Title

Behavioral context affects social signal representations within single primate prefrontal cortex neurons.

Permalink

Journal

Authors

Jovanovic, Vladimir
Fishbein, Adam
de la Mothe, Lisa
et al.

Publication Date

DOI

Peer reviewed

Behavioral context affects social signal representations within single primate prefrontal cortex neurons

Vladimir Jovanovic[1,2], **Adam Ryan Fishbein**[1], **Lisa de la Mothe**[3], **Kuo-Fen Lee**[4], **Cory Thomas Miller**[1,2,*]

[1]Cortical Systems and Behavior Laboratory, University of California, San Diego, La Jolla, CA 92093, USA

[2]Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA

[3]Department of Psychology, Tennessee State University, Nashville, TN 37209, USA

[4]Laboratory for Peptide Biology, Salk Institute, La Jolla, CA 92093, USA

Summary.

We tested whether social signal processing in more traditional, head-restrained contexts is representative of the putative natural analog – social communication – by comparing responses to vocalizations within individual neurons in marmoset prefrontal cortex (PFC) across a series of behavioral contexts ranging from traditional to naturalistic. Although vocalization responsive neurons were evident in all contexts, cross-context consistency was notably limited. A response to these social signals when subjects were head-restrained was not predictive of a comparable neural response to the identical vocalizations during natural communication. This pattern was evident both within individual neurons and at a population level, as PFC activity could be reliably decoded for the behavioral context in which vocalizations were heard. These results suggests that neural representations of social signals in primate PFC are not static, but highly flexible and likely reflect how nuances of the dynamic behavioral contexts affect the perception of these signals and what they communicate.

eTOC.

---

[*]Lead Contact: corymiller@ucsd.edu.

Author Contributions

VJ designed the experiment, conducted the experiment, analyzed the data and wrote the paper

ARF analyzed the neurophysiology data and wrote the paper

LDM analyzed the anatomy data

KFL analyzed the anatomy data

CTM designed the experiment, supervised the experiment and wrote the paper

Jovanovic and colleagues report that social signal processing in primate prefrontal cortex during traditional, head-restrained paradigms is not representative of the analogous process in natural communication. These findings suggest that some facets of primate social brain function can only be elucidated in the contexts in which they occur naturally.

Introduction.

Communication is an inherently interactive social process characterized by the active exchange of signals between conspecifics (Guilford and Dawkins, 1991). Yet because of practical constraints, experiments seeking to explicate the neural basis of social communication in the primate brain have traditionally employed paradigms in which social signals – e.g. faces and vocalizations – are presented as static stimuli divorced from the dynamic behavioral contexts in which they naturally occur. While these approaches have been notably prolific at revealing integrated networks for both face and voice processing in primate temporal and frontal cortex (Freiwald et al., 2016; Gifford et al., 2005; Hung et al., 2015; Perrodin et al., 2011; Petkov et al., 2008; Romanski et al., 2005; Schaeffer et al., 2020; Tsao et al., 2006; Tsao and Livingstone, 2008; Tsao et al., 2008), fundamental questions remain about the role these neurons play during active communication. Indeed, McMahon and colleagues (McMahon et al., 2015) found that although neurons in the AF face patch exhibited classic selectivity to face stimuli when subjects passively viewed images of faces, the responses of the same individual neurons were primarily driven by a myriad of different properties – including social proximity and movement – when subjects simply watched videos of monkeys engaged in natural social interactions. Such findings raise a critical question of whether neural responses to passively presented social stimuli in conventional primate head-restrained paradigms are representative of how the same individual neurons (or populations) will respond when subjects are actively participating in natural communication exchanges. Here we sought to test this critical issue by directly comparing the responses of individual neurons in marmoset PFC to conspecific vocalizations in a series of behavioral contexts – ranging from a more traditional, head-restrained paradigm to freely-moving monkeys engaged in interactive natural communication (Figure 1A). We hypothesized that if the responses of neurons to vocalizations in a more traditional context was representative of natural communication, similar patterns of neural activity would be evident within individual units and/or at a population level across behavioral contexts. If, however, neural responses to vocalizations in one context poorly predicted another, it would demonstrate a more dynamic system in which behavioral context affects the perception of social signals and related neural representations in primate PFC. Such a result would also highlight the caveats of using traditional head-restrained paradigms to elucidate the neural basis of natural primate social behaviors.

Results.

Vocalization Responses in Marmoset PFC.

We recorded the activity of 388 single neurons in the PFC of three marmoset monkeys (*Callithrix jacchus*) in response to vocalizations in three behavioral contexts – 'Restrained'

(R), 'Freely-Moving' (F) and 'Communication' (C) (Figure 1A). To test how behavioral context affected neural activity, identical phee calls produced by a single caller were presented to subjects in all three contexts within each daily test session. We found neurons that exhibited significant changes in firing rate in response to hearing marmoset phee calls in all areas of PFC examined here (Figure 1B,C) and all three contexts,

Within-Neuron Comparisons.

This study was designed to directly test whether two key behavioral features that differ between traditional and naturalistic studies of communication significantly affected PFC responses to vocalizations - subjects' mobility (head-restrained or freely-moving) and stimulus presentation (consistent timing interval or interactive). We focused the next set of analyses on the 247 single PFC neurons that maintained consistent isolation across all three behavioral contexts. We hypothesized that if subjects' mobility was a key factor modulating vocalization responsiveness within PFC neurons, neural activity in the R context would be distinct from the other two. By contrast, if stimulus presentation significantly affected responses to vocalizations, neural responsiveness in the consistent interval conditions (R and F) would be similar and differ from the interactive context (C). If, however, a broader suite of contextual features affects vocalizationresponsivity in primate PFC neurons, we would expect little consistency across the behavioral contexts.

Overall, data were consistent with the latter hypothesis. Neurons exhibited stimulus-driven responses to marmoset phee calls in all behavioral contexts (Figure 1D), but significant neural responses to phee calls in one behavioral context poorly predicted a comparable response in another (Figure 2A). Overall, 170 neurons demonstrated a statistically significant change in activity in response to phee calls. Of these neurons, 29 units were responsive only in R, 34 only in F, 38 only in C, 13 in both R and C, 15 in both R and F, 15 in both F and C, and 26 neurons exhibited a significant change in activity in all three behavioral contexts – i.e. R, F, and C – and described as RFC units subsequently (see Table S1). The probability that a vocalization responsive neuron in one behavioral context would exhibit the same response in another context was only 0.16 (SD = 0.418). Figure 2B plots exemplar neurons that exhibited a significant change in firing rate to phee calls in only a single context (C), two contexts (R and F) or all three contexts (i.e. RFC neuron). This pattern of acoustic stimulus responsivity and contextual heterogeneity was not limited to phee calls, as it was also evident when comparing neural responses to a broader corpus of marmoset vocalizations and noise stimuli between the R and F contexts (Figure S1). The exemplar neuron in Figure 2C illustrates this pattern. Lastly, Spearman's rank tests showed very weak correlation in changes in spike rate across contexts within the 170 vocalization responsive units further highlighting the cross-context heterogeneity across this population (Figure 2D).

Several factors other than the dynamic nature of behavioral contexts could explain the pattern of results here. One possibility is differences in attention or arousal might drive a linearly additive effect on neural responses across behavioral contexts (R < F < C). However, vocalization-driven activity was remarkably consistent (Figure 3A), and there was no evidence that firing rate changed linearly between the contexts ($F(1, 2.015) =$

4.017, p = 0.182, n = 29164 observations). A second possibility is that PFC response heterogeneity could result from spatial selectivity due to head-direction. Consistent with previous experiments Cohen et al., 2009), we failed to find evidence that spatial position affected PFC responses here (Figure S2A). Lastly because stimulus timing differed between the R/F contexts and C, it is possible that differences in the inter-stimulus interval inadvertently affected neural responses. However, analyses failed to find any evidence of such an effect, as the standard deviation in firing rate was similar across all behavioral contexts regardless of the interval between the phee stimuli (Figure S2B).

The heterogeneity of PFC responses was also evident at the population level. We applied a decoder to test if patterns of neural responses differed by context across the population. We independently tested two sets of neurons – RFC Units (n = 26) and All Units (n = 247) - as inputs across 500 simulations and used the Matthew's Correlation Coefficient (MCC, +1 perfect prediction, 0 for random, −1 total disagreement) to evaluate performance. We hypothesized that the decoder for all units would be highly accurate in predicting the context and would outperform the RFC units, consistent with a mixed-selectivity population level coding scheme (Bernardi et al., 2020; Blackman et al., 2016; Fusi et al., 2016; Parthasarathy et al., 2017; Rigotti et al., 2013). Indeed, All Units exceeded 90% accuracy and significantly outperformed RFC Units (Wilcoxon, n = 500, signed rank = 0, p < 0.001; Figure 3B). In fact, RFC neurons did not perform better than a random selection of an equal number of neurons (n = 26; Figure 3C). Notably, the high decoding performance for behavioral context was not ubiquitous. When applied to the five acoustic stimulus types presented in the R (n = 233 units) and F (n = 276 units) contexts, a neural decoder was able to classify noise stimuli with high accuracy, but did poorly with the vocalization stimuli (Figure S3). Interestingly, performance for both stimulus classes was significantly better in the R context than in F (Wilcoxon, n = 500, signed rank = 125000, p < 0.001; Figure S3A), suggesting that primate PFC population dynamics for stimulus coding are particularly sensitive to the behavioral context in which the stimuli are heard and may not code all potential variables with similar robustness.

The confusion matrices in Figure 3D illustrate other notable decoding differences between the RFC and All Units populations. First, the decoder for All Units had an overall substantial increase in accuracy for all contexts relative to RFC Units. Second, though more pronounced for the RFC Units, both decoders revealed higher false positives between the R and F contexts. The overlap between these populations is also evident through dimensionality reduction in a principal component analysis (Figure 3E). Although neural activity in these two contexts was accurately decoded at over 90% for All Units, it does suggest some similarity considering that the decoding accuracy for the Communication context was nearly perfect (99.3%).

Given the notably improved accuracy of the decoder that included All Units, we tested the decoding performance of individual units across the three contexts. For all 247 units, we ran 500 simulations and calculated the 99.99% confidence interval of their accuracy in classifying the three contexts. We found that 193 units (78%) had significantly better performance than chance classification (Figure 3F; H01: n = 114, H02: n = 77, E01: n = 2). Interestingly, the proportion of high accuracy units within each class of phee-responsive

neurons (R, F, C, etc.) was roughly similar across the population (Figure 3G) suggesting that the context-dependence of stimulus-driven activity may not be a significant factor for neural decoding.

Active vs Passive Listening during 'Communication'.

The next analyses examined whether task demands such as active vs. passive engagement with stimuli can be similarly decoded within a behavioral context. Indeed, task demands are known to affect PFC activity in primates, including those involving vocalizations (Cohen et al., 2009; Hwang and Romanski, 2015; Plakke et al., 2013). Here we took advantage of the fact that the C context comprises instances when marmosets passively listen to a conspecific call but do not produce a vocal response (Independent) and occasions when marmosets produce a vocalization response upon hearing a conspecific (Antiphonal) (Nummela et al., 2017). This analysis was performed on the 248 neurons recorded in Communication sessions that reached the behavioral threshold of at least 5 Antiphonal and 5 Independent events (E01: n = 48, H01: n = 200). Results indicated that although these contexts could be distinguished when considering only stimulus driven activity – Stimulus-Only - with modest reliability (Figure 4B Left), the decoding accuracy was notably lower than in the analyses above (Figure 3B). This suggests that the effect of passive and active engagement during communication may be less profound on PFC dynamics than between these other more distinct behavioral contexts. As reported in a previous study (Nummela et al., 2017), frontal cortex activity in the period of time immediately preceding and following a vocalization was highly correlated with the likelihood of a subsequent vocalization by the marmoset and may reflect the social state of the brain. We performed that same analysis for the dataset here – i.e. StateBased - and found a significant increase in performance over the Stimulus-Only analysis (Wilcoxon, n = 500, signed rank = 2550, p < 0.001; Figure 4B Left; Figure 4C). This result suggests that the state of PFC is more closely correlated to the probability of a subsequent vocal response than only stimulus driven activity.

The increase in decoding accuracy between the Stimulus-Only and State-Based decoders could either reflect a difference in PFC activity at the time subjects heard a phee call or because hearing the phee call elicits a distinct pattern of neural activity that persists past the offset of the stimulus. To distinguish between these alternatives, we tested the decoder with two separate data sets that included neural during the phees with either 1500ms prior to stimulus onset (PreStim) or 1500ms following stimulus offset (PostStim). Results indicated that PreStim significantly outperformed PostStim (Wilcoxon, n = 500, signed rank = 106000, p < 0.001; Figure 4B Right) suggesting that PFC activity prior to – rather than after - hearing a conspecific a phee call is more highly correlated with marmosets' propensity to engage in a natural vocal interaction.

We next analyzed the individual unit decoding performance between Antiphonal and Independent events within the C context for this latter State-Based analysis. Results revealed that 109 of 248 units (E01: n = 16; H01: n = 93) had a significantly higher accuracy than chance classification at the 99.99% confidence interval (Figure 4D).

Discussion.

Neurophysiological studies of social signal processing have typically relied on passive, head-restrained paradigms, but here we found that PFC activity in this context was not predictive of the putative natural analog – communication. Instead, we report that the context in which vocal signals are heard profoundly affects their neural representation in primate PFC. The pattern of results did not support a mechanism in which different neurons are responsive to vocalizations in different contexts but integrated at a population level for a consistent, context-independent representation. Rather, the results are consistent with a mixed-selectivity mechanism, in which task-relevant information is distributed across a neural population to support flexible behaviors (Bernardi et al., 2020; Blackman et al., 2016; Fusi et al., 2016; Parthasarathy et al., 2017; Rigotti et al., 2013). By holding the stimulus – phee calls - constant and manipulating the context in which they were heard, our findings reflect the fact that the perception of primate social signals is highly affected by the contextual nuances in which they occur (Cheney and Seyfarth, 2018). Our data indicate that the neural representations of social signals in traditional head-restrained paradigms is neither representative of natural communication nor reflects a baseline state upon which behavioral demands modulate activity. It is, however, likely that the sensitivity to behavioral context in PFC is not reflective of neural populations at all stages of the social processing hierarchy in the primate brain. Given its role in the myriad behavioral and cognitive processes, PFC may be particularly coupled to the nuances of social contexts during natural communication.

Differences in the neural representations of vocal signals here likely reflect how contextual variables affect the perception of social signals. Marmosets only perceive conspecifics to be communicative partners if they abide species-specific social rules during vocal interactions (Miller et al., 2009; Toarmino et al., 2017). Therefore, although marmosets did hear phee calls in each context, those calls were likely only perceived as being produced by a communicative partner in the C context. By contrast, phee call stimuli in both the R and F contexts were not broadcast in accordance with the temporal dynamics of marmoset conversations, but rather were randomized with other vocalizations and noise stimuli at a regular temporal interval. Neural responses at the individual unit and population levels were not uniform between these contexts, nor linearly additive (Figure 2). These results suggest that the properties distinguishing behavioral contexts and their impact on social signal perception likely exist in a high-dimensional space that mirror the dynamics of primate sociality. An important question in this regard is whether PFC, and other neural substrates (Gothard, 2020), are sensitive to such a wide range of contextual variables that they can be decoded with high accuracy. While our current dataset does not allow us to fully address this issue, the notably weaker performance to classify vocalizations (Figure S3) relative to classifying behavioral contexts (Figure 2) suggests that PFC may be more sensitive to the nuances of context because of its effect on social perception (Figure S3B).

During the C context marmosets heard conspecific phee calls in two behaviorally distinct contexts - phee calls that elicited a vocal response from conspecifics (Antiphonal) and those that did not (Independent). We previously found that these two behavioral contexts could be almost perfectly decoded from population activity in marmoset frontal cortex when including the pre- and post-stimulus firing rates, but not from the vocal signal stimulus

alone (Nummela et al., 2017). Applying the same decoding approach here revealed a similar pattern of results. Although the decoder did perform modestly well in classifying neural activity in the Antiphonal and Independent contexts in the Stimulus-Only analysis, inclusion of neural activity before and after the phee stimulus led to a significant increase in decoding accuracy. Subsequent analyses showed that a decoder including only pre-stimulus activity outperformed one including only post-stimulus activity (Figure 4B), suggesting that the state of PFC prior to a marmoset hearing a phee call is significantly correlated with the probability of a subsequent social interaction, a potential unique parallel process to social signal processing that would be difficult to replicate in more traditional paradigms.

Communication is social behavior for which signal processing is one component, rather than a system in which signal processing forms a base on top of which behaviors are built. The current study focused solely on PFC but there are reasons to think that the effects of social context on primate neocortical function are more widespread (Ainsworth et al., 2021; Cléry et al., 2021; Sliwa and Freiwald, 2017). Indeed, merely the presence of conspecifics can affect primate decision-making (Chang et al., 2013; Chang et al., 2011), while simply watching conspecific social interactions as a third-party observer has been shown to change the response properties of face cells (McMahon et al., 2015). Coordinated social and vocal interactions drive brain-to-brain coupling in humans and other mammalian brains that may be a unique neural signature to coordinated social interactions (Stephens et al., 2012; Zhang and Yartsev, 2019). By demonstrating - for the first time – that the paradigms routinely employed to explicate the neural basis of communication in the primate brain do not faithfully reflect PFC activity during the analogous natural social behavior, these results emphasize the significance of natural behaviors to understand the neural dynamics of key facets of primate social brain functions.

STAR Methods.

RESOURCE AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Cory Miller (corymiller@ucsd.edu)

Materials Availability.—This study did not generate new unique reagents.

Data and Code Availability.—All data and original code have been deposited at Dryad (https://doi.org/10.6076/D11P4N) and is publicly available as of the date of publication. DOI is listed in the key resources table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Three adult common marmosets (*Callithrix jacchus*) were used for this experiment. H01 and E01 were female while H02 was male. All subjects were at least 1.5 years old at time of implant. E01 had bilateral arrays implanted. H01 had a right hemisphere implant, while H02 had a left hemisphere. H01 contributed data to all analyses. H02 contributed data to all within-neuron analyses comparing (R)estrained, (F)reely-Moving and (C)ommunication contexts but failed to reach the behavioral threshold for analyses comparing Antiphonal

and Independent contexts within Communication. E01 contributed data to the context comparisons within the C context but had only a single session in which isolated units were maintained across the three behavioral contexts (R, F and C). All animals were group housed, and experiments were performed in the Cortical Systems and Behavior Laboratory at University of California San Diego (UCSD). All experiments were approved by the UCSD Institutional Animal Care and Use Committee.

METHOD DETAILS

Behavioral Contexts.—Subjects were tested on three behavioral contexts in each session: 'Restrained', 'Freely-Moving' and 'Communication'. The order of these contexts was randomized each session and counterbalanced between subjects. Within each recording session, subjects were presented with the same set of phee calls produced by a single caller from the UCSD colony and recorded previously following standard vocal recording procedures (Miller and Wang, 2006). A total of 10 adult marmoset were used to generate the phee call stimulus sets and were randomly selected for each test session.

- <u>Restrained (R)</u>. Marmosets were head-restrained in a standard chair used in previous research (Mitchell et al., 2014; Nummela et al., 2019) and a series of acoustic stimuli with a 1 s inter-stimulus interval that comprised phee calls and 1 s noise was broadcast. The stimulus sets broadcast to subjects H01 and H02 also included reversed phee calls, twitters, and reversed twitters. Stimuli were organized into four blocks and the stimulus order was randomized. Thirty exemplars of each stimulus type (i.e. phees, noise, etc) were broadcast in every test session.

- <u>Freely-Moving (F)</u>. The identical stimulus presentation protocol as described for 'Restrained' was used in this context. The only difference was that the animals were freely-moving in the test box rather than head-restrained in a primate chair.

- <u>Communication (C)</u>. Similar to the Freely-Moving context, subjects here were able to move freely around the test box. Rather than broadcast the same stimuli as the previous two conditions, only phee calls were presented using our interactive playback paradigm. Here, marmosets engaged with a computer-generated Virtual Marmoset (VM) designed to broadcast phee calls in response to subjects' calls and simulate natural vocal interactions. The identical paradigm has been used in several previous behavioral and neurophysiological studies of marmosets in our lab (Miller and Thomas, 2012; Miller et al., 2015; Nummela et al., 2017; Toarmino et al., 2017). Briefly, the software is designed to detect phee calls produced by subjects online. Whenever subjects produce a phee call, the VM emits a phee call within 2–4 s in response. If subjects do not emit a phee call for more than 45–90 s, the VM will also broadcast a phee call. Phee calls produced by subjects within 10 s of a VM phee call are classified as an 'antiphonal' response. Calls produced outside of that time period are classified as 'spontaneous' phee calls.

Test Procedures.—All recording sessions took place in a Radio-Frequency shielding room (ETS-Lindgren) in a 4 × 3 m room. Subjects were placed in a clear acrylic test box

with a plastic mesh on the front side (32 × 18 × 46 cm) or standard primate chair positioned on a table on one side of the room. A single speaker (Polk Audio TSi100, frequency range 40–22,000 Hz) was placed 2.5 m away from the test box on the opposite side of the room. A black cloth occluder was positioned equidistant between the table and speaker to eliminate subjects' ability to see the speaker. One microphone was placed in front of the subject and speaker each (Sennheiser, model ME-66). The speaker broadcast acoustic stimuli at an approximate 80–90 db SPL measured 1 m from the speaker. Subject and speaker calls were recorded simultaneously with the neurophysiological data on a data acquisition card (NI PCI-6254).

Subjects H01 and H02 also participated in a separate 'Orientation' test condition to determine whether neurons in marmoset PFC were spatially selective. Here subjects were positioned in a chair and head-restrained, similarly to the Restrained context, and tested in the same test room but the relative angle of the animal to the speaker was systematically manipulated in 45° angles. With 0° being directly facing the speaker, we would reposition the animal at random 45° position offsets from forward in a random sequence for each recording session until all eight positions were completed. We broadcast 30 exemplars of noise stimuli with a 30 s inter-stimulus interval at each of the eight positions. We also measured subjects head position relative to the speaker in a subset of Freely-Moving and Communication contexts. For these test sessions, two cameras (GoPro Hero Session) recorded the animal's position simultaneously from two locations. One was positioned to record the animal from the right side of the test box, while the other was positioned directly above the test box. During these sessions, an Arduino system flashed an LED visible to both cameras at 0.5 Hz. This signal was recorded along with the audio and neural streams as well to accurately align the video streams with the start of neural recording. Images at stimulus onset were captured for each session and the relative angle of the head to the speaker measured.

Neurophysiological Recording Procedures.—Subjects were surgically implanted with an acrylic head cap using previously described procedures (Courellis et al., 2019; Miller et al., 2015; Nummela et al., 2017). In a subsequent procedure, a 16ch Warp16 microelectrode array (Neuralynx) was implanted in prefrontal cortex. Each array had 16 independent guide tubes in a 4 × 4 mm grid with tungsten electrodes. Each array was implanted on the surface of the brain with each electrode in the guide tubes entering the laminar tissue perpendicularly when pushed by a Warp Drive pusher. The calibrated Warp Drive pusher would attach to the end of a guide tube to allow advancement of 10 to 20 μm per electrode twice a week.

Electrodes were recorded at 20,000 Hz with a prefilter at 1 Hz to 9000 Hz, and 20,000 gain. Subjects had a 1:1 gain headstage preamplifier connected to the Warp16 arrays that was attached to a tether to allow subjects to freely move around in the test box. A metal coil tightly wrapped around the tether prevented any interference by the subject during Freely and Communication contexts. Offline spike sorting was completed by combining across multiple sessions recorded in a single day, applying a 300 to 9000 Hz filter and thresholding subsamples across the entire recording session. After base thresholding, unit 1 ms waveforms were plotted in PCA space across the first three principal components and

time. DBSCAN was used to automatically cluster the units followed by manual curations of all units to ensure proper clustering. Units with 13 dB SNR or greater and <1% violation of the 1 ms refractory period for inter-spike intervals were included in our analyses. Overall, 388 isolated single units were identified that met or exceeded these thresholds, with some channels collecting multiple well isolated single units. Of the 388 single units, 247 units met these thresholds for all three test contexts in a single session. Typically, neurons that failed to meet these criteria in all contexts did so because of noise introduced into the recording for one of the contexts.

Perfusion, Tissue Processing and Reconstructions.—At the conclusion of the study, animals were anesthetized with ketamine, euthanized with pentobarbital sodium, and perfused transcardially with phosphate-buffered heparin solution followed by 4% paraformaldehyde. The brain was impregnated with 30% phosphate-buffered sucrose and blocked. The frontal cortex was cut at 40 μm in the coronal plane. Alternating sections were processed for cytochrome oxidase (Wong-Riley, 1979), nissl substance with thionin, and vesicular glutamate transporter 2 (vGluT2). Areas were determined using previously identified criteria (Paxinos et al., 2012). Electrode penetrations, and tracer injection sites were used to reconstruct the location of the electrode array with respect to the anatomical borders and confirm the location of electrode penetrations. Images of the tissue were acquired using a Nikon eclipse 80i. These methods are similar to those employed in previous anatomical studies of marmoset neocortex (de la Mothe et al., 2006, 2012) and our earlier neurophysiology experiments (Miller et al., 2015; Nummela et al., 2017).

QUANTIFICATION AND STATISTICAL ANALYSIS.

Stimulus Response Significance.—Each single unit was classified as exhibiting a significant response to an acoustic stimulus if the firing rate during one of the three following stimulus periods exhibited a statistically significant change in activity at $\alpha=0.05$ significance level with a Sign Rank test relative to the 500 ms prior to stimulus onset: [1] the entire duration of the stimulus, [2] the peak 500 ms firing rate of the unit within the duration of the stimulus, or [3] the first or second pulse of the phee call stimulus. These latter two criteria were adopted because multi-pulsed phee calls have a long duration (~2500–3000 ms) and the observation that although many single neurons did not exhibit a sustained change in firing rate for the duration of these vocal signals, the firing rate increased substantially for some period of the stimulus.

Array Channel Responsiveness.—To determine whether specific areas of marmoset PFC had a higher concentration of acoustic responsive neurons, we performed the following analysis: we divided the total number of well-isolated single units for each electrode channel that exhibited a significant response to at least one stimulus by the total single units recorded from that electrode. Each electrode location position in PFC is shown in Figure 1B along with a color representation of that ratio responsiveness.

PSTH Normalization.—Unit trials were binned at 100 ms intervals starting 1000 ms prior to stimulus onset and ending 1500 ms after stimulus onset for Noise, Twitter, and Reverse Twitter. Phee and Reverse Phees included 4000 ms after stimulus offset. The firing rate

for each bin within each trial for a given unit was calculated. Normalization occurred by calculating the mean firing rate and standard deviation of all the bins prior to stimulus onset to Z-score the firing rate for the bins following stimulus onset.

Firing Rate Normalization.—To compare changes in firing rate between two contexts, we compared neural activity in the 1000 ms prior to stimulus onset with either the first pulse of the phee call or entire duration of the Noise stimulus. The first pulse and Noise duration were roughly the same duration on average: 1250 ms and 1000 ms, respectively. The firing rates prior to stimulus onset were used to Z-score the subsequent mean after onset Firing Rates. For analyses that only considered phee call comparisons, we used the entire duration of the phees and Z-scored in the same manner.

Spatial Selectivity Analysis.—For the Orientation context outlined in the Test Procedures above, all 240 trials were analyzed using a 2-Way ANOVA. The firing rate in the 500 ms before and after onset were used for comparison across all trials, with trials organized into the 8 different 45° bins based on the relative angle of the subjects' head to the speaker. If the unit was identified as stimulus responsive according to the metrics outlined in the Stimulus Response Significance section above, the interactive effect of spatial orientation was tested and Tukey-Kramer corrected comparisons were used to determine any orientation that had significance compared to the others. If all orientations exhibited similar responses with a main effect, we counted that unit as being generally responsive (all 8 orientations). In the Freely-Moving and Communication contexts, a similar analysis was used for Noise and Phee stimuli. As noted in Behavioral Recording Procedures, the gaze direction from the front of the head relative to the speaker on the transverse plane was marked. These orientations were binned into 8 different 45° groups similarly to the Orientation condition and the same analysis performed. The results from these analyses are shown in Figure S2A.

Inter-stimulus Interval Analysis.—For each unit used in the analysis shown in Figure S2B, we calculated the firing rate of each trial in which a phee call was broadcast and the subsequent one with the same stimulus in each of the three behavioral contexts. The ratio of the firing rate for each of the two stimulus periods was calculated. Standard deviation was calculated by binning each 10 seconds and then finding the standard deviation 5 seconds before and after that bin time. 95% CI was calculated by the square root of the df of the calculated standard deviation divided by the χ^2 inverse of that df. Only pairs of phee stimuli with an inter-stimulus interval less than 60 s were used in this analysis.

Linear Model.—We conducted a linear mixed effect model (MATLAB function fitlme fit by maximum likelihood) to test for a linear effect of context (restrained < freely < communication) on normalized firing rate for all the maintained units. Context was coded as a numerical variable (restrained = 1, freely = 2, communication = 3) in order to test if it had a linear effect. The formula for the model was as follows using Wilkinson notation: FiringRate ~ 1 + Context + (1 | Subject) + (1 + Context | Subject) + (1 | Channel:Session:Subject) + (1 | Session:Subject) + (1 | Type), where Type is the selectivity type of the neuron. The random effects in the model account for the nested design of the

experiment where there were repeated observations for each channel within each session for each subject. The function anova using the Satterthwaite estimation of degrees of freedom was applied to the model to test for the significance of the effect of context.

Decoder Classification.—Five hundred Monte-Carlo simulations were created for each model by subsampling each unit's class of relevant trials with replacement. For single-unit decoders, 1000 samples were taken from each class as each class would have less than 100 samples per class. For multi-unit decoders that compared across multiple bins, we used 2161 samples calculated by multiplying and rounding the total number of possible units to include, the number of data points for each unit, and 1.25 to ensure enough representation. Equal numbers of samples were drawn for each class for each unit after 50% partition between training and test sets. Each simulation created new partitions for each unit. Each trial had 7 points of data (3 bins for each of the 2 pulses, and 1 bin for the inter-pulse-interval). The Samples × 7 matrix for each unit would then be combined with all other units with sufficient trials, for each of the classes within both training and test. For example, the 26 RFC units for Restrained, Freely, and Communication classes would produce a matrix of size 6483 × 182 for both training and test with no overlap of trials across those two. Each row is then assigned a value representing the class that all 182 values represent. PCA plots had difficulty representing the classes, so the MATLAB function tsne (t-distributed stochastic neighbor embedding) was used to represent the separability of classes. Only units with at least five events from each class of interest were included in these analyses.

The training data was fit using the "fitcecoc" function in MATLAB using the default settings. This creates multiclass support vector machines for each simulation, and then predicts the class of the test data. Each confusion matrix that results from the predictions is then stored, and the mean value shown. We also quantified the performance of each classifier by the Matthews Correlation Coefficient (MCC). For a K × K confusion matrix C (i.e. K = 3 for Restrained, Freely, and Communication classes):

$$MCC = \frac{\sum_k \sum_l \sum_m C_{kk} C_{lm} - C_{kl} C_{mk}}{\sqrt{\sum_k (\sum_l C_{kl})(\sum_{k' \mid k' \neq k} \sum_{l'} C_{k'l'})} \sqrt{\sum_k (\sum_l C_{lk})(\sum_{k' \mid k' \neq k} \sum_{l'} C_{lk'})}}$$

For K = 2, MCC ranges from −1 to +1. The +1 means perfect prediction, −1 means complete disagreement between predicted and actual, and 0 is no better than chance prediction. For K = 3, the lower limit is not at −1 and unique to any given classifier. Still, 0 means it is at chance prediction and +1 is perfect prediction. We conducted a null-hypothesis test for both classifiers running the same size data as the full data set (all 247 units). With a randomized assignment of class for each row in training and test data, the MCC was at 0 as expected.

State-based classification used only four points of data as previously done (Nummela et al., 2017) rather than the seven points applied to RFC context decoding. We knew from our data that there were different baseline firing rates across contexts for many units, which would give facile decoding. Thus, for state-based, within a single context, decoding, we looked at four bins of data. 1.5 sec before onset (Pre), the entire duration of the first pulse and then the second pulse (Pulse 1 & Pulse 2), and then 1.5 sec after offset (Post). To better compare across all stimulus sets for analyses in Figure S3, we split non-phee stimulus sets into first

half and second half, and likewise with the first pulse presented in phee and reverse phee sets. As a result, all stimuli were ~1s in duration. All stimulus sets within a given context would be binned according to the bins outlined above, and then z-scored within each bin. The subsequent classes would be split between training and test (50%) before Monte-Carlo simulations. If the decoder had three or more classes to classify, we used fitecoc as described before. For 2 class decoding (such as Antiphonal versus Independent), PCA was performed on the training set to get the first Principal Component (PC1). Test data was projected onto PC1. 2-means clustering of the training data within PC1 and the subsequent accuracy was used to determine whether the 2-means clustering of the test set would be "flipped." If accuracy was less than 50%, that would mean that the Antiphonal cluster would be greater than the Independent cluster and need to have the clustering assignments flip.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References.

Ainsworth M, Sallet J, Joly O, Kyriazis D, Kriegeskorte N, Duncan J, Schüffelgen U, Rushworth MFS, and Bell AH (2021). Viewing Ambiguous Social Interactions Increases Functional Connectivity between Frontal and Temporal Nodes of the Social Brain. The Journal of Neuroscience 41, 6070–6086. [PubMed: 34099508]

Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, and Salzman CD (2020). The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. Cell 183, 954–967.e921. [PubMed: 33058757]

Blackman RK, Crowe DA, DeNicola AL, Sakellaridi S, MacDonald AW, and Chafee MV (2016). Monkey Prefrontal Neurons Reflect Logical Operations for Cognitive Control in a Variant of the AX Continuous Performance Task (AX-CPT). The Journal of Neuroscience 36, 4067–4079. [PubMed: 27053213]

Chang SW, Gariepy JF, and Platt ML (2013). Neuronal reference frames for social decisions in primate frontal cortex. . Nature Neuroscience 16, 243–250. [PubMed: 23263442]

Chang SW, Winecoff AA, and Platt ML (2011). Vicarious reinforcement in rhesus macaques (Macaca mulatta). Frontiers in Neuroscience 5, 27. [PubMed: 21516263]

Cheney DL, and Seyfarth RM (2018). Flexible usage and social function in primate vocalizations. PNAS 115, 1974. [PubMed: 29432157]

Cléry JC, Hori Y, Schaeffer DJ, Menon RS, and Everling S.(2021). Neural network of social interaction observation in marmosets. eLife 10, e65012.

Cohen YE, Russ BE, Davis SJ, Baker AE, Ackelson AL, and Nitecki R.(2009). A functional role for the ventrolateral prefrontal cortex in non-spatial auditory cognition. PNAS 106, 10045–10050.

Courellis HS, Nummela SU, Metke M, Diehl G, Bussell R, Cauwenberghs G, and Miller CT (2019). Spatial encoding in primate hippocampus during free-navigation. PLOS Biology 17, e3000546.

de la Mothe LA, Blumell S, Kajikawa Y, and Hackett TA (2006). Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. J Comp Neurol 496, 27–71. [PubMed: 16528722]

de la Mothe LA, Blumell S, Kajikawa Y, and Hackett TA (2012). Cortical connections of the auditory cortex in marmoset monkeys: Lateral belt and parabelt regions. The Anatomical Record 295, 800–821. [PubMed: 22461313]

Freiwald W, Duchaine B, and Yovel G.(2016). Face Processing Systems: From Neurons to Real-World Social Perception. Ann Rev Neurosci 39, 325–346. [PubMed: 27442071]

Fusi S, Miller EK, and Rigotti M.(2016). Why neurons mix: high dimensionality for higher cognition. Curr Op Neurobiol 37, 66–74. [PubMed: 26851755]

Gifford GW, MacLean KA, Hauser MD, and Cohen YE (2005). The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. J Cog Neurosci 17, 1471–1482.

Gothard KM (2020). Multidimensional processing in the amygdala. Nature Reviews Neuroscience 21, 565–575. [PubMed: 32839565]

Guilford T, and Dawkins MS (1991). Receiver psychology and the evolution of animal signals. Anim Behav 42, 1–14.

Hung CC, Yen CC, Ciuchta JL, Papoti D, Bock NA, Leopold DA, and Silva AC (2015). Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. The Journal of neuroscience : the official journal of the Society for Neuroscience 35, 1160–1172. [PubMed: 25609630]

Hwang J, and Romanski LM (2015). Prefrontal Neuronal Responses during Audiovisual Mnemonic Processing. The Journal of Neuroscience 35, 960. [PubMed: 25609614]

McMahon DB, Russ BE, Elnaiem HD, Kurnikova AI, and Leopold DA (2015). Single-Unit Activity during Natural Vision: Diversity, Consistency and Spatial Sensitivity among AF Face Patch Neurons. J Neurosci 35, 5537–5548. [PubMed: 25855170]

Miller CT, Beck K, Meade B, and Wang X.(2009). Antiphonal call timing in marmosets is behaviorally significant: Interactive playback experiments. . Journal of Comparative Physiology A 195, 783–789.

Miller CT, and Thomas AW (2012). Individual recognition during bouts of antiphonal calling in common marmosets. Journal of Comparative Physiology A 198, 337–346.

Miller CT, Thomas AW, Nummela S, and de la Mothe LA (2015). Responses of primate frontal cortex neurons during natural vocal communication. J Neurophys 114, 1158–1171.

Miller CT, and Wang X.(2006). Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. Journal of Comparative Physiology A 192, 27–38.

Mitchell J, Reynolds J, and Miller CT (2014). Active vision in marmosets: A model for visual neuroscience. J Neurosci 34, 1183–1194. [PubMed: 24453311]

Nummela S, Jovanovic V, de la Mothe LA, and Miller CT (2017). Social context-dependent activity in marmoset frontal cortex populations during natural conversations. . J Neurosci 37, 7036–7047. [PubMed: 28630255]

Nummela SU, Jutras MJ, Wixted JT, Buffalo EA, and Miller CT (2019). Recognition Memory in Marmoset and Macaque Monkeys: A Comparison of Active Vision. J Cog Neurosci 31, 1318–1328.

Parthasarathy A, Herikstad R, Bong JH, Medina FS, Libedinsky C, and Yen S-C (2017). Mixed selectivity morphs population codes in prefrontal cortex. Nature Neuroscience 20, 1770–1779. [PubMed: 29184197]

Paxinos G, Watson CS, Petrides M, Rosa MGP, and Tokuno H.(2012). The marmoset brain stereotaxic coordinates. (New York, NY: Academic Press).

Perrodin C, Kayser C, Logothetis NK, and Petkov C.(2011). Voice cells in primate temporal lobe. Current Biology 21, 1408–1415. [PubMed: 21835625]

Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, and Logothetis NK (2008). A voice region in the monkey brain. Nature Neuroscience 11, 367–374. [PubMed: 18264095]

Plakke B, Ng CW, and Poremba A.(2013). Neural correlates of auditory recognition memory in primate lateral prefrontal cortex. Neuroscience 244, 62–76. [PubMed: 23583656]

Rigotti M, Barak O, Warden M, Wang X, Daw N, Miller E, and Fusi S.(2013). The importance of mixed selectivity in complex cognitive tasks. Nature 497, 585–590. [PubMed: 23685452]

Romanski LM, Averbeck BB, and Diltz M.(2005). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophys 93, 734–747.

Schaeffer DJ, Selvanayagam J, Johnston KD, Menon RS, Freiwald WA, and Everling S.(2020). Face selective patches in marmoset frontal cortex. Nature Communications 11, 4856.

Sliwa J, and Freiwald WA (2017). A dedicated network for social interaction processing in the primate brain. Science 356, 745–749. [PubMed: 28522533]

Stephens GJ, Silbert LJ, and Hasson U.(2012). Speaker-listener neural coupling underlies successful communication. PNAS 107, 14425–14430.

Toarmino C, Wong L, and Miller CT (2017). Audience affects decision-making in a marmoset communication network. Biology Letters 13, 20160934.

Tsao DY, Freiwald WA, Tootell RB, and Livingstone MS (2006). A cortical region consisting entirely of face-selective cells. Science 311, 670–674. [PubMed: 16456083]

Tsao DY, and Livingstone MS (2008). Neural mechanisms for face perception. Ann Rev Neurosci 31, 411–438. [PubMed: 18558862]

Tsao DY, Schweers N, Moeller S, and Freiwald WA (2008). Patches of face-selective cortex in the macaque frontal lobe. Nature Neuroscience 11, 877–879. [PubMed: 18622399]

Wong-Riley M.(1979). Changes in the visual system of monocularly sutured or enucleated cates demonstrable with cytochrome oxidase histology. Brain Res 1717, 11–28.

Zhang W, and Yartsev MM (2019). Correlated Neural Activity across the Brains of Socially Interacting Bats. Cell 178, 413–428.e422. [PubMed: 31230710]

Highlights.

- Primate PFC neurons were recorded in traditional to naturalistic contexts

- Neural responses to vocalizations were highly behavioral context specific

- Contextual selective responses evident within single neurons and across the population

- Facets of neural activity - social state – may only be evident in natural contexts
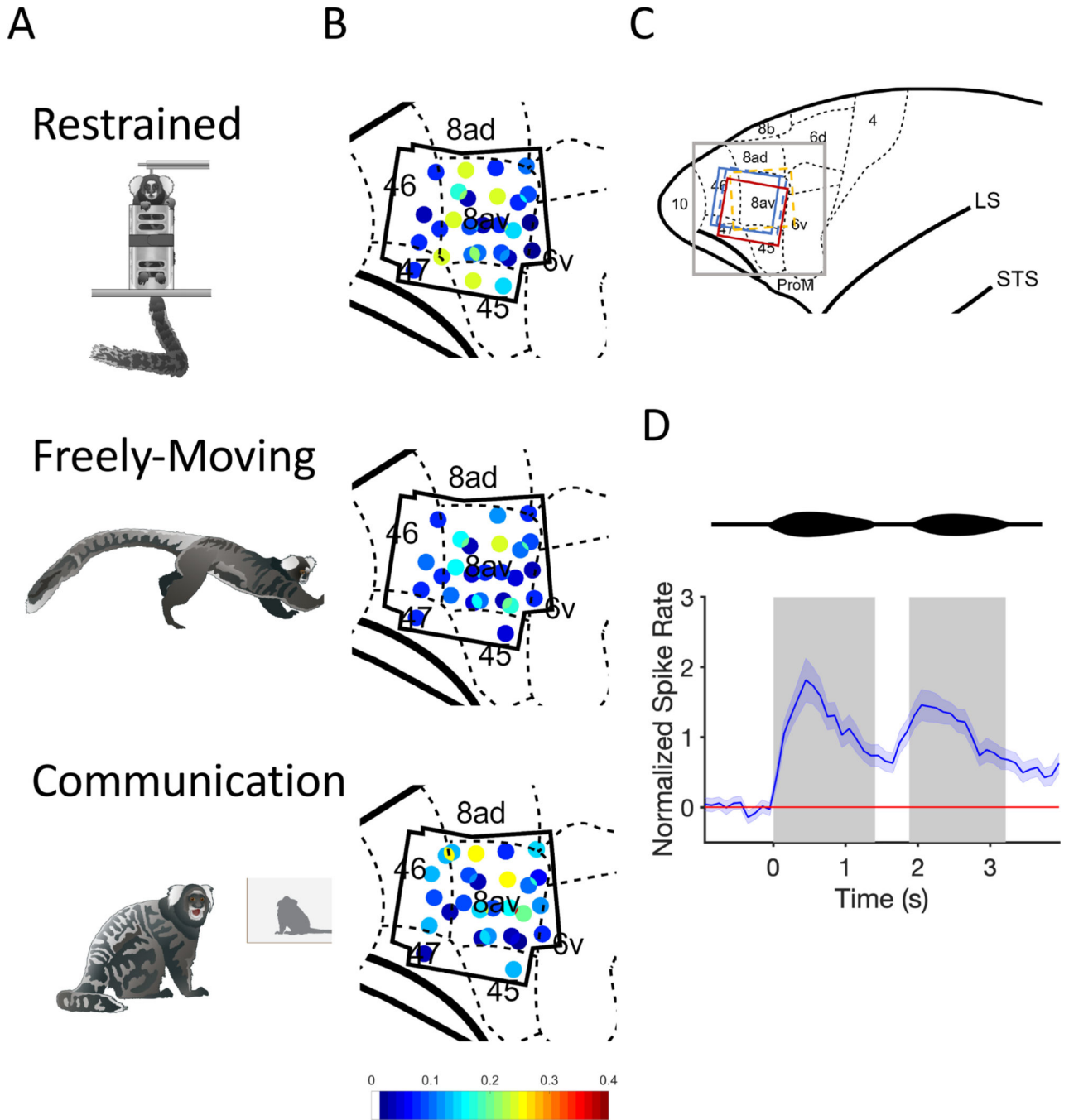
Figure 1. Behavioral Contexts and Recording Locations

(A) Drawings of the three behavioral contexts. (B) Distribution of phee responsive neurons in marmoset PFC for all three contexts. Black polygonal shape represents the outline of the four electrode arrays across three marmosets. Each dot represents one electrode channel and its ratio (indicated by the color bar) of phee responsive neurons relative to all single neurons recorded at that location. (C) Anatomical map of the frontal cortex. Gray square represents the zoomed in portion of PFC depicted in B. Colored squares represent the position and orientation of the four electrode arrays. Dashed lines = right hemisphere implants, Solid
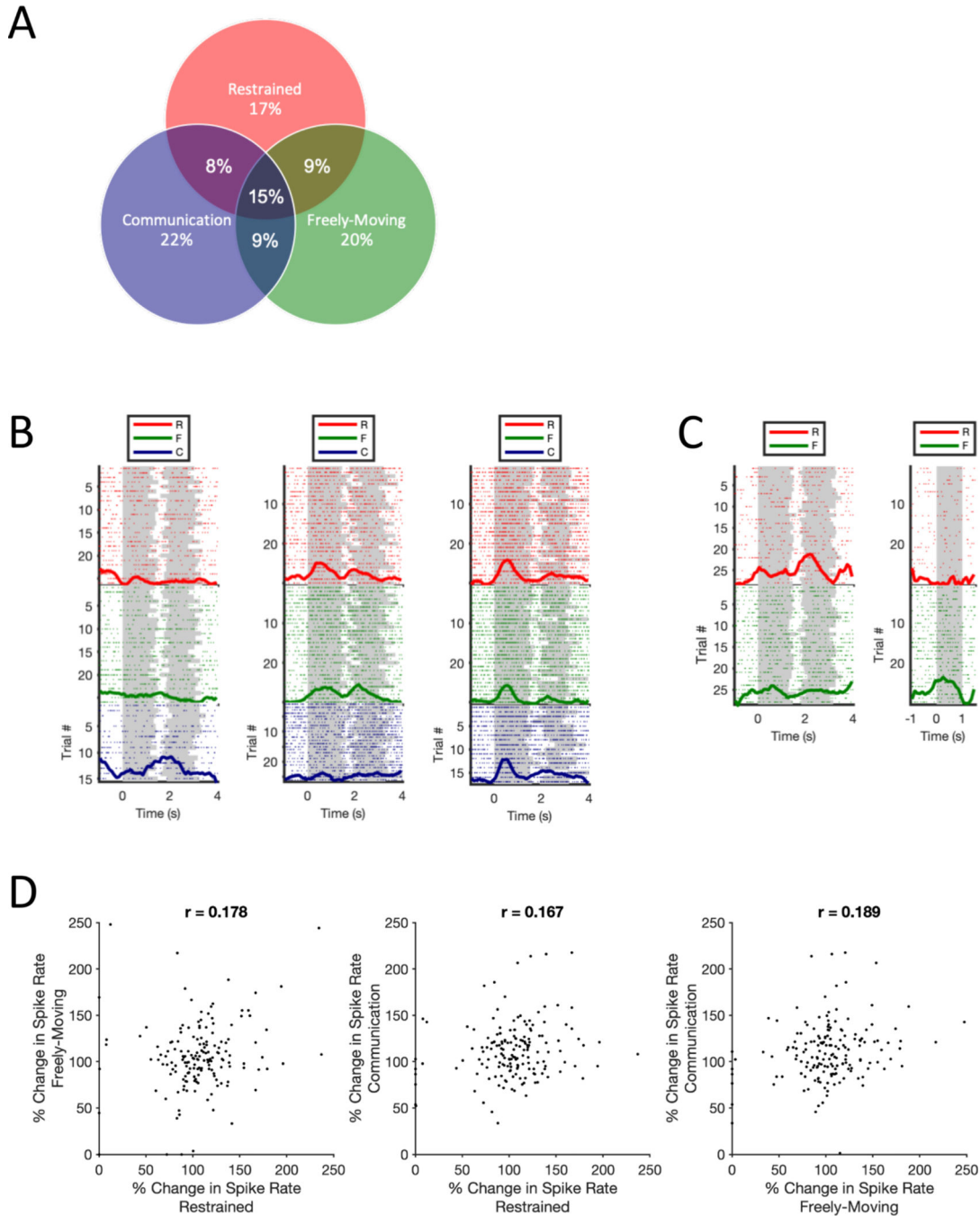
lines = left hemisphere. Color bar indicates the proportion of pheeresponsive neurons, which are shown in the colored dots in the anatomical map (D) Two-pulsed marmoset phee call is shown above a normalized PSTH with 95% Confidence Interval for all phee-responsive neurons. Gray boxes indicate the average duration of phee pulses.

Figure 2. Within-Neuron Context Comparisons

(A) Venn Diagram shows the percentage of significant phee responsive neurons (n=170) for all combinations of behavioral contexts. (B) Exemplar PFC neurons showing within-neuron contextual selectivity observed in this population. Raster plots show neural responses to individual phee call trial presentations on each row, while the normalized PSTH is shown as a thick dark line for each context. Grey shading on each trial indicates the onset/offset of each phee call pulse or noise stimuli. (C) Exemplar PFC neuron exhibits contextual and stimulus selectivity (D) Scatter plots show the percent change in neural activity to phee
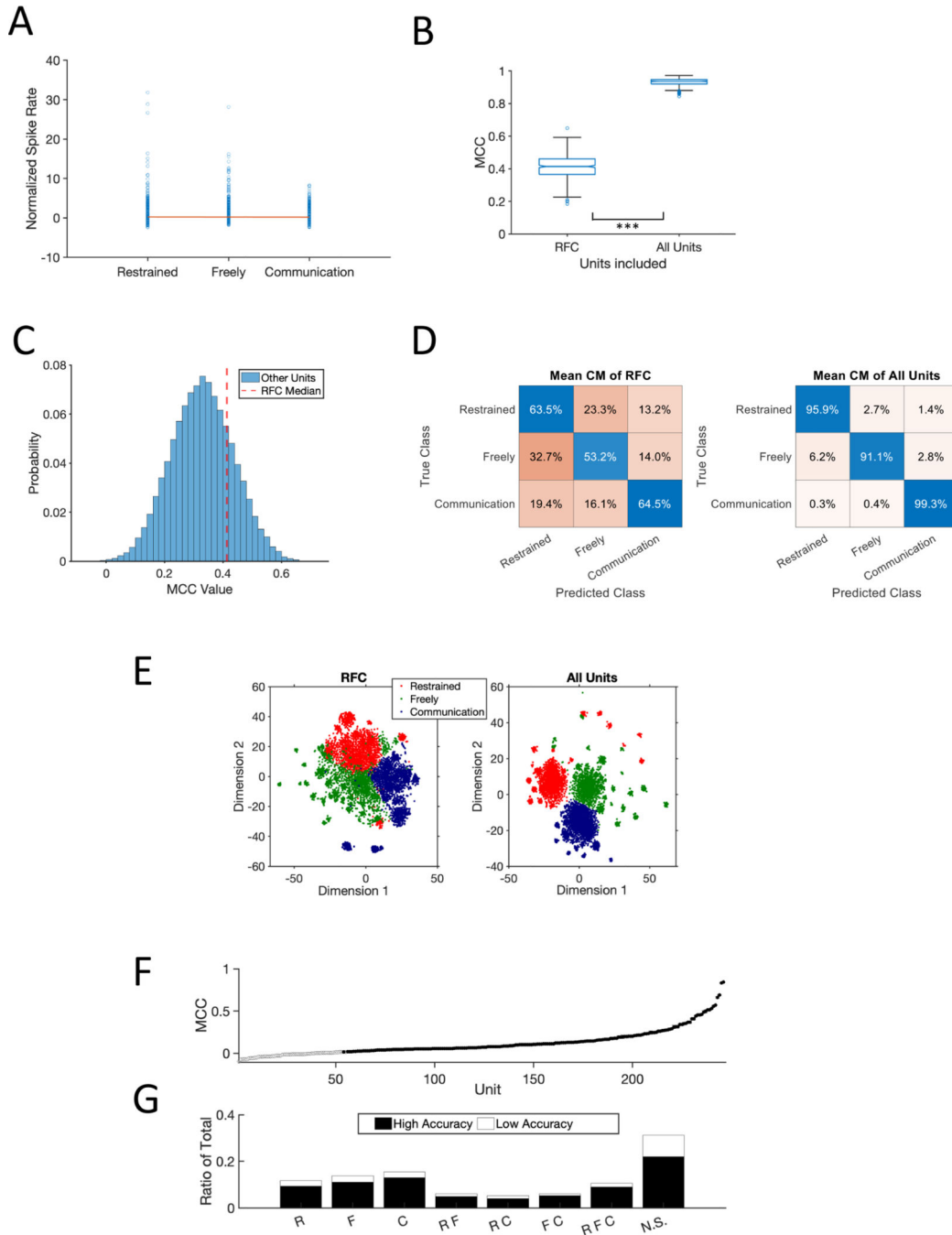
calls relative to baseline in each paired combination of contexts for the 170 phee responsive neurons. Median firing rates are based on 30 phee call presentations in R and F and mean 46 (range 15–93) in C.

Figure 3. Context Differences Across the Population

(A) Normalized firing rate to phee calls in each context (See STAR Methods). Single blue circles represent each individual phee-responsive neuron. The red line was constructed from the slope and intercept of the linear mixed effect model (B) Performance of neural decoder for RFC units and All Units in this population. (C) Distribution of MCC values for 100 randomly selected 26 non-RFC units from the population. The vertical red line plots the median MCC for RFC units. (D) Mean confusion matrix across 500 simulations decoders tested with only RFC units (left) and All Units (right). (E) t-SNE plot (t-Distributed

Stochastic Neighbor Embedding) created by inputting the median performing simulation's PCA. (F) Individual unit performance in decoding neural response in each behavioral context. 99.9% confidence intervals are shown but too small to be visible. Filled in points represent units with significantly higher response above chance in accuracy ($p = 0.001$). (G) Distribution of units that had significantly higher accuracy above chance (black) compared to below chance (white) binned across the category types from panel A. NS = not significant. * statistically significant.
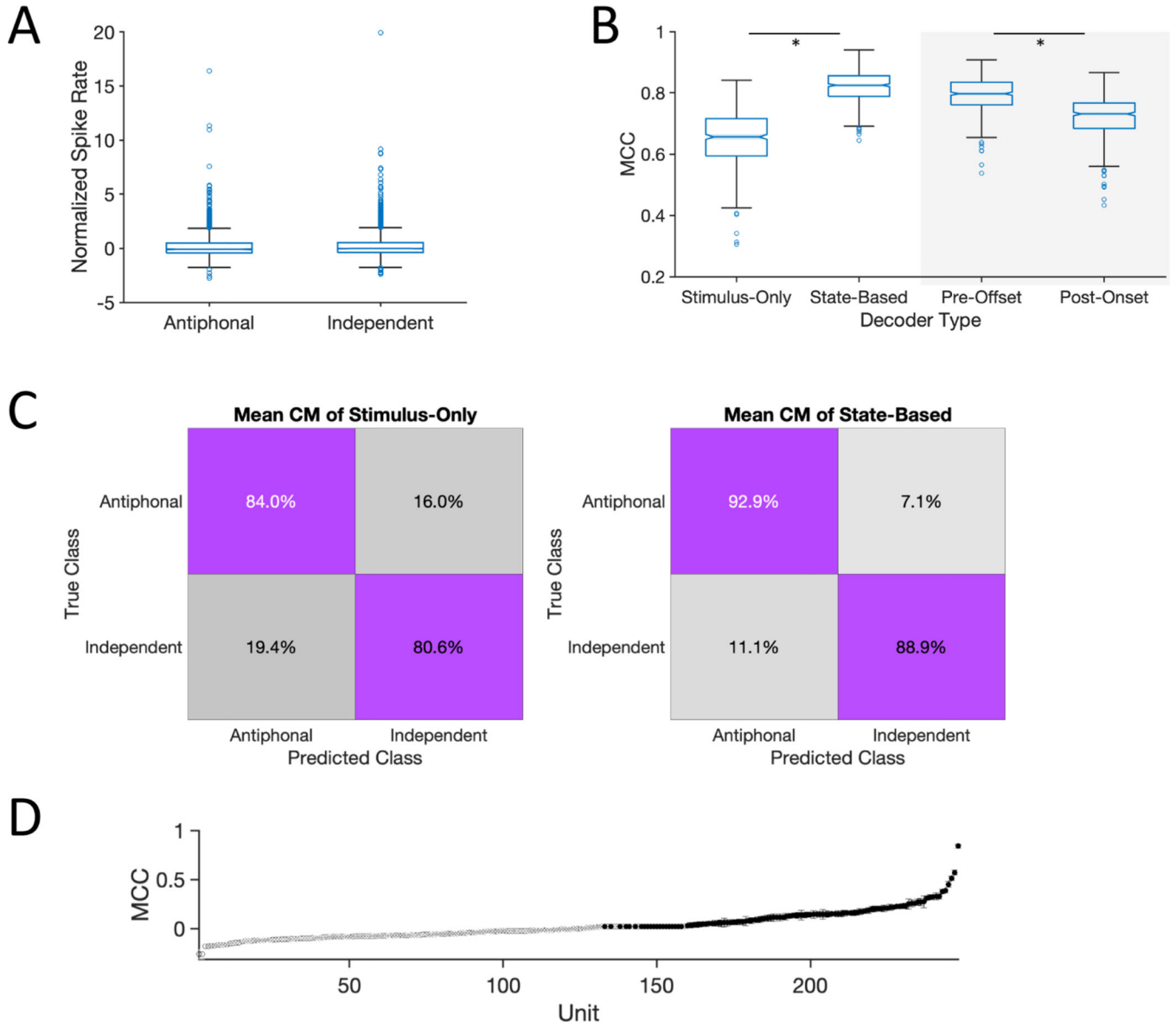
Figure 4. Social State Modulation of PFC
(A) Box plots show the normalized firing rate of neurons phee responsive neurons in the 'Antiphonal' and 'Independent' contexts. Single blue circles represent individual neurons. (B-Left) Performance of 'Stimulus-Only' and 'State-Based' neural decoders. (B-Right) Performance of PreStim and PostStim neural decoder. (C) Mean confusion matrix for decoder performance for Stimulus-Only (left) and State-Based (right). (D) Individual unit performance of units in the State-Based decoder. 109 neurons with accuracies above chance with 99.9% confidence are shown in black. 99.9% confidence interval bounds are plotted but are too small to be visible. * statistically significant.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Bacterial and virus strains		
Biological samples		
Chemicals, peptides, and recombinant proteins		
Critical commercial assays		
Deposited data		
Data and Code	Dryad	doi:10.6076/D11P4N
Experimental models: Cell lines		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Oligonucleotides		
Recombinant DNA		
Software and algorithms		
Other		