

Sensory and Motor Systems

# Reward Expectation Modulates Local Field Potentials, Spiking Activity and Spike-Field Coherence in the Primary Motor Cortex

Junmo An,<sup>1</sup> Taruna Yadav,<sup>1</sup> John P. Hessburg,<sup>2</sup> and Joseph T. Francis<sup>1</sup><https://doi.org/10.1523/ENEURO.0178-19.2019>

<sup>1</sup>Department of Biomedical Engineering, University of Houston, Houston, TX 77204 and <sup>2</sup>Department of Physiology and Pharmacology, Robert F Furchgott Center for Neural and Behavioral Science, State University of New York Downstate Medical Center, Brooklyn, NY 11203

## Abstract

Reward modulation (M1) could be exploited in developing an autonomously updating brain-computer interface (BCI) based on a reinforcement learning (RL) architecture. For an autonomously updating RL-based BCI system, we would need a reward prediction error, or a state-value representation from the user's neural activity, which the RL-BCI agent could use to update its BCI decoder. In order to understand the multifaceted effects of reward on M1 activity, we investigated how neural spiking, oscillatory activities and their functional interactions are modulated by conditioned stimuli related reward expectation. To do so, local field potentials (LFPs) and single/multi-unit activities were recorded simultaneously and bilaterally from M1 cortices while four non-human primates (NHPs) performed cued center-out reaching or grip force tasks either manually using their right arm/hand or observed passively. We found that reward expectation influenced the strength of  $\alpha$  (8–14 Hz) power,  $\alpha$ - $\gamma$  comodulation,  $\alpha$  spike-field coherence (SFC), and firing rates (FRs) in general in M1. Furthermore, we found that an increase in  $\alpha$ -band power was correlated with a decrease in neural spiking activity, that FRs were highest at the trough of the  $\alpha$ -band cycle and lowest at the peak of its cycle. These findings imply that  $\alpha$  oscillations modulated by reward expectation have an influence on spike FR and spike timing during both reaching and grasping tasks in M1. These LFP, spike, and spike-field interactions could be used to follow the M1 neural state in order to enhance BCI decoding (An et al., 2018; Zhao et al., 2018).

**Key words:**  $\alpha$  power; brain computer interface; mirror neurons; primary motor cortex; pulsed inhibition; reward

## Significance Statement

Knowing the subjective value of performed or observed actions is valuable feedback that could be used to improve the performance of an autonomously updating brain-computer interface (BCI). Reward-related information in the primary motor cortex (M1) may be crucial for more stable and robust BCI decoding (Zhao et al., 2018). Here, we present how expectation of reward during motor tasks, or simple observation, is represented by increased spike firing rates (FRs) in conjunction with decreased  $\alpha$  (8–14 Hz) oscillatory power,  $\alpha$ - $\gamma$  comodulation, and  $\alpha$  spike-field coherence (SFC), as compared to nonrewarding trials. Moreover, a phasic relation between  $\alpha$  oscillations and FRs was observed where FRs were found to be lowest and highest at the peak and trough of  $\alpha$  oscillations, respectively.

## Introduction

The brain is a highly adaptable learning machine, and at least partially learns via reinforcement learning (RL) mechanisms. As a main function of the brain is to move the individual through their environment in a manner that maximizes reward and minimizes punishment, we hypothesized that one would see neural dynamics reminiscent of various aspects of a RL machine (Tarigoppula et al., 2018), and we expected to see this at every level of the neural representation from mesoscopic/macroscale EEG (i.e., electroencephalography) down to single units (Marsh et al., 2015; McNeil et al., 2016; An et al., 2018), and conducted this work to explore this hypothesis. Another motivation for our work is that we hope to use reward related information from a user's brain, such as state value information, or reward prediction error, toward autonomously updating brain-computer interface (BCI) agents that use RL update rules. Thus, if we can better understand the neural signatures of reward expectation, we can better move toward such autonomously updating BCIs.

Recently, cross-frequency coupling (CFC) has been used to measure statistical correlations between different frequency bands of macroscopic (slow-frequency oscillations) and/or mesoscopic (high-frequency oscillations) local field potential (LFP) oscillations (Soltesz and Deschênes, 1993; Bragin et al., 1995). CFC has been considered an important measure of cognitive information processing in different brain regions (Canolty and Knight, 2010). Phase-to-phase coupling (Palva et al., 2005; Belluscio et al., 2012), phase-to-amplitude coupling (PAC; Mormann et al., 2005; Canolty et al., 2006; Tort et al., 2008; Colgin et al., 2009), and amplitude-to-amplitude coupling (Bruns et al., 2000) are different methods in which CFC can be analyzed (Jensen and Colgin, 2007; Canolty and Knight, 2010). In particular, PAC between the amplitude of high-frequency oscillations and the phase of low-frequency oscillations has been observed in cognitive tasks in both human (Canolty et al., 2006; Axmacher et al., 2010) and animal models (Buzsáki et al., 2003; Lakatos et al., 2005; Tort et al., 2008). In addition, the correlation between microscopic (spike trains) and macroscopic and/or mesoscopic (LFPs) scales has been considered to play an important functional role in neural processing

(Fries et al., 2001; Jarvis and Mitra, 2001). Several studies have employed spike-field coherence (SFC), the phase dependency between spikes and LFPs, to investigate the communication and synchronization within neuronal groups of the same or different cortical regions (Womelsdorf et al., 2006; Witham et al., 2007; Fries et al., 2008; Pesaran et al., 2008; Gregoriou et al., 2009; Jutras et al., 2009; Chalk et al., 2010).

For many years,  $\alpha$  (8–14 Hz) oscillations were thought to serve a non-functional purpose, where the power in this band reflected the opened or closed state of the eyes (Adrian and Matthews, 1934). However, several recent studies have shown the functional role of  $\alpha$  oscillations and their importance in cognitive processing. Specifically,  $\alpha$  oscillations reflect inhibitory activity in visual and auditory attention (Foxe et al., 1998; Thut et al., 2006; Rihs et al., 2009; Kerlin et al., 2010), perception (VanRullen and Koch, 2003), and working memory (Jensen et al., 2002; Sauseng et al., 2005; Jensen and Mazaheri, 2010) tasks. However, it is less clear how reward expectation affects these oscillations in the primary motor cortex (M1).

To better understand reward-related effects on M1, we conducted the present study where we simultaneously recorded neural spiking activity (single- and multi-units) and LFPs from contra/ipsilateral M1 in non-human primates (NHPs) while they performed cued trial value center-out reaching tasks and grip force tasks. We have previously shown that reward modulates single unit activity, population firing rates (FRs), and LFP power in M1 while NHPs either performed or observed a single target center-out reaching task (Marsh et al., 2015), work that has since been corroborated and extended by others (Ramkumar et al., 2016; Ramakrishnan et al., 2017). To further understand the impact of reward expectation on other aspects of M1 neural activity, we studied power spectral density (PSD), phase-to-amplitude comodulation, SFC, and phase correlation with spiking activity.

## Materials and Methods

### Surgery

A rhesus macaque [NHP P (female): *Macaca mulatta*] and three bonnet macaques [NHPs A (male), S (male), and Z (female): *Macaca radiata*] were implanted with 96-channel microelectrode Utah arrays (10 × 10 array consisting of 1.5 mm in length electrodes spaced 400  $\mu$ m, Blackrock Microsystems, LLC.) in the M1 region associated with their right hand and forearm. NHPs A, S, and P were implanted in the contralateral M1 with respect to the right arm, which all NHPs used to perform the manual tasks. NHP Z was previously implanted twice in the contralateral M1, therefore the array was implanted in ipsilateral M1 for this study. All studies and procedures were approved by the Institutional Animal Care and Use Committee at the State University of New York (SUNY) Downstate Medical Center and complied with the National Institutes of Health Guide for the Care and Use of Laboratory Animals guidelines. The surgical procedures used in the experiment were the same as those as described in our previous work (Chhatbar et al., 2010; Marsh et al., 2015). In brief, veterinary staffs from the SUNY Downstate

Received May 13, 2019; accepted May 30, 2019; First published June 6, 2019.

The authors declare no competing financial interests.

Author contributions: J.A., T.Y., and J.T.F. analyzed data; J.A., T.Y., and J.T.F. wrote the paper; J.P.H. and J.T.F. designed research; J.P.H. and J.T.F. performed research.

This work was supported by the National Institutes of Health Grant 1R01NS092894-01; the National Science Foundation Grant IIS-1527558; the Defense Advanced Research Projects Agency Grant N66001-10-C-2008; and New York State Spinal Cord Injury Research Board Grants C30600GG, C030838GG, and C32250GG.

Correspondence should be addressed to Joseph T. Francis at joey199us@gmail.com.

<https://doi.org/10.1523/ENEURO.0178-19.2019>

Copyright © 2019 An et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Division of Comparative Medicine performed general anesthesia and animal preparation. Aseptic conditions were maintained during the course of surgery. Anesthesia was induced with ketamine and maintained using isoflurane and fentanyl. To prevent inflammation, dexamethasone was used during the surgical procedure.

The first surgery for each animal was the implantation of a back post, a titanium post (Crist Instrument Co., Inc.) implanted onto the caudal aspect of the NHP's cranium to attach to the primate chair during training and recording. The post was placed on the caudal aspect of the skull, holes drilled and tapped (Synthes 2.0 mm drill bit, Synthes 2.0 mm tap) and affixed with 6- or 8-mm screws (Synthes, 2.7 mm in diameter, titanium) depending on the thickness of the NHP's skull.

For microelectrode implantation, the animal was prepared for surgery in the same way. A rostrally placed front post (titanium, Crist Instrument Co., Inc.) was affixed similarly to the back post, to serve as a platform for the electrode connectors. An  $\sim 2 \times 3$ -cm craniotomy window was then created using a dremel tool with conical tip (Dremel Multipro), over the cortical areas of interest. The dura mater was reflected, and the target locations were identified visually with cortical landmarks. To confirm the location of somatosensory cortex (S1), an electrode (Michigan probe, 4-shank 32 channel silicone array, NeuroNexus Technologies, Inc.) was lowered stereotactically into the cortex. A lab member then stimulated the animal by tapping the contralateral hand and arm, with the electrode output represented audibly through loudspeakers routed through a TDT recording system (Tucker Davis Technologies, Inc.). A 96-channel microelectrode array (Blackrock Utah array, platinum/iridium, 1.5 mm in length) was placed in the determined S1 location, and a pneumatic piston (pneumatic control box, Cyberkinetics Neurotechnology Systems, Inc.) used to secure the array into place in the cortex. The M1 implanted location was immediately across the central sulcus, and put into place in the same manner.

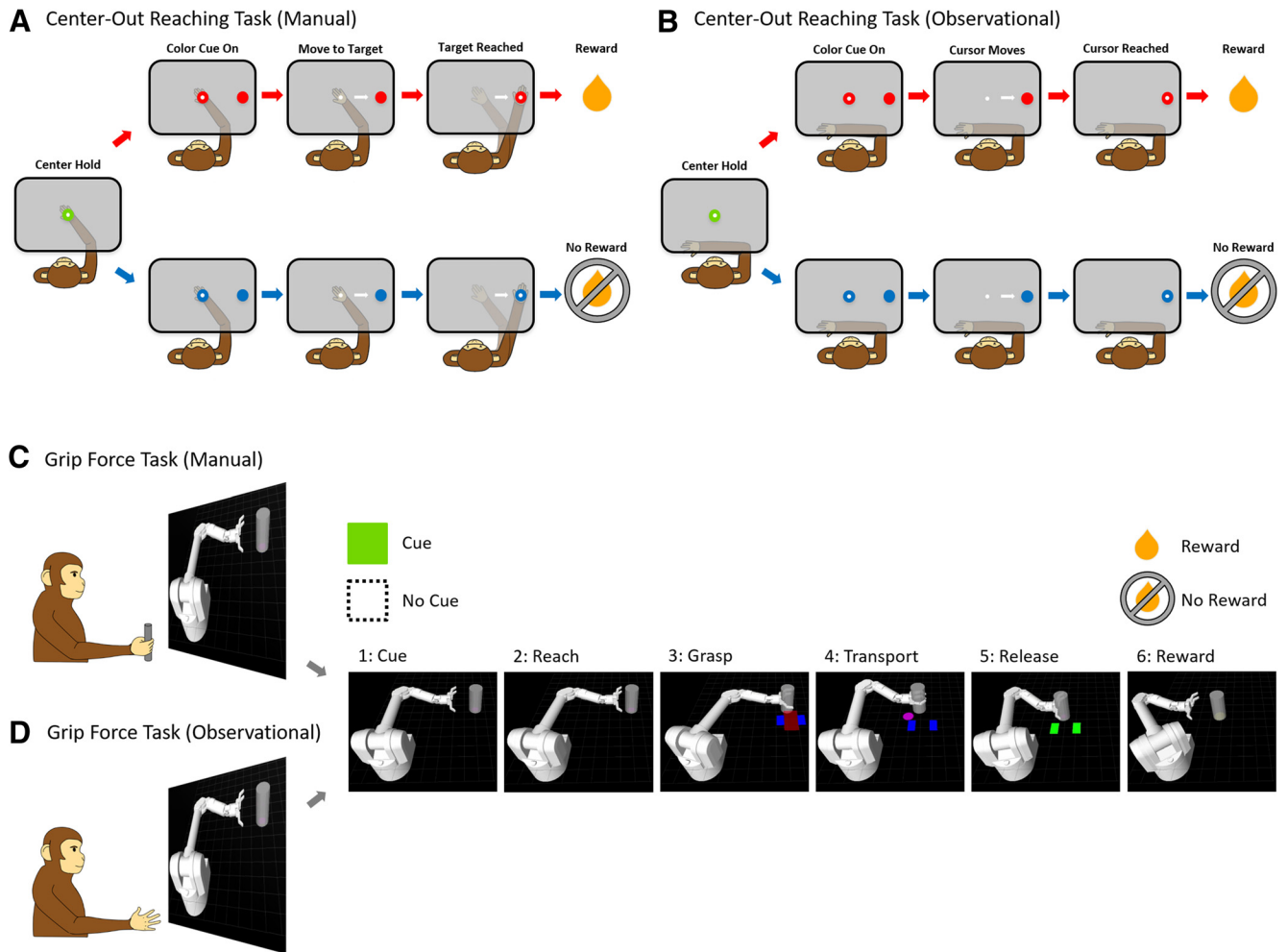
Electrode wires were gathered together and routed to one corner of the craniotomy, and then along the front post to the electrode connectors, which were placed into a plastic frame attached to the front post. Once all the connectors were affixed the dura was sutured back into place, and the bone fragment from the craniotomy window placed above. The bone was affixed with a titanium mesh and bone screws (1.9 mm in diameter, 4 mm in length titanium self-tapping screws, Bioplate Inc.) on the skull and on the bone fragment. Dental acrylic (Palacos, Zimmer Biomet) was used to attach the electrode wires to the skull, and additionally to create a protective layer between the craniotomy and the front post; so, the NHP could not interfere with the wires. NHPs were given a six-week rest period following back post implantation to allow for osseointegration before resuming training and recording, and two weeks following microelectrode implantation to allow for the site to heal and cortical inflammation to reduce.

### Cued center-out reaching task

Two NHPs (A and Z) were trained to perform a center-out delayed hold reaching task with their right arm using a two-link robotic exoskeleton (KINARM, BKIN Technologies Ltd.) for the manual task, as shown in [Figure 1A](#). They were also trained to observe a feedback cursor moving automatically with a constant speed toward the target without performing physical effort during the observational task, as shown in [Figure 1B](#). During the manual task, the NHPs would have to hold their hand on a central target for 325 ms, following a color cue period of 100–300 ms dependent on the NHPs temperament. During the color cue period the peripheral targets color and the hold target's color indicated the trials value that is rewarding or nonrewarding. The NHP had to wait another 325–400 ms until the go cue, which was the disappearance of the hold target. For a successful trial, the NHP had to reach and hold on the peripheral target for 325 ms. Every successful reach in rewarding trials resulted in a juice reward to the NHP, whereas reward was withheld on nonrewarding successful trials. If a nonrewarding trial was unsuccessful, it was repeated to encourage NHPs to make successful movements. During observational tasks, NHPs observed passively while a feedback cursor moved automatically from center to peripheral target at a constant speed ( $\sim 1$  cm/s). Visual color cues were similar to the manual task, and informed the NHP about the rewarding or nonrewarding trial value if successful. During manual and observational tasks, rewarding and nonrewarding trials were presented in a random order. There was one exception for NHP A during observational tasks where trials followed a set structure or rewarding followed by nonrewarding and repeating.

### Cued grip force task

Two NHPs (S and P) were trained to perform cued grip force tasks sitting comfortably in a primate-training chair (BKIN Technologies Ltd.) either manually ([Fig. 1C](#)) or observationally ([Figure 1D](#)). The task consisted of a virtual robotic arm, modeled as a Barrett WAM Arm and Hand (Barrett Technology) in Gazebo on ROS (Robot Operating System) that reached toward a virtual cylindrical object. When the robotic arm reached the object, the NHP was required to apply and maintain a visually indicated amount of force on a custom-made manual gripper (force transducer) which was composed of a metal bar, a cylindrical plastic frame, and a force transducer (FC2231-0000-0100-L, Measurement Specialties, Inc.) to measure grip force while the cylinder was transported to the target position by the WAM simulation. Once at the target position, the NHP would release the manual gripper, and the robotic hand would release the object, and the arm moved back to the neutral starting position. The target force was indicated with a pair of blue rectangles in the virtual environment, where the width of the rectangles represented the upper and lower bounds of the required grip force. The actual magnitude of force applied by the NHP was shown as a red rectangle that expanded as the force increased. The correct amount of grasping force was maintained by keeping the red rectangle within the



**Figure 1.** Center-out reaching tasks and grip force tasks. Schematic of (A) manual and (B) observational task during the cued center-out reaching tasks as well as for cued manual (C) and observational (D) grip force tasks.

upper and lower bounds of the blue rectangle. For a trial to be considered successful two conditions had to be met. The first was the application of appropriate force during the transport of the object to the target location, and the second condition was the release at the end of object’s transfer. Based on whether a trial was cued rewarding or nonrewarding, the NHP, respectively, received or did not receive a juice reward at the end of a successful trial. The visual color cue was shown at the beginning of each trial and remained visible throughout the trial. This trial progression was divided into six time-frames for analysis: (1) cue, when the cue was presented; (2) reach, when the virtual arm approached the target; (3) grasp, when the NHP applied force to the manual gripper so the virtual hand grasped the object; (4) transport, where the object automatically moved to the target location while the NHP maintained grip force; (5) release, when the gripper was released and the virtual hand released the object; and (6) reward or non-reward, when the NHP received a juice reward for successful completion of a rewarding trial, did not receive reward for completion of a nonrewarding trial, or did not receive reward for the unsuccessful completion of either trial type.

The structure of rewarding and nonrewarding trials was completely predictable in some recording blocks, where trials alternated between the two. The remainder of the blocks was partially predictable, where 50%, 75%, or 90% of the trials were rewarding, and the trial type was selected pseudorandomly with this bias. In the observational task, the mechanics of the task were the same, but the NHP had to passively observe the robot performing an automatic grasp and transport of the cylinder instead of manually applying force.

**Neural recording**

Spike trains and LFPs were recorded simultaneously from M1 cortices using a Multichannel Acquisition Processor recording system (MAP, Plexon Inc.) while the NHPs performed center-out reaching tasks and grip force tasks. The recorded neural signals were bandpass-filtered from 170 Hz to 8 kHz for spike trains and from 0.7 to 300 Hz for LFPs using a linear finite impulse response (FIR) filter, and sampled at 40 kHz for spikes and 2 kHz for LFPs using a MAP system. Offline spike sorting was performed using a custom-made sorting template in the commercial software, Offline Sorter (Plexon Inc.). We analyzed single/



multi-unit activity and LFPs recorded from the contralateral M1 area of NHPs A, S, and P as well as the ipsilateral M1 area of NHP Z.

**Analysis of PSD**

To probe whether reward expectation modulates the power of neural oscillations, PSD of the LFPs was estimated using the Welch periodogram method with 75% overlapping Hamming windows (Welch, 1967). LFPs recorded from 32 channels were preprocessed by first removing the line noise (60 Hz) from every LFP channel using a second-order Butterworth notch filter, following which each channel was z-scored. Then, LFP channels with signal-to-noise ratio (SNR) below 5 were excluded from further analysis. SNR was calculated as the ratio of peak-to-peak amplitude ( $A_{Pk-Pk}$ ) of the averaged LFP ( $LFP_{Avg}$ ) and twice the SD of the residual signal (given by the difference of LFP on  $i^{th}$  channel and  $LFP_{Avg}$ ) as in Equation 1, where  $i \in [1, 32]$  (Koralek et al., 2013)

$$SNR_i = \frac{A_{pk-pk}}{2 * SD(LFP_i - LFP_{Avg})} \tag{1}$$

This SNR calculation was performed separately for each NHP dataset. Based on SNR criterion, the number (average  $\pm$  SD) of LFP channels used for manual task datasets was  $27.5 \pm 1.0$ ,  $26.0 \pm 2.7$ ,  $30.7 \pm 0.6$ , and  $24.0 \pm 0.0$  in NHPs A, Z, S, and P, respectively. For observational task datasets,  $24.0 \pm 2.0$ ,  $24.3 \pm 3.5$ ,  $30.0 \pm 0.0$ , and  $27.0 \pm 5.8$  of LFP channels were used for NHPs A, Z, S, and P, respectively. Selected LFP channels were averaged and used for PSD estimation. For comparison between rewarding and nonrewarding trials, each PSD was then normalized by dividing the power at each frequency by the average of all power from 0.5 to 100 Hz (Colgin et al., 2009). The trial-averaged PSD in the  $\alpha$ -band (8–14 Hz) was employed to compare rewarding to nonrewarding trials while NHPs performed all tasks.

**Analysis of CFC**

To quantify the CFC in LFPs and test its relationship with reward expectation, the PAC method [Kullback–Leibler-based modulation index (MI)] was used as proposed by Tort et al. (2008). PAC analysis can be described briefly in the following manner (Tort et al., 2008). PAC appears when the amplitude of fast frequency oscillations is modulated by the phase of slow-frequency oscillations (Lakatos et al., 2005; Canolty et al., 2006; Jensen and Colgin, 2007; Tort et al., 2008). First, the averaged LFP was bandpass-filtered at a low-frequency band (8 Hz to 20 Hz) for phase and a high-frequency band (25–100 Hz) for amplitude. Second, the phase of the low-frequency band and amplitude of the high-frequency band were extracted from the above filtered LFP by applying Hilbert transform. Third, the strength of amplitude comodulation by phase was computed as the MI measure at each phase frequency and amplitude frequency pair. The MI was calculated in steps of 0.5 Hz for phase frequency and 5 Hz for amplitude frequency using the normalized entropy measure as shown in Equations 2–4 (for review, see Tort et al., 2008).

$$p_j = \frac{\langle A_{f_A} \rangle \phi_{f_p}(j)}{\sum_{j=1}^N \langle A_{f_A} \rangle \phi_{f_p}(j)} \tag{2}$$

$$H = - \sum_{j=1}^N p_j \log(p_j) \tag{3}$$

$$MI = \frac{H_{max} - H}{H_{max}} \tag{4}$$

Here,  $N$  is the total number of phase bins ( $N = 18$ ; each bin has  $20^\circ$  interval ranging from  $0^\circ$  to  $360^\circ$ ).  $\langle A_{f_A} \rangle \phi_{f_p}(j)$  indicates the mean of the instantaneous amplitude  $A_{f_A}$  value over the phase bin  $j$  of phase frequency  $\phi_{f_p}(j)$ .  $H_{max}$  denotes the maximum possible entropy value ( $H_{max} = \log N$ ).

Using this PAC method, we computed the strength of phase-to-amplitude comodulation in rewarding and non-rewarding trials for center-out reaching tasks and grip force tasks and compared them for statistically significant differences.

**Analysis of SFC**

To examine whether reward expectation modulates the coherence between spike trains and LFPs, we conducted SFC analysis. SFC is commonly used to compute phase synchronization between point process data (e.g., spike trains) and continuous data (e.g., LFPs). To rule out the possibility that spikes might have contaminated LFPs (Waldert et al., 2013), all LFPs from selected channels (higher SNR channels) were averaged, then the power spectra of the binned point processes (single unit activity, binned at 1 ms) and channel-averaged LFP were computed using the multitaper estimation method. For this purpose, we used a Chronux built-in function, *coherency*, with nine tapers and a time-bandwidth (TW) value equal to 5. The coherence between spikes and LFPs was computed using the formula as shown in Equation 5 (Fries et al., 2001; Jarvis and Mitra, 2001)

$$C_{SL} = abs\left(\frac{S_{SL}}{\sqrt{S_S * S_L}}\right) \tag{5}$$

Here,  $S_{SL}$  indicates the cross-spectrum between spikes and LFPs, and  $S_S$  and  $S_L$  indicate the autospectra of spikes and LFPs, respectively. An SFC value ( $C_{SL}$ ) of zero indicates the absence of phase synchronization between spikes and LFPs, and one indicates the perfect synchronization between them. The trial-averaged SFC was analyzed over the frequency band (0.5–100 Hz) and the trial-averaged SFC at the  $\alpha$  frequency band was used to compare rewarding trials to nonrewarding trials while NHPs performed the aforementioned tasks.

**Relation between  $\alpha$  cycle and neural spiking**

The interaction between the  $\alpha$ -band cycle and neural spiking activity was measured as follows (Haegens et al., 2011). First, we bandpass-filtered LFP oscillations from 8

to 14 Hz. Second, the phase of the  $\alpha$ -band LFP was obtained by applying the Hilbert transform. Third, we divided the  $\alpha$ -band cycle into six equally sized phase bins of 60° each. Fourth, each unit's FR was normalized across the six phase bins using the min-max scaling method, obtained by  $Normalized\ FR = [FR - \min(FR)] / [\max(FR) - \min(FR)]$ ; last, the average FR in each phase bin was computed over a population of units and represented with respect to the corresponding phase of  $\alpha$  oscillations.

### Statistical analysis

All statistical analyses were performed using MATLAB R2017b (MathWorks Inc.). The nonparametric two-sided Wilcoxon signed-rank (paired) test [significance level ( $p$ ) = 0.05] was used to evaluate significant differences of PSD, PAC, SFC, and spike FRs between rewarding and nonrewarding trials for all tasks. To test the statistical significance of changes in FR with respect to the  $\alpha$ -band phase bins, we performed one-way ANOVA with Bonferroni *post hoc* multiple comparison (significance level:  $\alpha$  = 0.05) for rewarding and nonrewarding trials individually. The  $F$ -statistic value (from one-way ANOVA) and  $p$  value computed between the phase bins of low and high FRs (from two-sample  $t$  test) were observed for different tasks and trial types. Our analysis described in this paper focused on the  $\alpha$ -band as this was the only band to show significant differences for all NHPs and all tasks using the two-sided Wilcoxon signed-rank test.

## Results

To investigate how cued reward expectation influences neural activity in M1 cortex, we performed PSD and PAC analyses with LFPs, computed spike FR, and quantified SFC on neural spikes and LFP data obtained from contralateral (NHPs A, S, and P) and ipsilateral (NHP Z) M1 across center-out reaching tasks and grip force tasks. In addition, we investigated how  $\alpha$ -band oscillations correlate with spike FR. As mentioned in the methods, task structure could either be completely predictable or random with a bias. For NHP A's observational task, rewarding and nonrewarding trials were presented as a sequence alternating between the two while all the other tasks for all NHPs were randomly ordered with a bias which varied based on the percentage of reward bias.

### Reward expectation modulates $\alpha$ power

For PSD analysis, we used a postcue-onset period of 800 ms for all tasks. Figure 2 displays the normalized PSD plots (left column in each subplot) and bar plots for the  $\alpha$ -band (right column in each subplot). Shown are significant differences of  $\alpha$  (8–14 Hz binned at 0.5 Hz) LFPs for rewarding (red) and nonrewarding (blue) trials across manual (upper row in each subplot) and observational (lower row in each subplot) tasks during both center-out reaching tasks (NHPs A and Z as shown in Fig. 2A,B) and grip force tasks (NHPs S and P as shown in Fig. 2C,D).

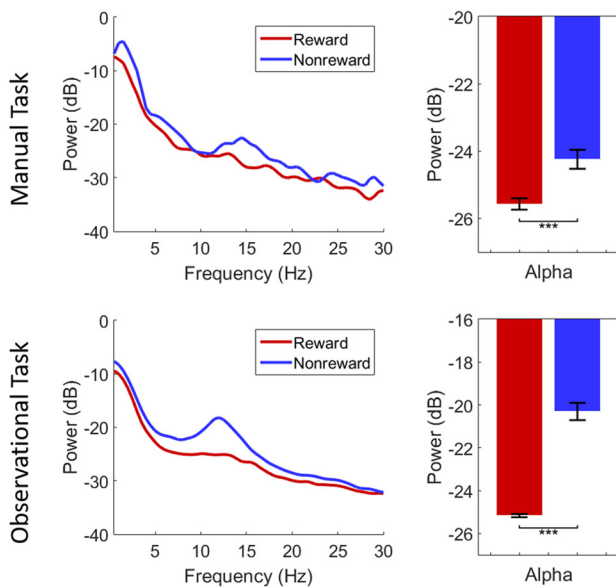
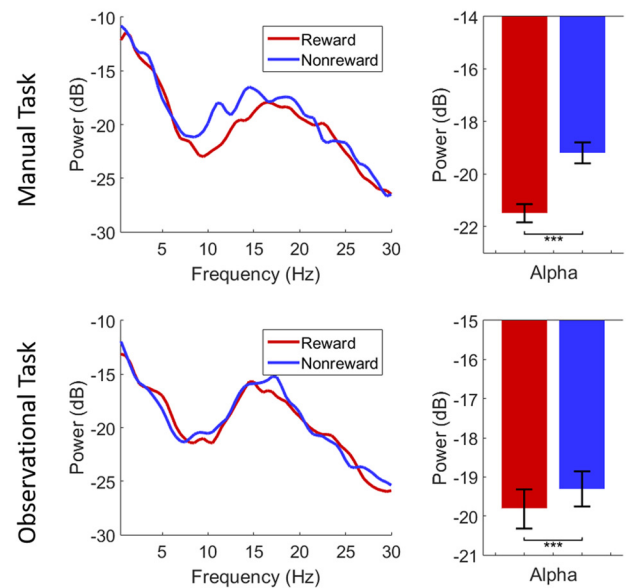
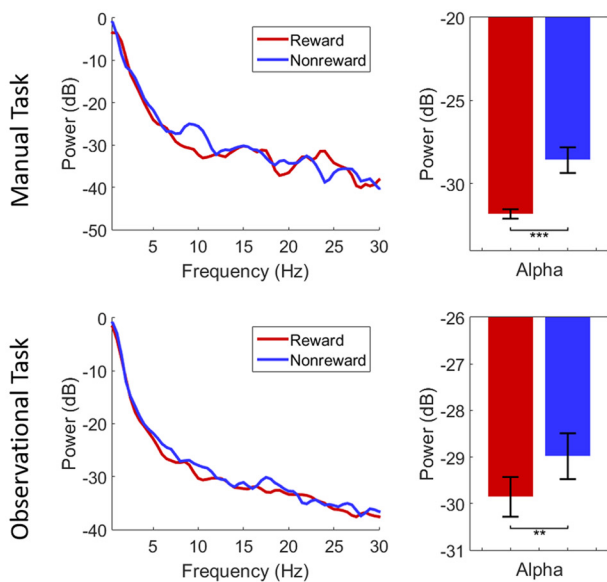
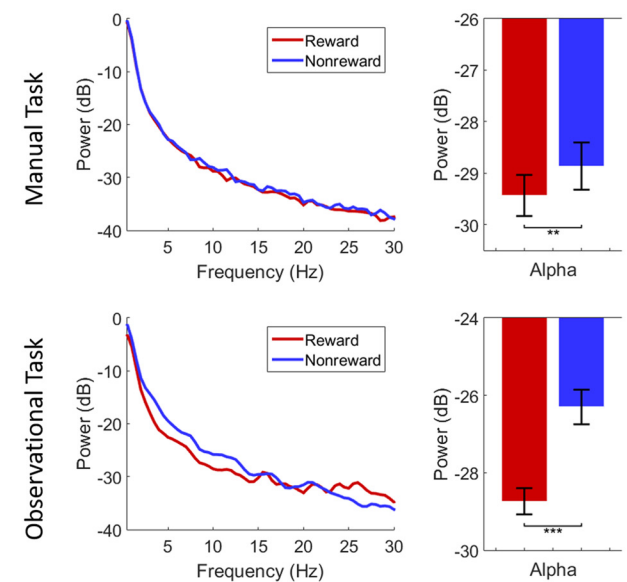
During the manual reaching task from NHP A's contralateral M1 (Fig. 2A, upper row), the PSD in the  $\alpha$ -band was significantly higher during nonrewarding trials than rewarding trials (Wilcoxon signed-rank test,  $p < 0.001$ ,  $z$ -statistic = 3.11,  $n = 13$ ). Furthermore, for the observa-

tional task (Fig. 2A, lower row), we found similar patterns of neural activation. Similar to the manual task,  $\alpha$ -band PSD was significantly higher during nonrewarding than rewarding trials for the observational task (Wilcoxon signed-rank test,  $p < 0.001$ ,  $z$ -statistic = 3.18,  $n = 13$ ). During the manual-reaching task from NHP Z's ipsilateral M1 (Fig. 2B, upper row), again the  $\alpha$  power was significantly higher during nonrewarding than rewarding trials (Wilcoxon signed-rank test,  $p < 0.001$ ,  $z$ -statistic = 3.18,  $n = 13$ ). During the observational task (Fig. 2B, lower row),  $\alpha$  PSD showed a significant increase during nonrewarding trials as well (Wilcoxon signed-rank test,  $p < 0.001$ ,  $z$ -statistic = 3.18,  $n = 13$ ).

Data from the manual grip force task from NHP S's contralateral M1 (Fig. 2C, upper row) showed that  $\alpha$  PSD was significantly higher for nonrewarding than rewarding trials (Wilcoxon signed-rank test,  $p < 0.001$ ,  $z$ -statistic = 3.11,  $n = 13$ ). Similarly, for the observational task (Fig. 2C, lower row),  $\alpha$  PSD was significantly higher for nonrewarding than rewarding trials (Wilcoxon signed-rank test,  $p < 0.01$ ,  $z$ -statistic = 2.76,  $n = 13$ ). During both manual (Fig. 2D, upper row) and observational (Fig. 2D, lower row) grip force tasks from NHP P's contralateral M1,  $\alpha$  PSD was significantly higher for nonrewarding than rewarding trials (Wilcoxon signed-rank test,  $p < 0.01$ ,  $z$ -statistic = 2.55,  $n = 13$  for manual task and  $p < 0.001$ ,  $z$ -statistic = 3.18,  $n = 13$  for observational task). The overall PSD results in Figure 2 indicate that the  $\alpha$  power of LFP oscillations were modulated by cued reward expectation. It also indicates that  $\alpha$  power was significantly increased for nonrewarding trials as compared to rewarding trials for both manual and observational variations of arm reaching tasks and hand grasping tasks.

### Reward expectation modulates $\alpha$ - $\gamma$ comodulation

To determine whether the phase of low-frequency oscillations was related to the amplitude of high-frequency oscillations in M1 during our reward cued experiments (Fig. 1), we computed phase-to-amplitude comodulation during a postcue-onset period of 800 ms both for rewarding and nonrewarding trials across all tasks. Figure 3 displays phase-to-amplitude comodulogram plots for rewarding trials (left column in each subplot) and nonrewarding trials (middle column in each subplot) across manual (upper row in each subplot) and observational (lower row in each subplot) tasks for contralateral (Fig. 3A,C,D) and ipsilateral (Fig. 3B) M1 cortices. In addition, the bar plot distributions (right column in each subplot) at phase frequencies of 8–14 Hz (binned at 0.5 Hz) and at amplitude frequencies of 30–100 Hz (binned at 5 Hz) in comodulograms showed significant differences of  $\alpha$ - $\gamma$  MI values for rewarding (red) and nonrewarding (blue) trials. During the manual reaching task from NHP A's contralateral M1 (Fig. 3A, upper row),  $\alpha$ - $\gamma$  PAC was significantly greater for nonrewarding trials than rewarding trials (Wilcoxon signed-rank test,  $p < 0.0001$ ,  $z$ -statistic = 11.12,  $n = 195$ ) when the phase is in the  $\alpha$  frequency at ~10 Hz. For the observational task (Fig. 3A, lower row),  $\alpha$ - $\gamma$  PAC was also significantly higher during nonrewarding trials than during rewarding trials (Wilcoxon signed-rank test,  $p$

**A** Monkey A Contralateral M1**B** Monkey Z Ipsilateral M1**C** Monkey S Contralateral M1**D** Monkey P Contralateral M1

**Figure 2.** PSD plots (left column in each subplot) of LFPs for rewarding (red) and nonrewarding (blue) trials across manual (upper row in each subplot) and observational (lower row in each subplot) tasks during center-out reaching (**A**, **B**) tasks and grip force (**C**, **D**) tasks. Bar graphs show significant differences of  $\alpha$  (8–14 Hz, right column in each subplot) band (\*\*\* $p$  < 0.001, \*\* $p$  < 0.01; Wilcoxon signed-rank test). Error bars in the bar plots represent SEM.

< 0.0001,  $z$ -statistic = 11.30,  $n$  = 195) at  $\alpha$  phase within the frequency band of 10–14 Hz. Similar trends were observed for NHP Z's ipsilateral M1 (Fig. 3B), where  $\alpha$ - $\gamma$  comodulation at phase frequencies of 8–10 Hz for manual task and at phase frequency of  $\sim$ 10 Hz for observational task during nonrewarding trials were significantly greater than rewarding trials (Wilcoxon signed-rank test,  $p$  < 0.0001,  $z$ -statistic = 10.03,  $n$  = 195 for manual task and

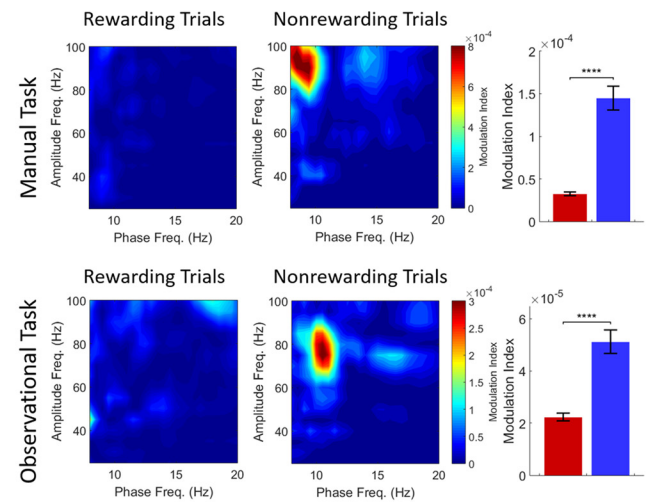
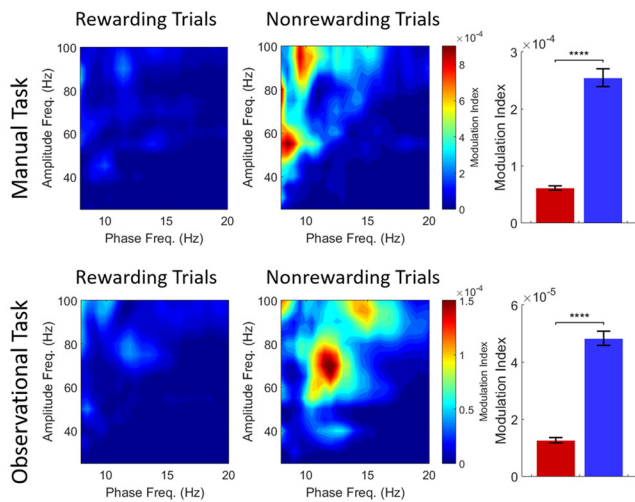
$p$  < 0.0001,  $z$ -statistic = 5.84,  $n$  = 195 for observational task).

The results for the center-out reaching tasks indicate that the strength of  $\alpha$ - $\gamma$  frequency PAC analyzed with the averaged LFP oscillations recorded from contralateral and ipsilateral M1 were modulated by reward expectation in both the presence (for manual) and absence (for observational) of arm reaching movements.



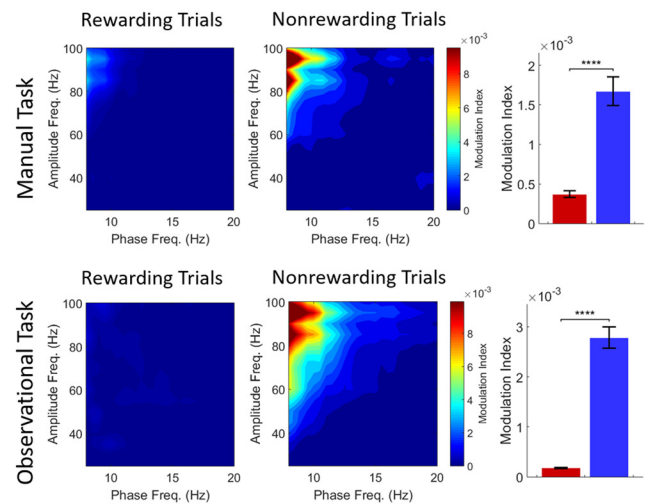
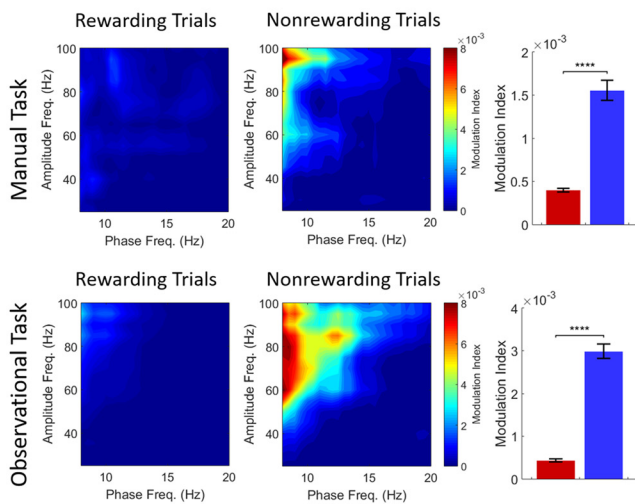
**A** Monkey A Contralateral M1

**B** Monkey Z Ipsilateral M1



**C** Monkey S Contralateral M1

**D** Monkey P Contralateral M1



**Figure 3.** Comodulograms showing the MI for rewarding (left column in each subplot) and nonrewarding trials (middle column in each subplot) across manual (upper row in each subplot) and observational tasks (lower row in each subplot) for contralateral (**A**, **C**, **D**) and ipsilateral (**B**) M1 cortices. Bar graphs (right column in each subplot) show significant differences of  $\alpha$ - $\gamma$  MI values for rewarding (red) and nonrewarding (blue) trials (\*\*\*\* $p < 0.0001$ ; Wilcoxon signed-rank test). Error bars in the bar plots represent SEM. **A**, **B** are for the center-out reaching tasks, and **C**, **D** are for the grip force tasks.

Qualitatively similar PAC results were observed in the grip force tasks. In **Figure 3C**,  $\alpha$ - $\gamma$  comodulations for both the manual (upper row) and observational (lower row) grip force tasks from NHP S's contralateral M1 were significantly higher during nonrewarding trials than rewarding trials (Wilcoxon signed-rank test,  $p < 0.0001$ ,  $z$ -statistic = 12.08,  $n = 195$  for manual task and  $p < 0.0001$ ,  $z$ -statistic = 12.11,  $n = 195$  for observational task). In **Figure 3D**, the PAC results of NHP P also showed significantly higher  $\alpha$ - $\gamma$  comodulation for nonrewarding trials than rewarding trials for both the manual (upper row) and observational (lower row) tasks (Wilcoxon signed-rank test,  $p < 0.0001$ ,  $z$ -statistic = 12.11,  $n = 195$  for manual task and  $p < 0.0001$ ,  $z$ -statistic = 12.11,  $n = 195$  for observational task). Similar to PAC results for the center-out

reaching tasks, our results demonstrate that the strength of  $\alpha$ - $\gamma$  comodulation was influenced by reward expectation during both manual and observational grip force tasks.

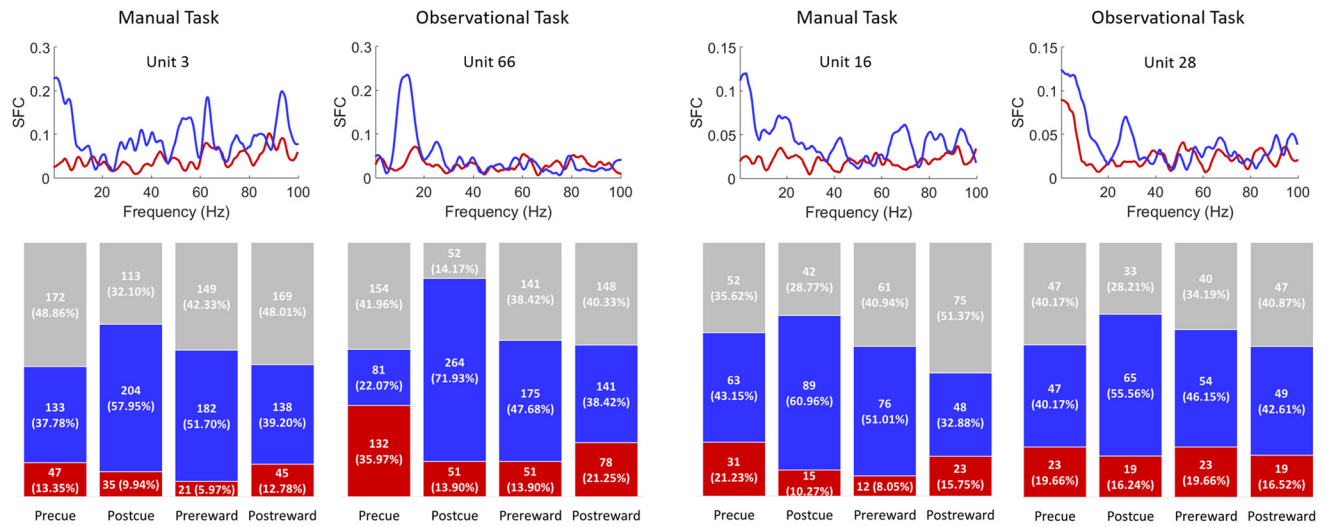
**Reward expectation modulates SFC**

To investigate whether  $\alpha$  oscillations are related to phase synchronization between spikes and LFPs, reward-related changes in the  $\alpha$ -band (8–14 Hz) SFC were estimated during the center-out reaching tasks and grip force tasks. **Figure 4** shows SFC plots (upper row in each subplot) for sample units (with specific unit number) during a postcue (after cue) period of 800 ms for rewarding (red) and nonrewarding (blue) trials for contralateral (NHPs A, S, and P) and ipsilateral (NHP Z) M1 cortices across all



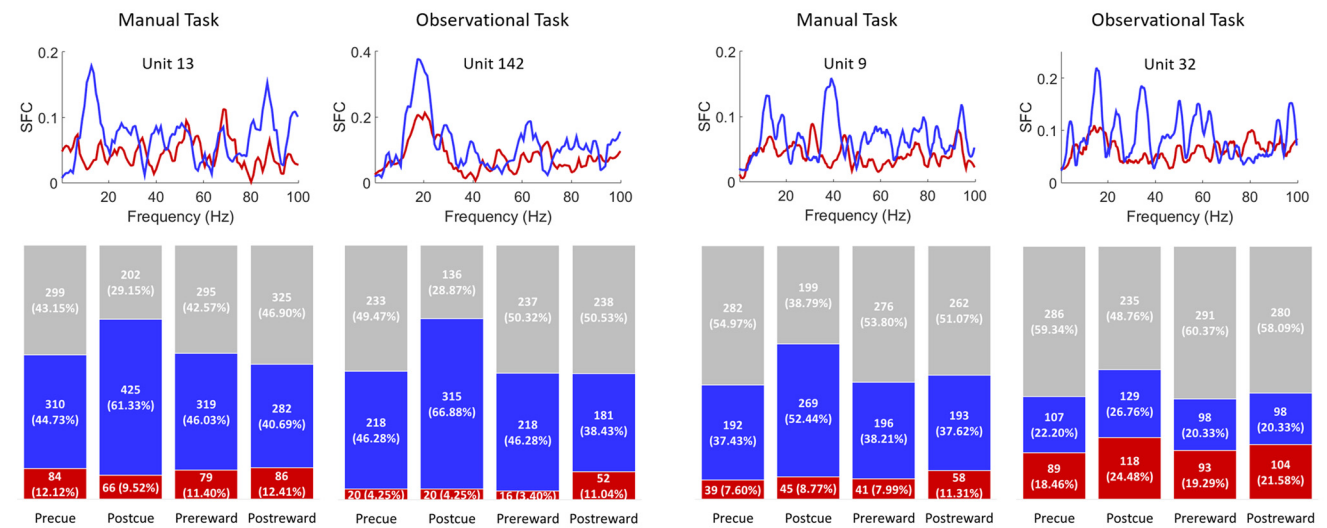
**A** Monkey A Contralateral M1

**B** Monkey Z Ipsilateral M1



**C** Monkey S Contralateral M1

**D** Monkey P Contralateral M1



● Higher SFC for Reward ● Higher SFC for Nonreward ● Neither

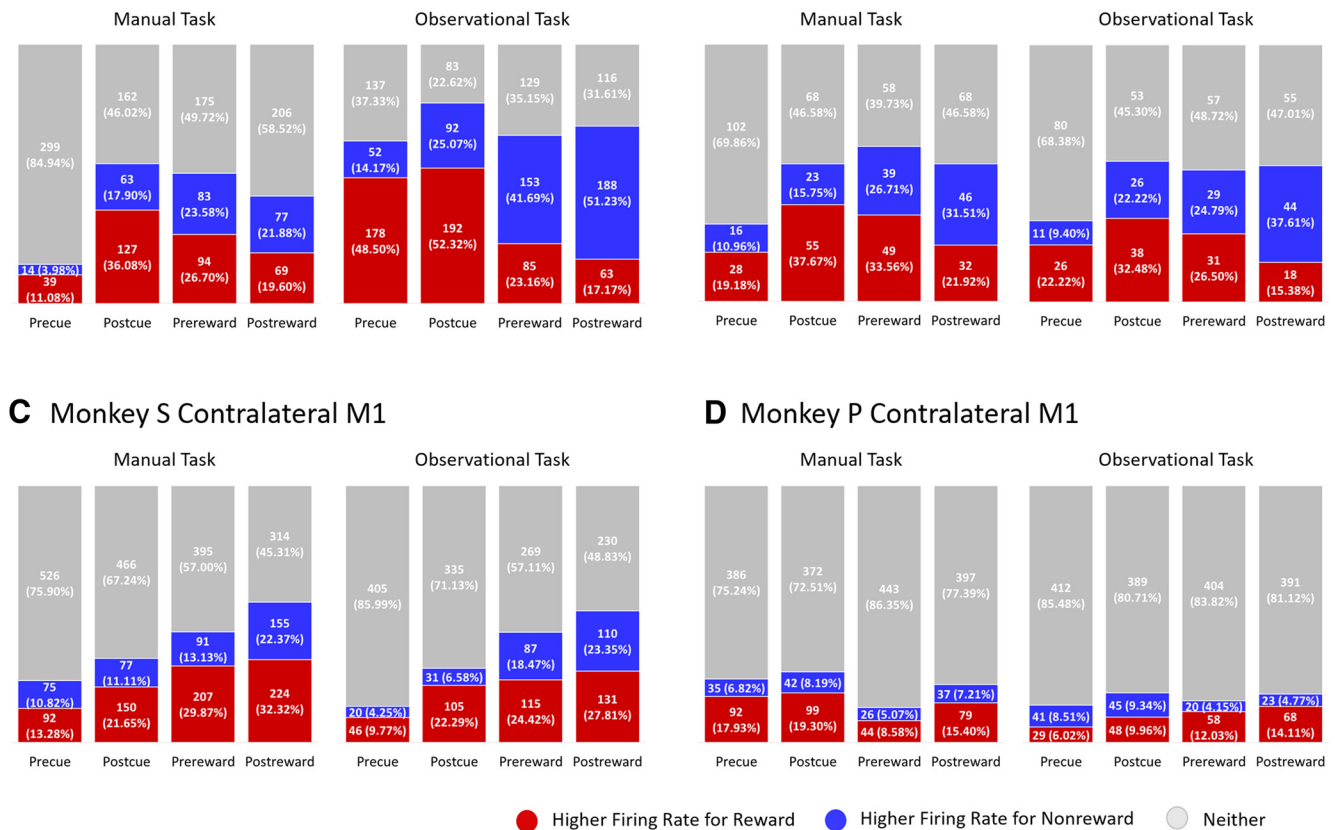
**Figure 4.** SFC plots (upper row in each subplot) for sample units for rewarding (red) and nonrewarding (blue) trials across manual (left column in each subplot) and observational (right column in each subplot) tasks for contralateral (**A, C, D**) and ipsilateral (**B**) M1 cortices. Bar charts (lower row in each subplot) represent the population of significantly different M1 units for SFC values in  $\alpha$ -band (8–14 Hz) during rewarding (red) and nonrewarding (blue) trials, and those with no significant difference (gray; Wilcoxon signed-rank test,  $p < 0.05$ ). Each column in the bar chart represents precue (before cue, 500 ms), postcue (after cue, 800 ms), prereward (before reward, 500 ms), and postreward (after reward, 500 ms) periods. **A, B** are for the center-out reaching tasks, and **C, D** are for the grip force tasks.

tasks. Additionally, **Figure 4** displays the population (lower row in each subplot) of significantly different units for SFC values in the  $\alpha$ -band during rewarding (red) and nonrewarding (blue) trials, or neither (gray; Wilcoxon signed-rank test,  $p < 0.05$ ). From left to right in the bar charts (lower row in each subplot), each column represents the precue (before cue, 500 ms), postcue (after cue, 800 ms), prereward (before reward, 500 ms), and postreward (after reward, 500 ms) time windows. In NHP A's contralateral M1 during manual center-out reaching

tasks (**Fig. 4A**, left column), 37.8% (133 of 352 for precue), 58.0% (204 of 352 for postcue), 51.7% (182 of 352 for prereward), and 39.2% (138 of 352 for postreward) of M1 units had significantly higher trial-averaged SFC during nonrewarding trials than rewarding trials. For observational center-out reaching tasks (**Fig. 4A**, right column), the percentages are 22.0% (81 of 367 for precue), 71.9% (264 of 367 for postcue), 47.7% (175 of 367 for prereward), and 38.4% (141 of 367 for postreward) of M1 units. Similar results were seen in NHP Z's ipsilat-

**A** Monkey A Contralateral M1

**B** Monkey Z Ipsilateral M1



**Figure 5.** Total population of M1 units that had significantly higher FRs for rewarding (red), nonrewarding (blue) trials, or neither (gray) across manual (left column) and observational (right column) tasks for contralateral (**A**, **C**, **D**) and ipsilateral (**B**) M1 cortices (Wilcoxon signed-rank test,  $p < 0.05$ ). Each column in the bar chart represents precue (500 ms), postcue (800 ms), prereward (500 ms), and postreward (500 ms) periods. **A**, **B** are for the center-out reaching tasks, and **C**, **D** are for the grip force tasks.

eral M1 during the center-out reaching tasks. The detailed breakdown of the results can be seen in Figure 4B, left column for manual task and right column for observational task.

As with NHPs A and Z, NHP S also showed significantly higher SFC percentages in nonrewarding trials for both manual and observational grip force tasks shown in Figure 4C. This pattern remains consistent in NHP P for manual tasks (Fig. 4D, left column), but with less significance for observational tasks (Fig. 4D, right column). Furthermore, the postreward (after reward delivery) period in observational tasks for NHP P showed a slightly higher percentage of SFC for rewarding than nonrewarding trials. Overall, the results in Figure 4 indicate that M1 units are significantly modulated by reward expectation for  $\alpha$ -band SFC during all reaching tasks and grasping tasks.

**Reward expectation modulates spike FR**

To examine the reward expectation modulation of neural spiking in M1, we computed the FR of M1 units during the center-out reaching tasks and grip force tasks. For each M1 unit, FR was computed using 50 ms bins during the following periods, precue (500 ms), postcue (800 ms), prereward (500 ms), and postreward (500 ms). Figure 5 displays the total percentage of M1 units that have sig-

nificantly higher average FRs for rewarding (red) trials and nonrewarding (blue) trials (Wilcoxon signed-rank test,  $p < 0.05$ ). In NHP A's contralateral M1 (Fig. 5A, left column), 11.1% (39 of 352 for precue), 36.1% (127 of 352 for postcue), 26.7% (94 of 352 for prereward) of M1 units for the manual task had significantly higher FRs during rewarding trials than nonrewarding trials; whereas, 19.6% (69 of 352 for postreward) of M1 units had significantly higher FRs for nonrewarding trials than rewarding trials. For the observational task (Fig. 5A, right column), 48.5% for precue and 52.3% for postcue of M1 units had significantly higher FRs for rewarding trials than nonrewarding trials; whereas, 41.7% for prereward and 51.2% for postreward of M1 units had significantly higher FRs for nonrewarding trials as compared to rewarding trials. The observation task performed by NHP A was a fully predictable sequence of rewarding trials followed by nonrewarding trials, which explains the precue activity pattern (Tarigoppula et al., 2018). In NHP Z's ipsilateral M1, during the manual task (Fig. 5B, left column), 19.2% for precue, 37.7% for postcue, and 33.6% for prereward had significantly higher spiking rates for rewarding trials than nonrewarding trials, but 31.5% for postreward period had higher FRs for nonrewarding trials than rewarding trials. Similar to the manual task, NHP Z showed a similar trend

of population FRs as for the observational task. The detailed population FRs are shown in [Figure 5B](#).

During grip force tasks, in NHP S's contralateral M1 results ([Fig. 5C](#)), rewarding (red) trials as a whole always showed significantly higher FRs than nonrewarding (blue) trials for both the manual task (left column) and the observational task (right column). In NHP P's contralateral M1 ([Fig. 5D](#)), during the manual task (left column), FRs for all periods were significantly higher for rewarding (red) trials than nonrewarding (blue) trials (Wilcoxon signed-rank,  $p < 0.05$ ). For the observational task (right column) from NHP P's contralateral M1 units, all periods except the precue period had higher FRs for rewarding trials than nonrewarding trials. The number of significant M1 units of NHP P's neural spiking rate during observational tasks was lower compared with significance of manual tasks. Overall, the results show that neural rates are significantly higher in M1 when reward is expected.

### **$\alpha$ Phase of LFP relates to neural spiking activity**

Previous work has shown a significant relationship between  $\alpha$  power and neural FR such that FRs were highest at the trough and lowest at the peak of the  $\alpha$  cycle ([Haegens et al., 2011](#)). Here we show some support for this observation, and demonstrate that it is strongest during the nonrewarding trials. [Figure 6](#) shows the relationship between the  $\alpha$ -band cycle and spike FR for rewarding (left column in each subplot) and nonrewarding (right column in each subplot) trials across manual (upper row in each subplot) and observational tasks (lower row in each subplot) for contralateral ([Fig. 6A,C,D](#)) and ipsilateral ([Fig. 6B](#)) M1 cortices. In the manual task (top row) for NHP A's contralateral M1 ([Fig. 6A](#)), the FR is high around the trough ( $3\pi/2$ ) of the  $\alpha$ -band cycle, and low around the peak ( $\pi/2$ ) of the cycle for nonrewarding trials (right column;  $F_{(5,660)} = 3.95$ ,  $p < 0.05$ ), whereas this was not observed for rewarding trials (left column;  $F_{(5,744)} = 1.36$ ,  $p > 0.05$ ). Similarly, for the observational task ([Fig. 6A](#), bottom row), FR was high around the trough of the cycle, whereas it was low around the peak for nonrewarding trials (right column;  $F_{(5,2040)} = 25.79$ ,  $p < 0.0001$ ). We did not see the same pattern for rewarding trials (left column;  $F_{(5,2040)} = 3.7$ ,  $p < 0.01$ ). In NHP Z's ipsilateral M1, during the manual task ([Fig. 6B](#), top row), the FR was highest around the trough of the  $\alpha$  cycle for nonrewarding trials (right column;  $F_{(5,534)} = 3.35$ ,  $p < 0.01$ ), but it was highest at the peak and lowest at the trough for rewarding trials (left column;  $F_{(5,564)} = 1.17$ ,  $p > 0.05$ ). During the observational task ([Fig. 6B](#), bottom row), highest FRs were associated with the trough of the  $\alpha$  cycle for nonrewarding trials (right column;  $F_{(5,276)} = 3.30$ ,  $p < 0.01$ ) but not for rewarding trials (left column;  $F_{(5,294)} = 1.63$ ,  $p > 0.05$ ).

We found similar results in the grip force tasks as seen in center-out reaching tasks. For the manual task (upper row), NHP S's contralateral M1 ([Fig. 6C](#)) showed that the FR was high around the trough, and low around the peak of the cycle for nonrewarding trials (right column;  $F_{(5,528)} = 2.60$ ,  $p < 0.05$ ) whereas it was lowest around the peak of the  $\alpha$  cycle for rewarding trials (left column;  $F_{(5,678)} = 2.43$ ,  $p < 0.05$ ). For the observational task ([Fig. 6C](#), bottom

row), FR was high at the trough of  $\alpha$  cycle, and it was low around the peak for nonrewarding trials (right column;  $F_{(5,1890)} = 10.68$ ,  $p < 0.0001$ ). The FR was lowest at the peak for rewarding trials (left column;  $F_{(5,2016)} = 5.95$ ,  $p < 0.0001$ ). For NHP P's contralateral M1 ([Fig. 6D](#)), manual task results (upper row) showed that FR was highest around the trough but not lowest around the peak for nonrewarding trials (right column;  $F_{(5,1050)} = 3.80$ ,  $p < 0.01$ ). The FR does not have the same trend for rewarding trials (left column;  $F_{(5,1020)} = 5.04$ ,  $p < 0.001$ ). For the observational task (bottom row), the highest and lowest FRs both for rewarding ( $F_{(5,1350)} = 1.72$ ,  $p > 0.05$ ) and nonrewarding trials ( $F_{(5,1344)} = 3.52$ ,  $p < 0.01$ ) were not associated with the peak and trough of the  $\alpha$  cycle.

## **Discussion**

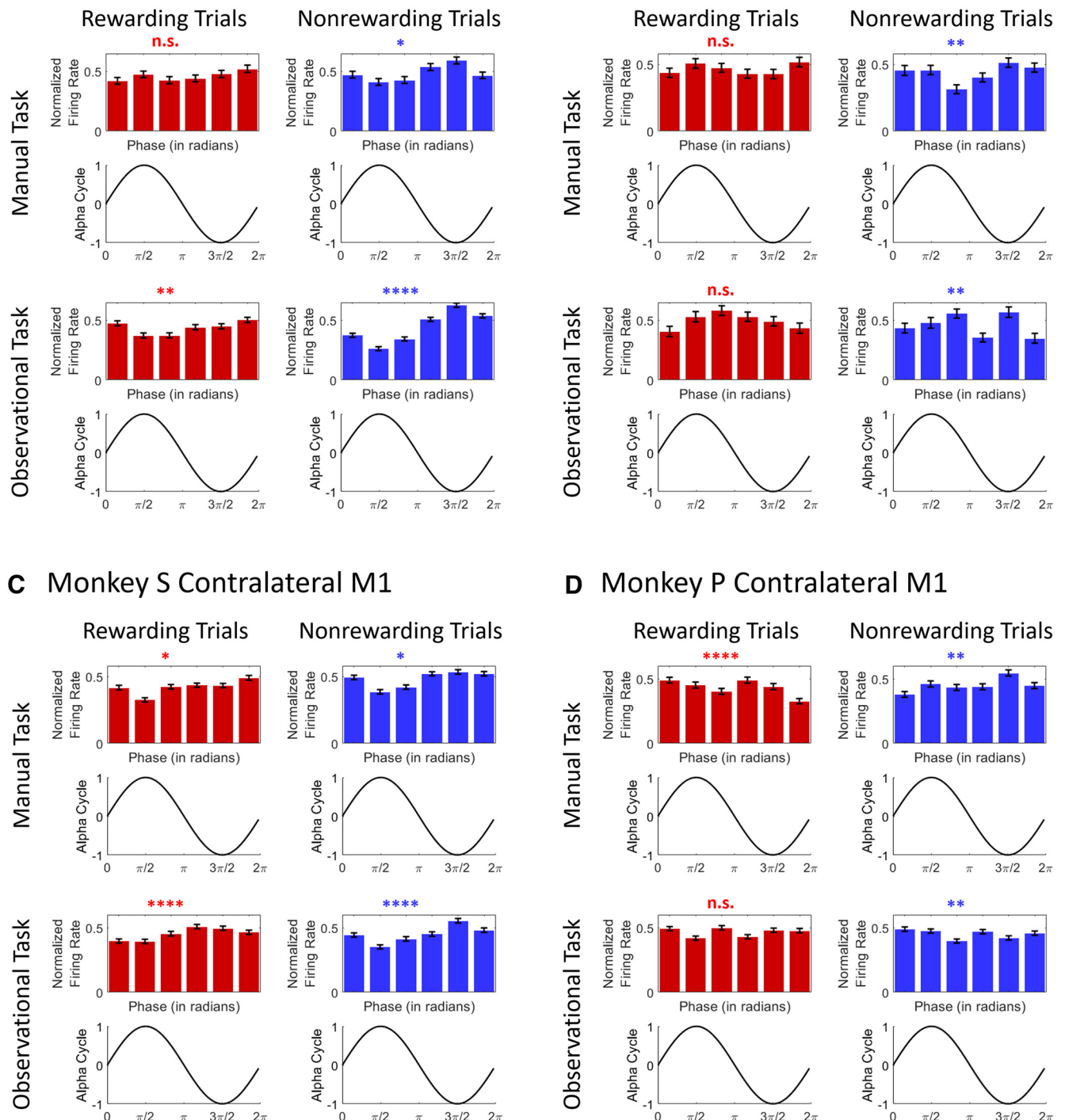
To determine the influence reward expectation has on M1, we recorded neural activity, single/multi-unit activity and LFPs from chronically implanted electrodes bilaterally in M1. We used multiple sensorimotor tasks where NHPs either made reaching or grasping movements, or observed such movements. We found several clear and reproducible patterns of activity between NHPs, cortical hemispheres and tasks. These patterns of activity included an increase in  $\alpha$  power during cued nonrewarding trials, although the manual tasks still required targeted movements by the NHPs, and thus some level of attention. Nonrewarding trials also had stronger  $\alpha$ -band SFC between spiking activity and the averaged LFP activity, and stronger  $\alpha$ -phase  $\gamma$ -amplitude coupling as compared to cued rewarding trials.

The power of LFP oscillations is modulated by visual and auditory attention ([Foxy et al., 1998](#); [Thut et al., 2006](#); [Rihs et al., 2009](#); [Kerlin et al., 2010](#)) and reward expectancy ([van Wingerden et al., 2010](#); [Lansink et al., 2016](#)). Our PSD results showed a consistent reward-related decrease in the mean  $\alpha$  power in bilateral M1 during postcue periods for rewarding as compared to nonrewarding trials for manual and observational tasks ([Fig. 2](#)). Evidence suggests that dopamine plays a critical role in the selection of targets for attention ([Rose et al., 2010](#)), and that injecting a dopamine D1-agonist into the prefrontal cortex of rats enhanced attentional accuracy, while a D1-antagonist led to decreased performance ([Granon et al., 2000](#); [Chudasama and Robbins, 2004](#)). Furthermore, dopamine depletion has an influence on attention-deficit/hyperactivity disorder (ADHD; [Jucaite et al., 2005](#); [Silvetti et al., 2013](#)). These studies suggest that changes in dopamine transmission or release could be responsible for changes in attention. Following this, it is expected that subjects pay more attention during rewarding trials to reach the targets ([Dalley et al., 2002](#); [Demiralp et al., 2007](#); [Rose et al., 2010](#)) compared to nonrewarding trials. Our findings are consistent with previous studies showing that dopamine depletion led to an increased power of LFP oscillations ([Cassidy et al., 2002](#); [Sharott et al., 2005](#); [Costa et al., 2006](#); [Kühn et al., 2008](#); [Mallet et al., 2008](#); [Lemaire et al., 2012](#)). These results provide a clue toward explaining the relationship that exists between dopamine,



**A** Monkey A Contralateral M1

**B** Monkey Z Ipsilateral M1



**Figure 6.** The relationship between neural firing and the  $\alpha$ -band cycle during rewarding (left column in each subplot) and nonrewarding (right column in each subplot) trials across manual (upper row in each subplot) and observational task (lower row in each subplot) for contralateral (**A**, **C**, **D**) and ipsilateral (**B**) M1 cortices (\*\*\*\* $p < 0.0001$ , \*\* $p < 0.01$ , \* $p < 0.05$ ; n.s., no significance; one-way ANOVA). Error bars in the bar plots represent SEM. **A**, **B** are for the center-out reaching tasks, and **C**, **D** are for the grip force tasks.

reward expectation (or attention/motivation), and LFP oscillations in M1 cortex.

The phase of low-frequency oscillations modulate with the amplitude of high-frequency oscillations (Canolty et al., 2006; Tort et al., 2008; Canolty and Knight, 2010).

Some evidence of phase-to-amplitudes comodulation has come from studies conducted across different frequency bands:  $\delta$ - $\gamma$  (Gross et al., 2013; López-Azcárate et al., 2013; Szczepanski et al., 2014),  $\theta$ - $\gamma$  (Bragin et al., 1995; Chrobak and Buzsáki, 1998; Canolty et al., 2006;

Tort et al., 2008; Voytek et al., 2010; Lisman and Jensen, 2013; Voloh et al., 2015),  $\alpha$ - $\gamma$  (Osipova et al., 2008; Cohen et al., 2009; Voytek et al., 2010; Spaak et al., 2012; Yanagisawa et al., 2012; van Kerkoerle et al., 2014; Bonnefond and Jensen, 2015; Park et al., 2016; Seymour et al., 2017; Tzvi et al., 2018), and  $\beta$ - $\gamma$  (de Hemptinne et al., 2013; Kim et al., 2015; Swann et al., 2015). In particular, some of these studies investigated  $\alpha$ - $\gamma$  comodulation in the visual cortices (Osipova et al., 2008; Voytek et al., 2010; Spaak et al., 2012; van Kerkoerle et al., 2014; Bonnefond and Jensen, 2015; Seymour et al., 2017), parietal-occipital areas (Tzvi et al., 2018), lingual gyrus (Park et al., 2016), and sensorimotor cortex (Yanagisawa et al., 2012). In our analysis, we found that reward expectation influenced comodulation between the phase of  $\alpha$ -band (8–14 Hz) oscillations and the amplitude of  $\gamma$ -band (30–100 Hz) oscillations during the postcue period in M1 (Fig. 3). We found a higher strength of phase-to-amplitude comodulation in nonrewarding trials during both manual tasks and observational tasks while performing either a reaching (NHPs A and Z) or grasping task (NHPs S and P). In addition, increased  $\alpha$  power led to stronger  $\alpha$ - $\gamma$  comodulation during all tasks; that is, the strength of  $\alpha$ - $\gamma$  comodulation was positively correlated with  $\alpha$  power (Tort et al., 2008, 2013). These results are consistent with previous studies where stronger  $\alpha$ - $\gamma$  comodulation occurred while  $\alpha$ -band activity increased (Osipova et al., 2008; Voytek et al., 2010; van Kerkoerle et al., 2014). Similar  $\alpha$ - $\gamma$  comodulation was modulated by different reward conditions in the nucleus accumbens of humans as well (Cohen et al., 2009).

As shown in Figure 4, a larger subpopulation of reward-modulated M1 units had significantly higher phase synchronization as measured with SFC between  $\alpha$  oscillations and neural spikes in nonrewarding trials than in rewarding trials for all tasks. This indicates that higher phase-synchronization is seen in the presence of stronger  $\alpha$  oscillations. Our SFC results are consistent with a previous study (Haegens et al., 2011), which showed that an increase of  $\alpha$  power was associated with an increase in  $\alpha$ -band SFC in premotor, motor, and somatosensory regions during a discrimination task. This suggests that stronger  $\alpha$  oscillatory activity in M1 may give rise to a suppression of neural spiking and modulates the inhibitory timing of spiking (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Mazaheri and Jensen, 2010; Jensen et al., 2012; Klimesch, 2012). Thus, a possible explanation for higher phase synchronization in the  $\alpha$ -band between spikes and LFPs is that the lack of reward expectation leads to a decrease in dopamine associated motivation and/or attention, which amplifies  $\alpha$ -band LFP oscillations, and as a result  $\alpha$  oscillations lead to the inhibition and timing of neural spiking. It should be kept in mind that these results were seen when the NHPs had to manually perform the tasks, and thus they still needed to attend, and thus some of these differences are likely due to decreased motivational intensity and not only attention.

M1 FR was modulated by cued reward expectation as we have previously shown (Marsh et al., 2015). Similar reward-related modulation in observational tasks and

manual tasks indicate that M1 units encode reward expectation (McNiel et al., 2016; Ramakrishnan et al., 2017; Tarigoppula et al., 2018). M1 units show mirror like modulation during the execution of reaching movements as well as observation of such actions (Tkach et al., 2008), and mirror neurons in the premotor cortex can be modulated by subjective value (Caggiano et al., 2012). Thus, there may be mirror like neurons modulated by value in M1 and S1 (Marsh et al., 2015; McNiel et al., 2016; Ramkumar et al., 2016; Ramakrishnan et al., 2017; An et al., 2018; Tarigoppula et al., 2018).

Figure 6 shows the possible existence of a relationship between  $\alpha$  phase and neural FRs. During nonrewarding trials for both manual and observational tasks in contralateral M1 (NHP A for reaching tasks, and S for grasping tasks), FRs were high around the trough and low at the peak of the  $\alpha$  cycle. These findings led to an initial conclusion that spiking was locked to the trough of  $\alpha$  oscillations. These findings are consistent with a previous study (Haegens et al., 2011) where the FR was highest at the trough and lowest at the peak of the  $\alpha$  cycle, and in general fit the pulsed-inhibition theory (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Mazaheri and Jensen, 2010; Haegens et al., 2011; Jensen et al., 2012; Klimesch, 2012) which is a mechanism through which  $\alpha$  oscillations suppress neural spiking activity. All of our results are in line with the idea that  $\alpha$  oscillations serve an inhibitory function and that stronger  $\alpha$  power reduces FRs (Haegens et al., 2011). However, unlike in Haegens et al., we saw slightly different  $\alpha$  phase to FR results from our female NHPs in contralateral M1 units from NHPs P and AC8 (data not shown) as well as ipsilateral M1 units from NHP Z, where the FR was highest at the trough of the  $\alpha$  cycle, but it was not lowest at the peak of the cycle. The reason for these results is not clear, but could be that ipsilateral M1 units were less significantly modulated by reward expectation than contralateral M1 units (Donchin et al., 1998; Cisek et al., 2003; Ganguly et al., 2009). Another interesting possibility is that this could be due to sex differences in dopamine receptor concentrations. Previous work has shown that female humans have much fewer dopamine receptors in the striatum as compared to males (Zaidi, 2010). These possibilities could have influenced the inhibition and timing of neural spiking with respect to the  $\alpha$  cycle, although this is speculative.

The differences shown here between rewarding and nonrewarding trials have potential use toward the development of an autonomously updating BCI that utilizes the detectable reward modulation described in this paper and a BCI agent that updates autonomously via RL update rules. Toward this goal, our lab recently showed that integrated features of PSD and SFC yielded near-perfect classification accuracy (An et al., 2018) between rewarding and nonrewarding trials and could be used as a neural critic in autonomously updating BCIs (Bae et al., 2011; Sanchez et al., 2011; Tarigoppula et al., 2012, 2018; Marsh et al., 2015; An et al., 2018). Recently, we and others have shown that M1 directional tuning is also modulated by reward during manual (Ramakrishnan et al., 2017) and BCI control (Zhao et al., 2018), and that taking

this into account could improve BCI control (Zhao et al., 2018). Further work is needed to incorporate our accurate classifier for an autonomously updating and accurate BCI system.

## References

- Adrian ED, Matthews BHC (1934) The Berger rhythm: potential changes from the occipital lobes in man. *Brain* 57:355–385.
- An J, Yadav T, Ahmadi MB, Tarigoppula VA, Francis JT (2018) Near perfect neural critic from motor cortical activity toward an autonomously updating brain machine interface. *Conf Proc IEEE Eng Med Biol Soc* 2018:73–76.
- Axmacher N, Henseler MM, Jensen O, Weinreich I, Elger CE, Fell J (2010) Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc Natl Acad Sci USA* 107:3228–3233.
- Bae J, Chhatbar P, Francis JT, Sanchez JC, Principe JC (2011) Reinforcement learning via kernel temporal difference. *Conf Proc IEEE Eng Med Biol Soc* 2011:5662–5665.
- Belluscio MA, Mizuseki K, Schmidt R, Kempter R, Buzsáki G (2012) Cross-frequency phase-phase coupling between  $\theta$  and  $\gamma$  oscillations in the hippocampus. *J Neurosci* 32:423–435.
- Bonnefond M, Jensen O (2015) Gamma activity coupled to alpha phase as a mechanism for top-down controlled gating. *PLoS One* 10:e0128667.
- Bragin A, Jandó G, Nádasdy Z, Hetke J, Wise K, Buzsáki G (1995) Gamma (40–100 Hz) oscillation in the hippocampus of the behaving rat. *J Neurosci* 15:47–60.
- Bruns A, Eckhorn R, Jokeit H, Ebner A (2000) Amplitude envelope correlation detects coupling among incoherent brain signals. *Neuroreport* 11:1509–1514.
- Buzsáki G, Buhl DL, Harris KD, Csicsvari J, Czeh B, Morozov A (2003) Hippocampal network patterns of activity in the mouse. *Neuroscience* 116:201–211.
- Caggiano V, Fogassi L, Rizzolatti G, Casile A, Giese MA, Thier P (2012) Mirror neurons encode the subjective value of an observed action. *Proc Natl Acad Sci USA* 109:11848–11853.
- Canolty RT, Knight RT (2010) The functional role of cross-frequency coupling. *Trends Cogn Sci* 14:506–515.
- Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT (2006) High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313:1626–1628.
- Cassidy M, Mazzone P, Oliviero A, Insola A, Tonalì P, Di Lazzaro V, Brown P (2002) Movement-related changes in synchronization in the human basal ganglia. *Brain* 125:1235–1246.
- Chalk M, Herrero JL, Gieselmann MA, Delicato LS, Gotthardt S, Thiele A (2010) Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1. *Neuron* 66:114–125.
- Chhatbar PY, von Kraus LM, Semework M, Francis JT (2010) A bio-friendly and economical technique for chronic implantation of multiple microelectrode arrays. *J Neurosci Methods* 188:187–194.
- Chrobak JJ, Buzsáki G (1998) Gamma oscillations in the entorhinal cortex of the freely behaving rat. *J Neurosci* 18:388–398.
- Chudasama Y, Robbins TW (2004) Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology* 29:1628–1636.
- Cisek P, Crammond DJ, Kalaska JF (2003) Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. *J Neurophysiol* 89:922–942.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE (2009) Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. *J Cogn Neurosci* 21:875–889.
- Colgin LL, Denninger T, Fyhn M, Hafting T, Bonnevie T, Jensen O, Moser MB, Moser EI (2009) Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* 462:353–357.
- Costa RM, Lin SC, Sotnikova TD, Cyr M, Gainetdinov RR, Caron MG, Nicoletis MA (2006) Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. *Neuron* 52:359–369.
- Dalley JW, Theobald DE, Pereira EA, Li PM, Robbins TW (2002) Specific abnormalities in serotonin release in the prefrontal cortex of isolation-reared rats measured during behavioural performance of a task assessing visuospatial attention and impulsivity. *Psychopharmacology (Berl)* 164:329–340.
- de Hemptinne C, Ryapolova-Webb ES, Air EL, Garcia PA, Miller KJ, Ojemann JG, Ostrem JL, Galifianakis NB, Starr PA (2013) Exaggerated phase-amplitude coupling in the primary motor cortex in Parkinson disease. *Proc Natl Acad Sci USA* 110:4780–4785.
- Demiralp T, Herrmann CS, Erdal ME, Ergenoglu T, Keskin YH, Ergen M, Beydagi H (2007) DRD4 and DAT1 polymorphisms modulate human gamma band responses. *Cereb Cortex* 17:1007–1019.
- Donchin O, Gribova A, Steinberg O, Bergman H, Vaadia E (1998) Primary motor cortex is involved in bimanual coordination. *Nature* 395:274–278.
- Foxe JJ, Simpson GV, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9:3929–3933.
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563.
- Fries P, Womelsdorf T, Oostenveld R, Desimone R (2008) The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci* 28:4823–4835.
- Ganguly K, Secundo L, Ranade G, Orsborn A, Chang EF, Dimitrov DF, Wallis JD, Barbaro NM, Knight RT, Carmena JM (2009) Cortical representation of ipsilateral arm movements in monkey and man. *J Neurosci* 29:12948–12956.
- Granon S, Passetti F, Thomas KL, Dalley JW, Everitt BJ, Robbins TW (2000) Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *J Neurosci* 20:1208–1215.
- Gregoriou GG, Gotts SJ, Zhou HH, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324:1207–1210.
- Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol* 11:e1001752.
- Haegens S, Nacher V, Luna R, Romo R, Jensen O (2011)  $\alpha$ -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc Natl Acad Sci USA* 108:19377–19382.
- Jarvis MR, Mitra PP (2001) Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Comput* 13:717–749.
- Jensen O, Colgin LL (2007) Cross-frequency coupling between neuronal oscillations. *Trends Cogn Sci* 11:267–269.
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186.
- Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex* 12:877–882.
- Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn Sci* 16:200–206.
- Jucaite A, Fernell E, Halldin C, Forsberg H, Farde L (2005) Reduced midbrain dopamine transporter binding in male adolescents with attention-deficit/hyperactivity disorder: association between striatal dopamine markers and motor hyperactivity. *Biol Psychiatry* 57:229–238.
- Jutras MJ, Fries P, Buffalo EA (2009) Gamma-band synchronization in the macaque hippocampus and memory formation. *J Neurosci* 29:12521–12531.



- Kerlin JR, Shahin AJ, Miller LM (2010) Attentional gain control of ongoing cortical speech representations in a “cocktail party.” *J Neurosci* 30:620–628.
- Kim JH, Chien JH, Liu CC, Lenz FA (2015) Painful cutaneous laser stimuli induce event-related gamma-band activity in the lateral thalamus of humans. *J Neurophysiol* 113:1564–1573.
- Klimesch W (2012)  $\alpha$ -Band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci* 16:606–617.
- Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev* 53:63–88.
- Koralek AC, Costa RM, Carmena JM (2013) Temporally precise cell-specific coherence develops in corticostriatal networks during learning. *Neuron* 79:865–872.
- Kühn AA, Brücke C, Schneider GH, Trottenberg T, Kivi A, Kupsch A, Capelle HH, Krauss JK, Brown P (2008) Increased beta activity in dystonia patients after drug-induced dopamine deficiency. *Exp Neurol* 214:140–143.
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* 94:1904–1911.
- Lansink CS, Meijer GT, Lankelma JV, Vinck MA, Jackson JC, Pennartz CM (2016) Reward expectancy strengthens CA1 theta and beta band synchronization and hippocampal-ventral striatal coupling. *J Neurosci* 36:10598–10610.
- Lemaire N, Hernandez LF, Hu D, Kubota Y, Howe MW, Graybiel AM (2012) Effects of dopamine depletion on LFP oscillations in striatum are task- and learning-dependent and selectively reversed by L-DOPA. *Proc Natl Acad Sci USA* 109:18126–18131.
- Lisman JE, Jensen O (2013) The  $\theta$ - $\gamma$  neural code. *Neuron* 77:1002–1016.
- López-Azcárate J, Nicolás MJ, Cordon I, Alegre M, Valencia M, Artieda J (2013) Delta-mediated cross-frequency coupling organizes oscillatory activity across the rat cortico-basal ganglia network. *Front Neural Circuits* 7:155.
- Mallet N, Pogosyan A, Sharott A, Csicsvari J, Bolam JP, Brown P, Magill PJ (2008) Disrupted dopamine transmission and the emergence of exaggerated beta oscillations in subthalamic nucleus and cerebral cortex. *J Neurosci* 28:4795–4806.
- Marsh BT, Tarigoppula VS, Chen C, Francis JT (2015) Toward an autonomous brain machine interface: integrating sensorimotor reward modulation and reinforcement learning. *J Neurosci* 35:7374–7387.
- Mazaheri A, Jensen O (2010) Rhythmic pulsing: linking ongoing brain activity with evoked responses. *Front Hum Neurosci* 4:177.
- McNiel DB, Choi JS, Hessburg JP, Francis JT (2016) Reward value is encoded in primary somatosensory cortex and can be decoded from neural activity during performance of a psychophysical task. *IEEE Eng Med Bio* 2016:3064–3067.
- Mormann F, Fell J, Axmacher N, Weber B, Lehnertz K, Elger CE, Fernández G (2005) Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus* 15:890–900.
- Osipova D, Hermes D, Jensen O (2008) Gamma power is phase-locked to posterior alpha activity. *PLoS One* 3:e3990.
- Palva JM, Palva S, Kaila K (2005) Phase synchrony among neuronal oscillations in the human cortex. *J Neurosci* 25:3962–3972.
- Park H, Lee DS, Kang E, Kang H, Hahn J, Kim JS, Chung CK, Jiang H, Gross J, Jensen O (2016) Formation of visual memories controlled by gamma power phase-locked to alpha oscillations. *Sci Rep* 6:28092.
- Pesaran B, Nelson MJ, Andersen RA (2008) Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453:406–409.
- Ramakrishnan A, Byun YW, Rand K, Pedersen CE, Lebedev MA, Nicolelis MAL (2017) Cortical neurons multiplex reward-related signals along with sensory and motor information. *Proc Natl Acad Sci USA* 114:E4841–E4850.
- Ramkumar P, Dekleva B, Cooler S, Miller L, Kording K (2016) Pre-motor and motor cortices encode reward. *PLoS One* 11:e0160851.
- Rihs TA, Michel CM, Thut G (2009) A bias for posterior alpha-band power suppression versus enhancement during shifting versus maintenance of spatial attention. *Neuroimage* 44:190–199.
- Rose J, Schiffer AM, Dittrich L, Güntürkün O (2010) The role of dopamine in maintenance and distractibility of attention in the “prefrontal cortex” of pigeons. *Neuroscience* 167:232–237.
- Sanchez JC, Tarigoppula A, Choi JS, Marsh BT, Chhatbar PY, Mahmoudi B, Francis JT (2011) Control of a center-out reaching task using a reinforcement learning brain-machine interface. 2011 5th International IEEE/EMBS Conference on Neural Engineering, pp 525–528. Cancun, Mexico:IEEE.
- Sauseng P, Klimesch W, Doppelmayr M, Pecherstorfer T, Freunberger R, Hanslmayr S (2005) EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Hum Brain Mapp* 26:148–155.
- Seymour RA, Rippon G, Kessler K (2017) The detection of phase amplitude coupling during sensory processing. *Front Neurosci* 11:487.
- Sharott A, Magill PJ, Harnack D, Kupsch A, Meissner W, Brown P (2005) Dopamine depletion increases the power and coherence of beta-oscillations in the cerebral cortex and subthalamic nucleus of the awake rat. *Eur J Neurosci* 21:1413–1422.
- Silvetti M, Wiersma JR, Sonuga-Barke E, Verguts T (2013) Deficient reinforcement learning in medial frontal cortex as a model of dopamine-related motivational deficits in ADHD. *Neural Netw* 46:199–209.
- Soltesz I, Deschênes M (1993) Low- and high-frequency membrane potential oscillations during theta activity in CA1 and CA3 pyramidal neurons of the rat hippocampus under ketamine-xylazine anesthesia. *J Neurophysiol* 70:97–116.
- Spaak E, Bonnefond M, Maier A, Leopold DA, Jensen O (2012) Layer-specific entrainment of  $\gamma$ -band neural activity by the  $\alpha$  rhythm in monkey visual cortex. *Curr Biol* 22:2313–2318.
- Swann NC, de Hemptinne C, Aron AR, Ostrem JL, Knight RT, Starr PA (2015) Elevated synchrony in Parkinson disease detected with electroencephalography. *Ann Neurol* 78:742–750.
- Szczepanski SM, Crone NE, Kuperman RA, Auguste KI, Parvizi J, Knight RT (2014) Dynamic changes in phase-amplitude coupling facilitate spatial attention control in fronto-parietal cortex. *PLoS Biol* 12:e1001936.
- Tarigoppula A, Rotella N, Francis JT (2012) Properties of a temporal difference reinforcement learning brain machine interface driven by a simulated motor cortex. *Conf Proc IEEE Eng Med Biol Soc* 2012:3284–3287.
- Tarigoppula VSA, Choi JS, Hessburg JH, McNiel DB, Marsh BT, Francis JT (2018) Primary motor cortex encodes a temporal difference reinforcement learning process. *bioRxiv*. Advance online publication. Retrieved January 21, 2018. doi:10.1101/257337.
- Thut G, Nietzel A, Brandt SA, Pascual-Leone A (2006) Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci* 26:9494–9502.
- Tkach D, Reimer J, Hatsopoulos NG (2008) Observation-based learning for brain-machine interfaces. *Curr Opin Neurobiol* 18:589–594.
- Tort AB, Kramer MA, Thorn C, Gibson DJ, Kubota Y, Graybiel AM, Kopell NJ (2008) Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *Proc Natl Acad Sci USA* 105:20517–20522.
- Tort AB, Scheffer-Teixeira R, Souza BC, Draguhn A, Brankačk J (2013) Theta-associated high-frequency oscillations (110–160Hz) in the hippocampus and neocortex. *Prog Neurobiol* 100:1–14.
- Tzvi E, Bauhaus LJ, Kessler TU, Liebrand M, Wöstmann M, Krämer UM (2018) Alpha-gamma phase amplitude coupling subserves information transfer during perceptual sequence learning. *Neurobiol Learn Mem* 149:107–117.

- van Kerkoerle T, Self MW, Dagnino B, Gariel-Mathis MA, Poort J, van der Togt C, Roelfsema PR (2014) Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci USA* 111:14332–14341.
- VanRullen R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213.
- van Wingerden M, Vinck M, Lankelma J, Pennartz CM (2010) Theta-band phase locking of orbitofrontal neurons during reward expectancy. *J Neurosci* 30:7078–7087.
- Voloh B, Valiante TA, Everling S, Womelsdorf T (2015) Theta-gamma coordination between anterior cingulate and prefrontal cortex indexes correct attention shifts. *Proc Natl Acad Sci USA* 112:8457–8462.
- Voytek B, Canolty RT, Shestyuk A, Crone NE, Parvizi J, Knight RT (2010) Shifts in gamma phase-amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Front Hum Neurosci* 4:191.
- Waldert S, Lemon RN, Kraskov A (2013) Influence of spiking activity on cortical local field potentials. *J Physiol* 591:5291–5303.
- Welch P (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans Audio Electroacoust* 15:70–73.
- Witham CL, Wang M, Baker SN (2007) Cells in somatosensory areas show synchrony with beta oscillations in monkey motor cortex. *Eur J Neurosci* 26:2677–2686.
- Womelsdorf T, Fries P, Mitra PP, Desimone R (2006) Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439:733–736.
- Yanagisawa T, Yamashita O, Hirata M, Kishima H, Saitoh Y, Goto T, Yoshimine T, Kamitani Y (2012) Regulation of motor representation by phase-amplitude coupling in the sensorimotor cortex. *J Neurosci* 32:15467–15475.
- Zaidi ZF (2010) Gender differences in human brain: a review. *Open Anat J* 2:37–55.
- Zhao Y, Hessburg JP, Asok Kumar JN, Francis JT (2018) Paradigm shift in sensorimotor control research and brain machine interface control: the influence of context on sensorimotor representations. *Front Neurosci* 12:579.