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

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

Correspondence: joey199us@gmail.com
Number of Pages: 41
Number of figures: 12
Number of words for abstract: 240
Number of words for introduction: 646
Number of words for discussion: 1406

Abbreviated Title: Normalization by Affect in the Sensorimotor Cortices

Acknowledgements: Research was supported by NIH 1R01NS092894-01, NSF IIS-1527558, DARPA REPAIR Project N66001-10-C-2008, NYS SCIRB contracts C30600GG, C030838GG, and DOH01-C32250GG-3450000.

[1] Department of Biomedical 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 listing determined by alphabetical order.
[3] Lead Contact - joey199us@gmail.com

Significance Statement
Brain computer interfaces (BCI) translate user's movement intentions from neural activity in sensorimotor cortices into cursor, or robotic motion. Reinforcers, such as reward and punishment, induces changes in our emotional/affective state, and how this influences the sensorimotor cortices has clinical meaning towards the development of stable (BCIs), as a trials reward value changes the directional tuning function of sensorimotor cortical neurons (Zhoa et al. 2018, Ramakrishnan et al. 2017). It has been shown that trial values can lead to divisive normalization within brain regions involved in decision making. Here we show divisive normalization to both valence and motivational intensity is prevalent in the sensorimotor cortex of non-human primates (PMd, M1 and S1), and should be considered towards affect agnostic BCIs.

Abstract
Our brain's ability to represent vast amounts of information, including continuous ranges of reward values spanning orders of magnitude, with neurons with limited dynamic range, may be possible through divisive normalization. Recently, our group and others have shown that the sensorimotor cortices are sensitive to reward value. In order to determine if divisive 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 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-channel electrode arrays in the somatosensory, motor and premotor cortices (S1, M1, and PMd, respectively). Utilizing divisive normalization models, we show that there are non-linear relationships between neural firing rates, valence, and motivational intensity in all three regions. We found three types of relationships between neural activity and the representation of reward and punishment: linear, sigmoidal and hyperbolic. We found that these cortical regions have both strong valence and motivational intensity representations, which if not utilizing divisive normalization models appear skewed toward motivation. We found that one NHP subject showed a stronger representation of these variables post-cue while the other was more sensitive to the post-result period, with these subject-specific differences consistent between the three brain regions and both the post-cue and post-feedback periods within subject, despite the between subject differences.

Introduction
Neural networks within the primary sensorimotor cortices need to function within a finite dynamic range while encoding multiple forms of information such as kinematics and dynamics (Chhatbar and Francis 2013; Georgopoulos et al. 1982; Li, Padoa-Schioppa, and Bizzi 2001; Georgopoulos et al. 1992). The nervous system could increase its information carrying capacity by utilizing a temporal code, which would be less sensitive to a limited dynamic range than a rate code (Rieke 1997). Another possibility is provided by divisive normalization (Carandini and Heeger 2011), where the input from another region, 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) to networks involved in decision-making (Louie, Grattan, and Glimcher 2011; Cai and Padoa-Schioppa 2012; Padoa-Schioppa 2009; Rangel and Clithero 2012; Zimmermann, Glimcher, and Louie 2018; Yamada et al. 2018).

Recently, it has been shown that the primary motor (M1) and somatosensory (S1) cortices encode reward in addition to kinematics and dynamics (Ramkumar et al. 2016; An et al. 2019; Zhao et al. 2018; An et al. 2018; McNiel, Bataineh, et al. 2016; McNiel, Choi, et al. 2016; Marsh et al. 2015; Ramakrishnan et al. 2017). 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 the punishment and obtain the reward, and a neutral stimulus would induce zero motivation.

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 (Ramkumar et al. 2016; An et al. 2019; Tarigoppula et al. 2018; Marsh et al. 2015), such as the fact that

reward level changes the directional tuning of units in M1 during both manual reaching movements (Ramakrishnan et al. 2017) and those produced by a brain machine interface (BMI) (Zhao et al. 2018). We therefore characterized neural activity due to simultaneously cued possible reward and 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).

Here we demonstrate that neural modulation in S1, M1, and PMd is associated with both cued reward and punishment in an operant conditioning task, where reward or punishment were delivered according to an individual's performance. Both neural firing rates and the time to complete a trial can be represented 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 and on the individual NHP subject, however within each subject brain regions responded similarly. These results are important towards the production of affective agnostic BMIs that do not change their behavior based on the user's affective state that changes as a function of reinforcers in the environment, such as rewards and punishments.

Methods

Behavioral task

Two non-human primates (NHPs), one male rhesus macaque (NHP S, Macaca mulatta) and one female bonnet macaque (NHP P, Macaca radiata), were trained to perform a 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, reaching, grasping, transporting, 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. 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 square was displayed, then no reward was delivered upon successful completion of the trial. Red squares indicated the level of time-out punishment periods 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 (see Fig.1) was displayed over the environment and the animal was forced to wait before attempting the next trial. For unsuccessful trials the trial was repeated at the same reward and punishment level until completed successfully, thus motivating subjects to complete even non-rewarding trials successfully, and not allowing animals to purposefully fail and skip low value trials.

During the grip force task, the virtual arm automatically reached to a target cylinder (Fig.1), and the NHP controlled the grasping motion of the robotic hand by manually squeezing 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 while force was maintained. When the hand reached the target location the NHP released the gripper, 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.

A variation of this grip force task was performed where instead of reward and punishment options, only reward was included. The trials were conducted identically, but instead of punishment being included, only reward levels from 0 to 3 were considered. The rewarding and punishing trials and reward-only trials were recorded in separate sessions.

Figure 1. Grip Force Task

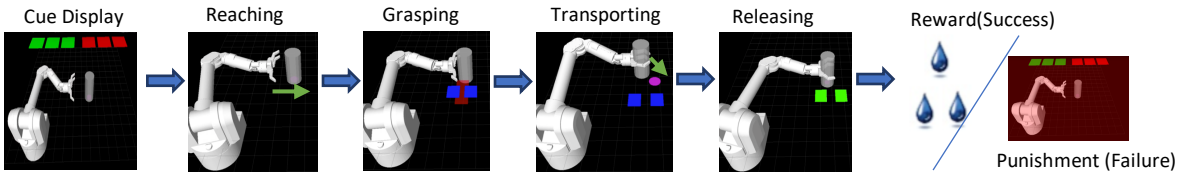

Figure 1. Cued grip force task. The behavioral task was composed of 6 scenes for each trial. First, cue and punishment levels were projected into the virtual environment during the cue display scene, with each green square indicating 0.5 seconds of juice reward delivery if the trial was completed successfully, 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. Force was applied manually during the grasping scene, represented by an expanding red rectangle that was then maintained within the blue force target rectangles during the transporting scene. Once the object reached the destination the target turned green, signaling that the gripper could be released. If these steps were completed successfully the task entered the success scene, and juice was delivered according to the cued amount. If at any point the NHP applied too much or too little force, the trial was deemed a failure and the failure scene was entered where the amount of time-out period punishment was delivered according to the cued amount as the task screen turned a transparent red, and then the trial was reset and could be attempted again.

Surgery

Once trained on the task, NHPs S and P were implanted in their left M1, S1 and PMd with chronic 96-channel platinum microelectrode arrays (Utah array, 10X10 array with 400 μm inter-electrode spacing, ICS-96 connectors, Blackrock Microsystems. 1.5 mm in length for M1 and PMd, 1.0 mm for S1). The hand and arm regions contralateral to their dominant (right) hands were implanted with the same technique as utilized in our previous work (Marsh et al. 2015, Chhatbar et al. 2010). All surgical procedures were conducted in compliance with guidelines set forth by the National Institutes of Health Guide for the Care and Use of Laboratory NHPs and were approved by the State University of New York Downstate Institutional Care and Use Committee. Briefly, NHP preparation and anesthesia were performed directly by members of the State University of New York Downstate Division of Comparative Medicine veterinary staff, and the researchers conducted the surgery under their supervision. Ketamine was used to induce anesthesia, and isofluorane and fentanyl were used for maintenance. Aseptic conditions were maintained throughout surgery as a craniotomy window was created over the target location, and a probing electrode was used to identify the hand and forearm region of S1. An electrode array was implanted there, and another immediately across the central sulcus for the corresponding M1 region. The PMd array was placed immediately dorsal to the spur of the arcuate sulcus. Dexamethasone was used to prevent inflammation during the procedure and diuretics such as mannitol and furosemide 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 modulation encoding models of the general form $f_{ij} = g(s_i)$ were fit to individual unit firing rate data from PMd, M1, and S1. f_{ij} indicates the jth unit's post-cue, or post-feedback firing rate in the ith trial, and g was a nonlinear function dependent on the trial stimulus s_i, defined as the reward and punishment level combination for the ith trial. For post-cue analysis the firing rate was averaged over the 500 ms window following cue presentation. In this time period the virtual arm was stationary or moving horizontally to the right during the reaching scene, approaching the cylinder at a consistent speed, and the NHPs were not yet applying 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. For reward and punishment trials, $f_{ij} = g(r_i, p_i)$, where f_{ij} was the post-cue or post-feedback firing rate for the jth unit during the ith trial, with the reward level r_i and punishment level p_i for that trial.

Reward-only encoding analysis

First, for both NHPs we fit a divisive normalization model between the time duration it took the NHPs to perform a given cued trial, defined as the time from cue display to result deliver, and the corresponding reward level r_i:

Model R1$_{dur}$: $t_i = \frac{A_1 r_i}{r_i + B_1} + C_1$ (1).

We then fit a reward-only neural firing rate encoding model $f_{ij} = g(r_i)$, where r_i is the reward level for the ith trial. Two divisive normalization models were used to design the nonlinear relationship g. For post-cue or post-feedback analysis, these models were:

Model R1: $f_{ij} = \frac{\alpha_{1j}(r_i)}{\delta_{1j} + (r_i)} + \beta_{1j}$ (2) and

Model R2: $f_{ij} = \frac{\alpha_{2j}(r_i)}{\delta_{2j} + (s_i)} + \beta_{2j}$ (3),

where f_{ij} is the firing rate for the jth unit at the ith trial, r_i is the reward level for the ith trial, and α_{1j}, β_{1j}, δ_{1j}, α_{2j}, β_{2j}, and δ_{2j} were parameters fit to the data (Matlab curve fitting app, function "fitnlm"). The difference between models R1 and R2 is the divisive term. Model R1 uses the reward level for the ith trial (r_i) as the divisive term, and model R2 uses s_i, a scalar based on the population firing rate in all trials when $r = r_i$. s_i was a scalar based on the population firing rates for the level of reward (r_i).

$$s_i = \sum_{j=1}^{N} \left| f_j^{r_i} - \mu_j^{00} \right| \quad (4),$$

where j is the neuron index, N is the total number of neurons, $f_j^{r_i}$ is the mean post-cue or post-feedback firing rate for the jth neuron on trials with reward level r_i, and μ_j^{00} is the mean firing rate of the jth neuron when the reward level was zero.

s_i is the mean absolute value firing rate difference compared to a baseline average population neural firing rate over all trials where r=0. We utilized four reward levels (R = [0,1,2,3]). The total trial number when the reward level $R = i$ is T_i. In our task, $T_i, i \in [0,1,2,3]$ were not always the same, therefore to avoid overfitting reward levels we used weighted least square (WLS) to fit firing rates instead of ordinary least square (OLS). OLS minimizes the sum squared error:

$$err = \sum_i (y_i - f(x_i))^2.$$

Here, y_i is the recorded data and $f(x_i)$ represents the fitted results. WLS minimizes the weighted sum squared error:

$$err = \sum_i w_i (y_i - f(x_i))^2.$$

w_i is the weight for the ith trial, based on the total trial number for a given reward level trial r_i. If $r_i = 0$, then $w_i = 1$. if $r_i = k$ then $w_i = \frac{T_0}{T_k}$, so that the total weights for each reward level were the same.

Reward and punishment encoding analysis

For reward and punishment trials $f_{ij} = g(r_i, p_i)$, where f_{ij} is post-cue or post-feedback firing rates for the jth unit and the ith trial, r_i is the reward level, and p_i is the punishment level for the ith trial. We designed one linear model (Model 1) and two divisive normalization models (models 2 and 3). As with the reward-only models we used WLS to avoid overfitting. The total trial number when the reward level $R = m$ and $P = n$ is T_{mn}. If $r_i = 0$ and $p_i = 0$, then $w_i = 1$. if $r_i = m$ and $p_i = n$ then $w_i = \frac{T_{00}}{T_{mn}}$, so that the total weights for each reward and punishment level were 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,

$$\text{Model 1: } f_{ij} = \alpha_{1j}(r_i + \gamma_{1j} p_i) + \beta_{1j} \quad (5),$$

where f_{ij} was the post-cue firing rate for the jth unit in the ith trial, and r_i and p_i are the reward and punishment levels, respectively, for that trial. There were four possible reward levels (R = [0,1,2,3]) and four punishment levels (P = [0,1,2,3]), for a total of 16 combinations of reward and punishment. For each trial the reward and punishment numbers were chosen pseudorandomly from a range of 0 to 3, so there was a slight variability in the exact number of trials for each reward level, punishment level, and combination thereof. Equation (5) was fit to firing rates from units from each region using Matlab's fitlm function. AIC values were calculated for all models to determine which best explained the data as well as adjusted R^2 (see below).

Model 2: The second model was a divisive normalization model incorporating reward and punishment levels, similar to equation (2) for the reward-only task.

$$\text{Model 2: } f_{ij} = \frac{\alpha_{2j}(r_i + \gamma_{2j} p_i)}{\delta_{2j} + (r_i + \gamma_{2j} p_i)} + \beta_{2j} \quad (6),$$

where f_{ij} was the post-cue or post-feedback firing rate for the jth unit in the ith trial, and r_i and p_i were the reward and punishment levels, respectively, for that trial. γ_{2j} was the weight between reward and punishment for the jth unit, included to allow for the comparison of reward and punishment types of different modalities, namely the juice reward and time-out

period punishment. To simplify this equation a stimulus was defined, which was a linear combination of the cued reward and punishment level:

$$\text{Affective Stimulus: } s_i = r_i + \gamma_{2j} p_i \text{ (7)},$$

where s_i was the stimulus, r_i was the cued reward level, and p_i the cued punishment level for the ith trial. Equation (6) from Model 2 can then be written as:

$$\text{Model 2: } f_{ij} = \frac{\alpha_{2j}(s_i)}{\delta_{2j}+(s_i)} + \beta_{2j} \text{ (8)}.$$

Equation (6) was used to determine if the jth unit encoded valence or motivation. If $\gamma_{2j} < 0$ (eq. 6) the unit was considered to encode valence, where reward was positive and punishment negative, that is 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 were positive, that is reward and punishment modulated the firing rate in the same direction. 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").

For a behavioral measurement, we also utilized a model similar to model 2 to fit the trial duration (Model 2_{dur}) time (t_i) to the corresponding reward (r_i) and punishment (p_i) level combination:

$$\text{Model } 2_{dur}\text{: } t_i = \frac{A_2(r_i+B_2 p_i)}{C_2+(r_i+B_2 p_i)} + D_2 \text{ (9)}.$$

Model 3: For the third model, the denominator from equation (5) was modified to incorporate neural population information:

$$\text{Model 3: } f_{ij} = \frac{\alpha_{3j}(r_i+\gamma_{3j} p_i)}{\delta_{3j}+s_i} + \beta_{3j} \text{ (10)},$$

where $r_i = [0,1,2,3]$ and $p_i = [0,1,2,3]$. 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 r_i and punishment level p_i , noted as f_{ij}, where used to fit model 3. s_i was a scalar based on the population firing rates for the combination of reward and punishment levels (r_i, p_i).

$$s_i = \sum_{j=1}^{N} \left| f_j^{r_i p_i} - \mu_j^{00} \right| \text{ (11)},$$

where j is the neuron index, N is the total number of neurons, $f_j^{r_i p_i}$ is the main post-cue or post-feedback firing rate for the jth neuron on trials with reward level r_i and punishment level p_i, and μ_j^{00} is the mean firing rate of the jth neuron when the reward level and punishment level are both zero. s_i is a representation of the difference between the mean population firing rate when reward level is r_i and punishment level is p_i compared to the affective baseline *r=p=0*. Using the raw mean firing rate data was not sufficient because individual units may modulate in different ways in response to reward or punishment, thus canceling each other out. 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.

Motivational intensity and valance encoding: First, we fit every unit to models 1, 2 and 3. The model with the lowest mean squared error (MSE) was defined as the best fitting model for that unit, and then we tested if that best model was significantly different from a constant model using the F-test ($p < 0.05$). The constant model was defined as $f_j = c_j$, where c_j is the mean

firing rates for all data points, and did not incorporate any modulation due to reward or punishment levels. We first calculated the F value by:

$$F = \frac{(ss_1 - s_{s2})/(par_2 - 1)}{ss_2/(n - par_2)},$$

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, par_2 was the number of parameters from the fit model, and n was the number of data points. Units that were significantly different were classified as motivational or valance encoding using that model's γ. 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. A given unit could show motivation in one model, but valence in another model (see figure 7, M1 units). Among the significant units, the number of motivational modulation or valence modulation were noted and presented in the results.

Post-feedback data analysis for reward and punishment task was conducted as with the post-cue methods described above. The feedback delivery period was the 500ms 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. A difference between the post-cue and the post-feedback analysis is that at the beginning of the trial (post-cue), the animals are more likely to incorporate both reward and punishment expectation, as there is uncertainty because they do not yet know the outcome of the trial. The post-feedback time period is after the animal has successfully completed or failed the trial, and we hypothesized that the animal at that point may only be encoding reward for successfully completed trials, or punishment for unsuccessfully completed trials. Therefore, 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:

$$\text{Model 4: } f_{ij} = \frac{\alpha_j r_i}{\delta_j + r_i} + \beta_j \quad (12),$$

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

$$\text{Model 5: } f_{ij} = \frac{\alpha_j p_i}{\delta_j + p_i} + \beta_j \quad (13),$$

where r_i and p_i were the reward or punishment level for 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. Equation (11) was fit using all data from all successful trials. Equation (12), was fit using all data from all failed trials. Similarly, to models 2 and 3, we designed models 6 and 7 using population firing rates as the divisive terms instead of the reward or punishment level for that trial. For successful trials only, we have:

$$\text{Model 6: } f_{ij} = \frac{\alpha_{6j}(r_i)}{\delta_{6j} + (s_i)} + \beta_{6j} \quad (14),$$

where s_i is the scalar based on population firing rates when reward level $r = r_i$, s_i is the mean absolute value firing rate difference compared to the baseline r=0, similar to model 3. Likewise, for failed trials we have:

$$\text{Model 7: } f_{ij} = \frac{\alpha_{7j}(p_i)}{\delta_{7j} + (s_i)} + \beta_{7j} \quad (15),$$

where s_i is the scalar based on population firing rates when punishment level $p = p_n$,

s_i is the mean absolute value firing rate difference compared to the baseline p=0. WLS was used for models 4-7 to avoid overfitting reward and punishment level combinations.
Comparing fitting results across models: The models had different numbers of parameters, therefore an adjusted R^2 value (R^2_{adj}) was calculated and used to compare models. This was determined by

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1}, \text{ (16)}$$

where *n* is the sample size for all models, and p is the total number of explanatory variables in the model. *p* = 2 for models 1, 4,5, 6 and 7, and *p* = 3 for models 2 and 3.

Results

Divisive normalization of neural and behavioral variables by cued reward

We started our work with a version of the grip force task (Fig.1) where we only cued the level of reward to be received if a successful movement was performed, with four reward levels (R = [0,1,2,3]). Thus, in this task valence and motivational intensity where fully congruent. This reward-only work was conducted to first determine the neural and behavioral relationships with a simple contextual environment, compared to the full reward and punishment environment considered in the subsequent tasks. For reward-only trials, NHP P completed 170 successful trials, out of a total of 218. Among these trials there were 55 R = 0, 40 R = 1, 42 R = 2, and 33 R = 3 trials. Similarly, NHP S completed 200 successful trials out of a total of 248. Among these there were 48 R = 0, 47 R = 1, 50 R = 2, and 55 trials where R = 3. As motivational intensity is known to influence a subject's reaction times and movement vigor we tested a divisive normalization model on the NHPs trial times as shown in Figure 2.1. The trial time was the time period between cue display and when the feedback began to be delivered. In Fig.2.1 we plotted the average trial time for each reward level for each NHP. These results demonstrate that mean trial time decreases as reward level increases, and that this change can be described using the divisive normalization model $R1_{dur}$ (see methods),

$$t_i = \frac{\alpha r_i}{r_i + \delta} + \beta$$

where t_i is the trial time for the ith trial with a given cued reward level r_i, and α, β, δ are parameters determined through weighted non-linear least squares fitting. For both NHPs, model $R1_{dur}$ had a sigmoidal form for trial durations, represented by the blue curves in Fig.2.1. Both trial time fitting results were significantly different from a constant model (F-test, p<0.05), with NHP S having a much smoother decrease in trial time with reward level, while NHP P appeared to take longer on non-rewarding (NR) trials as compared to any of the rewarding levels. Thus, NHP P was more sensitive to the cued reward level, which is consistent with the neural results seen below (see Fig. 8 for summary).

We utilized the same divisive model from equation one to fit the neural firing rate (f) of individual single- and multi-units in the three brain regions PMd, M1 and S1:

$$f_{ij} = \frac{\alpha_j r_i}{\delta_j + r_i} + \beta_j, \text{ model R1.}$$

Model R1 had one more parameter than the linear model $f_{ij} = \alpha_j r_i + \beta_j$. α_j, β_j and δ_j are parameters fit to the data. The mean firing rates (red squares) and model fit (blue lines) for example units are plotted in Fig.2.2. For each region (PMd, M1 and S1), we plotted an example unit from each of the three types of typical responses: linear, sigmodal, and hyperbolic. All example units in fig.2.2. are significantly different from the constant model (F-test, p<0.05).

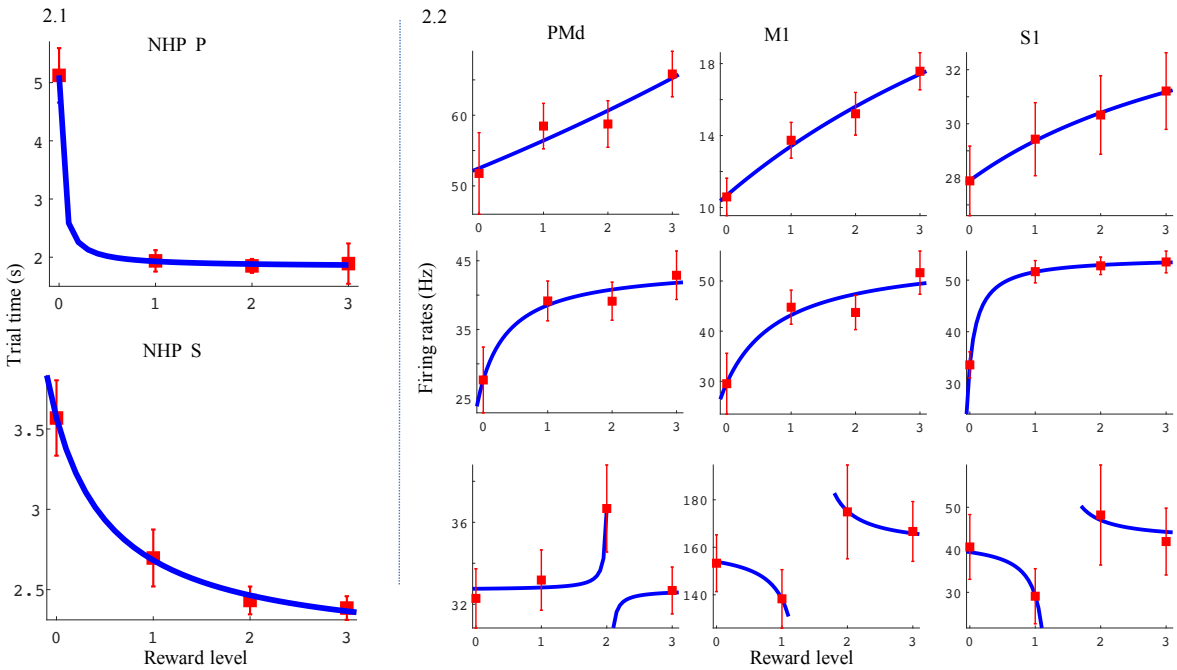

Figure. 2. Post-cue analysis for reward-only task with divisive normalization model (Models $R1_{dur}$ and R1, see methods). Figure 2.1 depicts the average and standard error of the mean (SEM) of the trial time (red) for different reward levels. The x-axis represents reward level and the y-axis represent trial time (s). Figure 2.2 shows example unit reward modulation for post-cue data. For each subplot, the x-axis represents reward level and the y-axis represents the post-cue firing rate (0-500ms after cue display). Each red point represents the mean post-cue firing rate for that reward level with the SEM, and the blue line represents model R1's 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. All example units and trial time fitting results are significantly different from a constant model (F-test, $p<0.05$).

In Fig.3 we show the population results for the divisive normalization models R1 and R2. The first, model R1, is $f_{ij} = \frac{\alpha_{1j}(r_i)}{\delta_{1j}+(r_i)} + \beta_{1j}$, where r_i is the reward level on the ith trial. Thus, model R1 uses the reward level labels from the task itself. Model R2 utilizes the population cortical activity for the divisive term s_i: $f_{ij} = \frac{\alpha_{2j}(r_i)}{\delta_{2j}+(s_i)} + \beta_{2j}$, where s_i is based on the population firing rates in all trials when $r = r_i$ (see methods eq. 4). As units can change their representation between the post-cue and post-feedback periods of time we have broken the data up into units that are either not significant in any time window, significant in just one of the time windows, or significant in both post-cue and post-feedback.

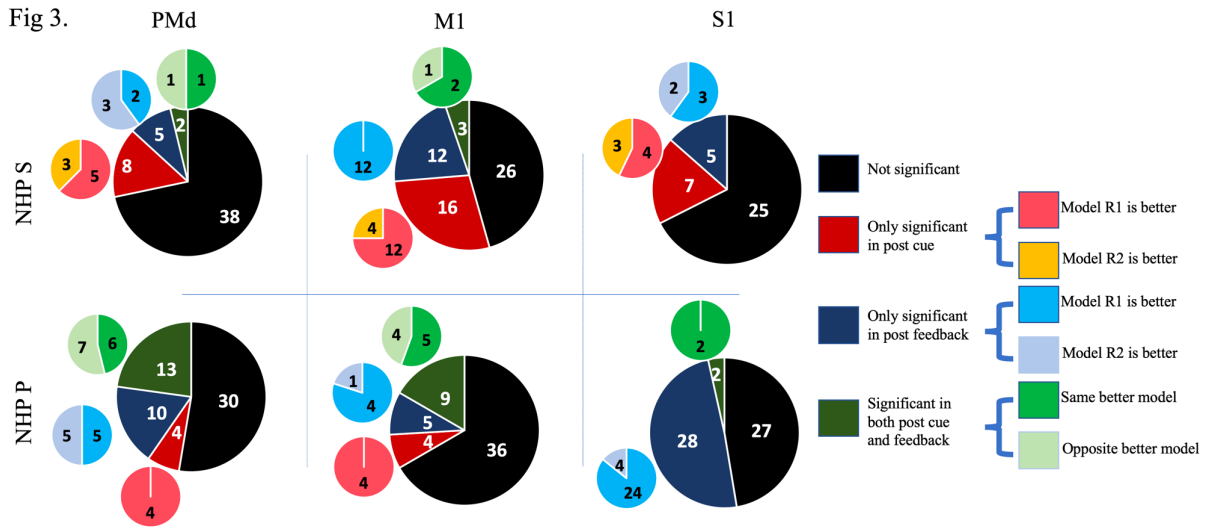

Figure.3 Population results for the number of units in a given NHP, brain region, and trial period, in the post cue and post feedback time periods. Model R1 utilizes reward levels from the task itself whereas model R2 utilizes information from the population's deviation from a baseline affective state where r=0 (see methods). Units were broken up into ones that only respond post cue (red), only in the post feedback period (dark blue) and during both periods (dark green). The number of units for each NHP and brain region were as follows for NHP S and P respectively, PMd 53, 57, M1 57, 54 and S1 37, 57 units.

Neural and behavioral variables modified by cued reward and punishment level

Fig.4 depicts simple peri-cue and peri-feedback time histograms for example single units. We plotted several example single unit responses from PMd, M1 and S1 for the two NHPs showing units that have either a motivational intensity or valence relationship with cued reward (r) and punishment (P). A valence-like relationship entails the neural firing rate increasing or decreasing with respect to valence, or reward level minus punishment level, where R0P3 has the lowest valence, and R3P0 the greatest. For a motivational intensity-like relationship, the neural firing rate varies with the motivational intensity, or reward level plus punishment level, where R0P0 has the least motivational intensity, and R3P3 the greatest. Motivational intensity is a symmetric function with respect to punishment and reward, whereas the valence function is linear from R0P3 to R3P0, in either an increasing or decreasing manner.

Post cued reward and punishment single-unit analysis

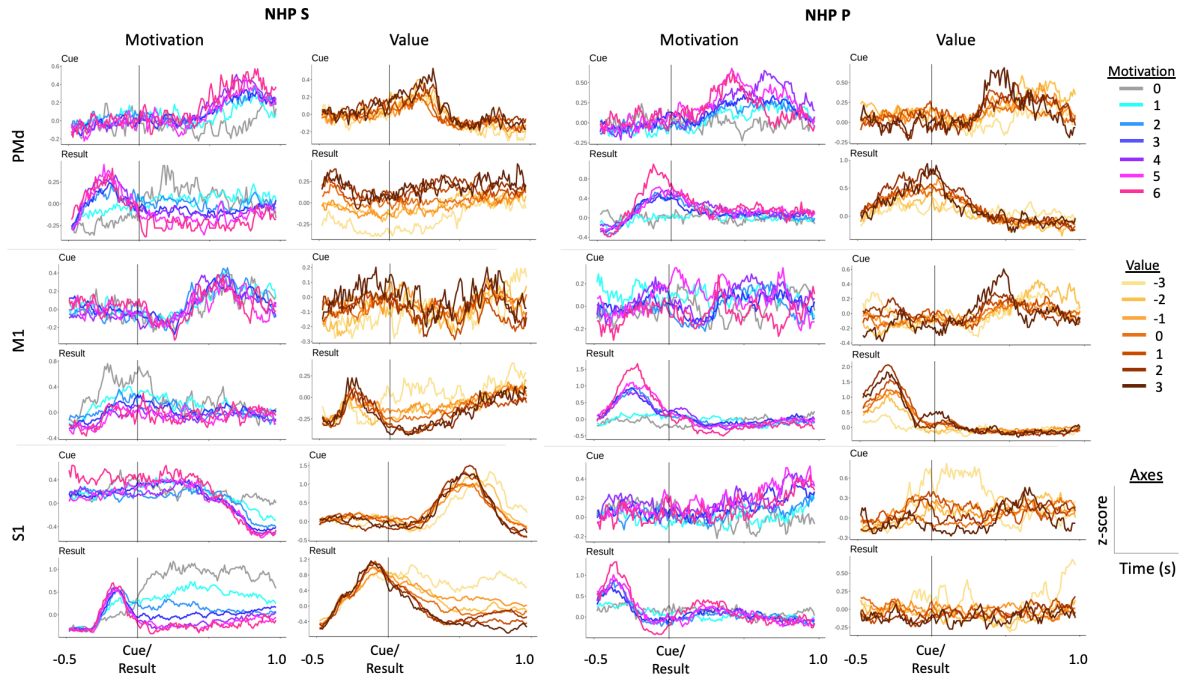

Figure. 4 shows the peri-cue and peri-feedback time histograms for several example single units from each of the brain regions recorded from both NHPs. Here we show units that preferentially represent either valence or motivational intensity. Units were binned and z-score normalized over the course of 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.

To further describe the influence of cued reward and punishment on our neural populations we analyzed the percentage of units with significant differences in mean firing rate between pre- and post-cue and pre- and post-feeback time periods. There are several clear trends that hold for both NHPs and all brain regions in this analysis. First, in the peri-cue period there is an increase in the percent of units that increase their activity in a manner that looks like a motivational intensity signal when viewed in this simple manner (differences in mean rate as compared to a divisively normalized rate). More units showed a significant increase in activity as the motivation level increased as compared to decreasing. Conversely, during the peri-feedback period both NHPs and all brain regions showed an increase in the percent of units with significant decreases in activity, however, the relationship for the increasing and decreasing units is simpler for NHP P than NHP S. NHP P shows an increase in the percent of units for both those that increase their activity and those that decrease their activity, while most of the units are decreasing their activity in the peri-feedback period that is the rate is lower post feedback as compared to pre-feedback. Note NHP P has lower relative percentages of units during the peri-cue period as compared to the peri-feedback period, and NHP S is the opposite. In addition, PMd for both NHPs is more evenly distributed between increasing and decreasing units in the peri-cue period as compared to M1 and S1 which are dominated by units increasing their rate post cue. If one were to stop their analysis at this point it might seem that the neural activity in these regions are representing motivational intensity and not valence. However, as we go on to show below, by using divisive normalization models we see that the relationship is more evenly split between a valence representation and motivational intensity.

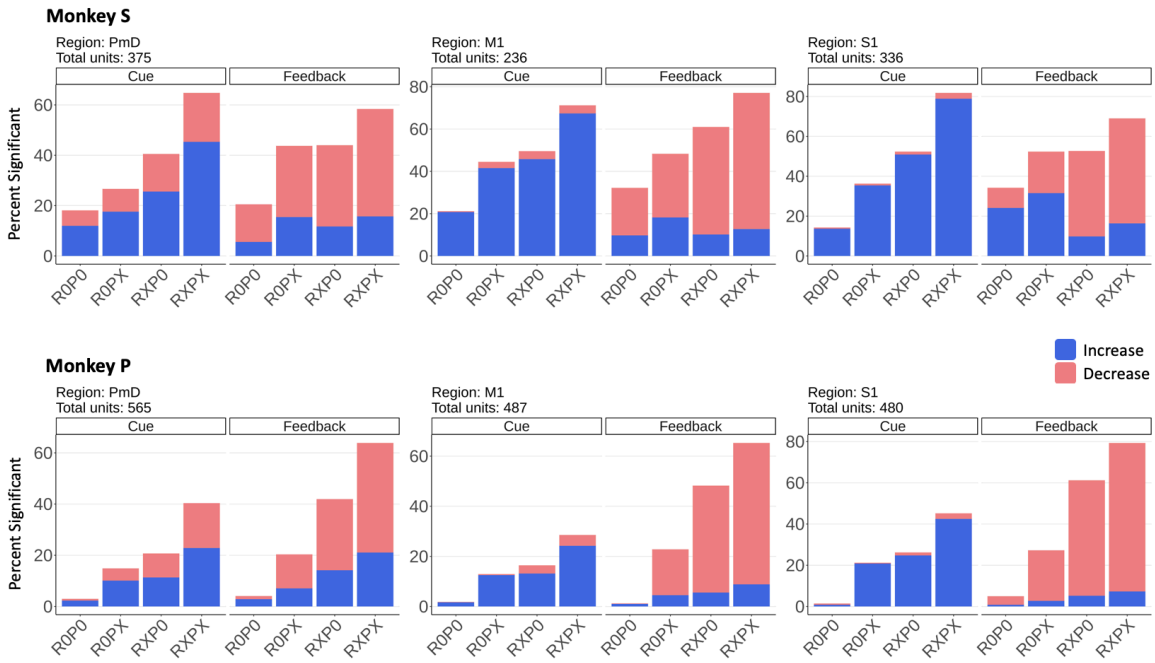

Figure.5. Percent of units showing increased or decreased firing rates during the post-cue and post-feedback period compared to the pre-cue or pre-feedback periods. The x-axis is ordered as motivation moving from the lowest motivation group R0P0 up to the highest RXPX, where here X is a place holder for any value other than zero.

Divisive normalization of neural and behavioral variables by cued reward and punishment level

It is clear from Figs.2 and 3 that some of the relationships between neural firing rate and cued valence and motivational intensity levels are not simply linear, as seen for the reward-only results. Therefore, we tested several divisive normalization models to explore this further. We started with two models, one that utilized the categorical reward and punishment levels from the task (model 2, see below and Figs.6 and 7) and another, model 3, that utilized the population activity in the divisive term and may be a more natural model form (see methods and Fig. 7). In Fig.6 we present the trial time results as seen in Fig.2, but now utilizing cued punishment and reward levels simultaneously (model 2_{dur}). As in Fig.2 we see that NHP P is very sensitive to differences between rewarding and non-rewarding trials. However, now we see that in addition to this there is an influence on trial duration due to the cued punishment level Fig.6.1, which is most clear when the reward level is zero. In these Fig.6 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. NHP S again had a smoother relationship between affective stimuli and trial time as seen in Fig.2. These (Fig.6.1) results demonstrate that the mean trial time can be explained with a motivational intensity encoding model as $B_2 > 0$ for both NHPs for Model 2_{dur}, which means that reward levels and punishment levels are being added, and tune to the same direction as expected for motivational intensity as compared to valence, where they would be subtracted. As in the reward-only case, the mean trial time decreases as the motivational intensity level increases. Trial duration as a function of reward and punishment levels has a sigmoidal form as it does for reward only. Individual unit mean firing rates as a function of the affective stimuli were investigated as well utilizing model 2, $f_{ij} = \frac{\alpha_{2j}(r_i+\gamma_{2j}p_i)}{\delta_{2j}+(r_i+\gamma_{2j}p_i)} + \beta_{2j}$ (fig 6.2), plotted as red squares

with the standard error of the mean (SEM) for each stimuli, of which there are 16. Also plotted (blue) are the outputs of model 2. With model 2, unit reward and punishment modulation as represented by the model falls into one of three categories: linear (Fig. 6.2 row 1), sigmoidal (row 2), and hyperbolic (row 3). These were seen in both animals and in all three cortical regions investigated (PMd, M1 and S1).

Fig 6.

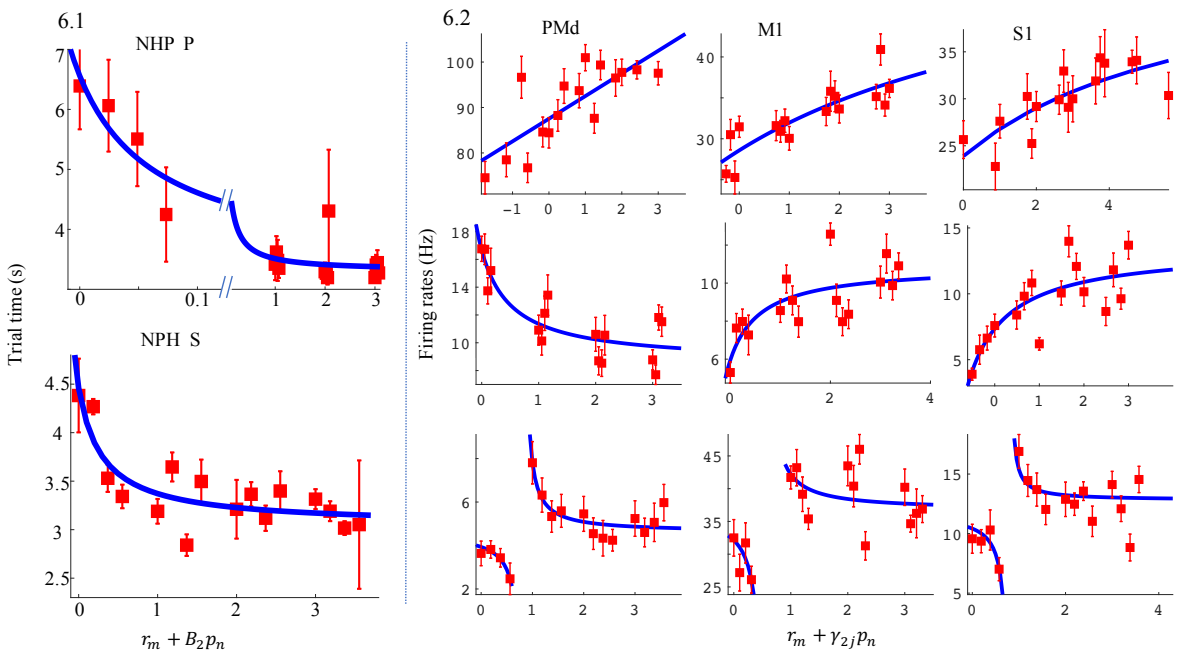

Figure 6. Post-cue reward and punishment analysis with divisive normalization model 2. Figure 6.1 depicts the average and SEM of the trial duration (red) for different stimuli (Model 2_{dur}). The x-axis represents the reward and punishment level in the form of an affective stimulus $r_m + B_2 p_n$, and the y-axis represents the trial duration (s). Plotted in blue are the output from model 2_{dur}. Figure 6.2 shows example unit reward and punishment modulation for post-cue data. Each unit has a reward and punishment encoding model 2 significantly different from the constant model (F-test, $p<0.05$). For each subplot, the x-axis represents the affective stimuli, a linear combination of reward and punishment where $s_i = r_m + \gamma_{2j} p_n$, and the y-axis represents the post-cue firing rate (Hz). Each red point represents the mean post-cue firing rate (0~500ms after cue display) for that reward level with the SEM, and the blue line represents model 2 fit to that unit. 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, sigmoid, and hyperbolic units.

Model 2 utilizes the categorical reward and punishment levels in the denominator, which act as the divisive, or normalizing term. We wished to see if a more natural divisive term could perform as well or better (model 3), and derived this term from the given cortical region's population activity (eqs. 10 and 11) as described in the methods section,

$$\text{Model 3 } f_{ij} = \frac{\alpha_{3j}(r_i + \gamma_{3j} p_i)}{\delta_{3j} + s_i} + \beta_{3j}.$$

We then compared the fitting results for units to models 2 and 3 (Fig. 7). Model 3 has a higher fitting accuracy for all example units shown, and visually better "explains" the modulation due to the stimulus. Units can have two general characteristics to their responses: focal tuning to a small reward and punishment range, and a more gradual response to a wider range of reward and punishment levels. Model 2 can only capture either the focal tuning (hyperbolic

units) or the broader modulation (linear or sigmoid units), but model 3 can capture both types of modulation as seen in Fig.7. Here we show the qualitative results for a set of example units and follow with the population results and the model comparisons.

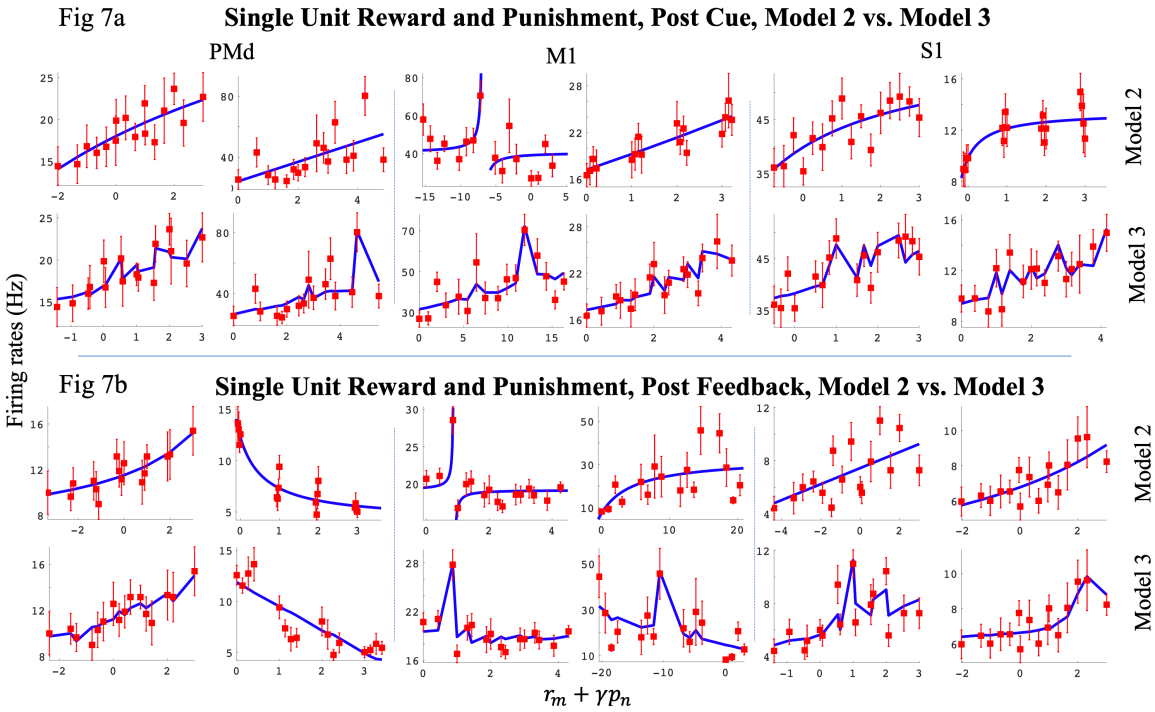

Figure 7. Post-cue and post-feedback example units for model 2, which utilizes task reward and punishment labels (top rows in each subplot) and model 3, which utilizes information from the local brain regions population activity in the divisive term (bottom rows in each subplot). The x-axis represents the affective stimuli, a linear combination of reward and scaled punishment $r_m + \gamma p_n$. The y-axis represents the post-cue (fig.7a) and post-feedback (fig.7b) firing rate (Hz). The mean and SEM of the firing rates (red) and 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 have a model fit significantly different from the constant model (F-test, $p<0.05$).

Figure 7b shows results for example units utilizing model 2 or 3 in the post-feedback window. 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 $f_{ij} = \frac{\alpha_{2j}(r_i+\gamma_{2j}p_i)}{\delta_{2j}+(r_i+\gamma_{2j}p_i)} + \beta_{2j}$ (6)

Model 3 is $f_{ij} = \frac{\alpha_{3j}(r_i+\gamma_{3j}p_i)}{\delta_{3j}+s_i} + \beta_{3j}$ (10)

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

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

Model 6 is $f_{ij} = \frac{\alpha_{6j}(r_i)}{\delta_{6j}+(s_i)} + \beta_{6j}$ for successful trials only (14)

Model 7 is $f_{ij} = \frac{\alpha_{7j}(p_i)}{\delta_{7j}+(s_i)} + \beta_{7j}$ for failure trials only (15)

The adjusted R-squared was utilized for cross-model comparisons. The results are shown in figure 8. Overall, NHP P actually had better model fits for group 1 (considering reward and punishment together for all trials), whereas NHP S showed the response we had anticipated, that is a better fit in the post feedback period with group 2 models (only considering reward for successful trials and only considering punishment for failed trials). Note that even though the two NHPs have different patterns of responses they are internally consistent between their own brain regions and thus these seem to be actual differences in the NHPs response to these affective manipulations. These results follow the trend that NHP P was more sensitive during the post cue presentation and NHP S was more sensitive to the post feedback period.

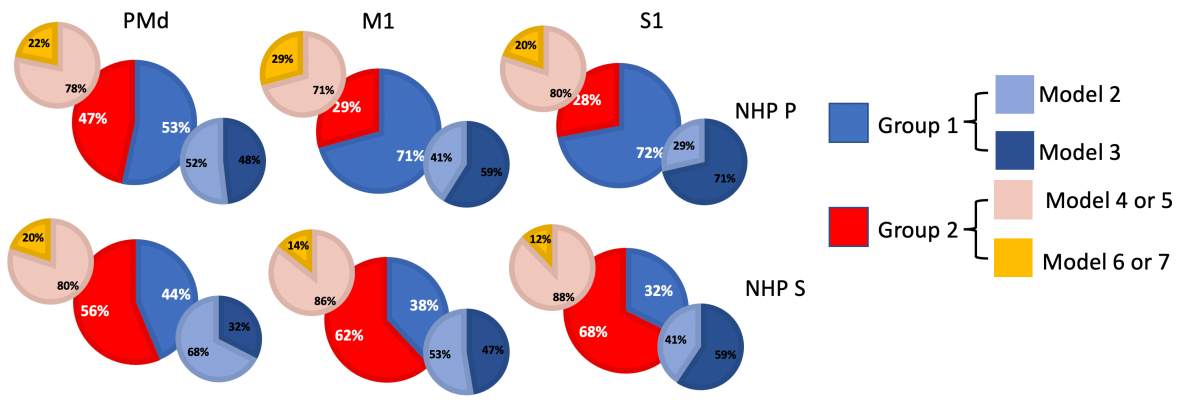

Figure 8. Percentage of units firing rate best fit for a given divisive normalization model. Group 1 models utilize the full reward and punishment stimulus information while group 2 models only utilize reward information and consider successful trials, or only utilize punishment information on failed trials. All data utilized here was during the post feedback period.

It has been suggested that premotor regions encode motivation more than value (Roesch and Olson 2004), however, this pervious work did not utilize divisive normalization, and as seen in Fig.5, under such analysis we also found what appears to be an overwhelming motivational intensity trend. In order to determine if divisive normalization could shed light on the cortical representation of motivational intensity and valence we have plotted the total number of units with significant fits to either a motivational intensity form or a valence form of either model 2 or 3, whichever was best for a given unit. In our model formulations we can determine if a unit is best fit to either motivational intensity or valence based on the sign of the γ parameter. If γ > 0 for a given unit, meaning that reward and punishment both modulated the firing rate in the same direction, the unit is considered to encode motivational intensity. However, if γ < 0 for a given unit we define this as a valence unit, meaning that reward and punishment had the opposite modulation for a given unit. We have plotted the results from this analysis in Fig.9. Again, we see that there are clear and consistent differences between the two NHPs while each NHP is consistent between brain regions. In Fig.9 we see that NHP S always has a stronger post feedback response as compared to NHP P, which generally has a stronger post cue response, except for in M1 for valence. The general trend is also consistent for each NHP for both their motivational and valence representations. The only brain region that does not follow the NHP specific patters is M1 in NHP P for valence, which is actually highest in the post feedback period.

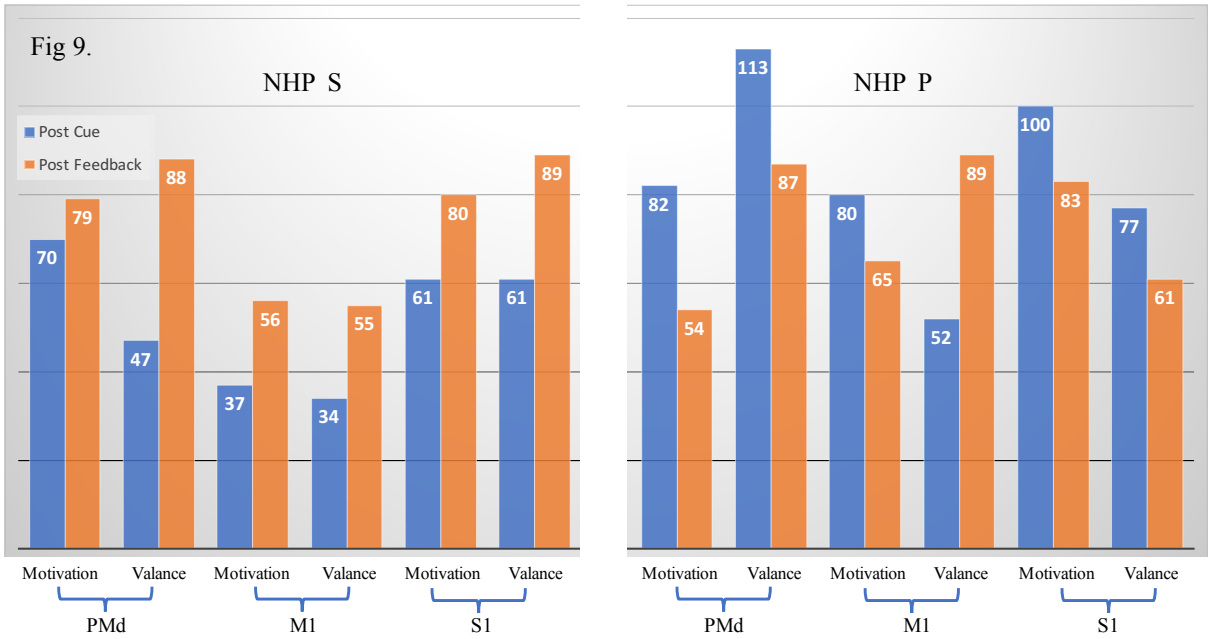

Figure 9. Number of units with significant divisive normalization representations of motivational intensity and valence. Plotted are the number of units for each NHP and brain region with significant fits to either model 2 or 3 (see Fig. 10 for distributions), for either motivational intensity or valence, during both the post-cue and post-feedback time periods.

In order to determine if and how units change their representation between motivational intensity and valence during different portions of the trials we looked at these distributions in the post-cue and the post-feedback time periods and plot this information in Figs.10 and 11. Here in Fig. 10 we have broken the population of all units into the following categories, units that were not significant, only significant in one period and significant in both periods. We then further broke the units down into motivational intensity and valence, and for units significant in both time periods whether they maintained the same tuning or not. More than 70% of the units show significance for at least one time period for all regions. In addition, in fig.11 we show the flow of these populations of units from the post-cue time period to the post-feedback time period. In Fig.11 one can see that for NHP S a large portion of the units that become significant in the post feedback period where in fact insignificant during the post-cue period. For NHP P a large portion of the population shows significance during the post-cue and during the post-feedback period. This trend is also seen in Fig. 10 for NHP S where a large portion of the units are seen to be only significant in the post feedback period (dark blue) as compared to NHP P's results.

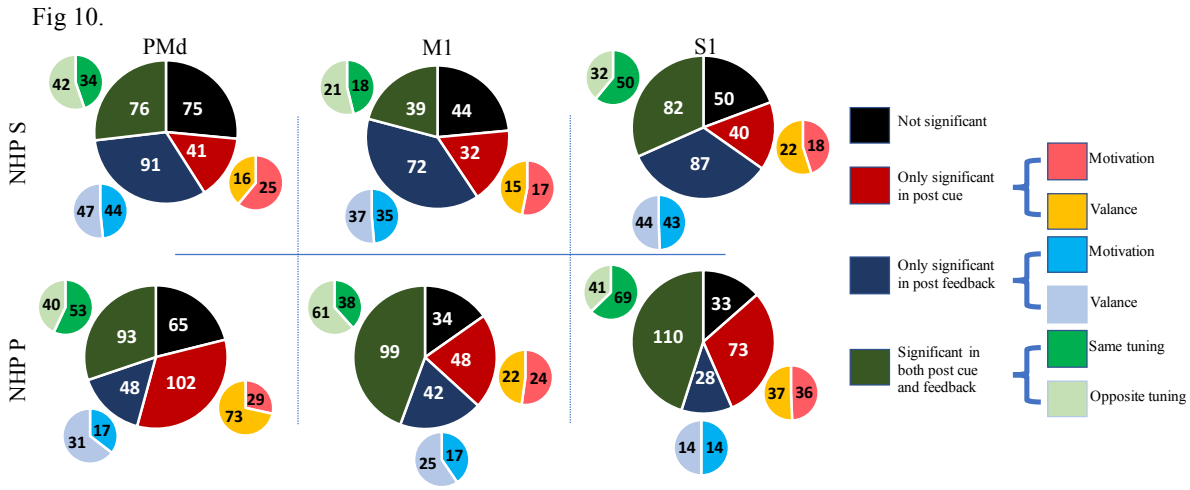

Figure. 10. Distribution of units from two NHPs for three sensorimotor brain regions fit to models 2 and 3 (see methods and text).

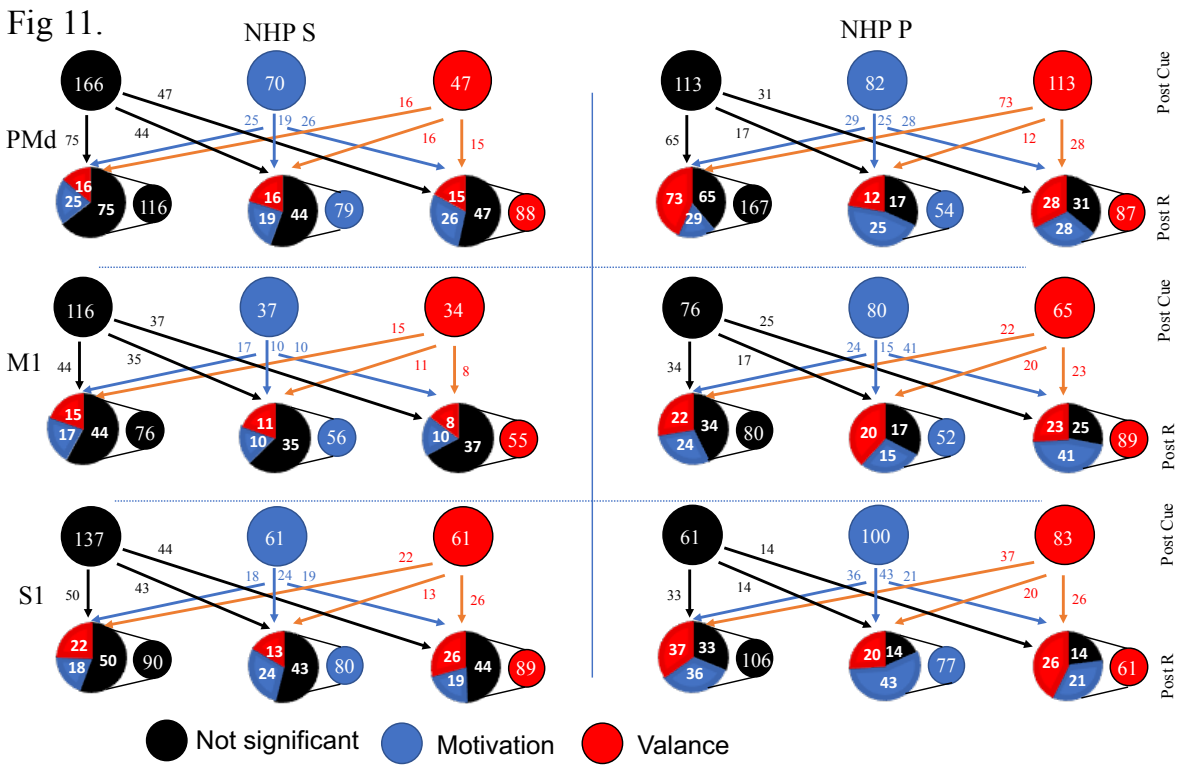

Figure 11. Flow of units from the post-cue to the post-feedback time periods. For each NHP and brain region we have plotted the number of units that where either insignificant (black), or significant for motivational intensity (blue) or valence (red). Top rows are for the post-cue period and bottom rows are for the post-feedback period. For the post feedback period we have color coded where the units came from during the post-cue period.

Model comparison

Akaike Information Criterion (AIC) 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). In addition, we have plotted the distribution of the adjusted R squared values for the best significant model for each unit from each NHP and brain region and show these distributions in Fig. 12. Note that although NHP S's responses are focused in the post-feedback period the model fits tend to be strongest for this NHP.

Table 1.

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 1. Percentage of units with better AIC (Akaike Information Criterion) for a given model comparison. PC = post cue, PF = post feedback.

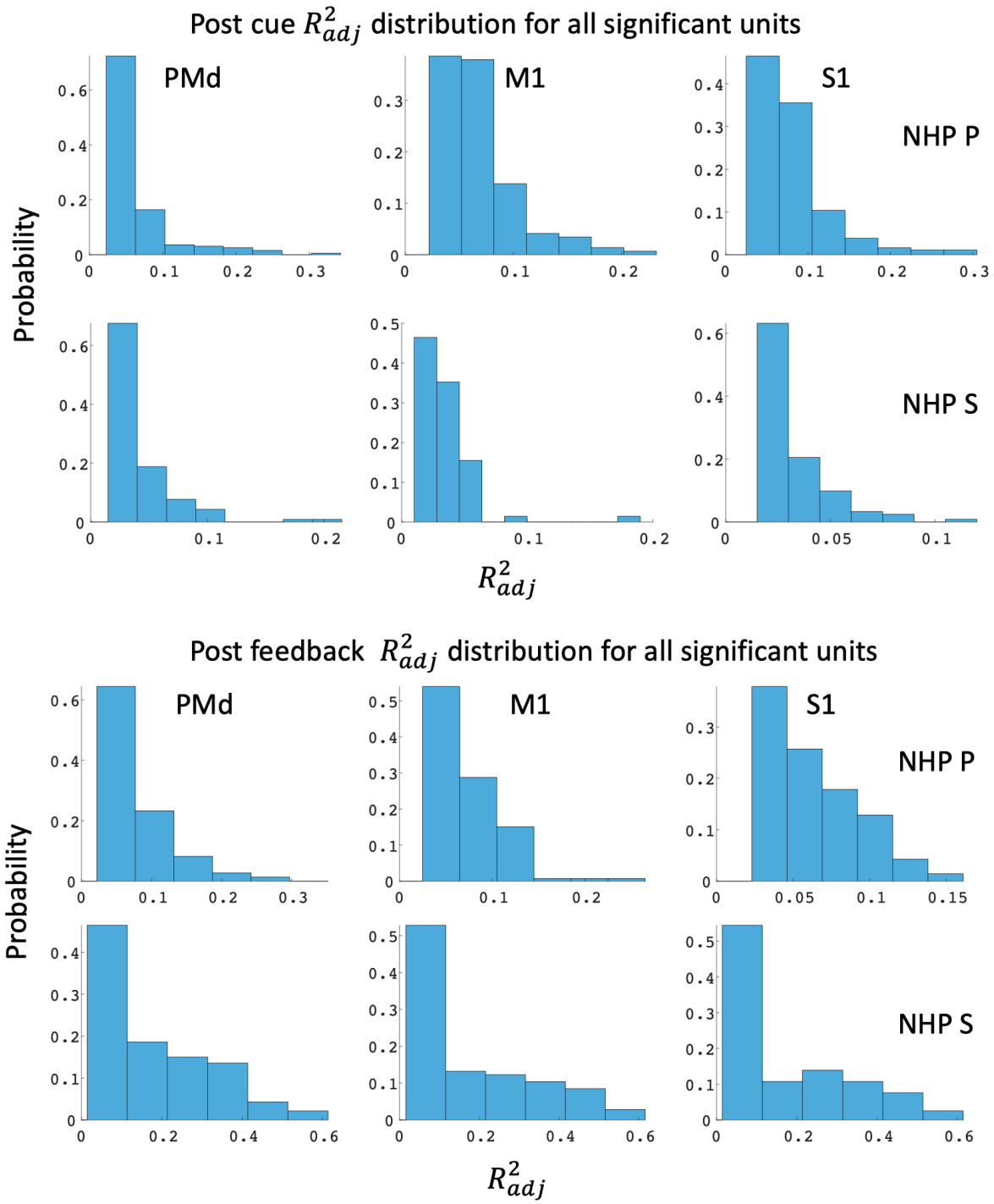

Figure. 12. Distribution of the adjusted R squared values for the best significant model fit for a given unit during the post-cue, top two rows, and the post-feedback, bottom two rows, for the two NHPs labeled on the right-hand side of the figure.

Discussion

Here we report evidence in support of widespread divisive normalization within the primary sensorimotor cortices (M1, S1 and PMd) of non-human primates by valence and motivational intensity. In this work we had two non-human primate subjects make targeted grip force movements. We modulated the level of cued valence, spanning from positive (reward) to

negative (punishment), and cued motivational intensity, which was modulated by simultaneously cueing the subjects to the possible level of reward they could receive if successful on a trial, and the possible level of punishment they would receive if unsuccessful. There were 4 levels of cued reward and 4 levels of cued punishment, leading to 16 different levels of valence and 16 different levels of motivational intensity, which we termed the affective stimuli in our models. We simultaneously recorded from 96 electrodes in each M1, S1 and PMd during this work for a total of 288 electrodes in each NHP and a total of 576 electrodes.

We found three prominent relationships between neural firing rate and affective stimuli in our divisive normalization analysis, regardless of whether we only modulated reward levels (see Fig.2), or modulated the more complete affective stimuli of valence and motivational intensity (see Figs.6 and 7). These relationships where either linear, sigmoidal, or hyperbolic. Sigmoidal units were characterized, as expected, by a two-part firing rate response to the affective stimuli (reward and punishment), consisting of a region of relatively high sensitivity to affect, and another that was less sensitive. These units allow affect to modulate the neural firing rate over a wide range of stimuli, and it is possible that our linear units become sigmoidal units if the affective stimuli have a broader range beyond what we investigated. This last point indicates that we suspect there is a distribution of units such that some have broad affective modulation ranges, and these looked linear in our tested affective space, while others have a more focal affective range, such as the sigmoidal units in our work. Hyperbolic units are particularly sensitive to a small range of the affective stimuli where the denominator in model #2 is close to zero, which is $\delta_j + r_i \approx 0$ for the reward-only task and $\delta_{2j} + (r_m + \gamma_{2j} p_n) \approx 0$ for the reward/punishment combination task. We could not determine the asymptotes for models 3, 6 and 7 as they utilize the population's activity in the divisive term, unlike model #2 that simply uses the R and P levels, and thus we could interpolate model #2 and find the asymptotes. The hyperbolic units from model #2 have their asymptotes concentrated around the extremes of the affective stimulus space, such as at R0P0 and R3P3, further work is needed to determine if this is a function of the particular set of stimuli or a more generalizable phenomenon. However, the model #2 observations hint that the sensorimotor cortices "know" where the boundaries are for the given affective context and future work will test this for the same population of units between context with differing numbers of affective stimuli.

In general, the population showed an increase in activity during the post-cue period and a significant suppression in activity in the post-feedback period. During both the post-cue and post-feedback time periods we found that the divisive normalization models (models 2, 3) clearly outperformed the linear model (model 1) as seen by the AIC comparisons in table 1. This suggests that for PMd, M1, and S1, reward and punishment modulation is nonlinear, and that this nonlinearity can at least in part be captured by divisive normalization models. Model 3 generally performed better than model 2 for NHP P, suggesting that the accuracy of divisive normalization models incorporating affect can be improved by including population firing rates in the normalization term, as model 2 used the task labels to define the affective stimulus, whereas model 3 utilized the deviation of the population activity for a given trial as compared to the "baseline" affective stimulus or r=p=0. This type of normalization, scaled to the response of a population of units, has also been described in the invertebrate olfactory system and the primary visual cortex (Carandini & Heeger 2012), further supporting the idea that divisive normalization is a computational mechanism performed widely, and in a variety of contexts. Also, of importance is the fact that this type of normalization by the population can be more easily utilized toward stabilizing BCIs in the face of changing affect as discussed below.

Previously, work on the influence of value in the sensorimotor stream has shown value based divisive normalization in the lateral intraparietal cortex (Louie, Grattan, and Glimcher 2011), where this activity was interpreted as encoding the action value within the context of choices between actions associated with different state values. In our study the subjects did not have options to choose from, other than the choice to not conduct the task, and simply had one type

of movement that they could perform successfully or not. Thus, in our work we were more likely studying state values, or state motivational intensity, as the actions were more or less the same in each trial. Outcome values could clearly have been represented in our work as well. Future work where we explicitly include choice should be helpful in further determining if both state-value and action-value are being represented in PMd, M1 and S1 (Tarigoppula et al. 2018; Ramakrishnan et al. 2017). 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, as we also see this relationship when looking at just the firing rate of units as seen in Fig.5. When we utilized divisive normalization models though, we see a more even representation between motivation and valence. 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 a more evenly distributed manner, and this distribution can depend on the time during the trial as well as on individual “personality” traits of the NHP subjects as seen in Fig.9.

Understanding modulation due to affect is relevant to biomedical engineers creating BCIs with activity from sensorimotor regions (Chhatbar and Francis 2013; Gilja et al. 2012; Ajiboye et al. 2017; Hochberg et al. 2006; Wodlinger et al. 2015; Carmena et al. 2003; Chapin et al. 1999; Taylor, Tillery, and Schwartz 2002; Sanchez et al. 2011). For control signals to be accurately decoded and encoded while retaining these additional levels of information, such as affect, it is important to determine if these units are utilizing an affective divisive normalization scheme. 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, Schwartz, and Chase 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 later study noted that grip force related tuning functions are likewise modulated by reward in M1. Such affect-related spillover into 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 (Choi et al. 2016; Marsh et al. 2015; Zhao et al. 2018; McNiel, Bataineh, et al. 2016; McNiel, Choi, et al. 2016; Bensmaia and Miller 2014). It is therefore vital to better understand the influence that affect has on the sensorimotor regions past our neuroscientific curiosity towards these iBCI engineering goals. Our future aim is to determine how to best separate movement related activity from these non-sensorimotor influences towards affect agnostic iBCIs.

Ajiboye, A. B., F. R. Willett, D. R. Young, W. D. Memberg, B. A. Murphy, J. P. Miller, B. L. Walter, J. A. Sweet, H. A. Hoyen, M. W. Keith, P. H. Peckham, J. D. Simeral, J. P. Donoghue, L. R. Hochberg, and R. F. Kirsch. 2017. 'Restoration of reaching and

grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration', *Lancet*, 389: 1821-30.

An, J., T. Yadav, M. B. Ahmadi, V. S. A. Tarigoppula, and J. T. Francis. 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., T. Yadav, J. P. Hessburg, and J. T. Francis. 2019. 'Reward Expectation Modulates Local Field Potentials, Spiking Activity and Spike-Field Coherence in the Primary Motor Cortex', *eNeuro*, June 6 2019.

Bensmaia, S. J., and L. E. Miller. 2014. 'Restoring sensorimotor function through intracortical interfaces: progress and looming challenges', *Nat Rev Neurosci*, 15: 313-25.

Cai, X., and C. Padoa-Schioppa. 2012. 'Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex', *J Neurosci*, 32: 3791-808.

Carandini, M., and D. J. Heeger. 2011. 'Normalization as a canonical neural computation', *Nat Rev Neurosci*, 13: 51-62.

Carmena, J. M., M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. Nicolelis. 2003. 'Learning to control a brain-machine interface for reaching and grasping by primates', *PLoS Biol*, 1: E42.

Chapin, J. K., K. A. Moxon, R. S. Markowitz, and M. A. Nicolelis. 1999. 'Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex', *Nat Neurosci*, 2: 664-70.

Chhatbar, P. Y., and J. T. Francis. 2013. 'Towards a naturalistic brain-machine interface: hybrid torque and position control allows generalization to novel dynamics', *PLoS One*, 8: e52286.

Choi, J. S., A. J. Brockmeier, D. B. McNiel, L. M. Kraus, J. C. Principe, and J. T. Francis. 2016. 'Eliciting naturalistic cortical responses with a sensory prosthesis via optimized microstimulation', *Journal of Neural Engineering*, 13: 056007.

Downey, J. E., L. Brane, R. A. Gaunt, E. C. Tyler-Kabara, M. L. Boninger, and J. L. Collinger. 2017. 'Motor cortical activity changes during neuroprosthetic-controlled object interaction', *Sci Rep*, 7: 16947.

Georgopoulos, A. P., J. Ashe, N. Smyrnis, and M. Taira. 1992. 'The motor cortex and the coding of force', *Science*, 256: 1692-5.

Georgopoulos, A. P., J. F. Kalaska, R. Caminiti, and J. T. Massey. 1982. 'On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex', *J Neurosci*, 2: 1527-37.

Gilja, V., P. Nuyujukian, C. A. Chestek, J. P. Cunningham, B. M. Yu, J. M. Fan, M. M. Churchland, M. T. Kaufman, J. C. Kao, S. I. Ryu, and K. V. Shenoy. 2012. 'A high-performance neural prosthesis enabled by control algorithm design', *Nat Neurosci*, 15: 1752-7.

Heeger, D. J. 1992. 'Normalization of cell responses in cat striate cortex', *Vis Neurosci*, 9: 181-97.

Hochberg, L. R., M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue. 2006. 'Neuronal ensemble control of prosthetic devices by a human with tetraplegia', *Nature*, 442: 164-71.

Li, C. S., C. Padoa-Schioppa, and E. Bizzi. 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., L. E. Grattan, and P. W. Glimcher. 2011. 'Reward value-based gain control: divisive normalization in parietal cortex', *J Neurosci*, 31: 10627-39.

Marsh, B. T., V. S. Tarigoppula, C. Chen, and J. T. Francis. 2015. 'Toward an autonomous brain machine interface: integrating sensorimotor reward modulation and reinforcement learning', *J Neurosci*, 35: 7374-87.

McNiel, D. B., M. Bataineh, J. S. Choi, J.P. Hessburg, and J. T. Francis. 2016. 'Classifier Performance in Primary Somatosensory Cortex Towards Implementation of a Reinforcement Learning Based Brain Machine Interface ', *IEEE Southern Biomedical Engineering Conference 2016*.

McNiel, David B, John S Choi, John P Hessburg, and Joseph T Francis. 2016. "Reward value is encoded in primary somatosensory cortex and can be decoded from neural activity during performance of a psychophysical task." In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, 3064-67. IEEE.

Padoa-Schioppa, C. 2009. 'Range-adapting representation of economic value in the orbitofrontal cortex', *J Neurosci*, 29: 14004-14.

Ramakrishnan, A., Y. W. Byun, K. Rand, C. E. Pedersen, M. A. Lebedev, and M. A. L. Nicolelis. 2017. 'Cortical neurons multiplex reward-related signals along with sensory and motor information', *Proc Natl Acad Sci U S A*, 114: E4841-E50.

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

Rangel, A., and J. A. Clithero. 2012. 'Value normalization in decision making: theory and evidence', *Curr Opin Neurobiol*, 22: 970-81.

Rasmussen, R. G., A. Schwartz, and S. M. Chase. 2017. 'Dynamic range adaptation in primary motor cortical populations', *Elife*, 6.

Rieke, Fred. 1997. *Spikes : exploring the neural code* (MIT Press: Cambridge, Mass.).

Roesch, M. R., and C. R. Olson. 2003. 'Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex', *J Neurophysiol*, 90: 1766-89.

Roesch, M. R., and C. R. Olson. 2004. 'Neuronal activity related to reward value and motivation in primate frontal cortex', *Science*, 304: 307-10.

Roesch, M. R., and C. R. Olson. 2007. 'Neuronal activity related to anticipated reward in frontal cortex: does it represent value or reflect motivation?', *Ann N Y Acad Sci*, 1121: 431-46.

Rolls, Edmund T. 2014. *Emotion and decision-making explained* (Oxford University Press: Oxford ; New York, NY).

Sanchez, J. C., A. Tarigoppula, J. S. Choi, B. T. Marsh, P. Y. Chhatbar, B. Mahmoudi, and J. T. Francis. 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)*: 525-28.

Tarigoppula, A., J. S. Choi, J.P. Hessburg, D. B. McNiel, B. T. Marsh, and J..T. Francis. 2018. 'Motor Cortex Encodes A Value Function Consistent With Reinforcment Learning ', *bioRxiv*.

Taylor, D. M., S. I. Tillery, and A. B. Schwartz. 2002. 'Direct cortical control of 3D neuroprosthetic devices', *Science*, 296: 1829-32.

Wodlinger, B., J. E. Downey, E. C. Tyler-Kabara, A. B. Schwartz, M. L. Boninger, and J. L. Collinger. 2015. 'Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations', *Journal of Neural Engineering*, 12: 016011.

Yamada, H., K. Louie, A. Tymula, and P. W. Glimcher. 2018. 'Free choice shapes normalized value signals in medial orbitofrontal cortex', *Nature Communications*, 9: 162.

Zhao, Y., J. P. Hessburg, J. N. A. Kumar, and J. T. Francis. 2018. 'Paradigm Shift in Sensorimotor Control Research and Brain Machine Interface Control: The Influence of Context on Sensorimotor Representations', *Frontiers in Neuroscience*, 12.

Zimmermann, J., P. W. Glimcher, and K. Louie. 2018. 'Multiple timescales of normalized value coding underlie adaptive choice behavior', *Nature Communications*, 9: 3206.