit 1 suppressed firing immediately after stimulus presentation, returning to baseline firing at the Go cue. Unit 2’s activity increased during the same period and normalized at the Go cue. This modulation lasted for the first 100 trials of the novel block, after which the spike rate showed less pervasive variability (dashed horizontal line). This was particularly visible after the animal (*Figure legend continues.*)



**Figure 3.** NBM unit modulation is specific to Novel blocks. Mean Spearman rank correlations between z-scored spike rate changes (relative to baseline activity 1 s before fixation) and trial-level variables using a sliding 50 ms window. Only Novel blocks showed median correlations significantly different from zero, as indicated by circles above the curves (Wilcoxon signed-rank test on the distribution of Spearman coefficients at each time bin, thresholded to FDR correction through time to  $p < 0.05$ , Bonferroni corrected for testing of multiple epochs). Both trial number and learning state correlated with spike rate modulation, suggesting that novelty of the stimulus drives modulation.  $n$  indicates the number of units per experiment in the NBM (orange). Shaded regions around lines indicate SEM. Curves show correlations between z-scored spike rate change from baseline and learning (purple), trial number (black), accuracy (green), or reaction time (burgundy). Columns are the different epochs of the task. Rows are the different types of blocks in the task. The Recall and Reversal blocks have fewer neurons due to loss of isolation on some cells during the final block of a session as well as experimental design (see Materials and Methods). In the Novel blocks, NBM firing modulating was significantly correlated to both learning and trial number, both of which are measures of stimulus novelty. No such encoding is present for any other block. Data are from 112 NBM units across two NHPs. Shaded regions around lines indicate SEM.

to compare the recollection of newly learned associations and contrasts with the long-term memory assessed in the Familiar block); and (4) the Reversal block, in which animals were presented with the same four images used during the Novel block, but the associated targets were changed, requiring animals to learn new associations. By reversing associations during this block, we could assess relearning of now-recognizable stimuli and dissociate between learning and novelty because as these stimuli were already presented during the Novel block (Fig. 1A; see Materials and Methods). Each experimental session contained combinations of three block types such as Novel/Familiar/Recall or Novel/Familiar/Reversal. Importantly, during Novel blocks, both the images and their associations were new to the

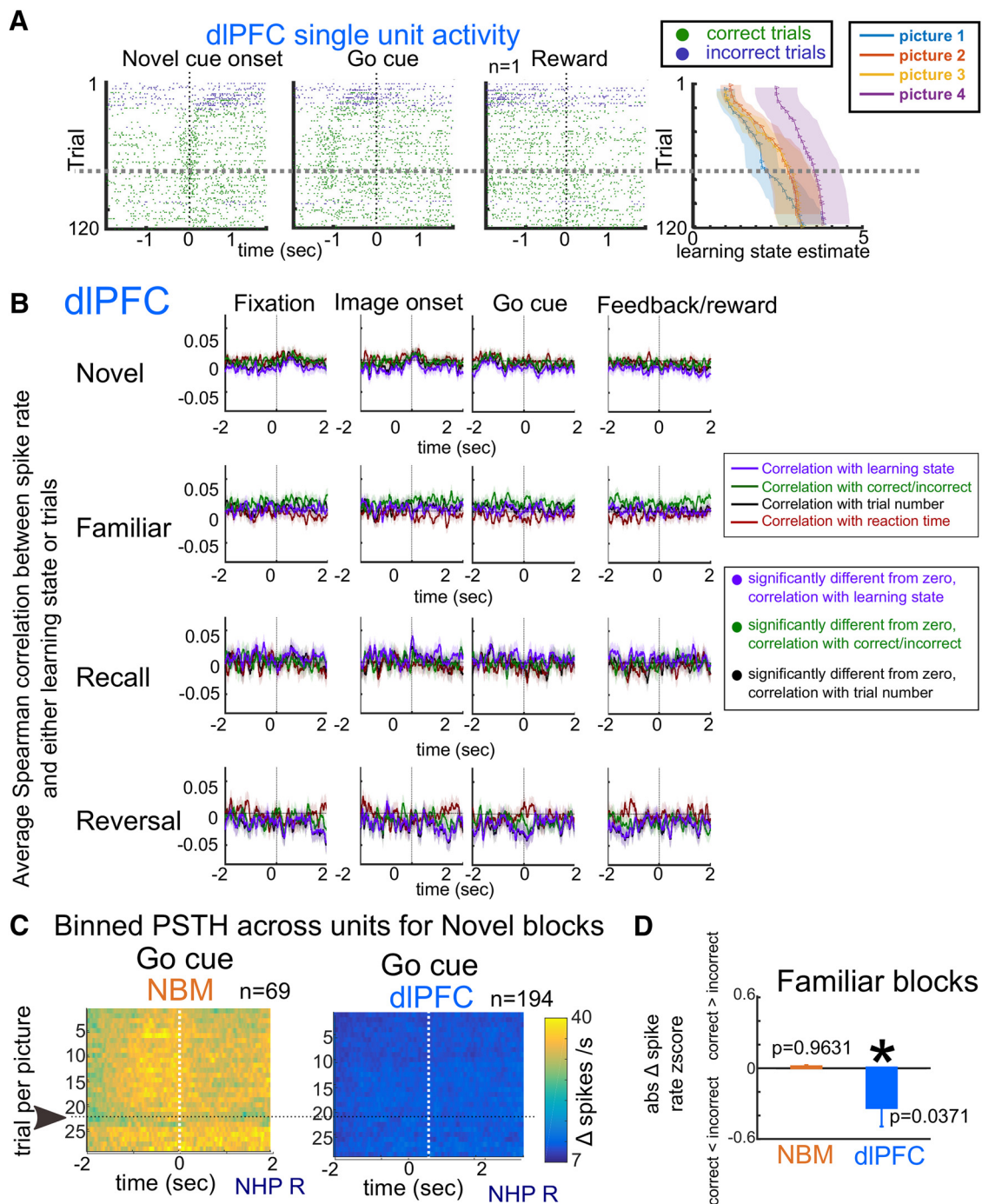
animal, whereas during Reversal, only the associations were new. Familiar blocks helped distinguish learning-related signals from signals more closely associated with reward. During Familiar associations, the animal could strongly anticipate reward on each trial, but basically no learning occurred. The Familiar block also served as a break between Novel and Recall blocks or Novel and Reversal blocks.

Both animals successfully learned, performed, and reversed associations between each of the four images and the associated target location (Fig. 1B). To summarize this, we integrated decision time and accuracy into a learning state (also called cognitive state) variable that was defined for each of the four images on every trial (Prerau et al., 2009). Having an estimate of learning for each possible association on every trial allowed for more accurate regression of neural activity against learning (Prerau et al., 2009) (Fig. 1C). As expected, the learning state estimate increased during Novel and Reversal blocks, as did the ratio correct (Asaad and Eskandar, 2011) (Fig. 1D,E). Because there was little further learning during Familiar and Recall blocks, the learning state remained high during these blocks (Fig. 1C–E).

**Single-neuron responses**

Because the NBM may influence learning through cortical projections (Mesulam et al., 1983; Irle and Markowitsch, 1986), we simultaneously recorded single neurons and LFPs from the NBM and dlPFC as animals performed the task. We confirmed the identity of NBM neurons both by anatomic mapping (Fig. 1Fi–

←  
(Figure legend continued.) learned the association, as represented in the learning state estimate to the right. Similarly, during Familiar (B, top) and Recall (B, bottom) blocks, again containing only well learned associations, Unit 1 from A maintains a constant firing rate. Unit 2, also shown in A (bottom), further demonstrates that novel associations alone are not sufficient to produce NBM modulation. This unit also does not modulate to task cues during the Familiar (C, top) block. Even during Reversal (C, bottom), when non-novel stimuli are paired with new associations, task cues do not change this unit’s firing rate. C (bottom) also illustrates that correct/incorrect performance on any given trial does not drive firing rate. D, After the Fixation cue of Novel blocks, when as-yet-unlearned images are presented, this unit from P increases its firing rate. As the associations are learned (increasing trial number), this modulation ceases, as evident after trial 100 in this plot. During a Familiar block in which stimuli are salient but associations are well known, this same unit does not modulate to task cues. Compare these responses with those shown in A–C, which is the same pattern in units from R.

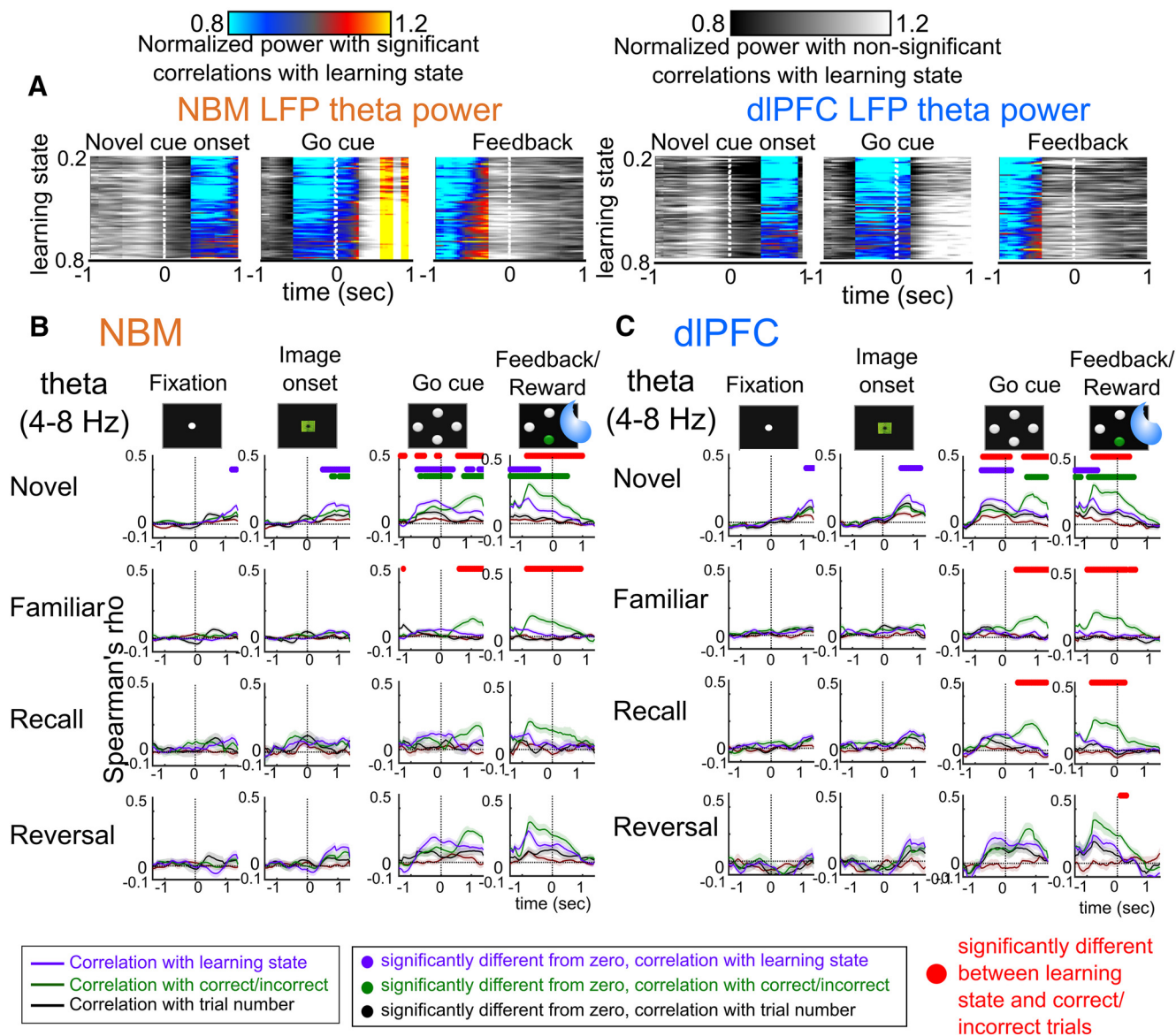


**Figure 4.** In contrast to NBM, dIPFC neurons do not respond strongly to stimulus novelty or unlearned associations. **A**, Example raster plots of a dIPFC unit (R). The unit did not show substantial cue-related modulation and this did not change as the learning state increased. This is further evident from **B**, which shows that dIPFC units did not demonstrate any significant correlations between learning state and spike rate changes for the same time windows. Mean Spearman rank correlations between z-scored spike rate changes (relative to baseline activity 1 s before fixation) and trial-level variables using a sliding 5 ms window are shown. *n* indicates the number of units per experiment in the dIPFC (blue). In the Novel blocks, dIPFC modulation was not significantly correlated to both learning and trial number, both of which are measures of stimulus novelty. Data are from 210 dIPFC units across two NHPs. Shaded regions around lines indicate SEM. **C**, Contrast between dIPFC and NBM activity is further illustrated in binned peristimulus time histograms (PSTHs) in the two regions. We took all units recorded during Novel blocks in R and averaged their firing around the time of the Go cue, normalizing and z-scoring activity as per Materials and Methods. When averaging peristimulus activity at the level of the individual picture, NBM neurons as a population show substantial modulation (absolute value of firing rate change vs pretrial baseline) just before the Go cue. As per Figure 2, this modulation diminishes with learning, seen here at approximately trial 20 (black arrow and dotted line). dIPFC neurons as a group do not show this modulation. **D**, Changes in spike rate with error and correct/incorrect. dIPFC neurons encoded error, especially unexpected error, more strongly than NBM. In the Familiar block, when animals had a high degree of confidence in their responses, dIPFC neurons had significantly higher ( $p = 0.0371$ ; Wilcoxon signed-rank test) firing rates immediately after Feedback in incorrect trials compared with correct. No such difference was found in NBM firing rates ( $p = 0.9631$ ; Wilcoxon signed-rank test). Error bars indicate SEM.

*Fiii*; see Materials and Methods) and waveform properties of each unit after spike sorting (Richardson and DeLong, 1990) (Fig. 1G,H). To be included in the analysis as NBM units, we used predetermined criteria (Richardson and DeLong, 1990). Two of

the following 3 criteria had to be met: (1) a spontaneous firing rate of 5–40 Hz; (2) a coefficient of variation of the interspike interval of  $<1$ ; and (3) a spike duration of  $>180 \mu\text{s}$  (initial negative phase, 200–10 kHz filtering) (Fig. 1G,H). Using this



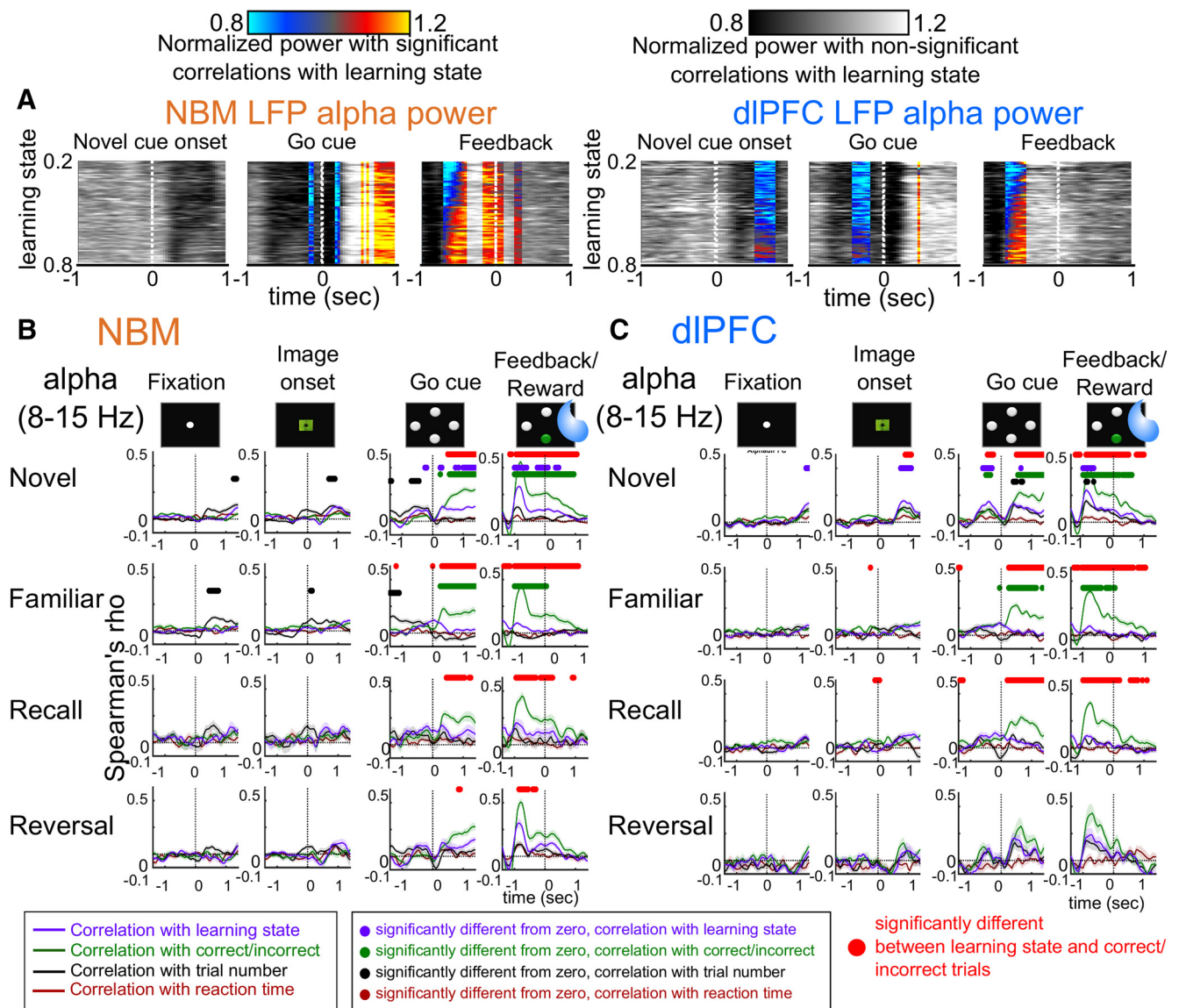


**Figure 5.** Low-frequency (theta-band) LFP power encodes both learning and reward expectation in NBM and dlPFC. **A**, Average changes in theta power from baseline across recordings and animals across learning states in the NBM ( $n = 44$  recordings across two animals; left) and dlPFC ( $n = 49$  recordings across two animals; right) for Novel blocks. Color versus gray designates 50 ms windows in which the correlation between theta power and learning state values is or is not significant, respectively, after FDR correction. For plotting, the learning state was binned from 0 to 1 in increments of 0.005 and trials were assigned to their nearest bin. This binning was not used for the statistical calculation. In both structures, theta power decreased after image presentation and returned to baseline just after the Go cue early in the learning process. This modulation dissipated as the animals learned. **B, C**, Spearman rank correlations among theta LFP power and learning state (purple), choice accuracy (green), and trial number (black). Red markers indicate points where the learning state and accuracy correlations are significantly different ( $p < 0.05$ , Bonferroni corrected for epoch and frequency band comparisons and FDR corrected through time, Wilcoxon rank-sum test). Other color markers indicate whether the mean correlation at each time point was significantly different from zero ( $p < 0.05$ , Bonferroni corrected for frequency bands and epochs and FDR corrected through time, Wilcoxon signed-rank test). The learning state was significantly correlated to theta power (and more strongly correlated than accuracy) from  $\sim 700$  ms before the Go cue (purple arrowheads) until just after Go cue in both NBM (**B**) and dlPFC (**C**). Accuracy was significantly correlated with theta power in the Feedback-locked analysis (green arrowheads) and this correlation exceeded the correlation with learning state both before and after Feedback. **B, C**, Shaded regions around lines indicate SEM.

method, we identified 322 neurons, with 112 NBM units and 210 dlPFC units.

We hypothesized that NBM neurons would respond to novel, salient stimuli (Mesulam et al., 1983; Rigdon and Pirch, 1986; Voytko and Lou, 1996; Wenk, 1997), whereas we have shown previously that dlPFC neurons signal reward expectation during learning (Asaad and Eskandar, 2011). Indeed, nearly half the NBM units (R: 33/69, P: 21/43) changed their firing rate in response to task cues for the first 50–100 trials of the Novel block, then stopped cue-linked modulation for the remainder of the block (Fig. 2A). These cells modulated their firing only when

novel and salient stimuli were presented and ceased to modulate when the stimuli were no longer novel. Units that modulated in Novel blocks did not modulate in other block types (Fig. 2B–D). Individual examples show that NBM neurons did not modulate during Recall blocks (Fig. 2B), which represent newly formed associations; during Reversal blocks (Fig. 2C), when the Novel object must be associated with a different location; and finally, no clear change in firing rate during Familiar blocks, representing established associations with highly familiar objects (Fig. 2B–D). Therefore, the combination of a Novel object with a new association appears necessary for these neurons to modulate their firing.

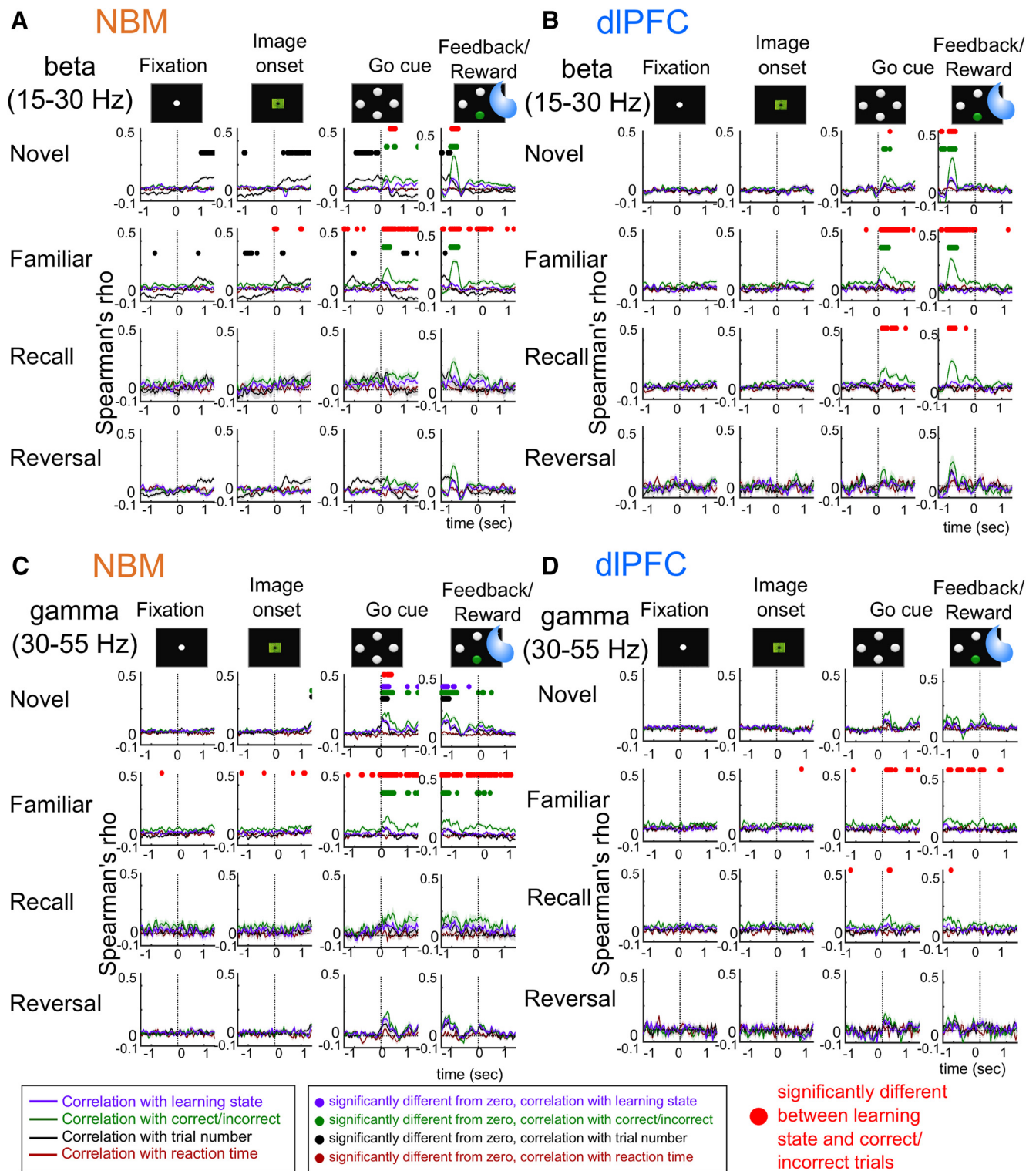


**Figure 6.** Low-frequency (alpha) LFP encoding of learning, but not reward anticipation, is specific to Novel blocks. **A**, Same metric and statistical approaches as in Figure 5A, but in the alpha (8–15 Hz) band. **B, C**, Same metrics and statistical approaches as in Figure 5, B and C, but in the alpha band in the NBM (**B**) and dIPFC (**C**). Post-image suppression was less prominent and the post-Go cue enhancement in NBM was stronger than the dIPFC. Alpha power generally encoded accuracy more strongly than it encoded learning and significant encoding was present as early as 1 s before Feedback. This was consistent with a reward anticipation signal. **B, C**, Shaded regions around lines indicate SEM. Data include NBM:  $n = 44$  recordings (sessions) across two animals and dIPFC:  $n = 49$  recordings (sessions) across two animals. Line and color markers follow the same schema as Figure 5, B and C. Shaded regions around lines indicate SEM.

The changes in firing rate could be due to novelty (number of exposures to any given image), learning, or reward anticipation from correct trials. To test this, we correlated the peri-event modulation of each unit in 5 ms bins with the learning state, trial number (testing novelty), decision accuracy (testing reward expectation or delivery), and reaction time (to separate the effect of reaction time from learning state). NBM unit modulation had significant negative correlations ( $p < 0.05$ , FDR controlled through time) to learning state and trial number, and not to reaction time, during Novel blocks. Significant correlations had begun shortly after image onset and abruptly offset at the Go cue (Fig. 3). NBM firing modulation did not correlate to learning, accuracy, trial number, or reaction time in any other block type, including Reversal (Fig. 3). This was consistent with a model in which NBM neurons modulated their firing when both cues and associations were new as the cues are presented and thus salient. Modulation ceased once the cues were learned (still salient but no

longer novel). Interestingly, this NBM spike rate modulation occurs immediately before and around the time point that the NHPs made a choice, meaning that the NBM activity could be tied to the decision making in Novel blocks.

In contrast, dIPFC neurons had little change in cue-linked modulation with learning (Fig. 4A). The novelty encoding observed with NBM neurons was not present in dIPFC neurons during any block (Fig. 4A–C). Consistent with our prior findings (Asaad and Eskandar, 2011), dIPFC neurons encoded prediction errors, significantly changing their firing rate relative to baseline for incorrect versus correct trials. This was most evident during Familiar blocks, in which errors occurred in the context of a high degree of confidence ( $p = 0.0371$ ; Wilcoxon signed-rank test;  $z$ -value = 2.0844; Fig. 4D). Specifically, errors during well learned associations induce changes in dIPFC activity. NBM neurons had no significant differences in spike rate modulation between correct and incorrect trials for the Familiar block type (Fig. 4C,D).

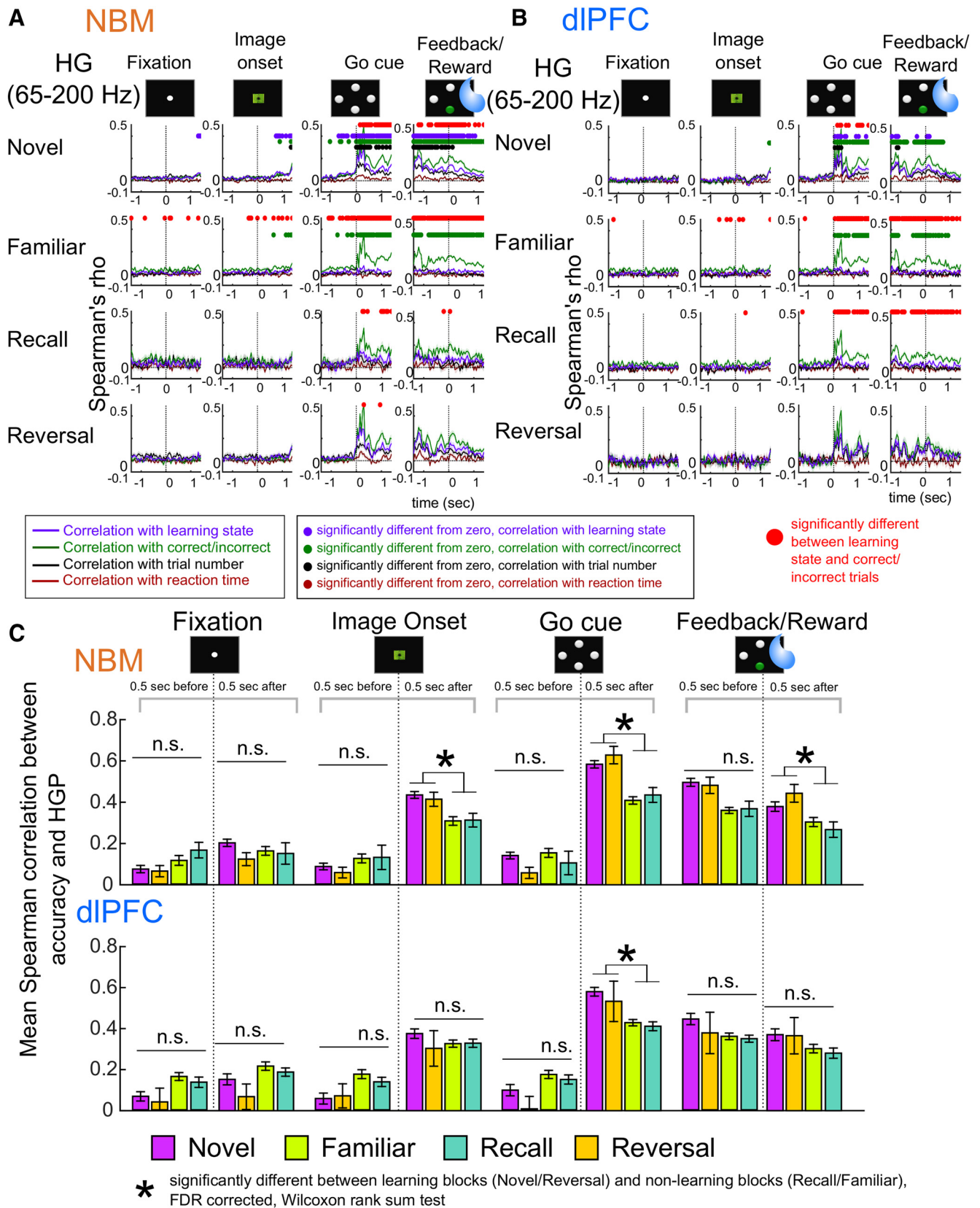


**Figure 7.** LFP power in beta (**A, B**) and gamma (**C, D**) bands does not encode learning processes. Correlations between LFP power and trial-level variables (learning state, accuracy, trial number, and reaction time) are plotted and statistically evaluated as in same metrics and statistical approaches as in Figure 5, **B** and **C**. Beta-band (15–30 Hz) correlation curves are plotted for the NBM (**A**) and dIPFC (**B**). Gamma-band (30–55 Hz) correlation curves are plotted for the NBM (**C**) and dIPFC (**D**). There are few time points where power is significantly correlated with the learning state to a smaller extent than in lower frequencies. Power correlations to learning state are always smaller than power correlations to accuracy (reward/anticipation). Data include NBM:  $n = 44$  recordings (sessions) across two animals and dIPFC:  $n = 49$  recordings (sessions) across two animals. Line and color markers follow the same schema as Figure 5, **B** and **C**. Shaded regions around lines indicate SEM.

**Multispectral band encoding in LFP power**

Because learning often involves changes in LFP oscillations in both cortical and subcortical structures (Brincat and Miller, 2015; Haque et al., 2015; Watrous et al., 2015), we examined

task-induced LFP power changes driven by learning, accuracy, reaction time, and trial number in the same sliding correlation framework. In sum, we found that theta (5–8 Hz) power early in the trial encoded learning, theta and alpha (8–15 Hz) late in the



**Figure 8.** HGP in the NBM encodes reward expectation, leading to an anticorrelation between spike and HGP modulation during learning. **A, B**, HGP (65–200 Hz) modulation is correlated with trial accuracy more strongly than with learning, although both show significant correlations just after the Go cue. This correlation between learning and HGP is present during novel learning such as Novel blocks, but not in Familiar blocks. Accuracy (which correlates with anticipated reward) showed significant correlations with HGP in both structures, continuing after Feedback when the reward was being consumed. This encoding reached statistical significance for most of the peri-feedback period in Novel and Familiar blocks in dIPFC and the NBM and was qualitatively present in Recall and Reversal blocks. Correlations between LFP power and trial-level variables (learning state, accuracy, trial number, and reaction time) are plotted and statistically evaluated as in same metrics and statistical approaches as in Figure 5, *B* and *C*. NBM ( $n = 44$  recordings across two animals). **C**, These correlations are present in the dIPFC, though with fewer (*Figure legend continues*.)

encoded reward feedback (or reinforcement through reward), and high gamma (65–200 Hz) encoded reward anticipation.

In early trials during Novel blocks, theta power decreased (relative to pretrial baseline) at image onset. Theta then increased after the Go cue in both the NBM and dlPFC (Fig. 5A). As seen with spike rate changes, both structures' theta power modulation attenuated as learning progressed. This was reflected in a significant correlation between theta power and the learning state in both NBM and dlPFC ( $p < 0.05$  after FDR correction), starting 650–700 ms before the Go cue (Fig. 5B,C). Unlike spike rate modulation, neither dlPFC nor NBM theta power was significantly correlated with trial number, suggesting that the theta-band effect was specific to the learning process and not to stimulus novelty. Pre-Go-cue theta power in the NBM (but not dlPFC) also significantly correlated with decision accuracy, but the correlation with learning state was consistently higher than the correlation with accuracy in both structures (Fig. 5B,C). The first time point of significant theta correlation with the learning state was earlier in the NBM than in the dlPFC (–700 vs –650 ms relative to the Go cue), suggesting that the learning-related theta band signal might originate in or near the NBM, particularly before the NHPs made their choices. The theta power/learning state correlation also exceeded the correlation with accuracy before the Go cue during Reversal blocks, but did not reach the prespecified significance threshold (Fig. 5B,C). This theta band activity, however, did not correlate as strongly with learning state during the Familiar or Recall blocks. Interestingly, however, the correlation to learning was slightly higher during Recall blocks than Familiar blocks, indicating a differentiation between newly acquired associations and well learned associations (Fig. 5B,C). This distinction in theta-band power preceding a choice could represent the NHPs making a decision as they were learning.

Alpha power modulation in NBM and dlPFC also showed periods of significant correlation to learning state, but correlated more strongly to reward anticipation, reward reinforcement (as the NHPs received a green target reinforcement cue before receiving a reward; Fig. 1) and consumption (Fig. 6A). Both before and after Feedback, in the NBM and dlPFC, theta and alpha power were significantly correlated with the accuracy of the animal's choice, with a peak around the time that the NHPs would receive the green target reinforcement cue ( $p < 0.05$ , FDR corrected). When analyzed locked to the Feedback cue, the accuracy correlation exceeded the power correlation with learning state (Fig. 6B,C). In contrast to the learning-related theta desynchronization (power decrease) that we observed before the Go cue, reward-correlated (or accuracy-correlated) theta/alpha power increased over the pretrial baseline and the increase amplified over the course of a block (Fig. 6). The reward encoding was stronger in the alpha band than in the theta band (NBM peak  $\rho = 0.42$  in alpha and 0.31 in theta, dlPFC peak  $\rho = 0.44$  in alpha and 0.29 in theta). After Feedback was delivered, the correlation between power and reward/accuracy sharply declined, showing

nonsignificant correlations by 450 ms (dlPFC) or 900 ms (NBM) after reward (Figs. 5B,C, 6B,C). This was consistent with a reward anticipatory-related signal that peaked when the animals were given the green target reinforcement cue that was no longer correlated with reward once it was consumed. Consistent with a reward anticipation signal, Feedback-locked alpha modulation was not specific to Novel blocks. Significant encoding also occurred during Familiar blocks, during Recall of well formed associations, and Reversal of those associations (Fig. 6B,C). Beta power, in contrast, correlated with accuracy only in a short time window around the Choice event and never encoded learning state more strongly than accuracy (Fig. 7A,B).

High gamma (65–200 Hz, HGP) and gamma (30–55 Hz) LFP power modulation also encoded reward anticipation and consumption (Figs. 7B,C, 8A,B). As with theta and alpha, HGP in the NBM and dlPFC encoded a correct trial, but modulated earlier in the trial, starting at the Go cue (Fig. 8A,B). This is consistent with reward anticipation, driven by confidence in a well learned decision. HGP modulation in NBM also significantly correlated with accuracy around the Feedback cue in Novel and Familiar blocks (reward consumption; Fig. 8). Further, neither structures' HGP showed higher correlation with learning state, reaction time, or trial number than with accuracy, suggesting a more "pure" reward encoding in HGP compared with the lower frequency bands. Encoding was stronger in the presence of active learning: correlations in NBM after Go and Feedback were significantly higher in Novel and Reversal blocks compared with Familiar and Recall (Fig. 8C). This result indicated there was a separation between blocks in the process of learning whether learning a Novel association or Reversing a newly learned association versus Recall or Familiar, when the association is recently acquired or well learned.

### Spike–LFP dissociation

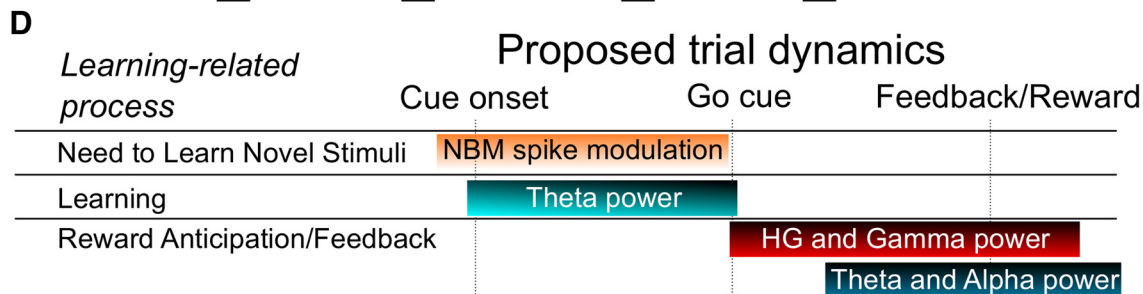
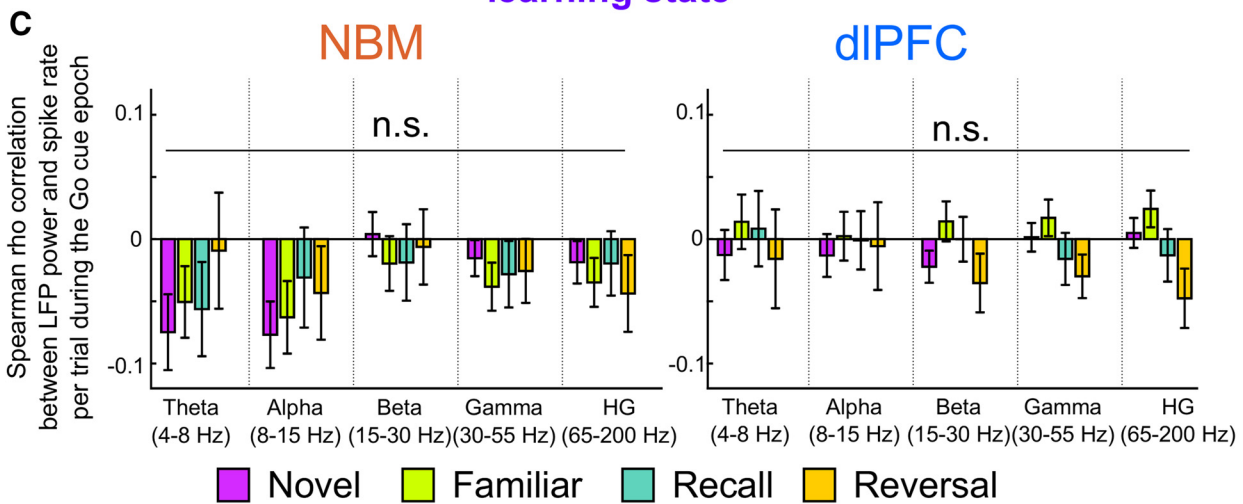
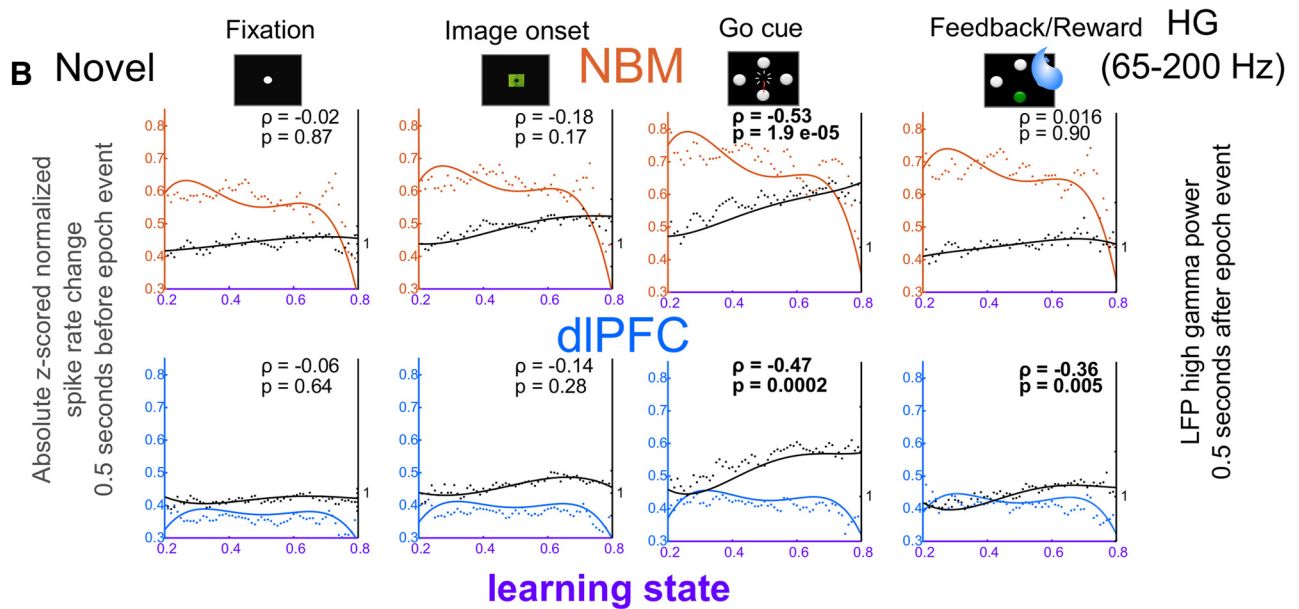
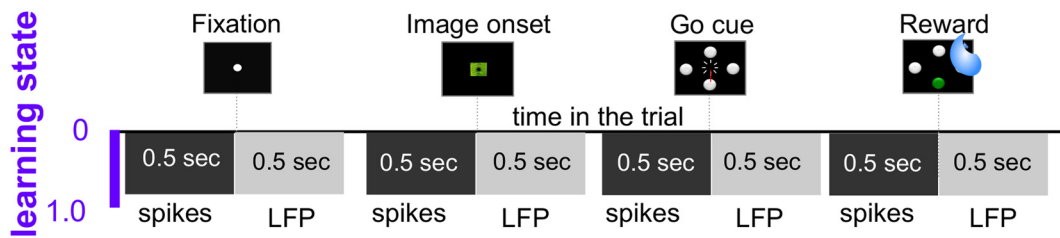
Together, these results suggested that spike rate and LFP power modulation in the NBM orthogonally encoded the learning process, particularly in the formation of Novel associations. Spike rate modulation increased before the Go cue and this modulation attenuated with learning (Fig. 2). HGP modulation increased after the Go cue and became stronger with learning (Fig. 8). To quantify the anticorrelation, we binned the learning state in steps of 0.01 and computed the mean HGP and spike rate modulation around Go for each bin. The binned HGP and spike modulation time-series had strong negative correlations in both NBM ( $\rho = -0.53$ ,  $p = 1.9e-5$ ) and dlPFC ( $\rho = -0.47$ ,  $p = 2.4e-4$ ) (Fig. 9A,B). Spike and HGP modulation were also anticorrelated in dlPFC around Reward ( $\rho = -0.35$ ,  $p = 0.0055$ ). At other trial epochs, neither structure showed strong correlations between spike and HGP modulation (Fig. 9A,B). HGP in cortex has been believed to track the level of local spiking and correlates with functional MRI signal (Ojemann et al., 2013; Yazdan-Shahmorad et al., 2013; Watrous et al., 2015). Our dissociation of these signals in NBM suggests that the spiking–HGP link is not universal.

In addition, we correlated spike rates with power per electrode per brain region during the Go cue epoch (when the most modulation occurs during the epochs) for all five frequency bands (Fig. 9C). We found no significant correlations both within blocks, within frequency bands, and between block types (for significant Spearman rho correlations:  $p < 0.0013$ , FDR controlled, Wilcoxon signed-rank test; For multiple comparisons between block types: NBM:  $\chi^2 = 16.91$ ;  $p = 0.596$ , dlPFC:  $\chi^2 = 16.03$ ;  $p = 0.655$ ; Kruskal–Wallis test; Fig. 9C). These results indicated that, outside of the effect of learning state, LFP power

←

(Figure legend continued.) significant points than in the NBM ( $n = 49$  recordings across two animals). Line and color markers follow the same schema as Figure 5, B and C. This encoding is also present before and after Feedback, implying that it represents both reward anticipation and consumption. Shaded regions around lines indicate SEM. C, Reward encoding in the perievent HGP modulation was also stronger during active learning. After the Go cue, in blocks that involve new learning (Novel/Reversal), the mean correlation between HGP and choice accuracy was larger than this same correlation in Familiar/Recall blocks without learning ( $p < 0.00013$ , Wilcoxon rank-sum test with FDR correction for testing across frequency bands, block types, epochs, and regions). Error bars represent SEM.

### A Changes across learning state



**Figure 9.** HGP is anticorrelated with spike activity modulation during learning. **A**, Analysis schematic for the correlation of spike rate and LFP modulation with learning state. The learning state was binned into ascending steps of 0.01 in each block (purple arrow). We then took the mean of spike rate modulation (dark gray boxes) or HGP (light gray boxes) during 0.5 s windows on each trial of that block. Each window was time locked to a specific event within the trial. We interpolated those mean values to the learning state bins, creating a standardized time course of event-locked neural modulation as a function of learning within a block. The learning state was binned into ascending steps of 0.01 in each block. We then took the mean (*Figure legend continues*.)

and NBM spike rate do not demonstrate a strong, significant anticorrelation. Instead, LFP power and NBM spike rate are anticorrelated in the context of Novel learning as the NHPs made their decisions in the task.

## Discussion

Our results suggest that single NBM neurons modulated strongly during early learning of novel cues and then ceased to respond late in learning. At the same time, low-frequency oscillations increased their modulation as the association was learned, possibly representing populations of neurons acting in increasing synchrony. We also found that novelty, learning stages, and reward were encoded by separable neural signals in the NBM and dlPFC (Fig. 9C). Although NBM spike activity encoded the need to form associations for novel stimuli, LFP encoded whether they had been formed and the anticipation of reward linked to correct association retrieval. Low- and high-frequency LFP encoding occurred in both NBM and dlPFC, generally with an earlier onset in the NBM. This may reflect a causal relationship in which NBM signals drive their counterparts in PFC, although proving that causality would require further experiments such as the use of neural stimulation and multielectrode arrays. From these current results, we hypothesize that this LFP–spike anticorrelation involves network effects of possible broadly projecting NBM neurons suppressing theta-band activity. This network effect could then change over time as the NBM neurons no longer modulate their firing in response to a Novel association being formed. However, without more electrodes per brain region, we did not feel that we could demonstrate this with the current dataset.

This concept of NBM neurons driving cortical targets is supported by the fact that NBM neurons project throughout the cortex and are the primary source of cholinergic innervation to the cortex in primates (Mesulam et al., 1983; Baxter and Chiba, 1999; Mesulam, 2013; Liu et al., 2015). NBM cholinergic efferents innervate the entire cortical mantle and olfactory bulb, whereas only the limbic and paralimbic areas (such as the cingulate gyrus, hippocampus, amygdala, and nucleus accumbens) have reciprocal connections back to the NBM (Gratwicke et al., 2013; Mesulam, 2013; Liu et al., 2015). It is therefore not surprising that stimulation of the NBM induces cortex-wide synchrony (Kilgard and Merzenich, 1998). Importantly, anatomically defined NBM subregions project to different parts of the brain (Liu et al., 2015).

←

(Figure legend continued.) of spike rate modulation (blue/orange) or HGP (black) during 0.5 s windows on each trial of that block. We interpolated those mean values to the learning state bins, creating a standardized time course of neural modulation relative to learning within a block. **B**, During Novel blocks, spike rate modulation around the Go cue is anticorrelated with HGP modulation around the same cue in both NBM (orange) and dlPFC (blue), as well as around the Reward/Feedback event in dlPFC. Points represent the mean modulation at each step of the learning state, whereas curves are a cubic spline fit to those points. Inset  $\rho$  and  $p$ -values are Spearman correlations and  $p$ -values between the interpolated spike and HGP modulation time courses. **C**, Correlation values between spike rate and LFP power for five frequency bands on a per trial basis for the different experiment types during the Go cue epoch. No mean Spearman correlation values were significantly different from zero ( $p < 0.0013$ , FDR controlled, Wilcoxon signed-rank test). In addition, the mean Spearman correlation values were not significantly different from one another (NBM:  $\chi^2 = 16.91$ ;  $p = 0.596$ , dlPFC:  $\chi^2 = 16.03$ ;  $p = 0.655$ ; Kruskal–Wallis test). **D**, Summary model of multimodal neural encoding of learning in NBM and dlPFC as aligned to trial events. NBM spike rates modulate during early learning, as does theta power, and both types of modulation diminish as learning proceeds. Theta, alpha, gamma, and HGP encode reward anticipation. Each of these LFP power bands shows significant correlation with trial accuracy even before the decision has been evaluated, suggesting that they encode the animal's confidence in its response. Both also continue to show that correlation after reward consumption.

Further studies and use of multichannel electrodes in multiple regions of the NBM, for instance, could be key for differentiating function in this structure, particularly to explain the counterintuitive anticorrelation between HGP LFP activity and the spiking activity that we found as the NHPs learned new associations (Fig. 9D).

Another important component to consider is that there are subsets of NBM neurons that release a diversity of neurotransmitters, including glutamate and GABA (Mesulam et al., 1983; Wenk, 1997; Semba, 2000; Liu et al., 2015). To overcome this heterogeneous distribution of neural types in the NBM, multichannel recording could sample more single-unit waveforms in the area and potentially identify separable populations of NBM neurons. A multichannel approach would also allow the examination of causal relationships between the NBM and cortex, particularly with LFP. With only single electrodes per region as in this study, we could not rule out volume conductance issues and thus could not report coherence or other synchrony/causality measures in this study (Bastos and Schoffelen, 2015).

Nevertheless, we found clear encoding of the learning state in theta-band modulation before the Go cue and not in other bands. In contrast, alpha, beta, and HGP correlated best with trial accuracy after the Go cue, reflecting reward anticipation. In particular, power in these frequency bands demonstrated significant correlations to trial accuracy across all block types. Our results show significant encoding of reward reinforcement and reward anticipation whether the animals had formed a new association (Novel blocks), were recalling a newly made association (Recall), were recalling a well formed association (Familiar), or were reversing a new association (Reversal). This could possibly indicate a more generalizable concept: that reward anticipation and reinforcement induce across-frequency changes in LFP power. In addition, these results point to an underlying mechanism in which different modalities of neural activity can encode and coordinate multiple aspects of learning, as shown in our proposed model (Fig. 9D). There is a growing body of evidence that different frequency bands serve to bind, coordinate, or otherwise enhance communication between different brain areas (Buzsáki and Watson, 2012; Watrous et al., 2015). For example, theta-band activity has long been associated with hippocampal memory function (Buzsáki and Watson, 2012) and recent evidence implicates this band in corticohippocampal coordination during learning (Benchenane et al., 2011; Brincat and Miller, 2015). More recently, gamma-power (30–50 Hz) and HGP (50–200 Hz) bands have been similarly proposed to support learning across species (Lee et al., 2014).

These separable encodings could guide the development of closed-loop stimulation to enhance learning and memory, potentially moving toward clinical therapeutics (Laxton and Lozano, 2013; Widge et al., 2017). Delivering stimulation during critical moments in memory formation has profound effects on learning (Williams and Eskandar, 2006; Katnani et al., 2016; Ezzayat et al., 2017). If NBM modulation signals the beginning of a novel learning process, then inducing that modulation at the right time might accelerate or induce novel learning by altering network activity. Deep brain stimulation of the NBM to treat Alzheimer's and Parkinson dementia has had variable results, suggesting that this is not a straightforward strategy (Kuhn et al., 2015; Mirzadeh et al., 2016). However, based on our observations, NBM stimulation could be targeted to specific phases of learning or NBM activity might identify time points when stimulation could be particularly effective. For instance, stimulation during the Go cue decision point (such as when the NBM single

units encode Novel learning) or during feedback (such as when the NBM LFP encodes reward anticipation) could differentially alter the underlying learning, or cognitive, state.

## References

- Asaad WF, Eskandar EN (2008) A flexible software tool for temporally-precise behavioral control in Matlab. *J Neurosci Methods* 174:245–258. [CrossRef Medline](#)
- Asaad WF, Eskandar EN (2011) Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *J Neurosci* 31:17772–17787. [CrossRef Medline](#)
- Bakin JS, Weinberger NM (1996) Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proc Natl Acad Sci U S A* 93:11219–11224. [CrossRef Medline](#)
- Bakker R, Tiesinga P, Kötter R (2015) The scalable brain atlas: instant web-based access to public brain atlases and related content. *Neuroinformatics* 13:353–366. [CrossRef Medline](#)
- Barefoot HC, Baker HF, Ridley RM (2002) Crossed unilateral lesions of temporal lobe structures and cholinergic cell bodies impair visual conditional and object discrimination learning in monkeys. *Eur J Neurosci* 15:507–516. [CrossRef Medline](#)
- Bastos AM, Schoffelen JM (2015) A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front Syst Neurosci* 9:175. [CrossRef Medline](#)
- Baxter MG, Chiba AA (1999) Cognitive functions of the basal forebrain. *Curr Opin Neurobiol* 9:178–183. [CrossRef Medline](#)
- Benchenane K, Tiesinga PH, Battaglia FP (2011) Oscillations in the prefrontal cortex: a gateway to memory and attention. *Curr Opin Neurobiol* 21:475–485. [CrossRef Medline](#)
- Brincat SL, Miller EK (2015) Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nat Neurosci* 18:576–581. [CrossRef Medline](#)
- Buzsáki G, Watson BO (2012) Brain rhythms and neural syntax: implications for efficient coding of cognitive content and neuropsychiatric disease. *Dialogues Clin Neurosci* 14:345–367. [Medline](#)
- Calabrese E, Badea A, Coe CL, Lubach GR, Shi Y, Styner MA, Johnson GA (2015) A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *Neuroimage* 117:408–416. [CrossRef Medline](#)
- DeLong MR (1971) Activity of pallidal neurons during movement. *J Neurophysiol* 34:414–427. [CrossRef Medline](#)
- Dubach MF, Bowden DM (2009) BrainInfo online 3D macaque brain atlas: a database in the shape of a brain. *Soc Neurosci Abstr* 35:199.5.
- Ezzayat Y, et al. (2017) Direct brain stimulation modulates encoding states and memory performance in humans. *Curr Biol* 27:1251–1258. [CrossRef Medline](#)
- Gibbs RB, Johnson DA (2007) Cholinergic lesions produce task-selective effects on delayed matching to position and configural association learning related to response pattern and strategy. *R. Neurobiol Learn Mem* 88:19–32. [CrossRef Medline](#)
- Gratwicke J, Kahan J, Zrinzo L, Hariz M, Limousin P, Foltynie T, Jahanshahi M (2013) The nucleus basalis of meynert: a new target for deep brain stimulation in dementia? *Neurosci Biobehav Rev* 37:2676–2688. [CrossRef Medline](#)
- Gratwicke J, Jahanshahi M, Foltynie T (2015) Parkinson's disease dementia: a neural networks perspective. *Brain* 138:1454–1476. [CrossRef Medline](#)
- Grothe MJ, Schuster C, Bauer F, Heinsen H, Prudlo J, Teipel SJ (2014) Atrophy of the cholinergic basal forebrain in dementia with lewy bodies and Alzheimer's disease dementia. *J Neurol* 261:1939–1948. [CrossRef Medline](#)
- Haque RU, Wittig JH Jr, Damera SR, Inati SK, Zaghoul KA (2015) Cortical low-frequency power and progressive phase synchrony precede successful memory encoding. *J Neurosci* 35:13577–13586. [CrossRef Medline](#)
- Ichihara-Takeda S, Funahashi S (2008) Activity of primate orbitofrontal and dorsolateral prefrontal neurons: effect of reward schedule on task-related activity. *J Cogn Neurosci* 20:563–579. [CrossRef Medline](#)
- Irle E, Markovitsch HJ (1986) Afferent connections of the substantia innominata/basal nucleus of meynert in carnivores and primates. *J Hirnforsch* 27:343–367. [Medline](#)
- Kahnt T, Heinze J, Park SQ, Haynes JD (2011) Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage* 56:709–715. [CrossRef Medline](#)
- Katnani HA, Patel SR, Kwon CS, Abdel-Aziz S, Gale JT, Eskandar EN (2016) Temporally coordinated deep brain stimulation in the dorsal and ventral striatum synergistically enhances associative learning. *Sci Rep* 6:18806. [CrossRef Medline](#)
- Kilgard MP, Merzenich MM (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279:1714–1718. [CrossRef Medline](#)
- Kilimann I, Grothe M, Heinsen H, Alho EJ, Grinberg L, Amaro E Jr, Dos Santos GA, da Silva RE, Mitchell AJ, Frisoni GB, Bokde AL, Fellgiebel A, Filippi M, Hampel H, Klöppel S, Teipel SJ (2014) Subregional basal forebrain atrophy in Alzheimer's disease: a multicenter study. *J Alzheimers Dis* 40:687–700. [Medline](#)
- Kuhn J et al. (2015) Deep brain stimulation of the nucleus basalis of meynert in Alzheimer's dementia. *Mol Psychiatry* 20:353–360. [CrossRef Medline](#)
- Laxton AW, Lozano AM (2013) Deep brain stimulation for the treatment of alzheimer disease and dementias. *World Neurosurg* 80:S28.e1–S28.e8. [Medline](#)
- Lee HS, Ghatti A, Pinto-Duarte A, Wang X, Dziejczapolski G, Galimi F, Huitron-Resendiz S, Piña-Crespo JC, Roberts AJ, Verma IM, Sejnowski TJ, Heinemann SF (2014) Astrocytes contribute to gamma oscillations and recognition memory. *Proc Natl Acad Sci U S A* 111:E3343–E3352. [CrossRef Medline](#)
- Liu AK, Chang RC, Pearce RK, Gentleman SM (2015) Nucleus basalis of meynert revisited: anatomy, history and differential involvement in Alzheimer's and Parkinson's disease. *Acta Neuropathol* 129:527–540. [CrossRef Medline](#)
- Masuda R, Fukuda M, Ono T, Endo S (1997) Neuronal responses at the sight of objects in monkey basal forebrain subregions during operant visual tasks. *Neurobiol Learn Mem* 67:181–196. [CrossRef Medline](#)
- Mesulam MM (2013) Cholinergic circuitry of the human nucleus basalis and its fate in Alzheimer's disease. *J Comp Neurol* 521:4124–4144. [CrossRef Medline](#)
- Mesulam MM, Mufson EJ (1984) Neural inputs into the nucleus basalis of the substantia innominata (Ch4) in the rhesus monkey. *Brain* 107:253–274. [CrossRef Medline](#)
- Mesulam MM, Mufson EJ, Levey AI, Wainer BH (1983) Cholinergic innervation of cortex by the basal forebrain: cytochemistry and cortical connections of the septal area, diagonal band nuclei, nucleus basalis (substantia innominata), and hypothalamus in the rhesus monkey. *J Comp Neurol* 214:170–197. [CrossRef Medline](#)
- Miasnikov AA, Chen JC, Gross N, Poytress BS, Weinberger NM (2008) Motivationally neutral stimulation of the nucleus basalis induces specific behavioral memory. *Neurobiol Learn Mem* 90:125–137. [CrossRef Medline](#)
- Miasnikov AA, Chen JC, Weinberger NM (2009) Behavioral memory induced by stimulation of the nucleus basalis: effects of contingency reversal. *Neurobiol Learn Mem* 91:298–309. [CrossRef Medline](#)
- Mirzadeh Z, Bari A, Lozano AM (2016) The rationale for deep brain stimulation in Alzheimer's disease. *J Neural Transm* 123:775–783. [CrossRef Medline](#)
- Ojemann GA, Ojemann J, Ramsey NF (2013) Relation between functional magnetic resonance imaging (fMRI) and single neuron, local field potential (LFP) and electrocorticography (ECoG) activity in human cortex. *Front Hum Neurosci* 7:34. [CrossRef Medline](#)
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869. [CrossRef Medline](#)
- Patel SR, Ghose K, Eskandar EN (2014) An open source 3-D printed modular micro-drive system for acute neurophysiology. *PLoS One* 9:e94262. [CrossRef Medline](#)
- Paxinos G, Huang XF, Toga AW (2000) The rhesus monkey brain in stereotaxic coordinates. San Diego: Academic.
- Poldrack RA, Packard MG (2003) Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41:245–251. [CrossRef Medline](#)
- Prerau MJ, Smith AC, Eden UT, Yanike M, Suzuki WA, Brown EN (2008) A mixed filter algorithm for cognitive state estimation from simultaneously recorded continuous and binary measures of performance. *Biol Cybern* 99:1–14. [CrossRef Medline](#)
- Prerau MJ, Smith AC, Eden UT, Kubota Y, Yanike M, Suzuki W, Graybiel AM, Brown EN (2009) Characterizing learning by simultaneous analysis of continuous and binary measures of performance. *J Neurophysiol* 102:3060–3072. [CrossRef Medline](#)
- Rabiei Z, Rafieian-Kopaei M, Heidarian E, Saghaei E, Mokhtari S (2014)



- Effects of zizyphus jujube extract on memory and learning impairment induced by bilateral electric lesions of the nucleus basalis of meynert in rat. *Neurochem Res* 39:353–360. [CrossRef Medline](#)
- Richardson RT, DeLong MR (1986) Nucleus basalis of meynert neuronal activity during a delayed response task in monkey. *Brain Res* 399:364–368. [CrossRef Medline](#)
- Richardson RT, DeLong MR (1990) Context-dependent responses of primate nucleus basalis neurons in a go/no-go task. *J Neurosci* 10:2528–2540. [Medline](#)
- Richardson RT, DeLong MR (1991) Electrophysiological studies of the functions of the nucleus basalis in primates. In: *The basal forebrain* (Napier TC, Kalivas PW, Hanin I, eds), pp 233–252. New York, NY: Springer.
- Ridley RM, Pugh P, Maclean CJ, Baker HF (1999) Severe learning impairment caused by combined immunotoxic lesion of the cholinergic projections to the cortex and hippocampus in monkeys. *Brain Res* 836:120–138. [CrossRef Medline](#)
- Rigdon GC, Pirsch JH (1986) Nucleus basalis involvement in conditioned neuronal responses in the rat frontal cortex. *J Neurosci* 6:2535–2542. [Medline](#)
- Rohlfing T, Kroenke CD, Sullivan EV, Dubach MF, Bowden DM, Grant KA, Pfefferbaum A (2012) The INIA19 template and NeuroMaps atlas for primate brain image parcellation and spatial normalization. *Front Neuroinform* 6:27. [CrossRef Medline](#)
- Semba K (2000) Multiple output pathways of the basal forebrain: organization, chemical heterogeneity, and roles in vigilance. *Behav Brain Res* 115:117–141. [CrossRef Medline](#)
- Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, Brown EN (2004) Dynamic analysis of learning in behavioral experiments. *J Neurosci* 24:447–461. [CrossRef Medline](#)
- Smith AC, Wirth S, Suzuki WA, Brown EN (2007) Bayesian analysis of interleaved learning and response bias in behavioral experiments. *J Neurophysiol* 97:2516–2524. [CrossRef Medline](#)
- Struble RG, Lehmann J, Mitchell SJ, McKinney M, Price DL, Coyle JT, DeLong MR (1986) Basal forebrain neurons provide major cholinergic innervation of primate neocortex. *Neurosci Lett* 66:215–220. [CrossRef Medline](#)
- Tsujimoto S, Sawaguchi T (2005) Neuronal activity representing temporal prediction of reward in the primate prefrontal cortex. *J Neurophysiol* 93:3687–3692. [CrossRef Medline](#)
- Voytko ML (1996) Cognitive functions of the basal forebrain cholinergic system in monkeys: memory or attention? *Behav Brain Res* 75:13–25. [CrossRef Medline](#)
- Voytko ML, Olton DS, Richardson RT, Gorman LK, Tobin JR, Price DL (1994) Basal forebrain lesions in monkeys disrupt attention but not learning and memory. *J Neurosci* 14:167–186. [Medline](#)
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081. [CrossRef Medline](#)
- Watrous AJ, Fell J, Ekstrom AD, Axmacher N (2015) More than spikes: common oscillatory mechanisms for content specific neural representations during perception and memory. *Curr Opin Neurobiol* 31:33–39. [CrossRef Medline](#)
- Weinberger NM (2003) The nucleus basalis and memory codes: auditory cortical plasticity and the induction of specific, associative behavioral memory. *Neurobiol Learn Mem* 80:268–284. [CrossRef Medline](#)
- Wenk GL (1997) The nucleus basalis magnocellularis cholinergic system: one hundred years of progress. *Neurobiol Learn Mem* 67:85–95. [CrossRef Medline](#)
- Widge AS, Ellard KK, Paulk AC, Basu I, Yousefi A, Zorowitz S, Gilmour A, Afzal A, Deckersbach T, Cash SS, Kramer MA, Eden UT, Dougherty DD, Eskandar EN (2017) Treating refractory mental illness with closed-loop brain stimulation: progress towards a patient-specific transdiagnostic approach. *Exp Neurol* 287:461–472. [CrossRef Medline](#)
- Williams ZM, Eskandar EN (2006) Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9:562–568. [CrossRef Medline](#)
- Williams ZM, Neimat JS, Cosgrove GR, Eskandar EN (2005) Timing and direction selectivity of subthalamic and pallidal neurons in patients with parkinson disease. *Exp Brain Res* 162:407–416. [CrossRef Medline](#)
- Wilson FA, Rolls ET (1990) Neuronal responses related to the novelty and familiarity of visual stimuli in the substantia innominata, diagonal band of broca and periventricular region of the primate basal forebrain. *Exp Brain Res* 80:104–120. [Medline](#)
- Yazdan-Shahmorad A, Kipke DR, Lehmkuhle MJ (2013) High gamma power in ECoG reflects cortical electrical stimulation effects on unit activity in layers V/VI. *J Neural Eng* 10:06002. [CrossRef Medline](#)