Cortical neurons multiplex reward-related signals along with sensory and motor information

Arjun Ramakrishnan[a,b], Yoon Woo Byun[b,c], Kyle Rand[c], Christian E. Pedersen[d], Mikhail A. Lebedev[a,b], and Miguel A. L. Nicolelis[a,b,c,e,f,g,1]

[a]Department of Neurobiology, Duke University, Durham, NC 27710; [b]Duke University Center for Neuroengineering, Duke University, Durham, NC 27710; [c]Department of Biomedical Engineering, Duke University, Durham, NC 27708; [d]Joint Department of Biomedical Engineering, University of North Carolina-Chapel Hill and North Carolina State University, Raleigh, NC 27695; [e]Department of Psychology and Neuroscience, Duke University, Durham, NC 27708; [f]Department of Neurology, Duke University, Durham, NC 27710; and [g]Edmund and Lily Safra International Institute of Neurosciences, Natal 59066060, Brazil

Rewards are known to influence neural activity associated with both motor preparation and execution. This influence can be exerted directly upon the primary motor (M1) and somatosensory (S1) cortical areas via the projections from reward-sensitive dopaminergic neurons of the midbrain ventral tegmental areas. However, the neurophysiological manifestation of reward-related signals in M1 and S1 are not well understood. Particularly, it is unclear how the neurons in these cortical areas multiplex their traditional functions related to the control of spatial and temporal characteristics of movements with the representation of rewards. To clarify this issue, we trained rhesus monkeys to perform a center-out task in which arm movement direction, reward timing, and magnitude were manipulated independently. Activity of several hundred cortical neurons was simultaneously recorded using chronically implanted microelectrode arrays. Many neurons (9–27%) in both M1 and S1 exhibited activity related to reward anticipation. Additionally, neurons in these areas responded to a mismatch between the reward amount given to the monkeys and the amount they expected: A lower-than-expected reward caused a transient increase in firing rate in 60–80% of the total neuronal sample, whereas a larger-than-expected reward resulted in a decreased firing rate in 20–35% of the neurons. Moreover, responses of M1 and S1 neurons to reward omission depended on the direction of movements that led to those rewards. These observations suggest that sensorimotor cortical neurons corepresent rewards and movement-related activity, presumably to enable reward-based learning.

reward | motor cortex | multichannel recording | primate | prediction error

Reward-based learning is a fundamental mechanism that allows living organisms to survive in an ever-changing environment. This type of learning facilitates behaviors and skills that maximize reward (1). At the level of motor cortex, synaptic modifications and formation of new synaptic connections that accompany learning have been reported (2). The reinforcement component of these plastic mechanisms is partly mediated by dopamine (DA), a key neurotransmitter associated with reward (3–5). DA inputs to cortical areas, including primary motor (M1) and somatosensory (S1) areas, arise from the midbrain (6–8). Recent studies demonstrating the presence of a higher concentration of DA receptors in the forelimb over the hind-limb representation in rodent motor cortex (9), and impairment of motor skill acquisition resulting from blocking these DA receptors in the forelimb representation (10), point to the role of DA receptors in learning skilled forelimb movements. In agreement with this line of reasoning, animals with a rich repertoire of skilled movements, such as monkeys and humans, show a disproportionate increase in the density of DA receptors in the cortical motor areas: M1, premotor, and supplementary motor cortex (6, 11, 12).

Although evidence points to the involvement of M1 and S1 circuitry in reward-related learning and plasticity (1, 13, 14), it is still not well understood how M1 and S1 neurons process both reward and motor information. A longstanding tradition in

neurophysiological investigations of M1 and S1 was to attribute purely motor and sensory functions to these areas. Accordingly, the majority of previous studies have focused on the relationship of M1 and S1 spiking activity with such characteristics as movement onset (15), movement direction (16, 17), limb kinematics (18, 19), muscle force (20–22), temporal properties of movements (23, 24), and progression of motor skill learning (25). Fewer studies have examined nontraditional representations in M1 and S1, such as spiking activity associated with rewards (26–31). Representation of reward in M1 was provoked by several studies that, in addition to motor-related neural activity, examined neuronal discharges in these areas that represented motivation, reward anticipation, and reward uncertainty (30–33). Although these studies have confirmed that M1 neurons multiplex representation of reward with the representation of motor parameters, a number of issues need further investigation. Particularly, it is unclear whether reward-related neuronal modulations in M1 (31, 33) represent rewards per se, their anticipation, and/or the mismatch between the actual and expected rewards. Furthermore, the relationship of these reward-associated signals with the processing of motor information is poorly understood. Understanding of these interactions will help elucidate mechanisms underlying reward-based learning (34).

To this end, we have developed an experimental paradigm in which reward and characteristics of movements were manipulated independently: (*i*) Movements and reward were separated by two different delays, (*ii*) multiple movement directions were examined for their impact on reward, and (*iii*) unexpected variability in reward amounts were introduced to produce their

---

Significance

The ability to learn highly skilled movements may depend on the dopamine-related plasticity occurring in motor cortex, because the density of dopamine receptors—the reward sensor—increases in this area from rodents to primates. We hypothesized that primary motor (M1) and somatosensory (S1) neurons would encode rewards during operant conditioned motor behaviors. Rhesus monkeys were implanted with cortical multielectrode implants and trained to perform arm-reaching tasks with different reward schedules. Consistent with our hypothesis, M1 and S1 neurons represented reward anticipation and delivery and a mismatch between the quantities of anticipated and actual rewards. These same neurons also represented arm movement parameters. We suggest that this multiplexing of motor and reinforcement information by cortical neurons underlies motor learning.

mismatch to the animal's expectation of reward. Neuronal ensemble activity was recorded chronically in rhesus monkeys using multielectrode arrays implanted in both M1 and S1. These experiments revealed a widespread representation of reward anticipation in M1 and S1 neuronal populations. Additionally, many M1 and S1 neurons responded to unexpected changes in reward magnitude [i.e., represented reward prediction error (RPE)]. Moreover, neuronal responses to RPEs were directionally tuned; they varied according to the direction of the movement planned by the animal. Taken together these results indicate that predictive encoding of reward is robust in both M1 and S1 and that reward-related signals interact with neuronal directional tuning in sensorimotor cortex. As such, reward-related neuronal modulations in the monkey primary sensorimotor cortices could contribute to motor learning (35).

Results

Our central goal was to test whether M1 and S1, at both the single-cell and the ensemble level, can mediate reward-based learning. More specifically, we sought to measure cortical neuronal activity that encoded reward anticipation, unexpected changes in reward presentation, and modulation of reward-associated signals by arm movements. For this purpose, we used a task in which animals moved a realistic image of a monkey arm (avatar arm) on a computer screen using a hand-held joystick. In any given trial, the avatar hand had to be moved and placed over a circular, colored target and held over it for 3 s. A passive observation version of the task (Fig. 1C) was also implemented. In this version, a computer controlled the movement of the virtual arm while the monkey, whose limbs were restrained from moving, was rewarded for simply observing the task displayed on the computer screen. The color of the target informed the monkeys about the time when a juice reward would be delivered.

In early-reward trials, cued by a blue target, reward was delivered 1 s after target acquisition, whereas in late-reward trials, cued by a red target, reward was delivered after 2 s (Fig. 1C). In the joystick control task, after receiving the juice reward (0.66 mL, 500-ms pulse) monkeys continued to maintain the avatar arm at the target location for the remaining hold period. Holding still at the target location was encouraged by small juice rewards (50 ms, 0.03 mL) that were delivered periodically (every 500 ms) during the entire 3-s hold period. Changes in position during the hold time precluded the remaining rewards. All of the monkeys learned the requirement to hold the virtual arm still. In a majority of the trials (>90% in monkeys M and R and 80–85% in monkey C) the monkeys did not move their hands. The SD of joystick position during the hold time was 0.014 cm in monkey M, 0.06 cm in monkey C, and 0.025 in monkey R.

Because the total hold time period (3 s) was the same irrespective of the early or late reward delivery, the reward rate was maintained irrespective of trial sequence. We expected that this homogeneity of the reward rate would result in similar movement patterns for both trial types even though the rewards were given at different times after the movement was completed. Indeed, the reaction times (RT; Fig. S1C) and movement times (MT; Fig. S1B), which are proxies for motivation levels, were similar for both trial types. The early- (blue trace in cumulative histograms) and late- (red trace) reward trials were not significantly different [$P > 0.05$, Kolmogorov–Smirnov (KS) test, $n = 3$ monkeys]. The same held true in a majority of behavioral sessions (24/25 sessions in monkey M, 12/13 in monkey C, and 7/7 in monkey R). It was found that monkeys only weakly favored early reward over the late reward as inferred by a small difference between the numbers of attempted early-reward trials and late-reward trials (Fig. S1A, monkey M: 0.6%, monkey C: 2.4%, and monkey R: 1%). In summary, although monkeys had a slight tendency to initiate early-reward trials more often, there were no significant differences in the movement parameters.
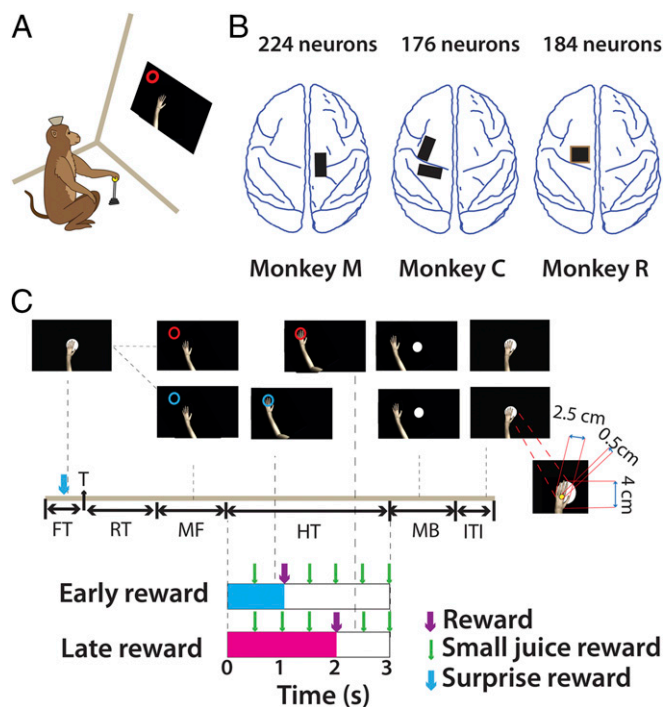

Fig. 1. Experimental setup and schematic of the paradigm. (A) Rhesus monkeys were trained to perform a center-out task using a custom-made hand-held joystick. Moving the joystick displaced an avatar monkey arm on the computer monitor. (B) Neural data were obtained from M1 neurons in three monkeys (monkeys M, C, and R) and S1 neurons in one monkey (monkey M) using chronically implanted arrays (black and brown rectangles). (C) Experiment schematic: Trial began when monkeys positioned the arm within the center target. Following fixation (FT), a circular target (T), red or blue, appeared at a distance of 9 cm from the center in one of eight target locations. Monkeys were trained to react quickly (RT) and acquire the target by moving the avatar arm (MF, movement forward). After reaching they held the arm at the target location for 3 s. A juice reward (R) was delivered either after 1 s (early reward) or after 2 s (late reward) into the hold time. Target color (blue vs. red) cued the monkey regarding reward time (early, cyan vs. late, magenta). Small juice rewards were given every 500 ms during the 3-s hold period (HT, hold time). At the end of the hold period, monkeys returned the arm to the center (MB, movement back). The next trial began after a brief intertrial interval (ITI). (Inset) The avatar monkey arm relative to a target-size disk. Note the cursor (0.5-cm yellow circle) has to be completely within the target to count as a successful "reach." In Exp. 2, in 10% of the trials out of the total a surprise reward (SR) was given to the monkey before the target appeared.

Neuronal Modulations During Different Task Epochs. A total of 485 neurons were recorded in the three monkeys (M1, monkey M: $n = 125$, monkey C: $n = 176$, monkey R: $n = 184$; S1 from monkey M: $n = 96$). Several representative neurons with different kinds of neuronal responses elicited during the task are illustrated in Fig. 2A. The data are separated into joystick control sessions (Fig. 2A, *Left*) and passive observation (Fig. 2A, *Right*), and early- (cyan) and late- (magenta) reward conditions. The associated arm position (joystick control session), the virtual arm position (passive observation session), and the electromyographies (EMGs) from biceps and triceps (passive observation session) are displayed above the perievent time histograms (PETHs). Some neurons were modulated by the position of the arm, that is, neuronal firing increased (e.g., RHA1_005a) or decreased (e.g., RHA3_008a) when the arm moved away from the center. Some neurons were modulated by the arm movement velocity: The neuronal firing increased as a function of the tangential velocity (e.g., RHA3_012b). Some neurons were modulated by arm position and reward, meaning that neuronal firing increased as the arm moved away from the center and decreased till the end of reward delivery epoch and increased right after, and, finally,
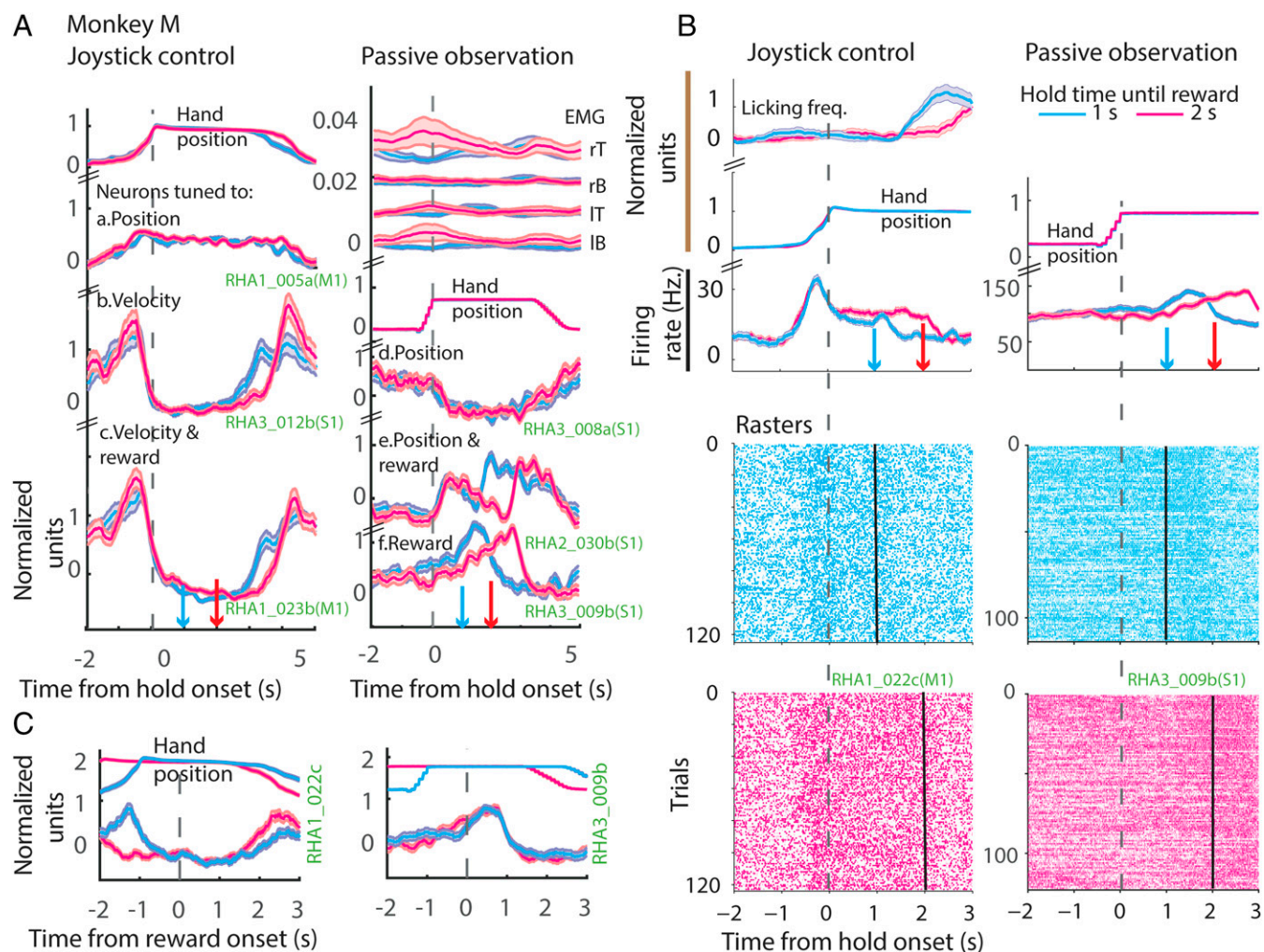
Fig. 2. Neuronal modulation during task epochs. (*A*) Example neurons from a joystick control session (*Left*) and passive observation session (*Right*) from Exp. 1. (*Left*) The changes in the mean (z-scored) arm position (at the top), followed by binned (50 ms), z-scored firing rate (PETH) of example neurons, all aligned to onset of hold time. (*Right*) A passive observation session; the EMG activations are displayed at the top. EMGs were measured from left (lT) and right triceps (rT), and left (lB) and right biceps (rB). Then, the virtual arm position is plotted. Neurons tuned to position, velocity, reward, or a combination of those is displayed next. Bands around the mean represent ±1 SE. Data are color-coded based on the time of reward: early (blue) and late (red). Time of reward is indicated in the upper panels by color-coded downward-pointing arrows. Note that the arm is stationary during the hold period. Weak muscle activations were seen around the time of movements in the passive observation session. (*B*) Example neurons plotted in the same format as in *A*; however, the *y* axis depicts the absolute firing rate modulation. The lower panels show spike raster plots. (*C*) PETH of the example neurons shown in *B* aligned to the time of reward onset. *y* axis depicts the normalized firing rate modulation.

spiking activity decreased again when the arm moved back to the center (e.g., RHA2_030b). Some arm velocity-related neurons also maintained spiking activity at a level higher than baseline till the end of the reward delivery epoch (e.g., RHA1_023b; see also RHA1_022c in Fig. 2*B*). Some neurons did not respond to position or velocity but increased their spiking activity till the end of the reward delivery epoch (e.g., RHA3_009b). Note that the neurons in a column, with different response properties, were all recorded simultaneously during the same session.

Fig. 2 *B* and *C* show two example M1 and S1 neurons whose spiking activity distinguished early from late reward. PETHs shown in Fig. 2*B* are aligned on hold onset and those in Fig. 2*C* are aligned on reward onset. The data are separated into joystick control sessions (*Left*) and passive observation (*Right*) and early- (cyan) and late- (magenta) reward conditions. The associated arm/hand position and licking frequency (joystick control session) and the virtual arm position (passive observation session) are displayed above the PETHs as in Fig. 2*A*. Spike rasters are also included in Fig. 2*B*. Although the arm position was stable (RHA1_023b in Fig. 2*A* and

RHA1_022c in Fig. 2*B*), and the EMGs (seen in Fig. 2*A*) did not show any significant muscle contractions around the time of reward, the M1 and S1 neurons exhibited differences in firing rate before reward onset ($P < 0.05$, two-sided *t* test).

Reward-Related Modulation in the M1/S1 Neural Population. The population average PETH for monkeys M, C, and R during joystick control are displayed in Fig. 3*A* and during passive observation conditions for monkeys M and C in Fig. S2*A*. Population PETHs, centered on hold onset, for the early- (cyan) and late-reward trials (magenta) show that the neuronal ensemble average firing rate is very high for movement (t < 0), except in monkey M for passive observation. The average spiking activity decays during the hold time up to the time of reward. Early-trial neural activity can be distinguished from late-trial activity, especially in monkey C (both joystick control and passive observation) and monkey M (joystick control). In all of the monkeys the average PETHs were not very different for the early and late trials before the onset of the hold period.

Simply averaging the spiking activity across the neural population may conceal some differences. Thus, to examine whether the neuronal ensemble can distinguish early- from late-reward trials, before the early reward (0–1 s of the hold period) we used a machine learning algorithm–the support vector machine (SVM). For that, we used a 500-ms-wide moving window along the trial time axis, while separate SVM classifiers were trained for each window position to decode the trial into early vs. late rewards. Different window durations, step sizes, and algorithms were also tested (*Materials and Methods*). To compare the representation by different cortical areas of the early- vs. late-reward conditions, the classifier was also trained on different pools of neurons (ipsilateral and contralateral M1, S1, etc.). Irrespective of whether the neurons were selected from the ipsilateral or the contralateral M1, the classifier could clearly dissociate early- from late-reward trials in the initial 1-s hold period preceding any of the rewards [shaded window in Fig. 3*B*, monkey C: contralateral M1, receiver operating characteristic (ROC): 0.77 ± 0.11; ipsilateral M1, ROC: 0.72 ± 0.09; *P* < 0.05, ROC post-SVM].
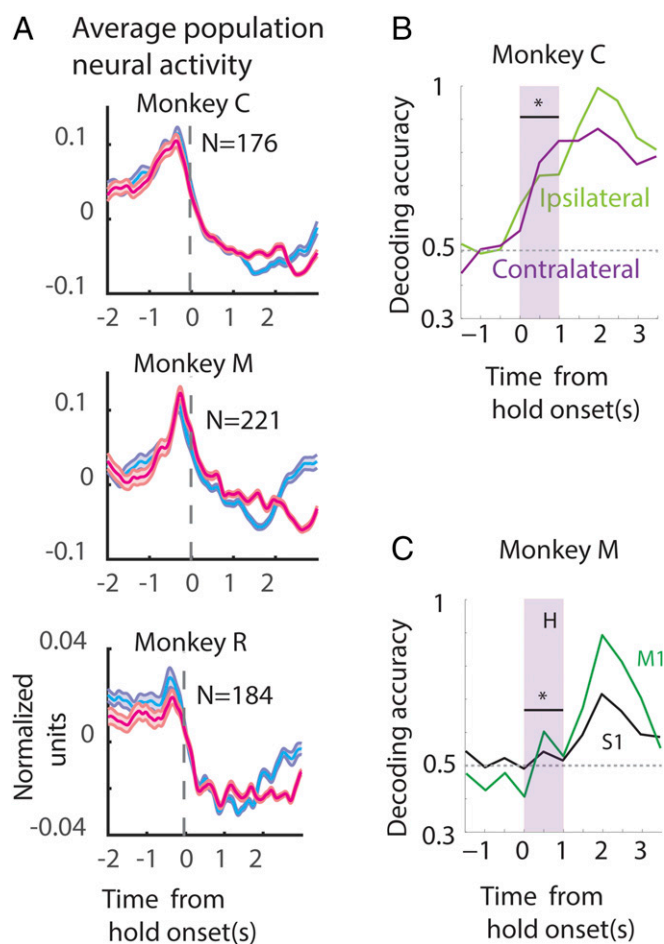
S1 neurons carried much less information to separate early- from late-reward trials because decoder performance was lower than that obtained with ensembles of M1 neurons (Fig. 3*C*, monkey M: M1, ROC: 0.6 ± 0.08; S1, ROC: 0.54 ± 0.04; ROC post-SVM).

As an additional test, we examined passive observation sessions where the animals did not produce any movements of their arms. Even in the passive observation task the SVM classifier could distinguish between early- and later-reward trials, based on the neuronal population activity recorded during the initial hold time period (0–1 s; Fig. S2*B*, monkey C: ROC: 0.58 ± 0.06; monkey M: 0.62 ± 0.07). These classification results suggest that reward time classification was not due to differences in arm movement patterns between the early- and late-reward conditions.

The association between the target color and neural representation of the expected time of reward (early vs. late) was reinforced with practice. If the difference in M1/S1 neural activity between early- and late-reward trials reflected this association, then predictions of reward timing based on neural ensemble activity should improve with practice. Consistent with this hypothesis, during the first session we observed that M1 neural ensemble activity, produced during the initial hold time period (0–1 s), did not clearly distinguish between early- vs. late-reward trials (Fig. S2*C*, orange trace; monkey C: ROC: 0.51 ± 0.1 and monkey M: 0.54 ± 0.07). However, after a few sessions of training (joystick control sessions: monkey C, nine sessions and monkey M, eight sessions), M1 neural ensemble activity could distinguish early- from late-reward trials (Fig. S2*C*, blue trace; monkey C: ROC: 0.77 ± 0.11 and monkey M: 0.62 ± 0.08). As a control, the target color information was withheld in the last session, by changing it to a neutral color (white). Immediately, the classifier performance dropped to chance levels during the initial hold time (0–1 s) (Fig. S2*C*, black trace for monkey C; ROC: 0.5 ± 0.1).

In summary, the neuronal population-based classifier analysis could distinguish the early- from late-reward trials before early reward. The improvement in classifier output with training suggests a learned association between the target color and the time of reward, which may be strengthened over time.

Modulation in Neuronal Activity Due to a Mismatch from Expected Reward Magnitude. We next asked whether M1/S1 neurons responded to unexpected changes in reward magnitude. To test this, in a small fraction (20%) of randomly interspersed trials reward was withheld (NR, no-reward trials). In these trials, the animals received a lower-than-expected amount of juice, referred to as a negative RPE (34, 36). In another subset of trials (20%), the animals received more than expected juice (0.9 mL, 750-ms pulse; DR, double-reward trials), referred to as a positive prediction error. In the remaining 60% of the trials the expected amount of reward was delivered (0.46 mL, 350-ms pulse; SR, single-reward trials).

When reward was withheld, negative RPE, a brief period of increased neuronal firing activity was observed in the PETHs of some neurons (RPE in Fig.4*A*, *Right*), whereas others, recorded during the same session, did not show this increase in spiking activity (Fig. 4*A*, *Left*). Fig. 4*B* shows the PETH of another example neuron (RHA_012a), along with the associated raster plot and licking frequency (see *SI Materials and Methods* for more information on licking frequency estimation). This M1 neuron displayed a significant increase in spiking activity when reward was omitted (cyan and magenta RPEs). More example neurons from monkeys M and R are displayed in Fig. 4*C*, *Top* and Fig. S3. These neurons exhibited a significant increase in spiking activity in response to reward omission (two-sided *t* test: *P* < 0.05). Such an increase in neuronal activity was recorded in many M1 and S1 neurons (monkey M, 113/221 neurons and monkey R, 38/189 neurons; two-sided *t* test: *P* < 0.05), shortly after the potential reward time [monkey M: 143 ± 5.6 ms (SD), *n* = 113 cells and monkey R: 165 ± 5.2 ms (SD), *n* = 38 cells]. As shown in Fig. 4*C*, *Bottom* (magenta trace), the average activity of the M1/S1 neuronal population also captured this transient neuronal excitation.

Fig. 3. Representation of reward by the neuronal population. (*A*) Population firing rate: The firing rate of each neuron was computed, z-scored, and averaged to plot the population PETH for each monkey for a joystick control session. The associated confidence bands are also shown (±1 SE). (*B* and *C*) Neuronal ensemble-based classifier analysis. Proportion of correct classification of trials, denoted by decoding accuracy, according to reward time (1 or 2 s) is shown on the *y* axis. The time aligned to hold onset is shown along the *x* axis. The classifier was trained using all of the contralateral M1 neurons (purple) or the ipsilateral M1 neurons (green). Note that the classification was significantly better (*P* < 0.05, bootstrapped ROC) than chance even before reward was delivered (at 0–1 s of hold period, shaded gray box) when using ipsilateral or contralateral M1 neurons (*B*) or S1 neurons (*C*).

When the reward magnitude was doubled-positive RPE a significant reduction in spiking activity was observed. Neurons in Fig. 4C, *Top* and Fig. S3 display this reduction in spiking activity (green trace; two-sided *t* test: $P < 0.05$) when the juice delivery epoch extended beyond the regular period of 350 ms. This effect, which was significant in many neurons (monkey M, 55/221 cells and monkey R, 30/189 cells), could also be captured by the average population activity of M1/S1 neuronal ensembles (Fig. 4C, *Bottom*, green trace).

In summary, a majority of M1 and S1 neurons responded to both positive and negative RPE by producing opposing changes in firing rate modulation: an increase to signal a reward withdrawal and a reduction for a reward increase.

M1 and S1 Neurons Multiplex Target Direction, Arm Velocity, and Reward Information. In a center-out arm movement task, M1/ S1 neurons are tuned to changes in the location of the target, direction of movement, and arm velocity, among other parameters. To assess the relative strength of the reward-related modulation in the same M1/S1 neuronal spiking activity, a multiple linear regression analysis was carried out by including reward magnitude, along with cursor velocity and target location, as regressors. This analysis was carried out independently in seven different, 400-ms-wide trial epochs [Fig. 5A: baseline (B), fixation (F), target (T), movement forward (MF), hold (H), reward (R), and movement back (MB)]. Single-reward and no-reward trials were included in one model; double- and single-reward trials were part of another model.

As expected, variability in firing activity in many M1/S1 neurons could be accounted for ($P < 0.05$) by changes in cursor velocity (monkey M: M1, 88%; $\beta = 0.18 \pm 0.006$. S1: 84%; $\beta = 0.17 \pm 0.008$. Monkey R: M1, 59%; $\beta = 0.08 \pm 0.0053$) and target location (Fig. 5A; monkey M: M1, 73%; $\beta = 0.067 \pm 0.016$. S1:68%; $\beta = 0.06 \pm 0.019$. Monkey R: M1, 71%; $\beta = -0.15 \pm 0.008$), especially during the movement phases (MF and MB in "cursor velocity" and "target" panels in Fig. 5A). However, reward magnitude significantly affected M1 neural activity in the reward epoch (R in "reward" panel of Fig. 5A): A sizeable population of M1 and S1 neurons robustly responded to a negative RPE (Fig. 5A; monkey M: M1: 82.6%; $\beta = -0.15 \pm 0.0069$. S1:61%; $\beta = -0.13 \pm 0.01$. Monkey R: M1, 63.4%. $\beta = -0.08 \pm 0.0058$).

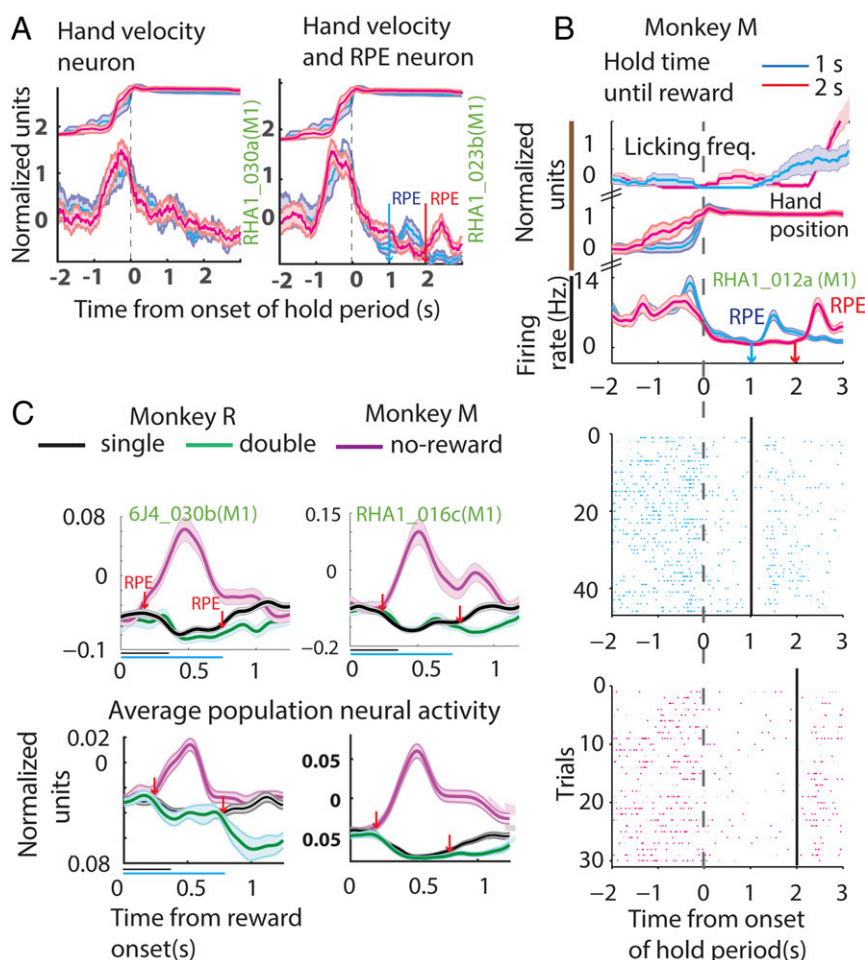**Fig. 4.** Neuronal modulation due to mismatch in the expected reward magnitude. (A) Plots showing hand movement trace and the neural activity of example neurons from the same session when a reward is omitted—a no-reward trial. The plot format is the same as that of Fig. 2A. Note the transient increase in spiking activity, indicative of RPE, after the potential reward onset time for the neuron in the right but not in the left panel. (B) Another example neuron, along with the absolute firing rate, licking frequency, and spike raster during a no-reward trial. The plot format is adapted from Fig. 2B. (C) (*Top*) Example neuron from monkey M (*Right*) and monkey R (*Left*). The normalized neural activity is plotted with respect to time of reward for single-reward trials (black), double-reward trials (green), and no-reward trials (magenta). (*Bottom*) The average population neural activity. Bands around the mean represent ±1 SE. The horizontal lines below the abscissa span the duration of the single reward (black, 350 ms) and the double reward (green, 700 ms). Note that when reward is withheld a transient increase in neural activity is observed, whereas when more than the expected reward is delivered a reduction in activity is witnessed. The arrows (downward-pointing, red) indicate when the neural activity was significantly different from the regular reward condition.

During reward delivery, when the flow of juice began to diminish, monkeys licked at the juice spout. A potential concern was whether the licking movements were represented in the neuronal modulations observed during the reward epoch. Although neural recordings were performed in the forearm areas of M1 (monkeys M and R) and S1 (monkey M), licking movements could be potentially represented by the M1/S1 neurons. In fact, a small percentage of neurons (<5%) showed spiking activity that corresponded to licking frequency. Licking frequency was determined using a custom methodology and added as a regressor in subsequent analysis. We observed that some neurons ("licking frequency" panel in Fig. 5A; monkey M: M1 = 13%, S1 = 7.5%; monkey R: 17%) were weakly tuned to changes in the frequency of licking (M1: $\beta = 0.0032 \pm 0.0032$; S1: $\beta = 0.01 \pm 0.002$, monkey R: 0.05 ± 0.007). More importantly, changes in reward magnitude continued to account for most of the variability in M1/S1 neuronal spiking activity, as they did

before introducing licking frequency as an additional regressor (monkey M: M1, 82.6%; $\beta = -0.15 \pm 0.0069$. S1:61%; $\beta = -0.13 \pm 0.01$. Monkey R: M1, 63.4%; $\beta = -0.08 \pm 0.0058$).

To isolate effects related to reward anticipation, the multiple regression analysis was performed by including reward anticipation as an additional regressor, along with other confounding variables, such as cursor velocity, target location, licking frequency, reward magnitude, and so on. Because reward anticipatory activity is expected to increase as the time of reward approaches (e.g., RHA3_009b in Fig. 2 A and B), the "time elapsed after target appearance" was included as the regressor that represents anticipation. By running this modified regression model we obtained the fraction of neurons significantly modulating ($P < 0.05$) their neural activity in accordance with the time elapsed after target onset. The fraction was represented as a function of time. We observed that, as the time of reward approached, a
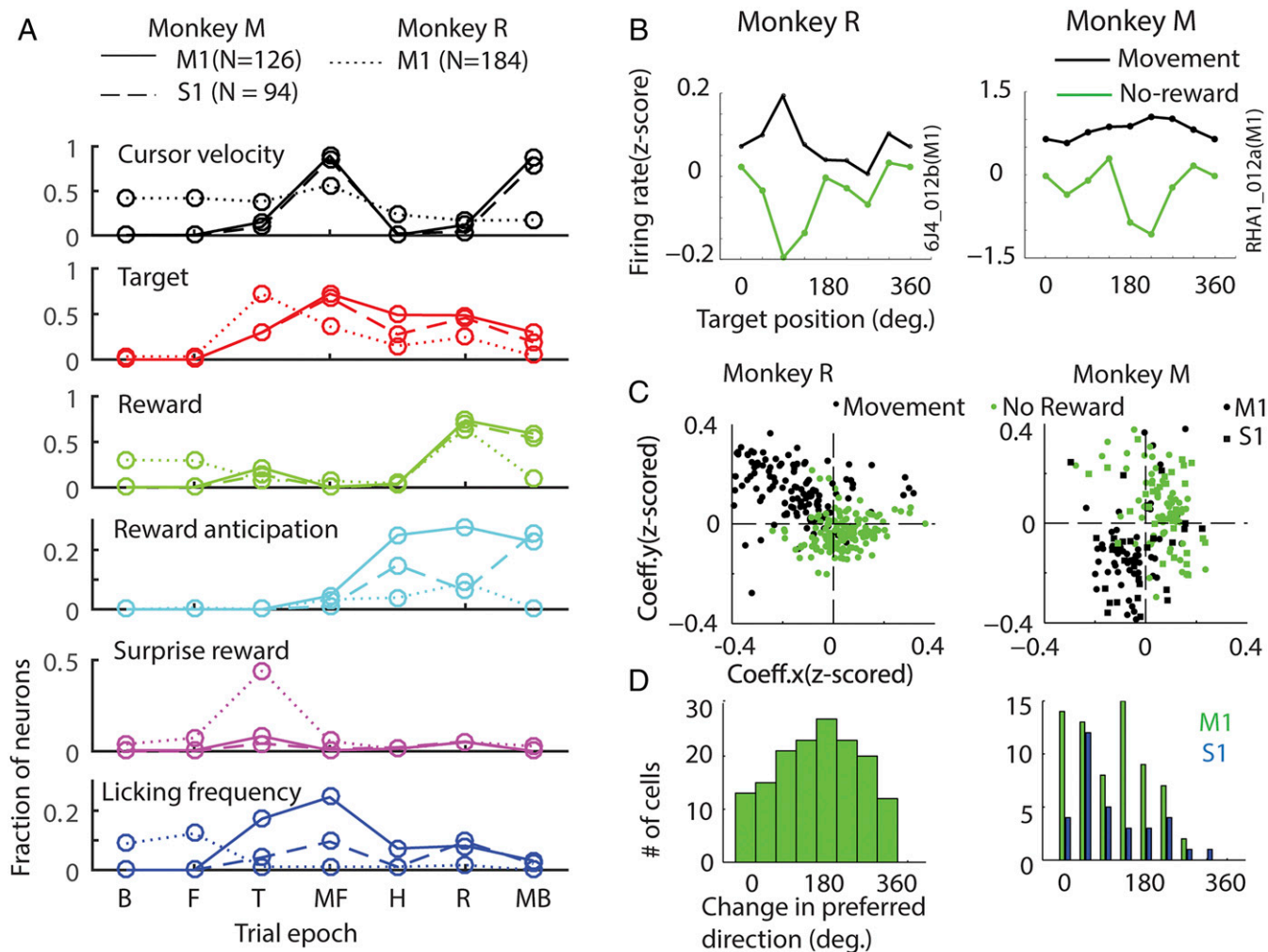

Fig. 5. Reward-related neural activity is multiplexed along with other task relevant parameters. (A) Epoch-wise multiple regression analysis. Each panel shows fraction of neurons whose variability in spiking activity could be explained by changes in cursor velocity (panel 1), target position (panel 2), reward (panel 3, single reward vs. no reward), reward anticipation (panel 4), surprise reward (panel 5), and licking frequency (panel 6), using a multiple regression model ($P < 0.05$), in different 400-ms-wide epochs during a trial: B, before fixation; F, fixation; T, target onset; MF, moving (forward) to target; H, hold period; R, reward delivery; and MB, moving back to center. (B) Spatial tuning of reward-related neural activity. Normalized firing rate of example cells is plotted for eight different arm movement directions. Average ± SEM (smaller than the markers) of the neural activity is shown for different epochs: during movement (300–700 ms after target onset, black) and after the expected time of reward (200–600 ms, green) for the no-reward trials. (C) The directional tuning of all of the neurons for the movement and reward epochs (in no-reward trials) was assessed using multiple regression. The directional tuning of all of the neurons is captured by the scatter plot. The model assessed neuronal firing rate with respect to different target positions along the x axis and y axis. By plotting the respective (x and y) coefficients (Coeff.x and Coeff.y, respectively) in 2D space, each neuron's preferred direction (angle subtended by the neuron with the abscissa) and tuning depth (radial distance of the neuron from the center) were determined. Tuning during the movement (black dots) and after the expected time of reward (200–600 ms, green) for the no-reward trials is shown. (D) The histograms show the number of neurons that underwent a change in spatial tuning from the preferred direction. Directional tuning of many neurons changed by 180 ± 45°.

greater fraction of cells reflected anticipation. For example, in monkey M the number of M1 anticipation-related neurons increased as the time of reward approached (Fig. 5A, "reward anticipation"): 0% around target appearance, 4% during movement, 25% during the hold time, and 27% around the time of reward. S1 neurons had a similar trend with the exception of a noticeable dip around the time of reward: 0% around target appearance, 1% during movement, 14% during the hold time, 9% around the time of reward, and 25% toward the end of hold time. In monkey R the effect of reward anticipation was weaker but nevertheless a significant population of M1 neurons showed the same trend: 0% around target appearance, 3% during movement, 4% during the hold time, and 9% around the time of reward. These results confirmed that M1/S1 spiking activity reflected the time elapsed from target onset. Notably, we did not see any significant changes in the representation of the other regressors (velocity, target location, etc.), which suggests that the time-dependent changes are specific to reward anticipation and not related to any movements.

To aid better visualization of the increase in reward anticipatory activity, trials were separated into early- and late-reward trials, and the regression analysis was performed in every time bin leading up to the reward. More specifically, we performed an independent regression on every 100-ms time bin, using the average of 100 ms of data. In this analysis we also included all of the previously used regressors: instantaneous velocity, target position, reward outcome, time after target appeared, surprise reward, and licking frequency. Trials for this analysis were aligned to target onset, instead of to each trial-relevant epoch. Because the reward time was staggered, ±2 SD of the distribution was determined for both early- and late-reward trials and indicated in the plots (black vertical lines, Fig. S4A). By running this continuous regression model, in monkey M, as expected from the previous analysis, gradually more and more M1/S1 neurons significantly modulated their firing rate as the time of reward approached (Fig. S4A). Importantly, the time course of this increase in the number of M1/S1 cells tuned to reward plateaued around the time reward was delivered. As a result, the time course was faster for the early-reward condition compared with the late-reward condition. A similar trend, albeit weaker, can be seen for monkey R also. Because monkey C did not participate in Exp. 2, and because licking movements were not monitored for monkey C, the regression model for this monkey only included target position, velocity, and reward anticipation. However, a similar trend could also be seen for reward anticipatory activity in monkey C despite the utilization of a reduced model. Interestingly, in monkey C some neurons were observed to be modulated right after target appearance, which may be indicative of the prospective reward. In all, these results reinforce the notion that M1/S1 neurons also modulate their firing rate as a function of reward anticipation.

Having examined the impact of reward omission (negative RPE) on spiking activity in M1 and S1, a similar multiple linear regression analysis was used to study the influence of an unexpected increase in reward (positive RPE). Several neurons modulated their firing rate significantly ($P < 0.05$) by positive RPE, although the effects were weaker than for negative RPE (monkey M: M1, 45.5%; $\beta = -0.05 \pm 0.0037$. S1:22%; $\beta = -0.05 \pm 0.005$. Monkey R: M1, 35.5%; $\beta = -0.037 \pm 0.0043$). We anticipated that a more salient positive RPE would be generated if the monkeys were rewarded when they least expected it. To do so, in a few (10%) randomly chosen trials animals received juice (surprise reward in Fig. 1C) while they waited for the trial to begin. Although some neurons (~10%) in monkey M modulated their firing rate to the unexpected reward, a much larger fraction of neurons responded in monkey R ("surprise reward" panel in Fig. 5A; M1, 35%; β: -0.04 ± 0.003). To assess the strength of the neural activity elicited by surprise reward we examined the regression coefficients generated by the regression model (continuous regression was performed on every 100 ms, as mentioned

previously). By doing so, we observed that in monkey M, despite the fact that only 10% of the neurons significantly responded to the surprise reward, these cells showed a clear increase in firing modulation strength (Fig. S4B). Both M1 ($\beta = 0.023 \pm 0.002$) and S1 ($\beta = 0.016 \pm 0.003$) neurons showed an increase in activity around the time of the surprise reward. In monkey R, a similar trend was observed, with the coefficient of strength increasing to reflect surprise reward ($\beta = 0.025 \pm 0.003$).

In summary, our results suggest that, in addition to target location and cursor movement velocity, RPE, reward anticipation, and, to a small extent, licking movements contributed to the spiking activity of the M1 and S1 neurons. M1/S1 neurons also modulated their firing rate as a result of both negative and positive RPEs. The effects of negative RPEs on M1/S1 neurons seemed to be stronger and more widespread than those of positive RPEs.

Reward-Related Modulations in the Spatial Tuning of Sensorimotor Neurons. It is well known that the firing rate of M1/S1 neurons varies in a consistent way according to arm movement direction (black trace in Fig. 5B), a property known as neuronal directional tuning. Thus, as expected, the normalized firing rate of the neuron 6J4_012b from monkey R shows clear directional tuning (Fig. 5B, Left, black trace, $P < 0.05$, bootstrap): a higher firing rate for the target located at 90°, therefore referred to as the "preferred" direction, and a lower firing rate for the target at 270°, the "nonpreferred" direction. An example neuron from monkey M (Fig. 5B, Right, RHA1_012b) is also directionally tuned ($P < 0.05$, bootstrap). To examine directional tuning in the neuronal population, a multiple regression model (Materials and Methods) was fitted to every neuron's spiking activity when the arm was moving to the target (MF, movement forward epoch in Fig. 5A). The model assessed neuronal firing rate with respect to different target positions along the x axis and y axis. By plotting the respective (x and y) coefficients in 2D space, each neuron's preferred direction (angle subtended by the neuron with the abscissa) and tuning depth (radial distance of the neuron from the center) were determined. The black dots in Fig. 5C and Fig. S5 represent the directional tuning of individual neurons. These graphs indicate that many neurons were directionally tuned (monkey M: 206/220; monkey R: 167/198 neurons $P < 0.05$, bootstrap; monkey M: quadrant 3; monkey R: quadrant 2).

We then asked whether the same neurons exhibited directional tuning around the time of reward. For that, a multiple regression model was fitted to every neuron's average spiking activity in the last 400 ms, before reward delivery began. The directional tuning of every neuron was plotted in Fig. S5A (blue dots). This analysis revealed that the directional tuning, calculated at the time of reward onset, was correlated with the tuning during the movement epoch in monkey M ($R = 0.7$; $P < 0.001$; $n = 220$). The tuning was weaker in monkey C ($R = 0.27$; $P < 0.001$; $n = 176$) and monkey R ($R = 0.18$; $P = 0.002$; $n = 184$). These results suggest that M1/S1neurons were weakly tuned to the target location at the time of reward onset. This result is consistent with the observations from Fig. 5A (target panel) that show neural activity modulated by target location when the target was visible (till the end of hold epoch).

We then measured directional tuning when the reward was omitted. As shown in Fig. 5B (green trace), both neurons illustrated showed strong directional tuning ($P < 0.05$, bootstrap). However, in both cases these neurons' spiking activity was suppressed in the preferred direction (90° in 6J4_012b and 180° in RHA1_012b). To further examine directional tuning in the same M1/S1 neuronal population, a multiple regression model was fitted to every average neuronal spiking activity during the first 400 ms of the potential reward epoch in these no-reward trials. Fig. 5C shows the preferred direction and tuning depth for each M1 and S1 neuron (green dots and squares) during this epoch. These graphs indicate that many neurons underwent a significant change in their directional tuning, from the movement to the reward epochs (monkey M: quadrant 3–1; monkey R: quadrant

2–4). To assess whether these changes in M1 neuronal preferred direction were significant, a nested regression model was designed and tested (*Materials and Methods*). This analysis revealed that many neurons in monkey R (M1: $n = 154/198$, $P < 0.05$, bootstrap nested regression) and in monkey M (M1: $n = 68/126$; S1: $n = 33/94$, $P < 0.05$, bootstrap nested regression) underwent a significant change in preferred direction during this transition, from movement to reward epochs. Indeed, a fraction of these neurons (monkey R, 48%; monkey M: M1, 46%; S1: 30%) exhibited a change in directional tuning on the order of $180 \pm 45°$ (Fig. 5D). These results indicated that, when reward was omitted, a large fraction of M1/S1 neurons exhibited suppression in firing activity in their preferred direction.

Next, we tested whether these changes in directional tuning could be influenced by the return arm movement produced by the animals. The return arm movement was offset by at least 1 s from the beginning of the reward delivery epoch in late-reward conditions and by 2 s in early-reward conditions. Further, during the postreward hold time, small rewards kept the monkeys motivated to hold the hand at the target. Moreover, the animals delayed the return movement (300–500 ms) following reward omission. The task design and the behavior would therefore minimize the influence of return arm movement on the neural activity at the time of reward delivery. However, we also ran a multiple regression analysis to compare neuronal directional tuning around the time of return movement with that around the time of the potential reward, in a no-reward trial (Fig. S5B). The directional tuning during the return movement (cyan dots in Fig. S5B) was not correlated with the tuning observed during the no-reward epoch (green dots; replicated from Fig. 5B) in monkey M (R = −0.05; $P = 0.28$; $n = 220$) but negatively correlated in monkey R (R = −0.35; $P < 0.001$; $n = 184$). The return movement had no influence on the neural tuning at the time of reward for one monkey, whereas for the second monkey a negative relationship between the neuronal activity during these two epochs was found, which by itself may not account for the firing rate modulations around the time of reward.

Discussion

In the present study we manipulated reward expectation using explicit instruction of reward timing by color cues. This experimental setting allowed us to study the representation of temporal intervals related to rewards by M1 and S1 neurons, as well as the representation of the reward, and the integration of this information with the traditional M1 and S1 neuronal firing modulations, such as encoding of movement direction and velocity. We observed that (*i*) more than 9–27% of M1/S1 neurons exhibited anticipatory neuronal activity related to reward timing, which was distinct enough to differentiate reward times that were 1 s apart, (*ii*) and modulated their neural activity as a function of errors in reward expectation, and (*iii*) M1/S1 neurons corepresented reward and motor parameters.

Reward-Related Activity in Sensorimotor Neurons. Modulation in neural activity was observed in many M1/S1 neurons around the time of reward, which we have interpreted as representing both expectation of reward and mismatch in reward expectation. These findings are consistent with reward anticipatory modulations previously seen in premotor cortical neurons (30, 32) and in motor cortical neurons (33) (see *SI Discussion* for more information). However, given that many of these M1/S1 neurons are involved in planning of arm movements, a potential concern is that the monkeys could be moving their arms, loosening their grip on the joystick, or performing other unmonitored hand movements, which could be misinterpreted as reward-related modulation. However, in our task design any arm movement produced during the hold time would preclude the animal from receiving the intended reward. To comply with this constraint, all three monkeys learned to hold their arms steady. This was confirmed by inspection of the joystick

position, during the 3-s hold time, which revealed that the arms of the monkeys remained stable on most trials during the hold time. Despite a steady position, one could still attribute the observed M1/S1 neural firing modulation to a change in arm or body posture, which led to no change in the arm end point. However, because the monkeys were operating a spring-loaded joystick, adjusting their arm/body posture while keeping the joystick stationary would be extremely demanding and logistically quite an unfeasible task for the animals. Notwithstanding these mitigating factors, we have conducted passive observation sessions, where monkeys did not operate the joystick at all but simply watched the task being performed on the screen. In these sessions, the monkey's hands were well-restrained to the sides of the chair, while EMGs were recorded to control for any inadvertent muscle contractions. A slight increase in EMG was recorded from the (long head of) triceps and the biceps as the virtual arm moved; however, no EMG modulations were observed around the time of reward. Under these conditions of complete absence of overt arm movements, clear reward-related M1/S1 neuronal firing modulations occurred in several neurons (e.g., RHA2_030b and RHA3_009b in Fig. 2). Furthermore, examination of neuronal firing rate modulations demonstrated that the spiking activity of some neurons represented arm position and velocity, but not reward-related activity, whereas other simultaneously recorded neurons displayed reward-related activity. If the reward-related activity was influenced by arm movements, this finding would be unlikely. Furthermore, a few M1/S1 neurons, recorded simultaneously during the same session, displayed only reward-related activity.

Finally, we identified licking-related neurons and the firing rate modulations in these neurons followed the changes in licking frequency. To determine the changes in neural activity that were correlated with licking movements, we have used a generalized linear model analysis (37–39). From this analysis we learned that spiking activity of only a small percentage of neurons was related to licking frequency (<15%), whereas a much higher number of our recorded M1/S1 neurons represented arm movements (>80%) and changes in reward (>50%). These findings are consistent with the fact that our multielectrode arrays were originally implanted in the so-called forearm representation of M1/S1. Based on all these controls, we propose that our results are totally consistent with the interpretation that the observed M1/S1 neuronal firing rate modulations described here were related to reward expectation, rather than emerging as the result of some unmonitored animal movement.

RPE Signals in M1/S1. Following "catch" trials when a reward was withheld, a transient increase in neural activity was observed in many M1 and S1 neurons. This observation is consistent with a recent study (31) that reported an increase in firing in M1 and dorsal premotor (pMd) neurons when reward was not delivered in one monkey. We have replicated those findings in our study and further extended them by demonstrating that M1 and S1 neurons reduced their firing rate for a higher-than-expected reward. Altogether, these results support the hypothesis that M1 and S1 encode bidirectional RPEs. Our findings are also consistent with those obtained in an imaging study (40) and a recent electrocorticography study (41) in which RPE-related activity was detected in sensorimotor brain regions.

Neuronal response to both positive and negative RPEs, indicated by distinct bidirectional changes in firing rates, was initially reported for midbrain DA neurons (34, 36, 42). Subsequently, bidirectional RPE signals have been observed in many other brain areas (43–47). The widespread nature of RPE signals may suggest a common general mechanism involved in learning from outcomes. Our results suggest that this mechanism may also exist in the monkey primary motor and sensory cortical areas as well.

Although the role of DA in reward processing is rather complex and still a subject of debate (48, 49), one line of research has

implicated the phasic DA response, typically seen in midbrain DA neurons, as the mediator of RPEs (36, 50–52). Typically, reinforcement-based learning models have used prediction errors to account for behavioral plasticity (53–55). Recent studies that looked at plasticity in the motor system, particularly regarding the role of reward in motor adaptation (13, 56) and facilitation of movement sequences (57), have suggested that M1 neurons may play a crucial role in this process. The fact that we observed here that M1 and S1 neurons exhibit bidirectional prediction error signals reinforces this hypothesis significantly.

Interestingly, a significant fraction of M1/S1 neurons, at the time of reward omission, exhibited directional tuning that was distinct from the one displayed by them during the arm movement epoch. Changes to the M1 neuron directional tuning, and hence, to the cell's preferred direction, has been observed previously, as a function of task context (58–60) or even during the time duration of a trial (61, 62). During a typical trial, no tuning could be detected before the target appeared; strong neuronal tuning emerged only after the target appeared and remained throughout the period the target was acquired, until the end of the hold period. However, when reward was withheld, many M1/S1 neurons changed their preferred directions by $180 \pm 45°$. We speculate that these changes are related to reward omission and may serve to facilitate reward-based learning. However, it still remains to be tested whether these reward-related modulations affect future M1/S1 neural activity.

Distributed Encoding in M1/S1 Neurons. Based on studies carried out in the last couple of decades, M1 neuronal activity has been correlated with several kinetic (20–22, 63, 64) and kinematic parameters (16, 17, 19). Even though many of these movement parameters are correlated with one another, in general one line of thinking in the field purports that cortical motor neurons are capable of multiplexing information related to different parameters (65–68), including reward (65, 69, 70). Simultaneous recording from several M1/S1 neurons in our study has shown that many of these neurons are associated with some aspect of movement—position, velocity, and so on—whereas a subset of the neurons also seem to be receiving reward-related signals. This latter group of M1/S1 neurons are perhaps the ones that displayed both movement and reward-related activity. These findings are consistent with those from recent studies that have examined neurons in M1 and pMd responding to arm position, acceleration, and reward anticipation (31).

In conclusion, we observed that M1/S1 cortical neurons simultaneously represent sensorimotor and reward-related modulations in spiking activity. According to this view, M1 and S1 ensembles can be considered as part of a distributed network that underlies reward-based learning.

Materials and Methods

Behavioral Task. The task was a modified version of the center-out task. A circular ring appeared in the center of the screen. When monkeys held the avatar arm within the center target, a colored (red/blue) circular target ring appeared in the periphery, in one of eight evenly spaced peripheral locations. Monkeys had to then move the arm to place it on the target and hold it there for 3 s. A juice reward was delivered after 1 s (early reward) or after 2 s (late reward). The target color cued the time of reward (blue, early and red, late). In Exp. 2, in 20% of trials reward was omitted (no reward). In another 20% the reward was doubled (double reward). Further, in a small subset of randomly chosen trials (10%) a surprise reward (gated by a 350-ms pulse) was delivered during the fixation period. More information on the task design can be found in *SI Materials and Methods*.

Neural activity was recorded using a 128-channel multichannel acquisition processor (Plexon Inc.) (71). All of the analyses were performed in MATLAB (The MathWorks, Inc.).

PETH of Neuronal Activity. The recorded action potential events were counted in bins of 10 ms, smoothed using a moving Gaussian kernel (width = 50 ms) for illustration purposes. Details regarding normalization procedures can be found in *SI Materials and Methods*.

Classifying Trials Based on Early vs. Late Reward. To investigate the strength of information about the time of reward provided by a single neuron, or provided by the entire population, a decoding analysis was performed based on a linear discriminant analysis algorithm or SVM algorithm, using the MATLAB statistics toolbox (see *SI Materials and Methods* for details).

Multiple Linear Regression. To model every neuron's changes in firing rate as a function of task contingencies (Fig. 5A) a multiple regression analysis was carried out with the following general linear models (GLM; see *SI Materials and Methods* for details).

Spatial Tuning. The normalized mean firing rate was determined during the movement epoch, the hold time epoch, reward epoch, and return movement epoch. The z-scored firing rate in each epoch was fit for each neuron with a GLM (see *SI Materials and Methods* for details.)

1. Wolpert DM, Diedrichsen J, Flanagan JR (2011) Principles of sensorimotor learning. *Nat Rev Neurosci* 12:739–751.
2. Guo L, et al. (2015) Dynamic rewiring of neural circuits in the motor cortex in mouse models of Parkinson's disease. *Nat Neurosci* 18:1299–1309.
3. Wise RA (1981) Intracranial self-stimulation: Mapping against the lateral boundaries of the dopaminergic cells of the substantia nigra. *Brain Res* 213:190–194.
4. Wise RA (2004) Dopamine, learning and motivation. *Nat Rev Neurosci* 5:483–494.
5. Fibiger HC, LePiane FG, Jakubovic A, Phillips AG (1987) The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *J Neurosci* 7:3888–3896.
6. Berger B, Gaspar P, Verney C (1991) Dopaminergic innervation of the cerebral cortex: Unexpected differences between rodents and primates. *Trends Neurosci* 14:21–27.
7. Williams SM, Goldman-rakic PS (1993) Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb Cortex* 3:199–222.
8. Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74:1–58.
9. Hosp JA, Nolan HE, Luft AR (2015) Topography and collateralization of dopaminergic projections to primary motor cortex in rats. *Exp Brain Res* 233:1365–1375.
10. Molina-Luna K, et al. (2009) Dopamine in motor cortex is necessary for skill learning and synaptic plasticity. *PLoS One* 4:e7082.
11. Lewis DA, Campbell MJ, Foote SL, Goldstein M, Morrison JH (1987) The distribution of tyrosine hydroxylase-immunoreactive fibers in primate neocortex is widespread but regionally specific. *J Neurosci* 7:279–290.
12. Gaspar P, Berger B, Febvret A, Vigny A, Henry JP (1989) Catecholamine innervation of the human cerebral cortex as revealed by comparative immunohistochemistry of tyrosine hydroxylase and dopamine-beta-hydroxylase. *J Comp Neurol* 279:249–271.
13. Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015) The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18:597–602.
14. Krakauer JW, Latash ML, Zatsiorsky VM (2009) Progress in motor control. *Learning* 629:597–618.
15. Laubach M, Wessberg J, Nicolelis MA (2000) Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* 405:567–571.
16. Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. *Science* 233:1416–9.
17. Mahan MY, Georgopoulos AP (2013) Motor directional tuning across brain areas: Directional resonance and the role of inhibition for directional accuracy. *Front Neural Circuits* 7:92.
18. Lebedev MA, et al. (2005) Cortical ensemble adaptation to represent velocity of an artificial actuator controlled by a brain-machine interface. *J Neurosci* 25:4681–4693.
19. Moran DW, Schwartz AB (1999) Motor cortical representation of speed and direction during reaching. *J Neurophysiol* 82:2676–2692.
20. Kakei S, Hoffman DS, Strick PL (1999) Muscle and movement representations in the primary motor cortex. *Science* 285:2136–2139.
21. Sergio LE, Hamel-Pâquet C, Kalaska JF (2005) Motor cortex neural correlates of output kinematics and kinetics during isometric-force and arm-reaching tasks. *J Neurophysiol* 94:2353–2378.

22. Addou T, Krouchev NI, Kalaska JF (2015) Motor cortex single-neuron and population contributions to compensation for multiple dynamic force fields. *J Neurophysiol* 113:487–508.
23. Lebedev MA, Doherty JEO, Nicolelis MAL (2008) Decoding of temporal intervals from cortical ensemble activity. *J Neurophysiol* 99:166–186.
24. Roux S, Coulmance M, Riehle A (2003) Context-related representation of timing processes in monkey motor cortex. *Eur J Neurosci* 18:1011–1016.
25. Costa RM, Cohen D, Nicolelis MAL (2004) Differential corticostriatal plasticity during fast and slow motor skill learning in mice. *Curr Biol* 14:1124–1134.
26. Arsenault JT, Rima S, Stemmann H, Vanduffel W (2014) Role of the primate ventral tegmental area in reinforcement and motivation. *Curr Biol* 24:1347–1353.
27. Hosp JA, Luft AR (2013) Dopaminergic meso-cortical projections to M1: Role in motor learning and motor cortex plasticity. *Front Neurol* 4:145.
28. Kapogiannis D, Campion P, Grafman J, Wassermann EM (2008) Reward-related activity in the human motor cortex. *Eur J Neurosci* 27:1836–1842.
29. Kunori N, Kajiwara R, Takashima I (2014) Voltage-sensitive dye imaging of primary motor cortex activity produced by ventral tegmental area stimulation. *J Neurosci* 34:8894–8903.
30. Roesch MR, Olson CR (2003) Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J Neurophysiol* 90:1766–1789.
31. Ramkumar P, Dekleva B, Cooler S, Miller L, Kording K (2016) Premotor and motor cortices encode reward. *PLoS One* 11:e0160851.
32. Roesch MR, Olson CR, Matthew R (2005) Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J Neurophysiol* 94:1469–1497.
33. 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.
34. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
35. Pleger B, et al. (2009) Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol* 7:e1000164.
36. Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
37. Stapleton JR, Lavine ML, Wolpert RL, Nicolelis MA, Simon SA (2006) Rapid taste responses in the gustatory cortex during licking. *J Neurosci* 26:4126–4138.
38. Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network* 12:199–213.
39. Pillow JW, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454:995–999.
40. Vickery TJ, Chun MM, Lee D (2011) Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron* 72:166–177.
41. Ramayya AG, Pedisich I, Kahana MJ (2015) Expectation modulates neural representations of valence throughout the human brain. *Neuroimage* 115:214–223.
42. Schultz W (2015) Neuronal reward and decision signals: From theories to data. *Physiol Rev* 95:853–951.
43. Ding L, Gold JI (2010) Caudate encodes multiple computations for perceptual decisions. *J Neurosci* 30:15747–15759.
44. Kim H, Sul JH, Huh N, Lee D, Jung MW (2009) Role of striatum in updating values of chosen actions. *J Neurosci* 29:14701–14712.
45. So N, Stuphorn V (2012) Supplementary eye field encodes reward prediction error. *J Neurosci* 32:2950–2963.
46. Asaad WF, Eskandar EN (2011) Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *J Neurosci* 31:17772–17787.
47. Tachibana Y, Hikosaka O (2012) The primate ventral pallidum encodes expected reward value and regulates motor action. *Neuron* 76:826–837.
48. Berridge KC (2007) The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berl)* 191:391–431.
49. Beeler JA (2012) Thorndike's Law 2.0: Dopamine and the regulation of thrift. *Front Neurosci* 6:116.
50. Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72:1024–1027.
51. Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309.
52. Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
53. Sutton RS, Barto G (1998) Reinforcement learning: An introduction. *IEEE Trans Neural Netw* 9:1054.
54. Recorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton- Century-Crofts, New York), pp 64–99.
55. Pearce JM, Hall G (1980) A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* 87:532–552.
56. Shmuelof L, Krakauer JW, Mazzoni P (2012) How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J Neurophysiol* 108:578–594.
57. Kawai R, et al. (2015) Motor cortex is required for learning but not for executing a motor skill. *Neuron* 86:800–812.
58. Scott SH, Kalaska JF (1997) Reaching movements with similar hand paths but different arm orientations. I. Activity of individual cells in motor cortex. *J Neurophysiol* 77:826–852.
59. Caminiti R, Johnson PB, Burnod Y, Galli C, Ferraina S (1990) Shift of preferred directions of premotor cortical cells with arm movements performed across the workspace. *Exp Brain Res* 83:228–232.
60. Wu W, Hatsopoulos N (2006) Evidence against a single coordinate system representation in the motor cortex. *Exp Brain Res* 175:197–210.
61. Sergio LE, Kalaska JF (2003) Systematic changes in motor cortex cell activity with arm posture during directional isometric force generation. *J Neurophysiol* 89:212–228.
62. Sergio LE, Kalaska JF (1998) Changes in the temporal pattern of primary motor cortex activity in a directional isometric force versus limb movement task. *J Neurophysiol* 80:1577–1583.
63. Evarts EV (1968) Relation of pyramidal tract activity to force exerted during voluntary movement. *J Neurophysiol* 31:14–27.
64. Cheney PD, Fetz EE (1980) Functional classes of primate corticomotoneuronal cells and their relation to active force. *J Neurophysiol* 44:773–791.
65. Hayden BY, Platt ML (2010) Neurons in anterior cingulate cortex multiplex information about reward and action. *J Neurosci* 30:3339–3346.
66. Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–458.
67. Johnson MTV, Ebner TJ (2000) Processing of multiple kinematic signals in the cerebellum and motor cortices. *Brain Res Brain Res Rev* 33:155–168.
68. Carmena JM, et al. (2003) Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol* 1:E42.
69. Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
70. Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
71. Nicolelis MA, et al. (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. *Proc Natl Acad Sci USA* 100:11041–11046.
72. Roesch MR, Olson CR (2005) Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J Neurophysiol* 94:1469–1497.
73. Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
74. Lebedev MA, O'Doherty JE, Nicolelis MA (2008) Decoding of temporal intervals from cortical ensemble activity. *J Neurophysiol* 99:166–186.
75. Romo R, Schultz W (1987) Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. *Exp Brain Res* 67:656–662.
76. Weinrich M, Wise SP (1982) The premotor cortex of the monkey. *J Neurosci* 2:1329–1345.
77. Mauritz KH, Wise SP (1986) Premotor cortex of the rhesus monkey: Neuronal activity in anticipation of predictable environmental events. *Exp Brain Res* 61:229–244.
78. Crammond DJ, Kalaska JF (1994) Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. *J Neurophysiol* 71:1281–1284.
79. Lebedev MA, Wise SP (2001) Tuning for the orientation of spatial attention in dorsal premotor cortex. *Eur J Neurosci* 13:1002–1008.
80. Lebedev MA, Messinger A, Kralik JD, Wise SP (2004) Representation of attended versus remembered locations in prefrontal cortex. *PLoS Biol* 2:e365.
81. Fuster JM (2001) The prefrontal cortex–an update: Time is of the essence. *Neuron* 30:319–333.
82. Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14:477–485.
83. Kim J, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2:176–185.
84. Schall JD (2001) Neural basis of deciding, choosing and acting. *Nat Rev Neurosci* 2:33–42.
85. Namboodiri VMK, Huertas MA, Monk KJ, Shouval HZ, Hussain Shuler MG (2015) Visually cued action timing in the primary visual cortex. *Neuron* 86:319–330.
86. Merchant H, Harrington DL, Meck WH (2013) Neural basis of the perception and estimation of time. *Annu Rev Neurosci* 36:313–336.
87. Hikosaka K, Watanabe M (2004) Long- and short-range reward expectancy in the primate orbitofrontal cortex. *Eur J Neurosci* 19:1046–1054.
88. Tremblay L, Schultz W (2000) Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J Neurophysiol* 83:1877–1885.
89. Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J Neurophysiol* 61:814–832.
90. Hassani OK, Cromwell HC, Schultz W (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J Neurophysiol* 85:2477–2489.
91. Nicolelis MAL, Lebedev MA (2009) Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. *Nat Rev Neurosci* 10:530–540.
92. Schwarz DA, et al. (2014) Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nat Methods* 11:670–676.
93. Ramakrishnan A, et al. (2015) Computing arm movements with a monkey brainet. *Sci Rep* 5:10767.
94. Ifft PJ, Shokur S, Li Z, Lebedev MA, Nicolelis MA (2013) A brain-machine interface enables bimanual arm movements in monkeys. *Sci Transl Med* 5:210ra154.
95. O'Doherty JE, et al. (2011) Active tactile exploration using a brain-machine-brain interface. *Nature* 479:228–231.
96. Weijnen JAWM (1998) Licking behavior in the rat: Measurement and situational control of licking frequency. *Neurosci Biobehav Rev* 22:751–760.