Normalization by Valence and Motivational Intensity in the Primary Sensorimotor Cortices (PMd, M1 and S1)

Abbreviated Title: Normalization by Affect in the Sensorimotor Cortices

Zhao Yao[2*], John P Hessburg[2*] and Joseph Thachil Francis[1,3]

Correspondence: joey199us@gmail.com

Number of Pages: 41

Number of figures: 11

Number of tables: 5

Number of words for abstract: 196

Number of words for introduction: 589

Number of words for discussion: 1499

Conflict of interest statement: None.

Acknowledgements: Research was supported by NIH 1R01NS092894-01, NSF IIS-1527558, DARPA REPAIR Project N66001-10-C-2008.

[1] Department of Biomedical Engineering and Electrical and Computer Engineering, Cullen College of Engineering at The University of Houston, Houston, TX77204, United States;

[2] Department of Physiology and Pharmacology, State University of New York Downstate Medical Center, Brooklyn, NY11203, United States;

* Co-First Authorship.

[3] Lead Contact - joey199us@gmail.com

Abstract

Our brain's ability to represent vast amounts of information, such as continuous ranges of reward spanning orders of magnitude, with limited dynamic range neurons, may be possible due to normalization. Recently our group and others have shown that the sensorimotor cortices are sensitive to reward value. In order to determine if normalization plays a role in the sensorimotor cortices when considering non-sensorimotor variables, such as valence and motivational intensity, we had two non-human primate (NHP) subjects (one male bonnet macaque and one female rhesus macaque) make cued grip-force movements while simultaneously cueing the level of possible reward if successful, or timeout punishment if unsuccessful. We recorded simultaneously from 96 electrodes in each the somatosensory, motor and dorsal premotor cortices (S1, M1, PMd). We utilized several normalization models for valence, and motivational intensity in all three regions. We found three types of divisive normalized relationships between neural activity and the representation of reward and punishment, linear, sigmodal and hyperbolic. The hyperbolic relationships resemble receptive fields in psychological affect space, where a unit is particularly sensitive to a small range of the valence/motivational space. We found that these cortical regions have both strong valence and motivational intensity representations.

Significance Statement: Brain machine interfaces (BMIs) are likely to make their way into the clinical setting in the future. Increasing stability of brain derived control of such BMI systems is one essential aspect towards user acceptance, and stability must be maintained no matter the emotional state of the user. However, it is well known that we move faster for rewards of higher magnitude, indicating that emotions influence the motor control system, where BMI control signals come from. Here we report widespread affective modulation of the sensorimotor regions (PMd, PMv, M1 and S1) by cued levels of possible reward if successful and time-out punishment if unsuccessful in non-human primates, and that affect divisively normalizes these regions activity.

Introduction

Neural networks within the primary sensorimotor cortices need to function within a limited dynamic range while encoding multiple forms of information such as kinematics and kinetics (Georgopoulos et al., 1982, 1992; Li et al., 2001; Chhatbar and Francis, 2013). The nervous system could in part increase its information carrying capacity by utilizing a temporal code, which would be less sensitive to a limited dynamic range than a rate code (Bialek et al., 1996). Another possibility is provided by divisive normalization (Carandini and Heeger, 2011), where the input from another region and/or from the local population, can dynamically normalize the network so neural responses remain within a finite range. Such divisive normalization has been suggested as a canonical neural computation, and has been seen in neural systems from primary sensory cortical areas (Heeger, 1992; Schwartz and Simoncelli, 2001; Sato et al., 2016), regions involved in multisensory integration (Ohshiro et al., 2017) to networks involved in decision-making (Louie et al., 2011, 2014; Gluth et al., 2020; Webb et al., 2020).

Recently, it has been shown that the primary motor (M1) and somatosensory (S1) cortices encode reward in addition to kinematics and kinetics (McNiel et al., n.d.; Marsh et al., 2015; Ramkumar et al., 2016; Ramakrishnan et al., 2017; An et al., 2019; Moore and Francis, 2020). However, there has been little mention of psychological affect more broadly in the neurophysiological literature on sensorimotor integration, control, and learning. Affect here refers to information related to the internal representation of feelings or emotions (Rolls, 2014). Valence is the value associated with a task or object that ranges from positive to negative, reflecting stimuli ranging from negative valence (punishing) to positive valence (rewarding). These are associated with approach and avoidance behaviors, respectively. In contrast to valence, the interaction between reward and punishment can also be represented as motivational intensity, or motivational salience. Here, both reward and punishment would increase motivation, as an individual is driven to avoid punishment and obtain the reward, and a neutral stimulus would have zero motivation in this sense.

While many brain regions have been studied in terms of how they encode various aspects of affect, this has not been explored extensively in the primary sensorimotor cortices, although many tasks involving sensorimotor control may be impacted by these variables (Marsh et al., 2015; Ramkumar et al., 2016; Ramakrishnan et al., 2017; Tarigoppula et al., 2018; Zhao et al., 2018; An et al., 2019). We therefore characterized neural activity due to cued possible reward and simultaneously cued possible punishment in S1, M1 and PMd. By cuing both reward and punishment in this manner we could modulate not only valence, but also motivational intensity (Roesch and Olson, 2004), or perhaps arousal.

Here we demonstrate that neural modulation associated with both cued reward and punishment in an operant conditioning task, where reward or punishment were delivered according to an individual's performance, are encoded in S1, M1, and PMd. The neural firing rates and time to complete a trial, can be modeled well by divisive normalization models. We obtained similar results using either a divisive term based on a combined variable comprised of the cued reward level and a scaled version of the cued punishment level, or simply a function of the brain region's population activity. We found a strong representation of both valence and motivational intensity, and that the dominant of these two variables depends on the time period in a trial, such as post-cue vs. post-feedback, and that brain regions responded similarly, indicating a possible broadcast like signal led to our results.

Methods

Behavioral task and rational

In this work we hypothesized that the sensorimotor cortices (PMd, M1 and S1) represent, or are at least modulated by, cued affective information. As these regions are involved in producing movement, we wanted to use a task that had no movement by the non-human primates (NHPs) during the immediate post affective cue period, or at any stage of the task.

These NHPs and this task were also used during brain machine interfacing experiments (Zhao et al., 2018), after the data under consideration here were taken. The periods of time used in our analysis presented here were not during any actual movement produced by the NHP, and as seen below we utilized an isometric task such that the NHPs had no kinematics to plan at any stage of the task.

Two NHPs, one male rhesus macaque (NHP S, Macaca mulatta) and one female bonnet macaque (NHP P, Macaca radiata), were trained to perform the above isometric grip force task (Fig. 1). In this task, subjects controlled certain aspects of a simulated anthropomorphic robotic arm (Barrett WAM) interacting with a simulated cylindrical object. Each trial consisted of 6 stages: cue display, autonomous robot reaching, manual isometric grip force output, autonomous robot transporting, manual isometric releasing, and feedback, which could be reward delivery if it was a successful trial, or a timeout period if not. At the start of a trial, affective cues were displayed at the top of the virtual environment. The number of green squares indicated the level of fruit juice reward that the NHP would receive upon successful completion of the task, with the number present (0 - 3) corresponding to the number of 0.5 second juice delivery periods. If no green squares were displayed, then no reward was delivered upon successful completion of the trial. Red squares indicated the level of time-out punishment the NHP would receive if the trial was completed unsuccessfully. The number of red squares (0 - 3) corresponded to the number of 5 second time-out periods, where a transparent red screen was displayed on the video monitor over the environment and the animal had to wait for the next trial to start. For unsuccessful trials the trial was repeated at the same reward and punishment level until completed successfully, thus further motivating the subjects to complete even non-rewarding trials successfully, and not allowing animals to purposefully fail and skip low value trials, which they did before we moved to this format.

During this isometric grip force task, the virtual arm reached to a target cylinder automatically (Fig.1), and the NHP controlled the robotic grasping motion via isometric force output on a force transducer with its right hand. The amount of force applied was represented in the virtual

environment by a red rectangle that increased in width proportional to the force output, and by the fingers of the robotic hand grasping the object as this force was applied. The subject had to apply, then maintain, a level of force within a range indicated by a pair of blue force target rectangles. The robotic arm then automatically moved the cylinder to a target location (purple circle in figure 1, transporting) while isometric force was maintained. When the arm reached the target location the NHP released the robotic gripper by zeroing its isometric force output, placing the cylinder at the target location resulting in a successful trial. If the NHP completed the trial successfully, they received a juice reward based on the number of green cues displayed at the beginning of the trial, which were present throughout. If the animal failed the trial by not applying force at the proper time, or by applying too much or too little force during the transport period, they received a time-out punishment based on the number of red cues, also initially displayed and visible throughout. After reward or punishment delivery, hereafter termed feedback, the robotic arm retreated automatically horizontally leftward to the starting position, and the next trial began.

Figure 1. Grip Force Task

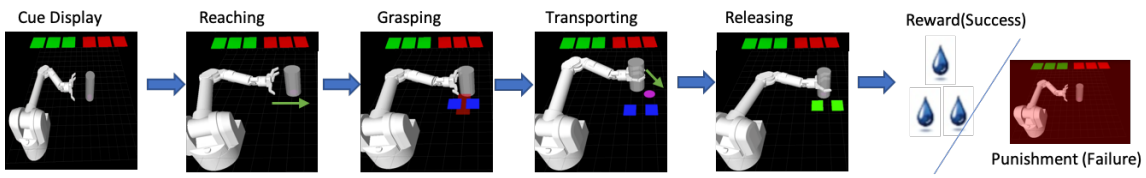


Figure 1. Cued grip force task. The behavioral task was composed of 6 scenes for each trial. First, cue and punishment levels were projected into the virtual environment during the cue display scene, with each green square indicating 0.5 seconds of juice reward delivery if the trial was completed successfully, and each red square indicating the number of 5 second time-out periods if the trial was not completed successfully. The virtual robotic arm then automatically moved to the cylindrical object during the reaching scene. Isometric force was applied manually during the grasping scene, represented on screen by an expanding red rectangle that was then maintained within the blue force target rectangles during the

154 transporting scene. Once the object reached the destination, purple circular target, the target
155 turned green, signaling that the NHP could zero its force output/release the virtual object. If
156 these steps were completed successfully the task entered the success scene, and juice was
157 delivered according to the cued amount. If at any point the NHP applied too much or too little
158 force, the trial was deemed a failure and the failure scene was entered where the amount of
159 time-out period punishment was delivered according to the cued amount, the task screen
160 turned a transparent red, and the trial was reset and could be attempted again.

161 **Surgery**

162 Once trained on the task, NHPs S (male) and P (female) were implanted in their left M1, S1,
163 and PMd with chronic 96-channel platinum microelectrode arrays (Utah array, 10X10 array
164 with 400 μm inter-electrode spacing, ICS-96 connectors, Blackrock Microsystems. 1.5 mm in
165 length for M1 and PMd, 1.0 mm for S1). The hand and arm regions contralateral to their
166 dominant (right) hands were implanted with the same technique as utilized in our previous
167 work (Chhatbar et al. 2010; Marsh et al. 2015a). All surgical procedures were conducted in
168 compliance with guidelines set forth by the National Institutes of Health Guide for the Care
169 and Use of Laboratory NHPs and were approved by the State University of New York
170 Downstate Institutional Care and Use Committee. Briefly, NHP preparation and anesthesia
171 were performed directly by members of the State University of New York Downstate Division
172 of Comparative Medicine veterinary staff, and the researchers conducted the surgery under
173 their supervision. Ketamine was used to induce anesthesia, and isofluorane and fentanyl were
174 used for maintenance. Aseptic conditions were maintained throughout surgery as a
175 craniotomy window was created over the target location, a probing electrode was used to
176 identify the hand and forearm region of S1. An electrode array was implanted there, and
177 another immediately across the central sulcus for the corresponding M1 region. The PMd
178 array was placed just dorsal to the spur of the arcuate sulcus. Dexamethasone was used to
179 prevent inflammation during the procedure and diuretics such as mannitol and furosemide
180 were available to further reduce cerebral swelling if needed. Both subjects were observed

hourly for the first 12 hours after implantation and were provided with a course of antibiotics (baytril and bicilin) and analgesics (buprenorphine and rimadyl).

Extracellular unit recordings

After a two- to three-week recovery period, spiking and local field potential (LFP) activities were simultaneously recorded with a multichannel acquisition processor system (MAP, Plexon Inc.) while the subjects performed the experimental task. Neural signals were amplified, bandpass filtered from 170 Hz to 8 kHz, sampled at 40 kHz, and thresholded to identify single units. Single- and multi-units were sorted based on their waveforms using principal component (PC)-based methods in Sort-Client software (Plexon Inc) followed by offline spike sorting (Plexon Inc) to identify primarily putative single units.

Divisive normalization modeling

Reward and punishment modulation encoding models of the general form $f_{ij} = g(r_i, p_i)$, were fit to individual unit firing rate data from PMd, M1, and S1, where f_{ij} was the post-cue or post-feedback firing rate for the jth unit during the ith trial, with reward level r_i and punishment level p_i for that trial, and g was a function. For post-cue analysis, the firing rate was averaged over the 500 ms window following cue presentation. In this time period the virtual robotic arm was stationary or autonomously moving horizontally to the right during the reaching scene, approaching the cylinder at a consistent speed, and the NHPs were not yet applying isometric grip force. For post-feedback analysis, firing rates were averaged over the 500 ms following the indication of the result (reward or punishment), thus there were no temporal confounds due to the different reward and punishment delivery time periods.

We designed one linear model (model 1) and two different divisive normalization models (model 2 and 3) for our data analysis seen below. We utilized four possible reward levels (R = [0,1,2,3]) and punishment levels (P = [0,1,2,3]). The total trial number when the reward level $R = m$ and punishment $P = n$ was T_{mn}. In our task, T_{mn} was not always the same for each

(R, P) group. To avoid overfitting some reward and punishment levels, instead of ordinary least squares (OLS), we used weighted least squares (WLS) to fit firing rates. OLS tries to minimize the sum squared error $err = \sum_i (y_i - f(x_i))^2$. Here, y_i is neural data, $f(x_i)$ is the fitted data. WLS tries to minimize the weighted sum squared error, $err = \sum_i w_i (y_i - f(x_i))^2$, where w_i is the weight for the ith trial type. w_i was based on the total trial number for that particular reward and punishment level in the ith trial r_i, p_i. If $r_i = 0$ and $p_i = 0$, $w_i = 1$, if $r_i = m$ and $p_i = n$, $w_i = \frac{T_{00}}{T_{mn}}$. We made the total weight for each reward and punishment level all the same.

Model 1: The first model assumed that a unit's firing rate was the result of a linear relationship between reward and punishment,

$$f_{ij} = \alpha_{1j}(r_i + \gamma_{1j} p_i) + \beta_{1j} \quad \text{(eq. 1)}, \qquad \text{Model 1}$$

where f_{ij} was the post-cue, or post-result firing rate for the jth unit in the ith trial, and r_i, p_i were the reward and punishment levels, respectively, for that trial, γ_{1j} scaled punishment, which was in a different domain than the reward, time vs. juice, and α_{1j} was a scaling factor as well. Note, model #1 is linear, but that linearity should not be confused with a linear relationship mentioned in our work below when considering models 2 and 3, which are divisive normalization model forms. There were four possible reward levels (R = [0,1,2,3]) and four possible punishment levels (P = [0,1,2,3]), for a total of 16 combinations of reward and punishment, or affective stimuli and induced neural states. For each trial the reward and punishment levels were chosen pseudo-randomly from a range of 0 to 3. Equation (1) was fit to firing rates from units for each region (PMd, M1 and S1). Adjusted R squared values were calculated for all models to determine which best explained the data (see below).

Model 2: The second model was a divisive normalization model incorporating reward and punishment levels,

$$f_{ij} = \frac{\alpha_{2j}(r_i+\gamma_{2j}p_i)}{\delta_{2j}+(r_i+\gamma_{2j}p_i)} + \beta_{2j} \text{ (eq. 2)}, \qquad \text{Model 2}$$

where f_{ij} was the post-cue, or post-result firing rate (500ms bin) for the jth unit in the ith trial, and r_i and p_i were the reward and punishment levels for that trial, respectively. γ_{2j} was the weight between reward and punishment for the jth unit and δ_{2j} was an additive factor. Equation (2) was used to determine if the jth unit encoded valence or motivation. If $\gamma_{2j} < 0$ (eq.2) the unit was considered to encode valence, where reward and punishment modulated the firing rate in opposite directions. If $\gamma_{2j} > 0$ the unit was considered to encode motivation, as both punishment and reward modulated in the same direction, as both reward and punishment are motivating factors. Estimated values of $\alpha_{2j}, \gamma_{2j}, \beta_{2j}$, and δ_{2j} were determined with a weighted non-linear least square fit (Matlab curve fitting app, function "fitnlm").

We utilized trial duration as a behavioral measurement and proxy for the NHPs affective state. We fit the trial duration data to a modified equation (2) for trial duration (t_i) for a given reward level (r_i) and punishment level (p_i) combination under the reasonable assumption that these sensorimotor brain regions are at least partially responsible for the behavioral output from the NHPs, and thus if the neural activity is normalized, so should be the behavior.

$$t_i = \frac{A_2(r_i+B_2p_i)}{C_2+(r_i+B_2p_i)} + D_2 \text{ (eq. 3)}, \qquad \text{Model 2.}$$

Model 3: For the third model, the denominator from equation (2) was modified to incorporate neural population information. We have

$$f_{mnj} = \frac{\alpha_{3j}(r_m+\gamma_{3j}p_n)}{\delta_{3j}+l_{mn}} + \beta_{3j} \text{ (eq. 4)}, \qquad \text{Model 3}$$

where $\mathrm{r_m} = [0,1,2,3]$ and $\mathrm{p_n} = [0,1,2,3]$. Similar to model 2, to avoid overfitting we used WLS. The post-cue or post-feedback firing rates for the jth unit with reward level $\mathrm{r_m}$ and punishment level p_n, noted as f_{mnj}, where used to fit model 3. l_{mn} was a scalar based on the population

firing rate for the combination of reward and punishment levels (r_m, p_n). l_{mn} was calculated using the following equation:

$$l_{mn} = \sum_{j=1}^{n} \left| \mu_j^{r_m p_n} - \mu_j^{00} \right|,$$

where j is the index of units, N is the total number of units. $\mu_j^{r_m p_n}$ is the mean post cue or post feedback firing rate for the jth unit on trials with reward level r_m and punishment level p_n, μ_j^{00} is the mean firing rate for the jth unit when reward level = punishment level = 0. Intuitively, l_{mn} is a representation of the difference between the mean population firing rate when reward level is $\mathrm{r_m}$and punishment level is $\mathrm{p_n}$ as compared to the affective "baseline" r=p=0. Using the raw mean firing rate data was not as robust, compared to the absolute value, as individual units may modulate in opposing directions in response to reward or punishment, thus canceling each other's changes out. Therefore, the baseline l_{00} = 0 was set for the combination r=p=0. For all other reward and punishment combinations, l_{mn} was defined as the absolute value of the firing rate differences between the mean unit responses when reward level was $r_m$and punishment level was p_n and the mean firing rate of that unit's baseline, r=p=0. Estimated values of $\alpha_{3j}, \gamma_{3j}, \beta_{3j}$ and, δ_{3j} were determined by fitting the model to the data as with model 2. We also tested the simple mean firing rates as normalization terms for comparison.

We compared fitting results using R^2 between mean firing rates and absolute differences. All significant units had higher R^2 values using the absolute value difference as compared to the simple average rate. We therefore continue in this paper considering only the absolute value difference form of model 3, as part of our goal is to find the best model form that could be used in a brain computer interface (Marsh et al., 2015; An et al., 2018; Zhao et al., 2018). Significant unit here means a unit had residuals with a normal distribution, passed an F-test and had at least one parameter that related to reward or punishment as significantly non-zero (please see next paragraph for more detail).

Motivational intensity and valance encoding: Motivation here is defined as $\gamma_{3j} > 0$, $\gamma_{2j} > 0$ or $\gamma_{1j} > 0$ for a given unit, meaning that reward and punishment both modulated the firing rate in the same direction. Valence is defined as $\gamma_{3j} < 0$, $\gamma_{2j} < 0$ or $\gamma_{1j} < 0$ for a given unit, meaning that reward and punishment had the opposite modulation for a given unit. Notice that one unit can show motivation in one model, but valance in another model (see figure 6, M1 units). We wanted to pick the best model for each unit and make sure it was significant and then determine if that model was for motivation or valance. For every unit, first, we obtained fitting results using models 1, 2 and 3. Second, we found the best model (the model with highest R^2_{adj}, see eq (10) for model comparisons). Third, we conducted statistical tests for the best model. The statistical tests contained the three following parts: first, we tested if the residuals for the best model had a normal distribution (JB test, p<0.05). If this model successfully passed the first test, then second, we ran an F-test for that model to see if that model was significantly different from a constant model (p<0.05). The constant model was defined as $f_j = c_j$, where c_j was the mean firing rate for all data points (without regard to r and p levels). We calculated the F value by:

$$F = \frac{(SS_1 - SS_2)/(P_{\mathrm{ram}} - 1)}{SS_2/(N - P_{\mathrm{ram}})} \quad \text{(eq. 5)}$$

where SS_1 was the weighted sum of the squared error from the constant model, SS_2 was the weighted sum of the squared error from the fit model, P_{ram} was the number of parameters from the fit model, and N was the number of data points. If this model successfully passed the F test in the second step, we then tested if the parameters that related to reward or punishment were significantly different from zero. If true, we defined this unit as significant. For example, in model 2, the parameters α_{2j} and γ_{2j} were tested. For model 3, α_{3j} and γ_{3j} were tested. A parameter is significantly different from zero when its 95% confidence interval does not include zero, if true, we then found if that unit was only significant for reward encoding (only parameters for reward level were significantly different form zero), only significant for punishment encoding (only parameters for punishment level were significantly different form

zero), or significant for both reward and punishment encoding. For a unit which was significance for both reward and punishment encoding, we then determine if that unit was motivational or valance encoding using that best model's γ, see above.

For all significant units, whose best model was model 2, additional analysis was conducted to determine if that unit was hyperbolic or not. A unit is hyperbolic when the fitted model 2 for the unit had a singularity and that singularity fell inside our stimulus range (R from 0 to 3 and P from 0 to 3). Thus, we are most likely underreporting the number of such units here.

Post-feedback data analysis: Models 2 and 3 were fit to post-feedback data as well as the post-cue time period. Equations (2) and (4) were used, where $r_m = [0,1,2,3]$ and $p_n = [0,1,2,3]$, and f_{ij} were the post-feedback firing rate for the jth unit at the ith trial with reward level r_i and punishment level p_i. This feedback delivery period was the 500 ms window following the onset of reward or punishment delivery, or for the zero-reward, zero-punishment trials, the point in the trial at which reward, or punishment would have been delivered. Similarly, to the post-cue analysis, we picked the best model using the above outlined methods, then ran F-tests for each unit to see if that model was significantly different from a constant model ($p<0.05$). For each significant unit, we then classified it as motivational encoding or valence encoding as with the post-cue analysis.

A difference between the post-cue and the post-feedback analysis was that at the beginning of the trial (post-cue), the animals were more likely to incorporate both reward and punishment expectation, as they did not yet know the outcome of the trial, thus there was uncertainty. The post-feedback time period was after the animal had successfully completed or failed the trial, and the animal at that point may only be encoding reward for successfully completed trials, or punishment for unsuccessfully completed trials, at least this was our hypothesis. Thus, we fit encoding models for successful trials and failed trials separately to test this hypothesis. For successful trials, the model was based on reward-only modulation for post-reward data:

$$f_{ij} = \frac{\alpha_j r_i}{\delta_j + r_i} + \beta_j \text{ (eq. 6)}, \qquad \text{Model 4,}$$

for successful trials only, and for failed trials the model was based on punishment-only modulation,

$$f_{ij} = \frac{\alpha_j p_i}{\delta_j + p_i} + \beta_j \text{ (eq. 7)}, \qquad \text{Model 5.}$$

Where r_i or p_i were the reward or punishment level for the ith trial, and f_{ij} was the post-feedback firing rate for the jth unit in the ith trial. δ_j and β_j were parameters fit with non-linear least squares fitting as before. Equation (6) was fit using all data from all successful trials. Equation (7), was fit using all data from all failure trials. Similarly, to models 2 and 3, we could design models 6 and 7 using information from population firing rates as the divisive terms instead of the current reward or punishment level. For successful trials we have

$$f_{mj} = \frac{\alpha_{6j}(r_m)}{\delta_{6j} + (l_m^r)} + \beta_{6j} \quad \text{(eq. 8)}, \quad \text{Model 6,}$$

where l_m^r is the scalar based on reward trial population firing rates when $r = r_m$, l_m was the mean absolute value firing rate difference compared to the baseline, which was r=0, similar to Model 3. Likewise, for failed trials we have

$$f_{mj} = \frac{\alpha_{7j}(p_m)}{\delta_{7j} + (l_m^p)} + \beta_{7j} \quad \text{(eq. 9)}, \quad \text{Model 7,}$$

where again l_m^p is the scalar based on punishment population firing rates when $p = p_m$, l_m^p is the mean absolute value firing rate difference compared to the baseline, which was p=0. WLS was used for models 4~7 to avoid overfitting certain reward or punishment levels.

Comparing fitting results across models

Models we used did not always have the same numbers of parameters, therefore the adjusted R_{adj}^2 values (10-fold cross validation) were calculated and used to compare the models. 10-

fold cross validation is a procedure where the original dataset is split into 10 equal partitions, and from this 9 are used for fitting the model and the 10th is used for determining how well the model generalizes to unseen data. This is then repeated 10 times such that every data partition is used once as the held-out data to determine the generalizability of the model. The R^2_{adj} was determined by,

$$R^2_{adj} = 1 - (1 - R^2)\frac{N-1}{N-Pram-1}, \text{ (eq. 10)}$$

where *N* is the sample size for all models, and $Pram$ is the total number of explanatory variables/parameters in the model, which was *n* = 2 for models 1, 4, 5, 6 and 7, and *n* = 3 for models 2 and 3. During model comparison, we first pick the best model by finding the model with the highest R^2_{adj} (see eq. 10). Then for this model, we ran an F-test (see eq. 5) to determine if that unit was significant or not.

Results

Before describing the divisive normalization analysis on cued reward and punishment trials we show qualitative results in Fig.2 with simple peri-cue and peri-feedback time histograms, for example single units from the three brain regions (PMd, M1 and S1) for the two NHPs showing units that either had a qualitative motivational intensity like relationship with cued reward and punishment, or a valance like relationship, that is the neural rate either increased or decreased in a simple manner as trial valence increased, such as R0P3 with the lowest valence and R3P0 with the highest. Where R stands for reward value and P for punishment value. Likewise, motivational intensity would be highest for R3P3 and lowest for R0P0 and is a symmetric function as compared to the valence function that would be linear from R0P3 to R3P0, in either an increasing or decreasing manner. Note the distribution of response types in Fig.2.

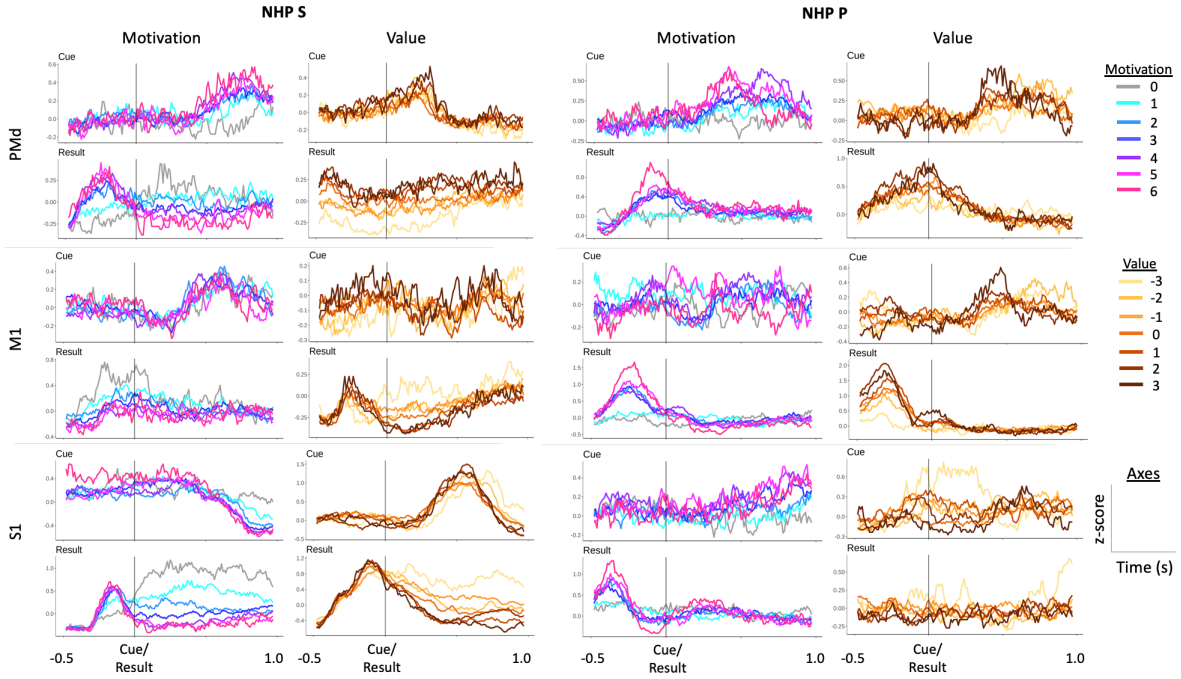


Figure. 2 shows the peri-cue and peri-feedback time histograms for several example single units from each of the brain regions recorded from both NHPs. These units were chosen qualitatively. Here we show units that qualitatively represent either valence-like or motivational-intensity-like response patterns. Units were binned and z-score normalized over the course of a session. This normalized firing rate was then averaged for each valence level, defined as the reward level minus the punishment level for a given trial, or each motivational intensity level, defined as the reward level plus the punishment level for a given trial. Individual units demonstrated a range of responses following cue presentation, and before and after feedback.

We recorded 3 blocks of data, from two different days, for each of the two NHPs. For each block, the number of trials for each R (reward) and P (punishment) level, and the number of units for every region and NHP are shown in Table.1. We tested several normalization models, one linear (model 1), and the rest divisive normalization models. As the divisive normalization (DV) models fit the data better 99-100% of the time, we focus on the DV models. Below we first show results from a DV model that utilized the categorical R and P levels from the task

(eq. 2), which was termed model 2 (Fig.3) and second, results for model 3 (eq. 4), shown in Fig.6. Model 3 utilized the population activity within a brain region in the divisive term (see methods). In Fig.3 we present the trial time results, utilizing cued R and P levels simultaneously (model 2, eq.3). The rational for fitting the trial times to the same divisive normalization model 2, was that we hypothesized the motor output from the sensorimotor regions are producing the isometric force output, at least in part, and therefore, the duration of a trial is dependent on the neural activity within these regions to an extent. By using the same model form we could determine if the neural activity follows the same form as a behavioral measure governed at least in part by the neural activity within the regions under study. We found that trial durations always followed a motivational intensity form more clearly than a valance form. This behavioral measure indicates the NHPs understood the cues, as our task did not involve any choice, other than the NHP simply deciding not to work.

In Fig.3.1 we see that NHP P was very sensitive to differences between rewarding (RX) and non-rewarding (R0PX) trials, where X is a placeholder for any level (0-3). In addition to this, there was an influence on trial time due to the cued punishment level Fig.3.1, which was most clear when the reward level was zero. In these Fig.3 plots the x-axis is termed the affective stimulus and is a combination of the cued reward level and a scaled version of the cued punishment level. This scaling was necessary as the punishment was in timeout periods while the reward levels were in juice volume. In addition, reward and punishment are known to be represented by differing neural circuits. NHP S had a smoother relationship between affective stimulus and trial time. These results demonstrate, as expected, that the mean trial time can be explained with a motivational intensity model, as $B_2 > 0$ for both NHPs in equation 3, which means that reward levels and punishment levels were added together, as expected for motivational intensity, as compared to valence, where reward and punishment levels would be subtracted. The mean trial time decreases as the motivational intensity level increases. In Fig.3.2 we plotted the neural firing rates vs. the affective stimulus, which is itself derived from the data for a given model that determined the relative scaling of punishment compared to

reward. Red squares ± the SEM for each category, of which there were 16, are plotted. In blue we have plotted the fit to the data from model 2,

$$\text{Model 2 } f_{ij} = \frac{\alpha_{2j}(r_i+\gamma_{2j}p_i)}{\delta_{2j}+(r_i+\gamma_{2j}p_i)} + \beta_{2j} \qquad \text{(eq. 2)},$$

with fits that falls into one of three categories: ~linear (Fig. 3.2 row 1), sigmoidal (row 2), and hyperbolic (row 3). These were seen in both animals, and all three cortical regions investigated (PMd, M1 and S1). Note that the aforementioned “linear response” is from the above divisive model 2 that is plotted against the affective stimulus (see Fig.3.2 x-axis) and should not be confused with the simple linear relationship between a unit’s rate and the stimulus, such as that described for model 1 in the methods section. In over 99% of all units model 1 was inferior to the DV models, and thus we do not focus on model 1, see table 3 for best model %.

439 **Table 1: Trial number for each cued reward (R) and timeout punishment (P) Level**

Section 1 for NHP P				
	R = 0	R = 1	R = 2	R = 3
P = 0	10	3	8	10
P = 1	11	14	14	13
P = 2	11	12	8	8
P = 3	22	13	6	6
Section 2 for NHP P				
	R = 0	R = 1	R = 2	R = 3
P = 0	41	5	16	7
P = 1	22	8	13	6
P = 2	7	5	11	5
P = 3	3	10	11	3
Section 3 for NHP P				
	R = 0	R = 1	R = 2	R = 3
P = 0	18	16	21	12
P = 1	13	17	4	12
P = 2	11	16	10	12
P = 3	18	11	12	13

Section 1 for NHP S				
	R = 0	R = 1	R = 2	R = 3
P = 0	18	18	17	21
P = 1	28	18	20	28
P = 2	15	21	21	19
P = 3	18	18	12	16
Section 2 for NHP S				
	R = 0	R = 1	R = 2	R = 3
P = 0	12	17	12	17
P = 1	14	21	14	17
P = 2	13	15	13	14
P = 3	15	13	15	21
Section 3 for NHP S				
	R = 0	R = 1	R = 2	R = 3
P = 0	26	18	12	14
P = 1	17	25	19	34
P = 2	24	14	17	22
P = 3	14	11	18	17

440

441 **Table 2: Number of Units for each recorded region for the two NHPs**

	PMd	M1	S1
NHP P Section 1	140	77	75
NHP P Section 2	83	67	78
NHP P Section 3	85	77	91
NHP S Section 1	85	57	81
NHP S Section 2	87	58	85
NHP S Section 3	109	72	93

442

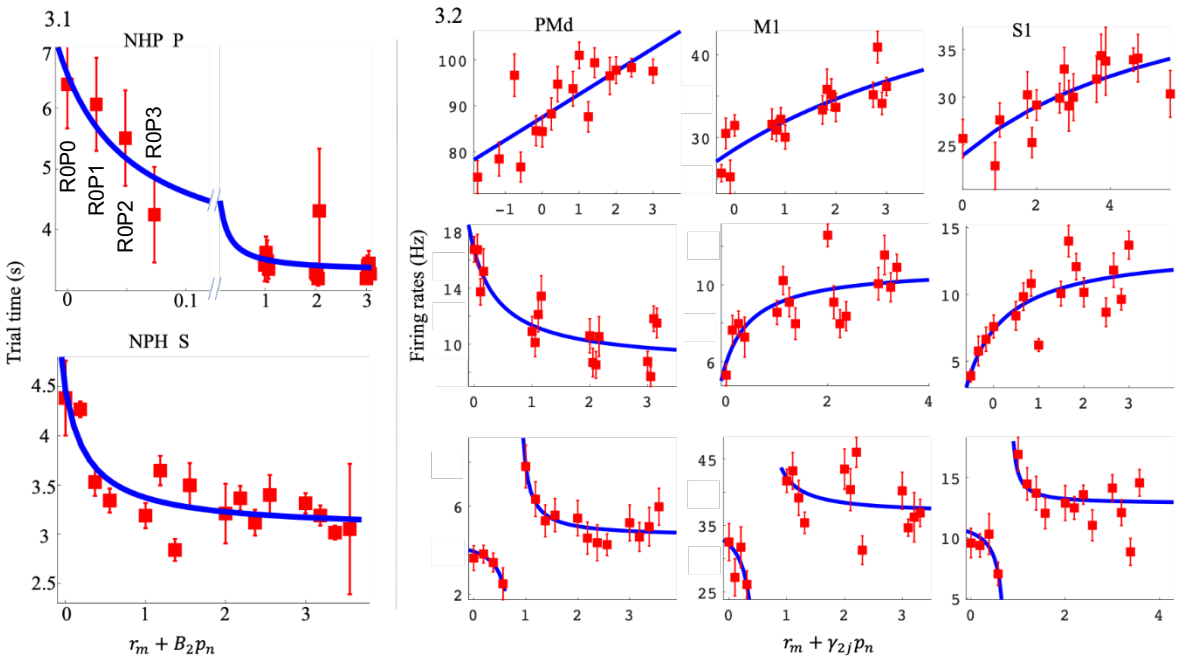


Figure 3. Post-cue reward (R) and punishment (P) analysis with divisive normalization model 2. Figure 3.1 depicts the average and SEM of the trial time (red) for different affective stimuli. The x-axis represents the R and P level in the form of an affective stimulus $r_m + B_2 p_n$, and the y-axis represents the trial time (s). We have labeled the R and P levels for the R = 0 group. Figure 3.2 shows example unit R and P modulation for post-cue neural spike data. Each unit has a R and P encoding model 2 significantly different from a constant model (F-test, $p<0.05$). For each subplot, the x-axis represents the affective stimulus as above, and the y-axis represents the post-cue firing rate (Hz). Each red point represents the mean post-cue firing rate (0-500ms post cue onset) for that R and P level ± the SEM, and the blue line represents model 2 fit to that unit's data. The first column includes units from PMd, the second column M1, and the third column S1. The rows from top to bottom show examples of linear, sigmoidal, and hyperbolic units.

Comparing linear (1) and divisive (2-3) models of valence and motivation

We picked the best model from models 1 (linear, eq.1), 2 and 3 (divisive, eqs.2, 4) for each unit. There were many units whose best model was model 2 (see table 3 and supplementally

material). For these units, we analyzed if they had hyperbolic relations to affective stimuli or not. The results are shown in figure 4. Overall, there were 36% - 100% of the units for a given brain region that were hyperbolic, thus there was a large portion of nonlinear modulation for units whose best model was model 2, with PMd having a clear majority of hyperbolic units in both NHPs post-cue and 100% in NHP P during the post cue time period.

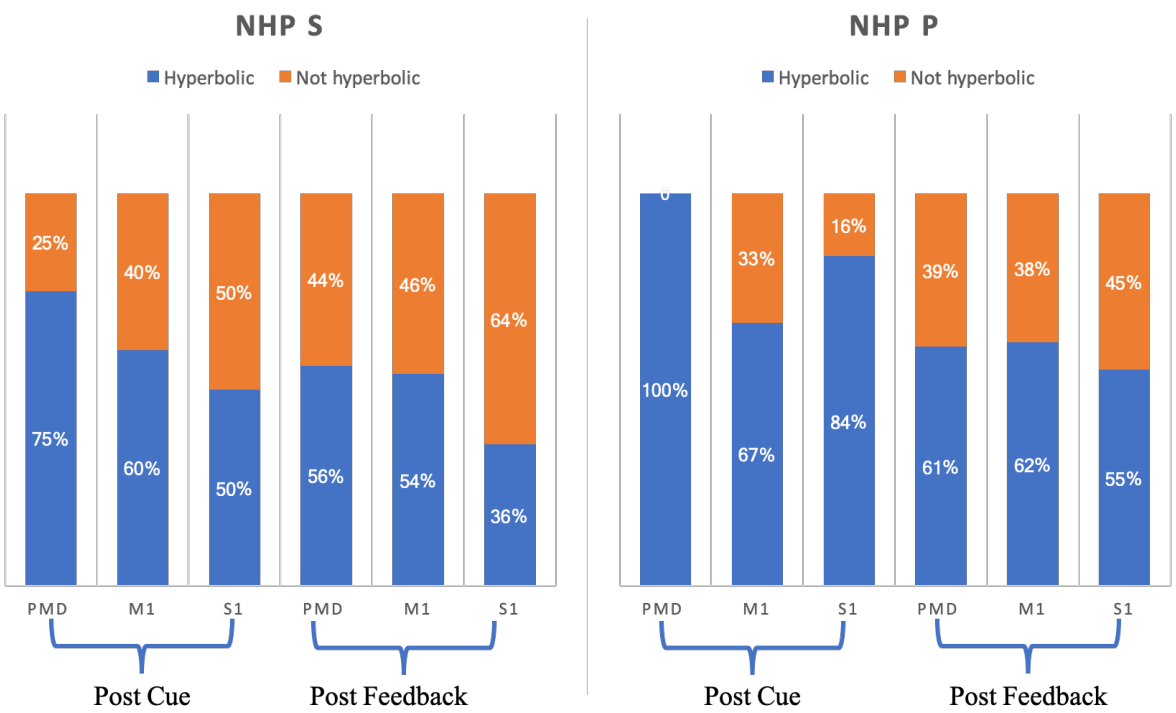


Figure. 4. Number of significant units from model 2 that were hyperbolic**.**

For non-hyperbolic units seen in Fig.4.red, we asked, how many units were sigmoidal vs. linear? To classify the distribution of these curve shapes for linear and sigmoidal units, we first found all significant units whose best model was model 2, $f_{ij} = \frac{\alpha_{2j}(r_i+\gamma_{2j}p_i)}{\delta_{2j}+(r_i+\gamma_{2j}p_i)} + \beta_{2j}$. We analyzed all of the non-hyperbolic unit curve shapes by comparing the two slopes at the start and end points of the curve as seen in examples (Fig.5). We then took the angle of the tangents at A and B, noted as θ_A, θ_B respectively. We calculated the angle difference $\Delta\theta = |\theta_A - \theta_B|$ and used this as a proxy for the curves' shape, linear or sigmodal. If $\Delta\theta$ was "large" (> 30°), the shape was arbitrarily considered sigmoidal, based on the curves seen in Fig.5.a.

If $\Delta\theta$ was < 30° to zero, the shape of the curve was considered linear. We combine all data for each NHP together so as to have enough datapoints for a distribution of the $\Delta\theta$. For NHP S, about 25% of the non-hyperbolic units had a $\Delta\theta$ larger than 30°. For monkey P, about 37% of the non-hyperbolic units had $\Delta\theta$ larger than 30° as seen below in Fig.5. Thus, adding the hyperbolic and the sigmodal unit groups to determine the amount of non-linear responses leads to a majority of units being classified as non-linear in the affective stimulus vs. unit rate space. Keep in mind, even the linear responses are still from a divisive normalization model and plotted in the aforementioned affective stimulus space.

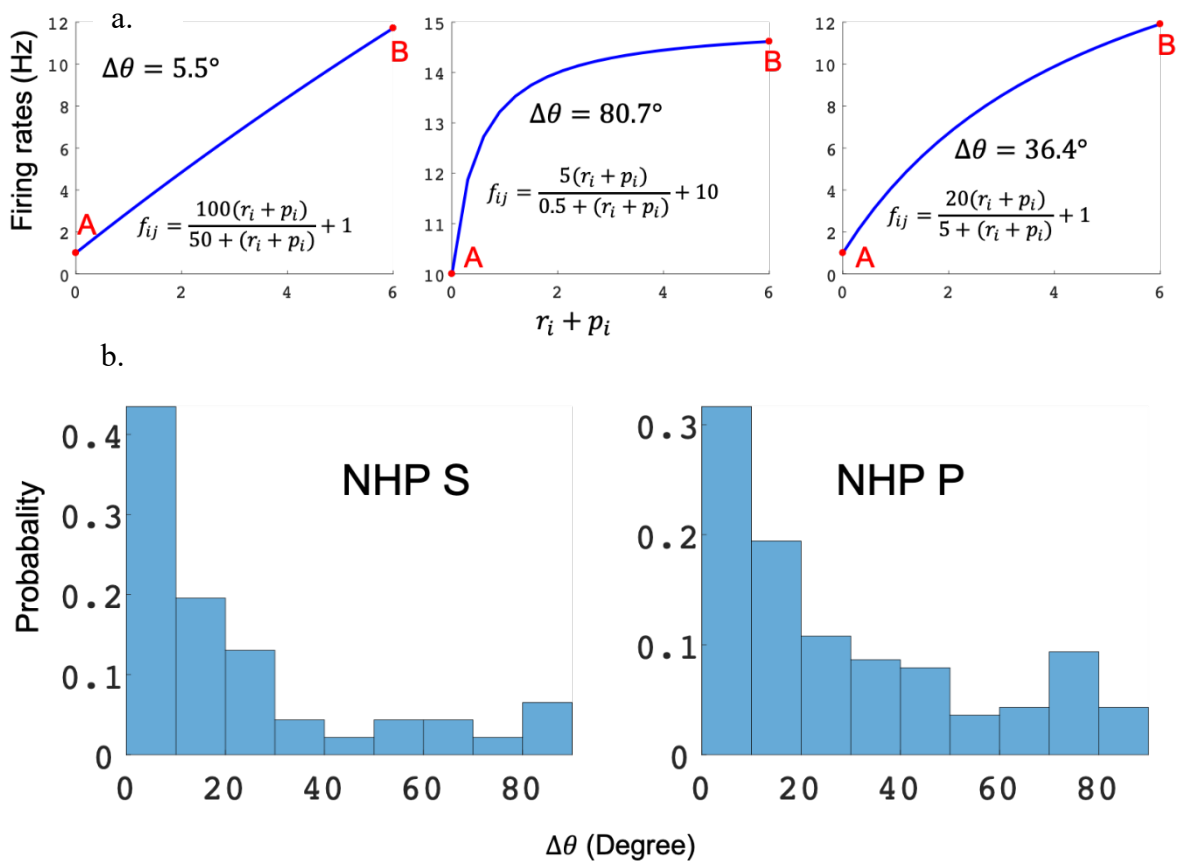


Figure. 5. (a.) Description of method used to determine % of population with sigmodal or linear response curves from the population that were not hyperbolic. (b.) Distribution of $\Delta\theta$ for the non-hyperbolic units using model 2. Note anything above 30° was considered sigmodal, also note the different ranges used in the plots.

As model 2 utilized the categorical R and P levels in the denominator, which act as the divisive term, we wished to determine if a more “natural” divisive term could perform as well or better. We derived the divisive term for model 3 from the given cortical regions population activity as described in the methods section, Model 3 is $f_{mnj} = \frac{\alpha_{3j}(r_m+\gamma_{3j}p_n)}{\delta_{3j}+l_{mn}} + \beta_{3j}$ (eq.6). In Fig.6 we show a comparison for example units between models 2 and 3. Model 3 had a higher fitting accuracy for all example units shown, and visually better “explained” the modulation due to the affective stimuli. Units could have two general characteristics to their responses: focal tuning to a small reward and punishment range, termed hyperbolic, or a more gradual response to a wider range of reward and punishment levels, sigmodal or linear. Model 2 could only capture either the focal tuning (hyperbolic) or the broader modulation (linear or sigmodal units), however, model 3 could capture both types of modulation as seen in Fig.6 (see units in green circles). Here we show the qualitative results for a set of examples and follow with the population results and the model comparisons. Note the x-axis, which represents the affective stimulus fit to the data can change between models 2 and 3, as can the unit’s apparent representation, valence vs. motivation.

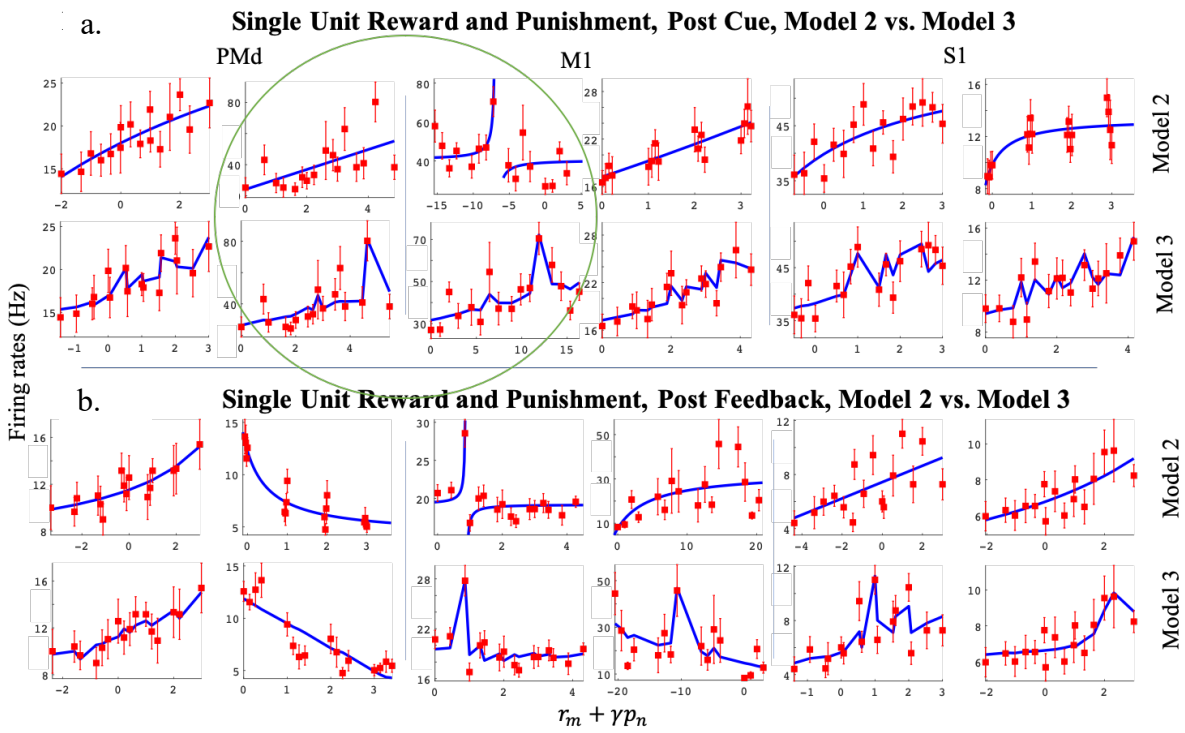


Figure 6. Post-cue (a) and post-feedback (b) example units for model 2 (top rows in each subplot) and model 3 (bottom rows in each subplot). The x-axis represents the affective stimuli, a linear combination of reward and scaled punishment $r_m + \gamma p_n$, where γ is fit to the data. The y-axis represents the post-cue (a.) and post-feedback (b.) firing rate (Hz). The mean ± SEM of the firing rates (red) and the mean for the model (blue) are shown. Each column represents one unit, fit to model 2 (top) and model 3 (bottom) rows. First two columns show units from PMd, third and fourth columns from M1, and last two columns from S1. All units shown here were significant (see methods).

For each unit, we first found the best model among models 1, 2 and 3, where model 1 is a linear non-divisive form, Model 1 is $f_{ij} = \alpha_{1j}(r_i + \gamma_{1j} p_i) + \beta_{1j}$ (eq.1). Second, we tested for each best model, if the parameters related to reward and punishment were significantly different from zero, such that their confidence intervals for the given parameter did not overlap zero (see methods). This led to four cases: 1, parameters related to reward and punishment were both not significant. 2, only the parameter related to reward was significant. 3, only the parameter related to punishment was significant. 4, Parameters related to reward and punishment were both significant. Group 2, 3 and 4 were considered significant units. Only units from group 4 (red pie charts in figure 7.b) could have motivation or valance encoding. Overall, we found approximately 25% - 48% of the units from the full population of a given brain region where significantly modulated by reward, punishment or both. For NHP S, we had 127 significant units among 283 total units in PMd (45%), 63/187 units in M1 (34%) and 125/259 units in S1 (48%). For NHP P, we had 106 significant units among 308 total units in PMd (34%), 67/221 units in M1(30%) and 61/244 units in S1 (25%).

In order to determine if and how units changed their representation of reward and punishment over the course of a trial we looked at the units' responses in the post-cue and the post-feedback time periods and plot this information in Fig.7. Here we have broken the population

of all units into the 4 groups mentioned above. We then further broke the units in group 4 down into motivational intensity and valence. From Fig.7, we can see that, in general units modulate more for reward compared with punishment during the post feedback period. It also appears that more units represent valance in these data sets as compared to motivation, however, the number of units that we could use for this analysis was small due to the units having to pass significance tests for multiple parameters simultaneously as compared to the other categories. Nonetheless, we still obtained 106 total units from 3 brain regions in 2 NHPs that passed such significance criterion, of which 84, or 80% showed a significant valance representation as compared to 20% that showed motivational intensity. As seen in table 4 NHP P experienced a greater number of failed trials as compared to NHP S, and this could have led to NHP P showing a larger percentage of units with punishment only representations as seen in Fig.7.b.

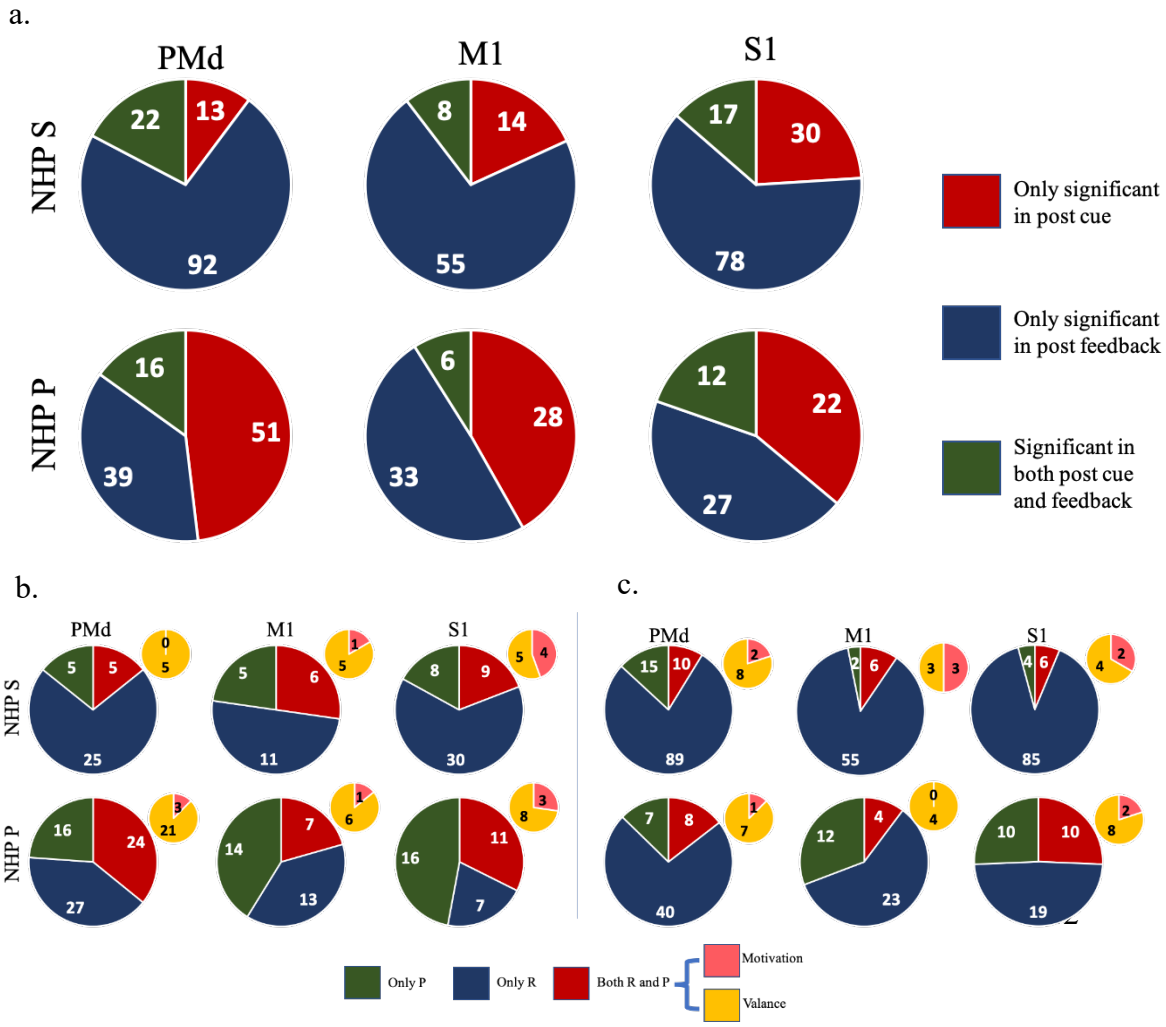
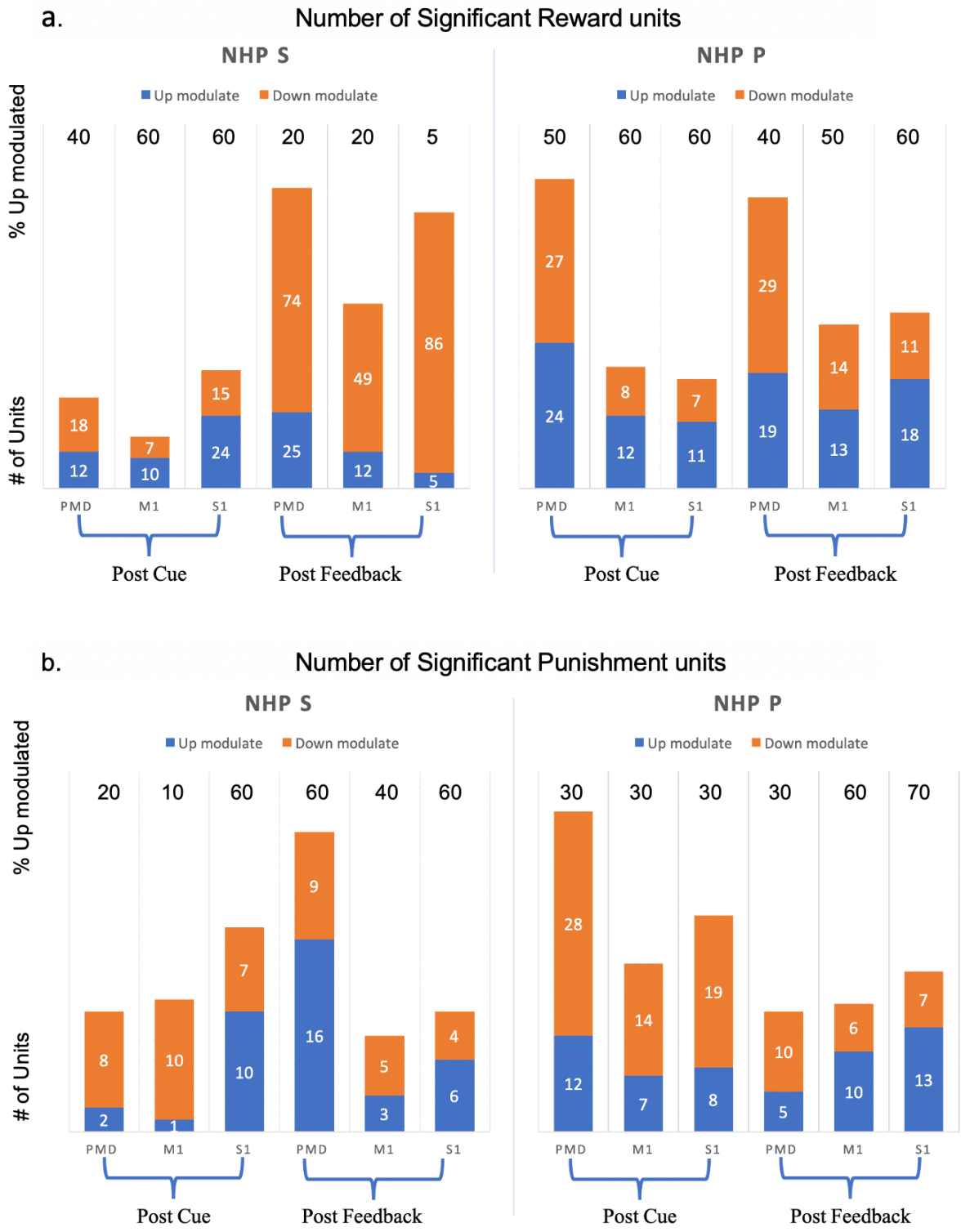


Figure.7. (a.) Time period distribution of significant units from two NHPs for three sensorimotor brain regions (PMd, M1, S1) fit to each unit's best model. (b.) Post cue reward and punishment encoding distributions of significant units using the units' best model (see methods and text). (c.) Same as (b.), but during the post feedback period.

We wished to determine if units up modulate or down modulate their neural activity for the different levels of reward, as well as punishment. A unit was significant only if the parameters related to reward or punishment were significantly different from zero (see methods). Thus, for a unit with a significant reward parameter we tested if that unit up modulated for reward (firing rates for that unit increased when the reward level increased) or down modulated (firing rates decreased when reward level increased), and similarly for the significant punishment units. The results are shown in figure 8.a for reward analysis and 8.b for punishment analysis. Notice that the fraction of upmodulated units for reward modulation in Fig.8.a decrease in general from post-cue to post-feedback, while for punishment the trend is reversed with more units up modulated in the post-feedback as compared to the post-cue time period. These trends are seen for both NHPs and most cortical regions.



562 **Figure. 8.** Modulation (up or down) for units with significant divisive normalization from two
563 NHPs for three sensorimotor brain regions. (a.) shows population analysis for reward
564 modulation. (b.) shows punishment modulation. Orange bars show the number of units with

down modulation. Blue bars show number of units with up modulation. Numbers at the top of each column are the percentage of units with up modulation.

Adjusted R-squared values for every unit using models 1, 2 and 3 for all regions have been compared. In all regions and time periods, in essence all units have better fits for the divisive normalization models 2 and 3 as compared to the linear model 1. All data are shown in table 1 for both NHPs and the three brain regions (PMd, M1 and S1). Please see supplementary material for significant units only S.Fig.3.

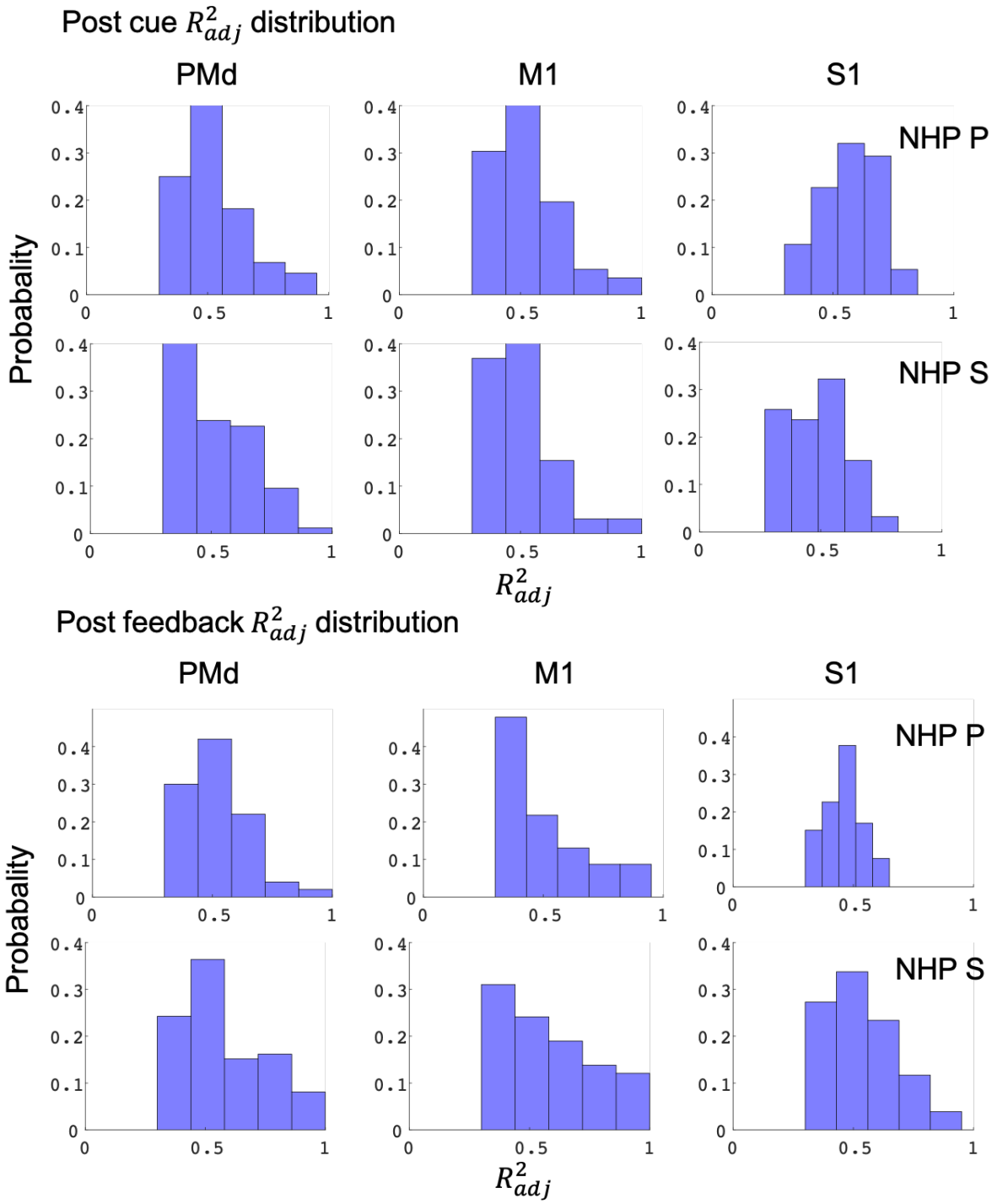
NHP P

	PC PMd	PC M1	PC S1	PF PMd	PF M1	PF S1
Model 2 better than 1	100%	100%	100%	100%	100%	100%
Model 3 better than 1	100%	100%	100%	100%	100%	100%
Model 3 better than 2	53%	66%	87%	38%	62%	67%

NHP S

	PC PMd	PC M1	PC S1	PF PMd	PF M1	PF S1
Model 2 better than 1	100%	100%	100%	99%	100%	100%
Model 3 better than 1	99%	100%	100%	99%	100%	100%
Model 3 better than 2	49%	64%	61%	41%	35%	49%

Table 3. Percentage of all recorded units with better adjusted R-squared for a given model comparison. PC = post cue, PF = post feedback.



582 **Figure. 9**. Distribution of the adjusted R squared values for the best significant model (model
583 2 and 3) fit for a given unit during the post-cue, top two rows, and the post-feedback, bottom
584 two rows, for the two NHPs labeled on the right-hand side of the figure. Here we have fit the
585 models to the mean firing rates for all 16 (R, P) levels, see supplementary figures S.Fig.2 for
586 more on the fit distributions.

Furthermore, we conducted statistical tests to determine if two medians from two R^2_{adj} distributions were significantly different from the data shown in Fig.9. The results are, for NHP S's data: Median R^2_{adj} from PMd during post feedback was significantly higher than all 3 regions during the post cue time period. The median R^2_{adj} from M1, post feedback, was significantly higher than all 3 regions during the post cue period. The median R^2_{adj} from S1, post feedback, was significantly higher than S1 post cue. For NHP P's data: The median R^2_{adj} from PMd post cue was significantly higher than S1, post feedback. M1, post cue was significantly higher than PMd, post cue, and M1/S1, post feedback. S1 post cue was significantly higher than all the other 5 brain regions. PMd post feedback was significantly higher than S1 post feedback. Medians and ranges for all R^2_{adj} distributions are shown in figure 10.a. The corresponding proportion of significant units for all brain regions are shown in figure 10.b. Overall, we can see that for NHP S, there were more significant units in the post feedback period compared with post cue. Also, medians from R^2_{adj} distributions in post feedback period were higher than post cue.

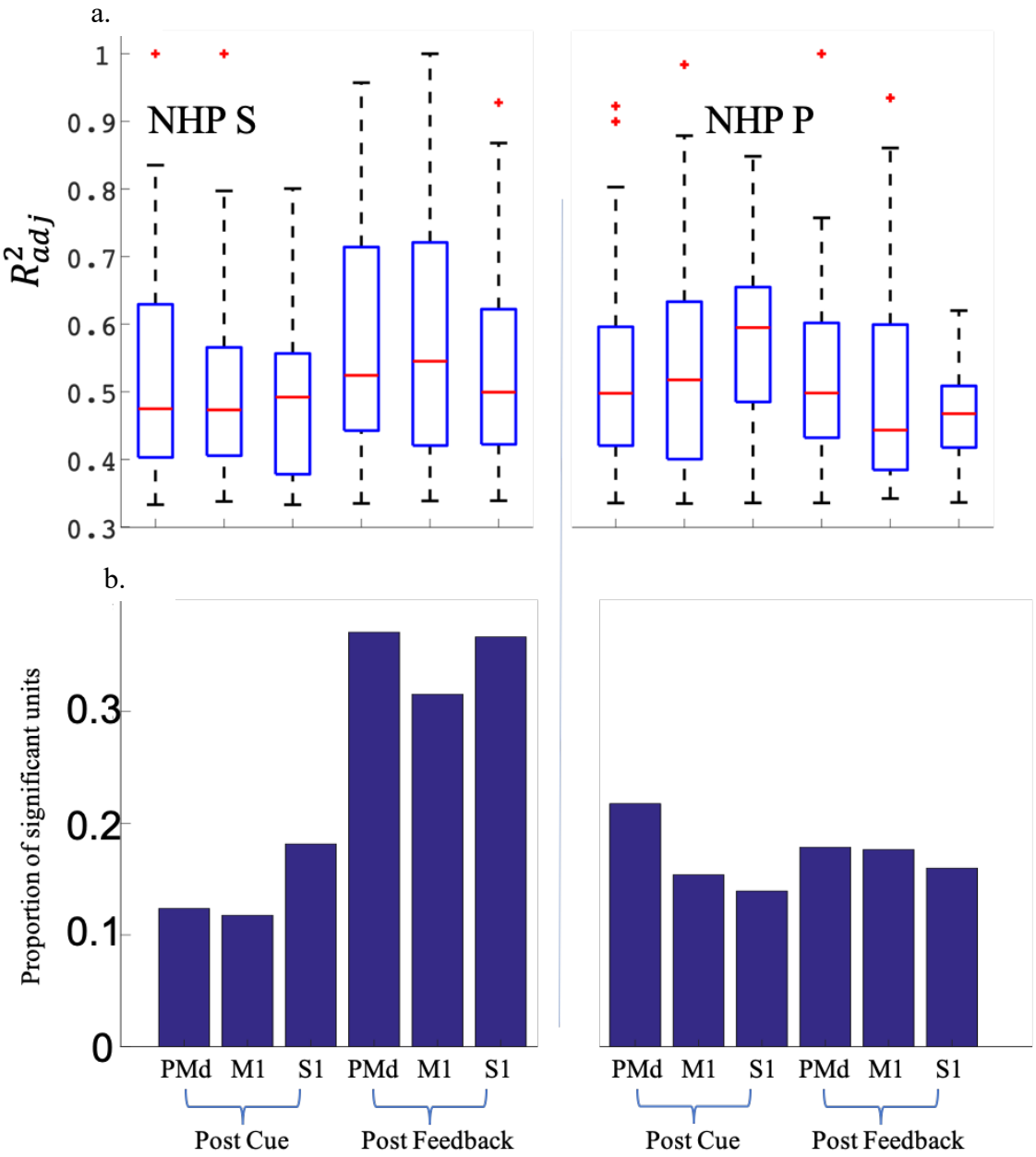


Figure. 10. Medians and ranges for the adjusted R squared distributions (a.), see main text for statistical differences between regions and time periods, and the proportion of significant units for all brain regions (b.).

Comparing divisive models (2-7) for reward and punishment feedback periods

Figure 7.b shows results for example units utilizing model 2 or 3 in the post-feedback window. For model 2 results, we plotted examples of sigmoidal and hyperbolic units. Post feedback

responses could be different from the post cue responses for a given unit. Our hypothesis for this post feedback period was that the units would be encoding the reward or punishment that the NHP was actually going to receive as compared to during the post cue period when the outcome was not yet known, and thus there was still uncertainty. Therefore, as an example, on successful trials in the post feedback period we expected a unit to have the same response when the reward and punishment levels were R = 1, P = 0 or R = 1, P = 1, since the NHP would have obtained the same reward in either case, and therefore if the neural activity was representing valence it should not change under these two example conditions, assuming limited history dependence of course. To test this assumption, we studied 4 different models for the post feedback analysis:

Model 2 is $f_{mnj} = \frac{\alpha_{2j}(r_m+\gamma_{2j}p_n)}{\delta_{2j}+(r_m+\gamma_{2j}p_n)} + \beta_{2j}$ (2)

Model 3 is $f_{mnj} = \frac{\alpha_{3j}(r_m+\gamma_{3j}p_n)}{\delta_{3j}+s_{mn}} + \beta_{3j}$ (6)

Model 4 is $f_{mj} = \frac{\alpha_{4j}(r_m)}{\delta_{4j}+(r_m)} + \beta_{4j}$ for successful trials only (12)

Model 5 is $f_{nj} = \frac{\alpha_{5j}(p_n)}{\delta_{5j}+(p_n)} + \beta_{5j}$ for failure trials only (13)

Model 6 is $f_{mj} = \frac{\alpha_{6j}(r_m)}{\delta_{6j}+(s_m)} + \beta_{6j}$ for successful trials only (14)

Model 7 is $f_{nj} = \frac{\alpha_{7j}(p_n)}{\delta_{7j}+(s_n)} + \beta_{7j}$ for failure trials only (15)

Table 4: Trial numbers for each R and P level

Successful trial numbers				
	R = 0	R = 1	R = 2	R = 3
NHP P Section 1	42	36	25	25
NHP P Section 2	30	31	36	26
NHP P Section 3	33	45	41	39
NHP S Section 1	70	74	67	83
NHP S Section 2	58	62	54	59
NHP S Section 3	77	63	57	77

Failure trial numbers				
	P = 0	P = 1	P = 2	P = 3
NHP P Section 1	10	13	8	10
NHP P Section 2	10	28	8	4
NHP P Section 3	19	14	10	15
NHP S Section 1	9	5	3	4
NHP S Section 2	6	7	4	4
NHP S Section 3	7	10	6	6

The adjusted R-squared was utilized for cross-model comparisons. The results are shown in figure 11 and table 5. Overall, NHP P actually had slightly more units with their best model from group 1 (considering reward and punishment together for all trials), whereas NHP S showed the response we had anticipated, that is better fits in the post outcome period with group 2 (only considering reward for successful trials and only considering punishment for failed trials). Note that even though the two NHPs have slightly different patterns of responses they are internally consistent between their own brain regions and these differences line up with the differences seen during unsuccessful trials. NHP S had 6% failures (out of 872 trials) while NHP P saw 27% (out of 558 trials). NHP S seemed to care more about the reward it would receive at the end of the task perhaps because it could finish most trials successfully. NHP P had a more balanced exposure to successful and unsuccessful trials, and also had more units that best fit with models 2 and 3 that included both types of trials.

Table 5: Number of significant units with best model # (2-7), all were better than the linear model 1.

	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
NHP S PMd	24	10	48	7	2	6
NHP S M1	10	3	48	0	0	0
NHP S S1	12	4	74	3	0	0
NHP P PMd	22	12	37	0	1	0
NHP P M1	14	11	11	3	1	0
NHP P S1	13	15	16	2	2	0

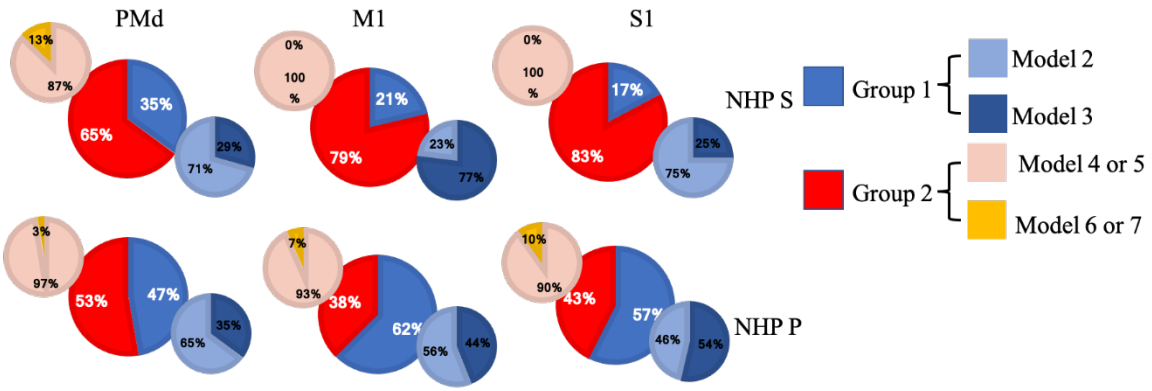


Figure 11. Percentage of units firing rate best fit for a given divisive normalization model. Group 1 models include both successful and unsuccessful trials, where model 2 utilized task R and P levels and model 3 utilized the population activity. Group 2 models include successful trials only (models 4, 6), or unsuccessful trials only (models 5, 7), with models 4 and 6 utilizing task R or P levels respectively, and models 5 and 7 using population activity in the divisive term from R or P trials respectively.

663 **Discussion**

664 Here we report widespread divisive normalization by affect within the primary sensorimotor
665 cortices (M1, S1 and PMd) of NHPs by valence and motivational intensity. Two NHP subjects
666 made cued isometric grip force "movements", while we modulated the level of cued valence,
667 spanning from negative (punishment) to positive (reward), and cued motivational intensity,
668 which were simultaneously modulated by cueing the level of reward they would receive if
669 successful, and level of time-out punishment if unsuccessful. There were 4 levels of cued
670 reward and cued punishment, leading to 16 levels of valence and motivational intensity termed
671 affective stimuli. We simultaneously recorded from 96 electrodes in each M1, S1 and PMd
672 for a total of 288 electrodes per NHP, total of 576 electrodes.

673 We found three prominent relationships between neural firing rates and affective stimuli, linear,
674 sigmoidal, or hyperbolic. Sigmoidal units had a two-part firing rate response, one high
675 sensitivity, and one lower sensitivity region. Sigmoidal and linear units allow affect to modulate
676 the neural rate over a wide range of stimuli, and it is possible that the linear units become
677 sigmoidal if the affective stimuli have a broader range. We suspect there is a distribution of
678 units, some having broad affective modulation ranges, while others have focal ranges, such
679 as the sigmodal and hyperbolic units. Hyperbolic units were very sensitive to a small range of
680 affective stimuli, where the denominator in model 2 was close to zero. We could not determine
681 the asymptotes for models 3, 6 and 7 as they utilized the population's activity in the divisive
682 term, unlike model 2 that utilized the R and P task levels, and thus we could interpolate model
683 2 and find the asymptotes. Model 2 may seem less biomimetic than models 3, 6 and 7,
684 however, as we do not know exactly where the divisive information is determined model 2
685 allowed us to ask questions about such normalization without making the assumption that the
686 local neural activity is performing this normalization. In this sense, model 2 is more powerful
687 as it makes fewer assumptions.

688 In general, the population of units showed an increase in activity during the post-cue period
689 and a significant suppression during the post-feedback period. During both the post-cue and
690 post-feedback periods we found the divisive normalization models (models 2-3) clearly
691 outperformed the linear model (model 1) as seen in table 3. This suggests that for PMd, M1,
692 and S1, reward and punishment modulation is nonlinear, and this nonlinearity can be captured
693 by divisive normalization models. Model 3 performed better than model 2 for several regions
694 and times in NHP P, suggesting that the accuracy of divisive normalization models
695 incorporating affect can be improved by including population firing rates in the normalization
696 term. Divisive normalization, scaled to the response of a population of units, has also been
697 described in the invertebrate olfactory system and the primary visual cortex (Carandini &
698 Heeger 2012), further supporting the idea that divisive normalization is a computational
699 mechanism performed widely, and in a variety of cortices. Also, of importance is the fact that
700 this type of normalization by the population can be more easily utilized toward stabilizing brain
701 computer interfaces in the face of changing affect as discussed below (Zhao et al., 2018).
702 However, more significant units had model 2 as their best model, which used the task labels
703 to define the affective stimuli, as compared to model 3, when including both NHPs as seen in
704 table 5. Model 4 had the most significant units with it as their best model, where model 4 only
705 looked at successful trials and reward levels (R = 0-3) while also utilizing the task R levels in
706 the divisive term.

707 **Valance and motivation**

708 Previously, work on the influence of value in the sensorimotor stream has demonstrated value
709 based divisive normalization in the lateral intraparietal cortex {Citation} (Louie, Grattan, and
710 Glimcher 2011), where this activity was interpreted as encoding the action value within the
711 context of choices between actions associated with different state values. Later work argued
712 against this interpretation in favor of motivational salience (Leathers and Olson, 2012).
713 Subsequent work suggested the action value neural interpretation holds true for behavior in
714 general, including human subjects as well as NHPs (Louie et al., 2013). However, (Gluth et

al., 2020) found that the behavioral responses were better fit with value-based attention when increasing the number of participants used in the previous study (Louie et al., 2013), however see (Webb et al., 2020). We have shown in the current paper that we see both a valence-like signal, and a motivational intensity-like signal. However, further work is needed to determine the exact non-movement related variables being encoded within these sensorimotor cortices. In one influential latent leviable model of affective space two latent dimensions, by which all affective states can be represented, are arousal and valence (Russell, 1980). It is possible that these two latent dimensions are what we were seeing in our data.

In our study the subjects did not have options to choose from, and simply had one type of movement that they could perform successfully or not. Future work where we explicitly include choice should be helpful in further determining if both state-value/motivation and action-value/motivation are being represented in PMd, M1 and S1. In a set of papers Roesch and Olson studied activity in prefrontal and frontal cortex including pre-motor regions and supplementary motor regions. They found increasing activity of these brain regions in line with increased motivational intensity as one moved more caudal (Roesch and Olson 2003) when comparing a small and large reward context. However, as only reward was modulated, they could not determine if it was valence encoding or motivational intensity. In later work they included possible reward and punishment, the paradigm that inspired ours, and found results in the premotor region that was best fit by motivational intensity (Roesch and Olson 2007, 2004). However, this could partially be due to the analysis methods used, or task specific aspects. Their task included a memory component and choice, whereas ours did neither, but focused on the influence that these affective stimuli had on neural activity during a simple isometric grip force task. By utilizing divisive normalization models, we see a stronger valence representation while still seeing evidence of motivational intensity. We are therefore claiming that both valence and motivational intensity are being represented in PMd, M1 and S1 during our grip force task, and in our data, valence is more strongly represented, see Fig.7.b. However, while looking at the larger picture from others work it seems task constraints may

be influencing the outcomes, and thus we aim to conduct a comprehensive study in the future including both tasks with and without choice. In addition, the range of the reward and punishment space should also be modulated in blocks in order to determine if the contextual boundaries of R and P being currently represented changes the neural relationship to individual R and P levels, such that, as we have suggested, our linear units are in fact sigmodal and to ask questions about the distribution of the hyperbolic units over differing affective spaces.

Towards affect agnostic brain computer interfacing (BCI)

Understanding neural modulation due to affect is relevant to biomedical engineers creating brain computer interfaces (BCIs) with activity from sensorimotor regions (Sanchez et al., 2011; Marsh et al., 2015; An et al., 2018; Zhao et al., 2018). For BCI control signals to be accurately decoded from neural activity it is important to determine if these units are utilizing a divisive normalization scheme when it comes to affective stimuli, or naturally changing moods and emotional states. Recently the impact of context on decoding neural information has been addressed in the BCI literature when dealing with kinematics under BCI control switching from two dimensional to three dimensional space (Rasmussen et al., 2017), as well as between reaching movements with and without expected object contact (Downey et al., 2017). Directional tuning in the primary motor cortex is modulated by reward during both manual tasks (Ramakrishnan et al., 2017) as well as during BCI control (Zhao et al., 2018). The latter study noted that grip force related tuning functions are likewise modulated by reward expectation in M1. Such affect information in the sensorimotor system could have profound implications when such neural activity is being used to control the movement of a BCI-controlled robotic limb, to drive the motions of a computer cursor towards a communications-based BCI, and possibly even when putting in sensory information as part of a bi-directional iBCI (Marsh et al., 2015; Choi et al., 2016; McNiel et al., 2016b, 2016a; Zhao et al., 2018). Affective information also influences correlational structure at the single unit level (Moore and Francis, 2020), local field potential (LFPs) level and between these two (An et al., 2018, 2019).

As LFPs can also be utilized for BMIs (Ince et al., 2010; Flint et al., 2013) , affect's influence on such neural activity is vital to understand in order to produce affect agnostic BCIs towards the restoration of movement control.

An J, Yadav T, Ahmadi MB, Tarigoppula VSA, 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.

An J, Yadav T, Hessburg JP, Francis JT (2019) Reward Expectation Modulates Local Field Potentials, Spiking Activity and Spike-Field Coherence in the Primary Motor Cortex. eNeuro June 6 2019.

Bialek W, de Ruyter van Steveninck R, Rieke F, Warland D (1996) Spikes: Exploring the Neural Code.

Carandini M, Heeger DJ (2011) Normalization as a canonical neural computation. Nat Rev Neurosci 13:51–62.

Chhatbar PY, Francis JT (2013) Towards a naturalistic brain-machine interface: hybrid torque and position control allows generalization to novel dynamics. PLoS One 8:e52286.

Choi JS, Brockmeier AJ, McNiel DB, Kraus LM, Principe JC, Francis JT (2016) Eliciting naturalistic cortical responses with a sensory prosthesis via optimized microstimulation. J Neural Eng 13:056007.

Downey JE, Brane L, Gaunt RA, Tyler-Kabara EC, Boninger ML, Collinger JL (2017) Motor cortical activity changes during neuroprosthetic-controlled object interaction. Sci Rep 7:16947.

Flint RD, Wright ZA, Scheid MR, Slutzky MW (2013) Long term, stable brain machine interface performance using local field potentials and multiunit spikes. J Neural Eng 10:056005.

Georgopoulos AP, Ashe J, Smyrnis N, Taira M (1992) The motor cortex and the coding of force. Science 256:1692–1695.

Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT (1982) On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J Neurosci 2:1527–1537.

Gluth S, Kern N, Kortmann M, Vitali CL (2020) Value-based attention but not divisive normalization influences decisions with multiple alternatives. Nature Human Behaviour 4:634–645.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–197.

Ince NF, Gupta R, Arica S, Tewfik AH, Ashe J, Pellizzer G (2010) High accuracy decoding of movement target direction in non-human primates based on common spatial patterns of local field potentials. PLoS One 5:e14384.

Leathers ML, Olson CR (2012) In Monkeys Making Value-Based Decisions, LIP Neurons Encode Cue Salience and Not Action Value. Science 338:132–135.

Li CS, Padoa-Schioppa C, Bizzi E (2001) Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. Neuron 30:593–607.

Louie K, Grattan LE, Glimcher PW (2011) Reward value-based gain control: divisive normalization in parietal cortex. J Neurosci 31:10627–10639.

Louie K, Khaw MW, Glimcher PW (2013) Normalization is a general neural mechanism for context-dependent decision making. PNAS 110:6139–6144.

Louie K, LoFaro T, Webb R, Glimcher PW (2014) Dynamic Divisive Normalization Predicts Time-Varying Value Coding in Decision-Related Circuits. J Neurosci 34:16046–16057.

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.

McNiel DB, Bataineh M, Choi JS, Hessburg JP, Francis JT (2016a) Classifier Performance in Primary Somatosensory Cortex Towards Implementation of a Reinforcement Learning Based Brain Machine Interface. IEEE Southern Biomedical Engineering Conference 2016.

McNiel DB, Choi JS, Hessburg JP, Francis JT (2016b) Reward value is encoded in primary somatosensory cortex and can be decoded from neural activity during performance of a psychophysical task. In, pp 3064–3067. IEEE.

McNiel DB, Choi JS, Hessburg JP, Francis JT (n.d.) Reward value is encoded in primary somatosensory cortex and can be decoded from neural activity during performance of a psychopyisical task. IEEE EMBS 2016 Conference proceding (Aug 2016).

Moore B, Francis JT (2020) Noise Correlation is Modulated by Reward Expectation in the Primary Motor Cortex Bilaterally During Manual and Observational Tasks in Primates. Front Behav Neurosci 14 Available at: https://www.frontiersin.org/articles/10.3389/fnbeh.2020.541920/abstract [Accessed November 10, 2020].

Ohshiro T, Angelaki DE, DeAngelis GC (2017) A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate Cortex. Neuron 95:399-411.e8.

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 U S A 114:E4841–E4850.

Ramkumar P, Dekleva B, Cooler S, Miller L, Kording K (2016) Premotor and Motor Cortices Encode Reward. PLoS One 11:e0160851.

Rasmussen RG, Schwartz A, Chase SM (2017) Dynamic range adaptation in primary motor cortical populations. Elife 6.

Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304:307–310.

Rolls ET (2014) Emotion and decision-making explained, First edition. Oxford ; New York, NY: Oxford University Press.

Russell JA (1980) A Circumplex Model of Affect. Journal of Personality and Social Psychology 39:1161–1178.

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 (Ner):525–528.

Sato TK, Haider B, Häusser M, Carandini M (2016) An excitatory basis for divisive normalization in visual cortex. Nature Neuroscience 19:568–570.

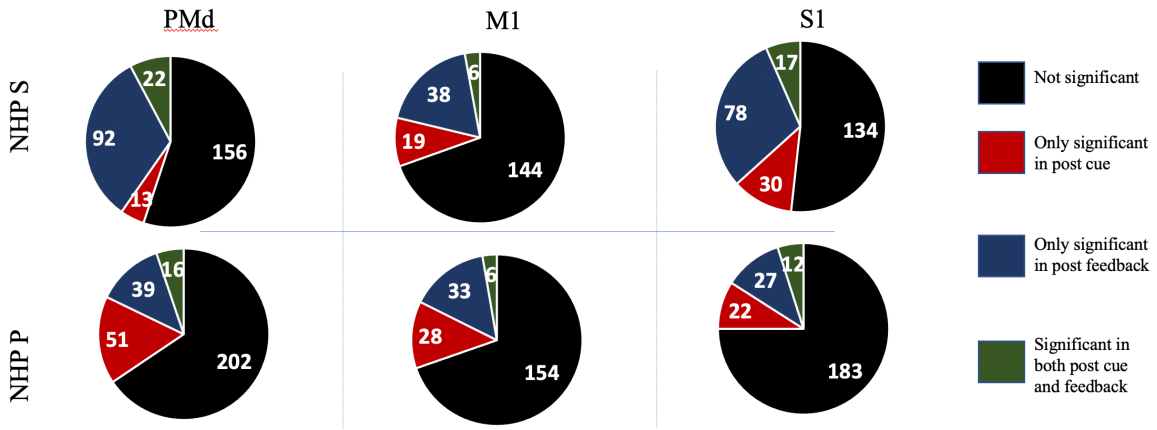
Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. Nature Neuroscience 4:819–825.

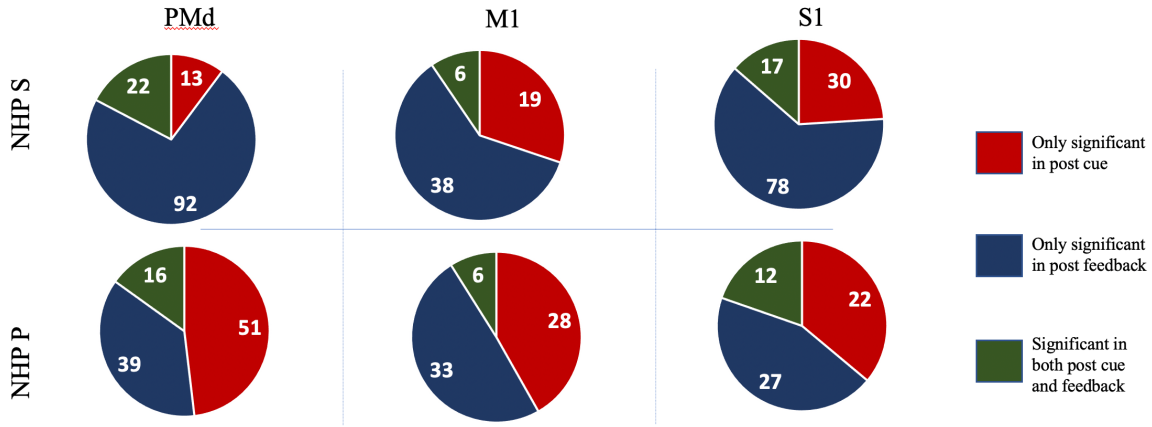
Tarigoppula VSA, Choi JS, Hessburg JH, McNiel DB, Marsh BT, Francis JT (2018) Motor Cortex Encodes A Value Function Consistent With Reinforcement Learning. bioRxiv:257337.

Webb R, Glimcher PW, Louie K (2020) Divisive normalization does influence decisions with multiple alternatives. Nature Human Behaviour 4:1118–1120.

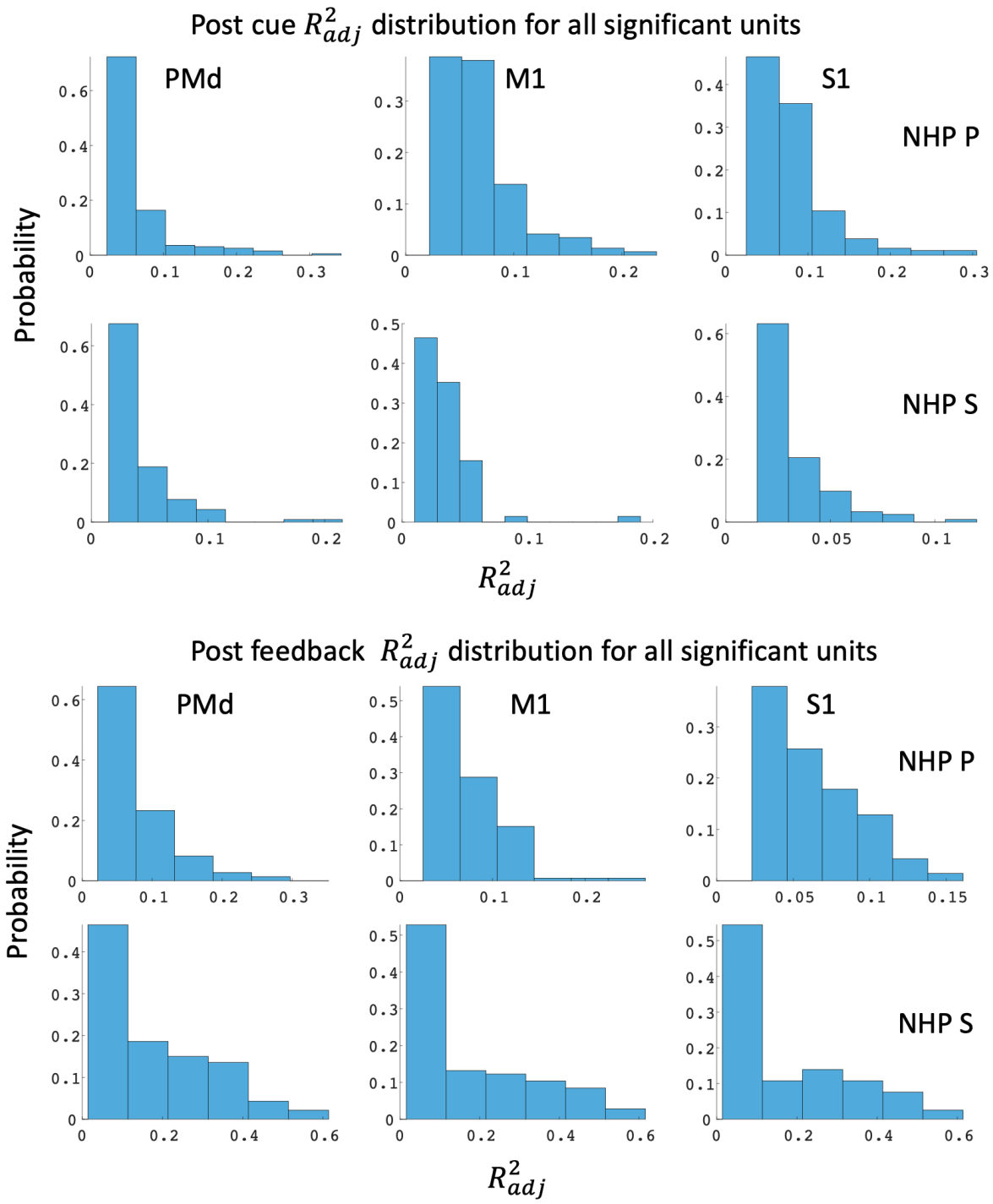
Zhao Y, Hessburg JP, Kumar JNA, Francis JT (2018) Paradigm Shift in Sensorimotor Control Research and Brain Machine Interface Control: The Influence of Context on Sensorimotor Representations. Front Neurosci-Switz 12.

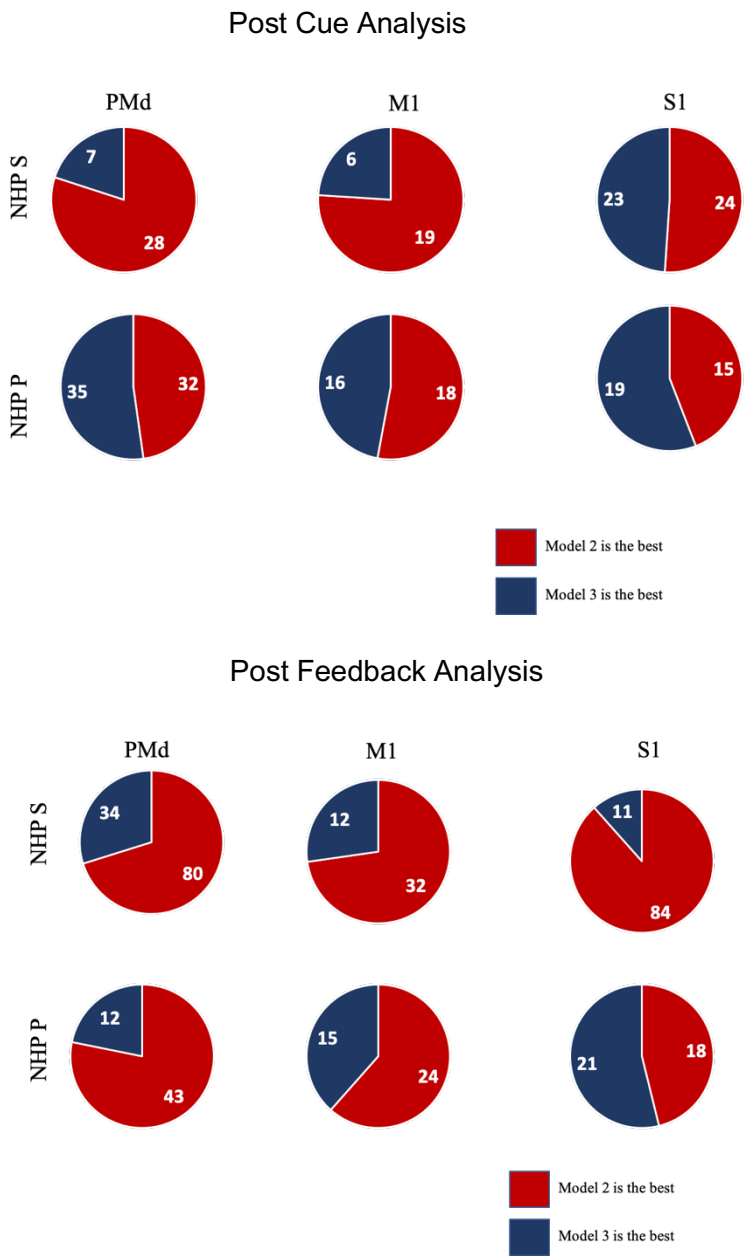
Supplementary Information:





S.Fig.1: Here we show not only the significant units and their makeup, but the perspective of the larger non-significant population (Top subplot). For convenience we have copied the main text figure 7 below it.



S.Fig.2. Distribution of the adjusted R squared values for the best significant model fit for a given unit during the post-cue, top two rows, and the post-feedback, bottom two rows, for the two NHPs labeled on the right-hand side of the figure. Fitting firing rates from every trial as compared to Fig.9 in the main text that comes from the model fits to the mean for each of the 16 different affective stimuli.

S.Fig.3. showing only significant units as compared to all units as seen in the main text table.3. Often, but not always, model #2, which utilized the R and P levels from the task in the denominator, outperformed model #3, which used the local population activity.