J Neurophysiol 112: 1775–1789, 2014.
First published July 9, 2014; doi:10.1152/jn.00533.2012.

Modulation of neural activity by reward in medial intraparietal cortex is sensitive to temporal sequence of reward

Rishi Rajalingham,[1,3] Richard Greg Stacey,[2] Georgios Tsoulfas,[2] and Sam Musallam[1,2]
[1]*Department of Electrical and Computer Engineering, McGill University, Montreal, Quebec, Canada;* [2]*Department of Physiology, McGill University, Montreal, Quebec, Canada; and* [3]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts*

Submitted 20 June 2012; accepted in final form 29 June 2014

Rajalingham R, Stacey RG, Tsoulfas G, Musallam S. Modulation of neural activity by reward in medial intraparietal cortex is sensitive to temporal sequence of reward. *J Neurophysiol* 112: 1775–1789, 2014. First published July 9, 2014; doi:10.1152/jn.00533.2012.—To restore movements to paralyzed patients, neural prosthetic systems must accurately decode patients' intentions from neural signals. Despite significant advancements, current systems are unable to restore complex movements. Decoding reward-related signals from the medial intraparietal area (MIP) could enhance prosthetic performance. However, the dynamics of reward sensitivity in MIP is not known. Furthermore, reward-related modulation in premotor areas has been attributed to behavioral confounds. Here we investigated the stability of reward encoding in MIP by assessing the effect of reward history on reward sensitivity. We recorded from neurons in MIP while monkeys performed a delayed-reach task under two reward schedules. In the variable schedule, an equal number of small- and large-rewards trials were randomly interleaved. In the constant schedule, one reward size was delivered for a block of trials. The memory period firing rate of most neurons in response to identical rewards varied according to schedule. Using systems identification tools, we attributed the schedule sensitivity to the dependence of neural activity on the history of reward. We did not find schedule-dependent behavioral changes, suggesting that reward modulates neural activity in MIP. Neural discrimination between rewards was less in the variable than in the constant schedule, degrading our ability to decode reach target and reward simultaneously. The effect of schedule was mitigated by adding Haar wavelet coefficients to the decoding model. This raises the possibility of multiple encoding schemes at different timescales and reinforces the potential utility of reward information for prosthetic performance.

decode; medial intraparietal area; neural prosthetics; reach; reward

NEURONS IN THE PARIETAL REACH region (PRR)—an area that overlaps the medial intraparietal area (MIP) and V6A in monkeys (Batista et al. 1999; Snyder et al. 2000)—are activated by reach plans to targets in space (Snyder et al. 1997). The firing rate of these neurons also varies with the reward associated with the planned reach (Musallam et al. 2004). Successful decoding of reach goals and reward from PRR activity established this region as a viable source of control signals for neural prosthetic systems (Andersen et al. 2010). Similarly, many studies have demonstrated the utility of invasive and noninvasive signals from other brain areas to drive prosthetic devices (Birbaumer 2006; Hatsopoulos and Donoghue 2009). Despite this progress, the dearth of information extracted from brain signals has hindered the development of prosthetic systems that can emulate natural movements (Gilja et al. 2011; Tehovnik et al. 2013). The use of reward to bias or reinforce an action or to assess a user's cognitive state, motivation, and preferences may improve the performance of prosthetic systems (Andersen et al. 2004; Zander and Jatzev 2009). Birbaumer recently emphasized the significance of reward for prosthetic systems by noting that severely paralyzed patients cannot learn to control a prosthetic device after the disengagement of the reward system (Birbaumer 2006). Reward has been used to improve the classification of motor signals recorded from rat motor cortex (Mahmoudi et al. 2011). Similarly, decoding preferred rewards from neurons in PRR increased our ability to decode reach goals (Musallam et al. 2004). Very few studies, however, have investigated the potential utility of reward for prosthetic performance. We set out to elucidate the stability of reward sensitivity in MIP neurons. The ability of reward to directly modulate the activity of reach neurons has been disputed and is also addressed (Maunsell 2004).

Reward can influence the activity of neurons in cortical and subcortical structures throughout the human and monkey brain (Barraclough et al. 2004; Ikeda and Hikosaka 2003; Minamimoto et al. 2005; Musallam et al. 2004; O'Doherty 2004; Padoa-Schioppa 2011; Platt and Glimcher 1999; Roesch and Olson 2004; Schultz 2000, 2006; Sugrue et al. 2004; Wallis and Miller 2003). Neural activity in the lateral intraparietal area (LIP)—an area in the posterior parietal cortex (PPC) implicated in planning and executing saccades and shifts of attention—is correlated with decision-making variables including the choice probability based on a leaky integration of acquired reward history (Bisley and Goldberg 2003; Boucher et al. 2007; Dorris and Glimcher 2004; Gold and Shadlen 2001, 2007; Gottlieb et al. 1998; Kiani et al. 2008; Kiani and Shadlen 2009; Peck et al. 2009; Platt and Glimcher 1999; Schall and Thompson 1999; Shadlen and Newsome 2001; Sugrue et al. 2004). Neural activity correlated with reward history can potentially track the acquired reward per unit time (Shadmehr et al. 2010) but may lead to unstable reward representation (Estepp et al. 2011; Kacelnik 1997; Krusienski et al. 2011; Shadmehr et al. 2010).

The ability of reward to modulate neural activity in motor planning areas is controversial (Leathers and Olson 2012; Padoa-Schioppa and Assad 2006; Roesch and Olson 2004). Reward-related modulation of oculomotor activity in LIP and prefrontal cortex has been attributed to changes in movement speed or movement reaction time or to changes in attention or

Address for reprint requests and other correspondence: S. Musallam, Dept. of Electrical and Computer Engineering, McGill Univ., 3480 University St., Rm. 645, Montreal, QC, H3A 2A7, Canada (e-mail: sam.musallam@mcgill.ca).

motivation (Hare et al. 2008, 2011; Leathers and Olson 2012; Maunsell 2004; Padoa-Schioppa and Assad 2006). Much less is known about reward encoding in the reach system. It is unclear whether reward sensitivity in MIP reflects a complex, time-dependent signal. Similarities with LIP are expected, but differences may also arise as the reach system must consider the added metabolic cost and added risk of limb movements (Izawa et al. 2008; Liu and Todorov 2007; O'Sullivan et al. 2009; Scott 2004).

We investigated the sensitivity of MIP neurons to identical rewards presented in different sequences. We hypothesize that the sensitivity of neurons to identical rewards will vary according to context, presented here as reward history. This is consistent with many studies that have shown that animals prefer a low-variance reward source when reward magnitude is varied (Buchkremer and Reinhold 2010; Kacelnik and Bateson 1997). Reach trials ending with the delivery of small and large rewards appeared in separate blocks of trials or in randomly interleaved trials. We found MIP neurons with variable reward sensitivities dependent on the temporal sequence of reward. The presence of the reward schedules degraded our ability to decode reward reliably. However, decoding using models based on Haar wavelet coefficients improved decode performance. We also confirmed that the reward response in MIP is not due to changing behavioral metrics.

MATERIALS AND METHODS

Subjects and Preparatory Surgery

Data were collected from three awake, behaving male monkeys (*Macaca mulatta*, *monkeys F*, *H*, and *M*), weighing 5.6, 6.5, and 11.9 kg, respectively. All surgical and experimental procedures complied with Canadian Council of Animal Care guidelines and were approved by the McGill University Animal Care Committee.

All surgical procedures were performed under sterile conditions. Monkeys were first implanted with a MRI-compatible head post (Rogue Research) and trained on the experimental paradigm until they were thoroughly familiar with the task. We then performed a second surgery to implant a 2-cm circular recording chamber (Crist Instrument, IAC series) on the contralateral side of the reaching right arm of *monkeys H* and *F*. Chamber placement was guided stereotaxically and centered at posterior 7 mm, lateral 5 mm in *monkey F*. For *monkey H*, we placed the chamber in a position that maximized access to the medial bank of the intraparietal sulcus as confirmed by Brainsight, a neuro-navigation system (Rogue Research). We secured the chamber in place with MRI-compatible ceramic screws (Rogue Research) and Simplex P bone cement (Stryker, Hamilton, ON, Canada). For *monkey M*, we implanted two 16-electrode arrays (MicroProbe, Gaithersburg, MD) in the left MIP, guided by Brainsight (Rogue Research). The length of the electrodes ranged from 3 to 6 mm. Monkeys were given glycopyrrolate intramuscularly at a dose of 0.005 mg/kg to eliminate excessive saliva. General anesthesia was induced with ketamine at 10 mg/kg and maintained with isoflurane at 1–4% delivered through an intubation tube. After the surgical procedures, monkeys received medicine for analgesia for 10 days. The monkeys were given a minimum of 14 days to recover. Monkeys were weighed and their health and growth monitored daily. Animals were also pair-housed and had constant access to exercise via a jungle gym on all days, including experimental days. Each monkey had its own recording schedule that ensured that experiments started at the same time every session. Monkeys were given fresh fruits in the lab after all completed sessions.

Behavioral Task

The monkeys performed the experiment in a dimly lit Faraday chamber. Each animal was seated in a chair within arm's reach of a touch screen (ELO Touch) that was coupled to a monitor. The only light in the chamber emanated from visual cues displayed on the monitor and the monitor's low ambient output. Monkeys reached the touch screen through an opening in the chair on the side of their reaching hand (right arm in all animals). The nonreaching hand was obstructed. Behavior was controlled by a real-time system (LabVIEW RT, National Instruments). Eye position was monitored with an infrared reflection camera (ISCAN, Boston, MA), and hand position was monitored with the acoustic touch screen. Monkeys received juice rewards for correctly performing the task.

We trained the monkeys on the behavioral task, illustrated in Fig. 1*A*, for several months before collecting data. To initiate a trial, monkeys touched a central green circular target and visually fixated a red target located in the center of a touch screen for 500 ms (fixation period). A cue (circular green target) was then flashed for 300 ms in one of four possible locations in the periphery. The cue indicated the target of the upcoming reach, while the size of the cue indicated the reward magnitude for the current trial. A randomized delay period (1.2–1.7 s) followed (memory period), during which monkeys had to maintain the visual fixation and the hand contact established in the fixation period. The central green target was then extinguished, instructing the monkeys to reach to the location previously indicated by the extinguished cue flash ("go"). Monkeys had to touch the screen inside a 2.5-cm circle centered on the target for 400 ms without breaking visual fixation. If successful, monkeys were rewarded with juice. The association between cue size and reward magnitude was randomized across different days. Monkeys were thoroughly trained on only two cue-reward associations (amounting to reversal of cue-reward association).

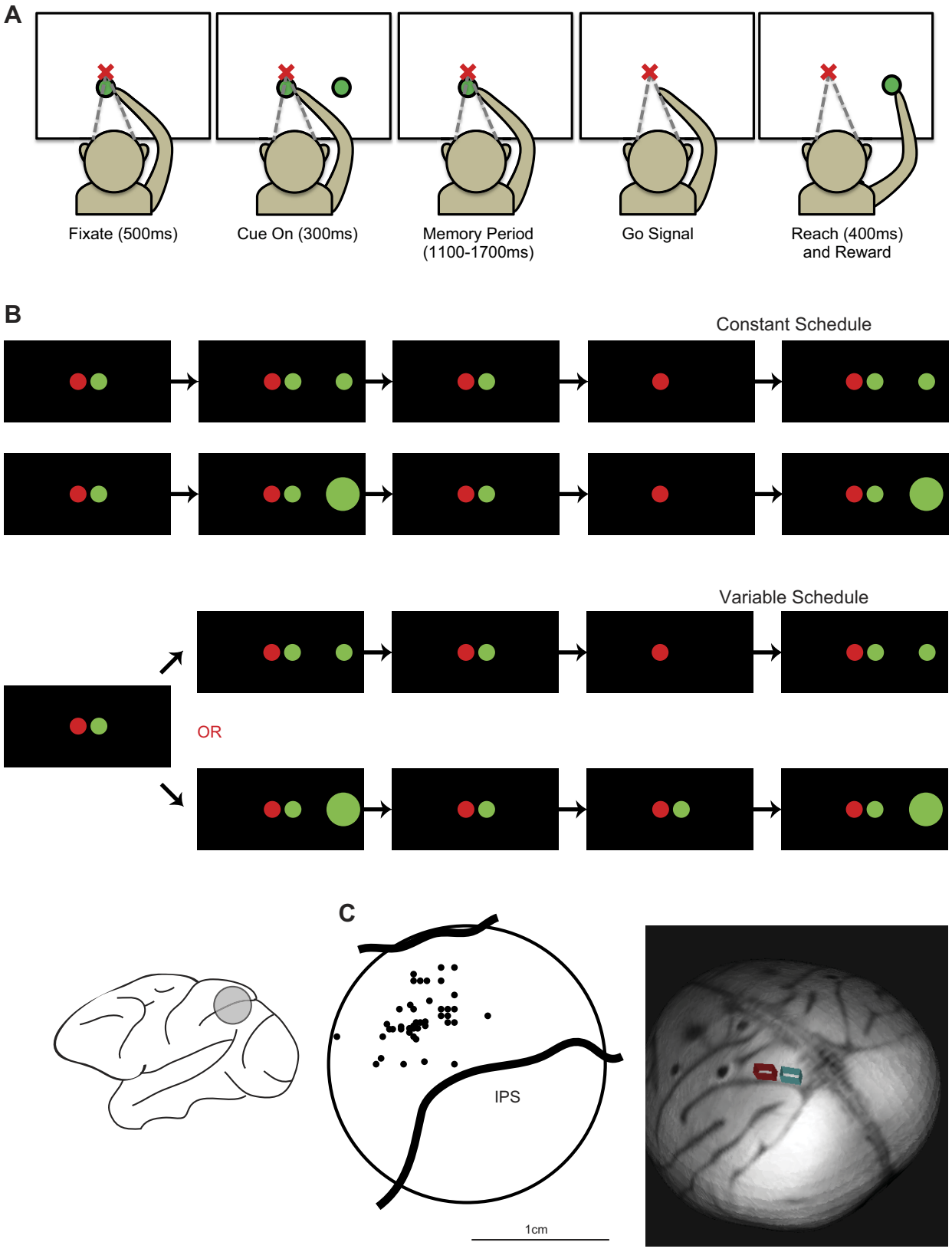
Monkeys did not always reach their daily minimum fluid intake during the experiment, as we stopped the session after a full data set was recorded even if satiety was not reached (equal to the minimum daily total fluid intake established with procedures outlined by the McGill University Animal Care Committee). This strategy reduced the effect of prolonged experiments that could change the subjective value of reward within a schedule. Monkeys were supplemented with water to meet their individualized daily minimal total volume or more after all recording sessions. All monkeys received twice their minimal total volume at least 1 day per week.

Fig. 1. *A*: sequence of memory reach task. *1*) Fixate: to initiate the trial, the animal must fixate the red target while touching the green target. *2*) Cue: a peripheral cue is flashed while the animal fixates; magnitude of the expected reward is reflected in the size of the cue. *3*) Memory: after the extinction of the peripheral cue, the animal must fixate the red target while touching the green target. *4*) Go: extinction of the green target instructs the animal to initiate a reach toward the remembered cue location. *5*) Reward: upon successful completion of a reach, the cued reward value is delivered. *B*: delayed-reach task under 2 conditions: In the constant-reward schedule, either a small (*top*) or a large (*middle*) reward was indicated and delivered for a set of trials. In the variable reward schedule (*bottom*), small and large rewards were randomly interleaved trial by trial for a set of trials. Each recording session consisted of a block of trials with a variable-reward schedule as well as blocks with both small and large constant-reward schedules. The expected value of reward received in either constant- or variable-reward schedules was identical. *C*, *left*: approximate position of recording chamber on rhesus brain. *Center*: position of recorded neurons in *XY* plane relative to chamber (circle) for *monkeys F* and *H*. The intraparietal sulcus (IPS) is labeled for reference. *Right*: position of chronically implanted microelectrode arrays in *monkey M*.

Reward Schedules

In this study, we investigated the sensitivity of reach neurons in MIP to identical rewards presented in different schedules. The volume of small rewards was half the volume of large rewards. The monkeys performed a delayed-reach task under two reward schedules.

Constant-reward schedule. Only one reward magnitude (small or large) was indicated and delivered for a set of trials (see Fig. 1*B*, *top*).



Variable-reward schedule. An equal number of small and large reward magnitudes were randomly interleaved trial by trial for a set of trials (see Fig. 1*B*, *bottom*).

Each recording session consisted of a block of trials with a variable-reward schedule and blocks with both small and large constant-reward schedules. The mean volume of reward received in the constant- and variable-reward schedules was identical. For example, a variable block of 40 trials consisting of 20 small- and 20 large-reward trials in random order was followed by a block of 20 small-reward trials and 20 large-reward trials. The number of trials was the same across reach directions. Blocks with trials interrupted by pauses or that affected time to completion were not analyzed. To avoid capturing nonstationary effects, the schedules were randomly ordered across recording sessions; some days started with a constant small-reward block and others started with a constant large- or variable-reward block. The mean (±SD) number of trials was 439 (±165), consisting of 2.47 (±0.75) schedule cycles for each session. The association between cue size and reward size was randomly interchanged across recording sessions: 34% of all neurons were recorded when a small target cue indicated a large reward.

Electrophysiology

Extracellular recordings in *monkeys F* and *H* were performed with a single channel from a multichannel micromanipulator system (NAN Drive, NAN Instruments). Before each recording session, a 23-gauge stainless steel guide tube containing a 120-mm tungsten microelectrode (FHC, Pl/Ir impedance: 1 MΩ) was placed over the appropriate cortical location (in the *XY* plane) and lowered onto the dura mater surface. Thus the guide tube served as the electrical ground. The electrode was then lowered into the cortex to a desired depth ranging from 2 to 8 mm and left to rest from 15 to 60 min before any subsequent movements. Single-unit data were collected during a 4-mo period. We did not map the spatial response fields of neurons or selectively record from neurons that were modulated by reward or direction but recorded from all neurons stabilized for recording. In total, 89 cells—40 from *monkey F* and 49 from *monkey H*—were recorded over 45 separate recording sessions.

We recorded an additional 46 neurons over one session, using a multielectrode array implanted in *monkey M*. Chronically implanted microelectrode array recordings allowed us to record from neurons that are stable over an experimental session. This recording provided a control for possible nonstationary effects associated with isolation of neurons. Additionally, by sampling multiple neurons simultaneously, this data set was used for offline decoding of behavioral variables from the neural population. Figure 1*C*, *center*, depicts the location of the electrode tip for all cells relative to the implanted chamber for *monkeys F* and *H*. Figure 1*C*, *right*, shows the location of the chronically implanted arrays for *monkey M*.

Single-channel/multichannel recordings were filtered (250–8,000 Hz for spikes) and amplified with commercially available systems (Plexon). Each channel was digitized (at sampling rates of 40 kHz for spikes) and continuously recorded to disk for further analysis. Spike waveforms were sorted online with commercial software (Plexon) and offline with custom software before being visually inspected.

Data Analysis

Reward modulation. Only successful trials were analyzed. We focused our analysis on the neural activity during the memory period because this epoch is separated from sensory and motor signals. We define the memory period firing rate as the number of spikes per second in an 800-ms window beginning 200 ms after the cue offset. We used a three-way ANOVA to test the effect of direction, reward magnitude, reward schedule, and their interactions on neural activity. To quantify the strength of tuning to these factors, we used an information theoretic metric, mutual information (MI) (Grunewald et al. 1999), to measure the amount of reward information contained in the neural data. The mutual information between two random variables X and Y is defined as $I(X;Y) = H(X) - H(X|Y)$, in which $H(.)$ is entropy. Thus, intuitively, $I(X;Y)$ measures the decrease in uncertainty about X that results from knowing Y. In addition, MI is a nonparametric measure of codependence, in contrast to regression analyses that assume linear or other parametric priors. Because we express MI as the number of bits required to encode $I(X;Y)$, we normalize mutual information by $H(X,Y)$, the joint entropy of X and Y, to get a scale-invariant metric of information. Furthermore, conditional mutual information (CMI) extends this notion to MI of two random variables X and Y conditioned on a third variable, Z: $I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$. This metric is the decrease in uncertainty about X that results from knowing Y, given that Z is already known. This metric is a measure of the amount of new information about X present in Y, independent of Z. These information theoretic measures are well suited for extracting stimulus-response relationships (Doya 2007).

To infer the significance of these metrics for each neuron, we used an exact test, or permutation test (Fisher 1955). We computed a null distribution of the metric from randomly shuffled data (1,000 times) and estimated the empirical probability of a result greater than or equal to the observed metric. For example, if 20 of 1,000 MI estimates in the null distribution were greater than the MI estimate computed on the original data, the P value was equal to 0.02.

A paired t-test on the MI distribution over all neurons tested for changes in MI at the population level. To infer significance of changes in MI between schedules at the single-neuron level, we compared bootstrapped distributions of the MI estimates, using a two-sample t-test. We obtained bootstrap distributions by recalculating the MI of resampled data (with replacement) 1,000 times.

Modeling and identification. Reward modulation of neural activity has previously been modeled with a priori parametric models and nonparametric regression analyses (Bayer and Glimcher 2005; Lau and Glimcher 2005; Lee and Seo 2007; Seo and Lee 2007; Sugrue et al. 2004). We combined these approaches by selecting the optimal function of reward with a nonparametric regression approach from systems identification and validated the function by comparing it to a parametric leaky integration of reward model.

To model firing rate as a function of reward magnitude alone, the effects of directional tuning were first removed via normalization of firing rates by direction, by computing the z score of firing rates per direction. This normalization step is justified, as the neural responses to reward and direction were separable (see RESULTS). For both parametric and nonparametric modeling analyses, the input $u(t)$ was a time series of rewards and the output $y(t)$ was a time series of neuronal firing rates. Specifically, $u(t)$ was assigned a value of 1 and −1 for preferred and null rewards (preference was determined from neural activity in the constant schedule). The z score of the firing rate signal was used for $y(t)$ to account for a constant bias (by forcing the output to have zero mean).

We also fit the data with parametric regression analysis previously used as a reward-based model of neural firing rate in area LIP (Sugrue et al. 2004). The output of this model (representing firing rate) is a convolution of the input (reward) with a decaying exponential kernel parameterized by a single value, a time constant τ. As an existing model of the dependence of firing rate on reward, the purpose of this model was to provide a benchmark for comparison and validation of the nonparametric fits. Figure 2*A* depicts a conceptual diagram of this model for two values of τ; convolving the time series of reward magnitudes (Fig. 2*A*, *left*) with the exponential kernels shown in Fig. 2*A*, *center*, yields the curves depicted in Fig. 2*A*, *right*. A larger τ results in more a slowly decaying kernel yielding an output constructed from a long memory of past rewards (see Fig. 2*A*, *right*, solid black trace). In contrast, a small τ results in a quickly decaying kernel and a fast-varying model output (see Fig. 2*A*, *right*, dashed gray trace). We refer to this model as EXP to describe its exponential kernel.

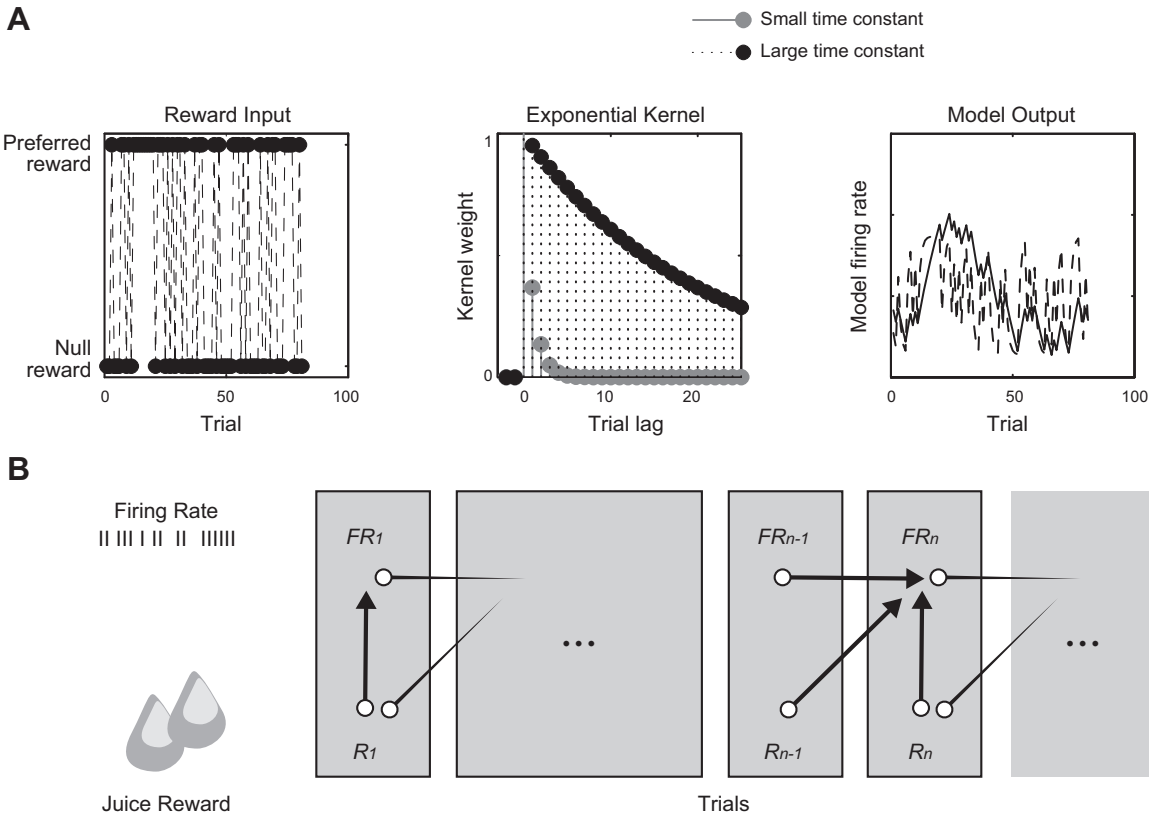


Fig. 2. *A*: conceptual diagram of the EXP model. *Left*: a time series of reward magnitudes, the input to the model. *Center*: 2 example exponential kernels, differing in their time constant τ. By convolving the reward input with each kernel, we get the corresponding model outputs shown at *right*. Note that a larger τ results in more slowly decaying kernel and, consequently, a long memory of past rewards (solid black trace). In contrast, a small τ results in a quickly decaying kernel and a fast-varying model output (dashed gray trace). We refer to this model as EXP to describe its exponential kernel. For each neuron, we simulated 1,000 choices of τ [log distributed, between 0 and log(1,000)] and selected the optimal τ based on the fit of the resulting EXP model to the experimental data. *B*: conceptual schematic showing dependence relationship between firing rate and reward. *x*-Axis corresponds to trial index; the blocks illustrate each trial's juice reward and measured firing rate. Arrows indicate that current trial firing rate FR_n depends on current trial reward R_n, as well as past trial firing rate and reward FR_{n-1}, R_{n-1}. This dependence can extend to more past trials, but arrows are not illustrated for simplicity. The weights of past trial rewards and firing rates are estimated with an autoregressive approach with exogenous inputs (ARX), and the extent into past trials is determined with the Akaike information criterion (AIC). The introduction of feedback of past firing rates does not change the relationship between the firing rate and reward. Instead, it creates an infinite impulse response (IIR) filter that estimates the weights of infinitely many past inputs fitting a small number of weights. Thus the implication of the ARX model is that the firing rate of the current trial is a weighted sum of past trial rewards.

For each neuron, we simulated 1,000 choices of τ [log distributed, between 0 and log(1,000)] and selected the optimal τ based on the fit of the resulting EXP model to the experimental data.

In addition to the parametric EXP model, we used a nonparametric systems identification approach to model neurons as systems whose outputs (firing rates) depend on past values of outputs and inputs (rewards). Similar autoregressive frameworks have been used to model behavioral and neural processes elsewhere (Scheidt et al. 2001; Truccolo et al. 2005). The modeling assumptions are that the system accounting for reward dependence is linear and that its properties are static over time. The resulting data were fit to an autoregressive with exogenous terms (ARX) model, described by

$$y(t_n) + a_1 y(t_{n-1}) + \ \ldots \ + a_K y(t_{n-K}) = b_{0u}(t_n) + b_{1u}(t_{n-1}) + \ \ldots \ + b_{Lu}(t_{n-L}) + e$$

where a_i, b_j are weights to be estimated. The additional signal e is the system disturbance and accounts for noise or other variables not considered. The parameters K and L are the orders of the delays and indicate how many past values of the output and input are included in the regression (Fig. 2*B*). The optimal orders were selected by minimizing the Akaike information criterion (AIC), which provides a conservative trade-off between model error and model complexity (Akaike 1974). To prevent overfitting, the data were split into two contiguous halves for model order estimation and validation. Thus the selected ARX model is as an optimal linear system that does not overfit the data. We chose the ARX model because the resultant optimal weight vectors sometimes oscillated and did not always peak at the first lag time. Because the impulse responses have infinite lengths but can decay quickly, we estimated the conservative effective duration of an impulse response function (IRF) as an approximate time constant, using $(1 - r)^{-1}$, where r is the radius of the largest pole (Orfanidis 1996).

For both EXP and ARX models, we used all trials to fit the model parameters for each neuron. Because of insufficient data per neuron (~2.5 cycles through the experiment on average), we did not cross-validate the model fitting procedure. We quantified the fit of both models to neural data with Pearson's correlation coefficient and imposed a criterion of $P < 0.01$ for statistically significant model correlations.

Classification. To assess the effects of reward schedule on decode performance, we performed offline classification of reach direction (4 classes) and reward magnitude (2 classes) for constant and variable schedules. We use the notation Direction*Reward to indicate simultaneous classification of reward and direction from a group with eight unique classes formed from all possible directions and rewards (chance = 0.125). The number of trials per class was balanced across classes and schedule conditions. The data were randomly partitioned into five subsets. The classifier was trained with four subsets, leaving

the remaining subset as the test set. The classification was repeated five times so that each subset was the test set at least once. This fivefold cross-validation procedure generated five classification estimates. The number of neurons in a set used in the classification was randomly selected with replacement from 32 MIP neurons recorded from a multielectrode array. The sampling of neurons for each set was repeated 10 times. For example, when classifying with 15 neurons, we randomly chose 15 neurons from our set of 32 neurons, classified the test data using fivefold cross-validation as described above, and repeated this procedure 10 times. Unless noted otherwise, the decode success in percentage points is reported as mean ± SD. All comparisons between decode performance were performed with a two-way *t*-test with unequal variances with $n = 50$ and a significance threshold set at $P = 0.05$.

We performed the classification with memory period firing rates, defined as the number of spikes in 800 ms of the memory period starting 200 ms after cue offset, using a linear discriminant analysis with uniform priors (Duda et al. 2001). We used this simple model to uncover the effect of reward schedule on performance. Classification was also performed with Haar wavelet coefficients projected onto 512 ms of the memory period starting 200 ms after cue offset. Haar wavelet coefficients capture the mean number of spikes in temporal windows of arbitrary length, can detect local changes or dynamic structure in the firing rate (Cao 2003; Musallam et al. 2004), and are favorable when analyzing nonstationary data (Mallat 1999). The 0th Haar coefficient corresponds to the mean firing rate. The method used to generate wavelets is described in the supplementary materials of Musallam et al. (2004). Only the coefficients from the training data, and not the test data, were tested for significance with a sorting algorithm based on information theory (Cao 2003).

Controls. Because we recorded from all neurons encountered, we first verified that the memory period spiking activity of neurons was significantly different from baseline for one or more conditions. We defined baseline activity as the number of spikes per second in 500 ms of neural data prior to cue onset. For the majority of recorded neurons (114/135, 84%), the memory period firing rate was significantly different from the baseline ($P < 0.05$, *t*-test), confirming that the memory period is a meaningful trial epoch for analysis. We also confirmed that neural modulation was not a response to the visual stimuli by randomly switching the association between cue size and reward size for different recording sessions. Preferred reward was not significantly different between cue size associations ($P = 0.3034$, $n = 135$, *t*-test); neurons did not systematically prefer a larger visual cue.

RESULTS

To identify neurons sensitive to reach direction, reward magnitude, or reward schedule, we performed a three-way ANOVA ($P < 0.05$) on the memory period firing rate of all neurons (Table 1). Reward magnitude, reward schedule, and interaction between magnitude and schedule modulated the firing of 109 of 135 neurons. The activity of 74 neurons was sensitive to reward magnitude (column R under "All Cells," Table 1) while 72% of these neurons (53/74) also showed significant reward magnitude and schedule interaction effects (column R×S under "Reward Magnitude Selective," Table 1). Conversely, the activity of 72 neurons was modified by reward schedule (column S under 'All Cells," Table 1), while 56% of these neurons (40/72) showed significant reward magnitude and schedule interaction effects (column R×S under "Reward Schedule Selective," Table 1).

Reward magnitude and reach direction independently modulated the firing of most neurons. Of the 67 and 74 neurons selective for reach direction and reward, only 15 neurons were sensitive to interactions between these two variables. Furthermore, 47% (35/74) and 57% (41/72) of neurons that responded to changes in reward magnitude and schedule in the memory period also encoded the reach direction in this period (columns D under "Reward Magnitude Selective" and Reward Schedule Selective," Table 1). Neural activity selective for reach direction was also present in the cue and motor periods of many neurons. The proportion of reward neurons and schedule neurons sensitive to direction increased to 75% (56/74) and 82% (59/72) when considering neural activity recorded in the motor or cue epochs. The large number of neurons sensitive to reward magnitude may be explained by our decision to record from all stable cells encountered in MIP.

Effect of Schedule on Neural Sensitivity to Reward

The interaction values in Table 1 suggest that the neural response to reward magnitude was not static but was dependent on the reward schedule. Figure 3*A* shows the peristimulus time histogram for an example neuron that responded to reward. For both schedules, large-reward trials elicited higher firing rates than small-reward trials. Even though reward magnitudes were the same for both schedules, the range between the firing for large and small rewards was greater during the constant schedule (Fig. 3*B*). The population showed a similar trend across all monkeys and recording procedures (Fig. 3*C*). The activity of 73% of reward- and schedule-sensitive neurons (80/109) [or 59% (80/135) of all recorded neurons] differed for small- and

Table 1. *Numbers of neurons that responded to task variables and their interactions*

		All Cells (135)								Reward Magnitude Selective (74)		Reward Schedule Selective (72)	
	Total	*D*	*R*	*S*	*D*×*R*	*D*×*S*	*R*×*S*	*D*×*R*×*S*	*R S R*×*S*	*D*	*R*×*S*	*D*	*R*×*S*
Monkey H	49	32	19	23	7	5	19	7	35	11	12	17	11
Monkey F	40	13	24	24	7	2	24	5	35	9	18	9	14
Monkey M	46	22	31	25	1	1	25	0	39	15	23	15	15
Total	135	67	74	72	15	8	68	12	109	35	53	41	40

Values are numbers of neurons that responded ($P < 0.05$, 3-way ANOVA) to task variables and their interactions, separated by monkey. The "Total" column indicates the total number of neurons recorded from each monkey. The columns under "All Cells" indicate the number of cells with significant modulation of memory period firing rate by task variables. The task variables *D*, *R*, and *S* correspond to reach direction, reward magnitude, and reward schedule, respectively. We use "×" and " " to denote interaction effects and the logical disjunction (OR) operator, respectively. For example, the *R S R*×*S* column highlights the number of cells with significant modulation by reward magnitude or schedule, or an interaction of the two. Ten cells did not vary their activity with any of these. Additional characterizations of directional tuning and reward interaction effects are provided for the subset of reward magnitude and schedule selective cells in the "Reward Magnitude Selective" and "Reward Schedule Selective" columns, respectively.

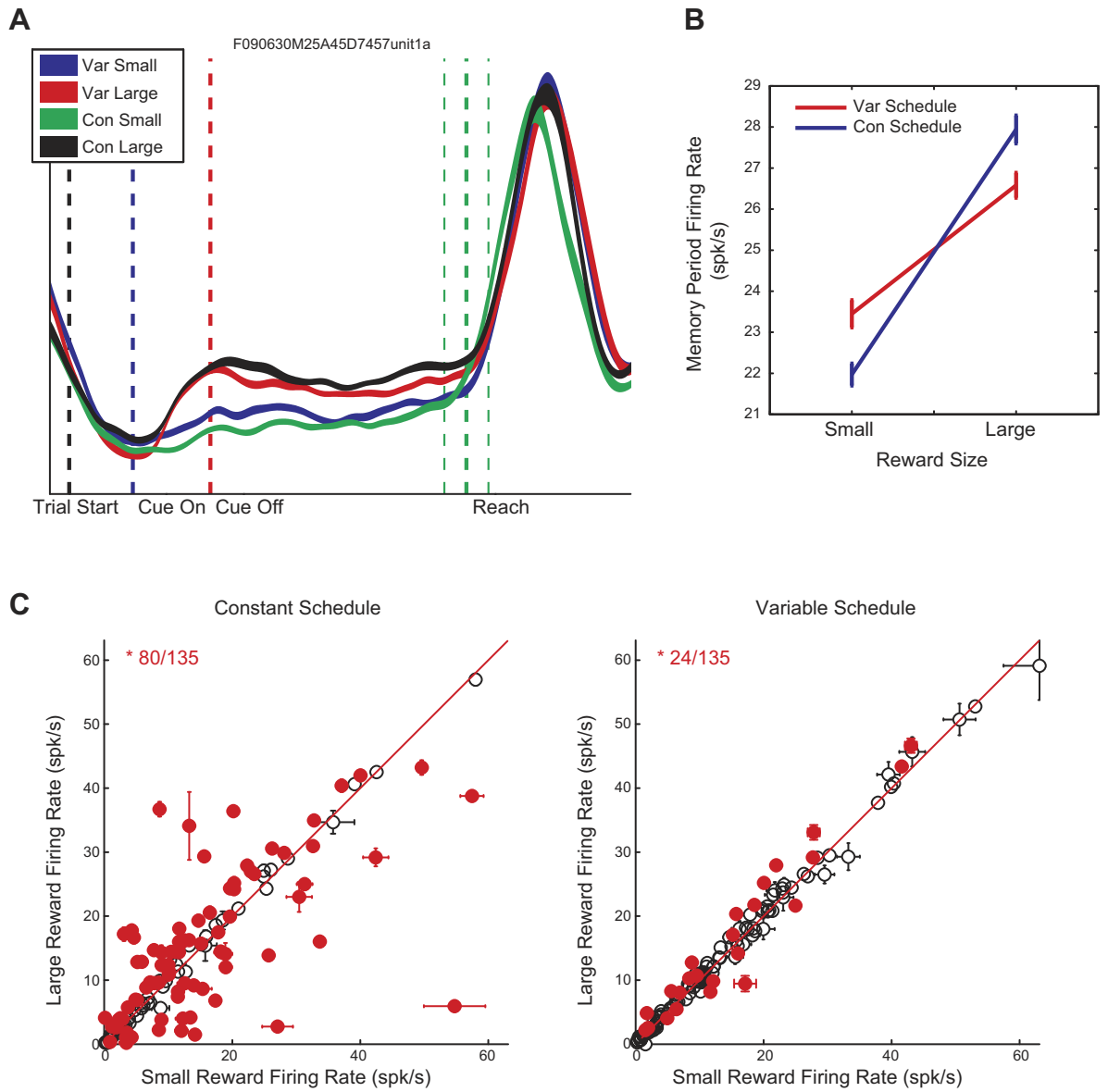


Fig. 3. *A*: example peristimulus time histogram (PSTH): PSTH (mean ± SE) for an example neuron for all combinations of reward magnitudes (small/large) and schedules (constant/variable), aligned to trial start. Vertical gridlines indicate important trial events. This neuron increased the firing rate for trials with large rewards over small rewards, especially in the memory period. The neuron's response to equal reward magnitudes varied significantly between the 2 reward schedules. *B*: reward tuning curve. For the example neuron in *A*, the reward tuning curve shows that the memory period firing rate of this cell has significant reward modulation but the strength of this reward tuning varies with the reward schedule; there is greater discrimination between small and large rewards in the constant schedule than in the variable schedule. *C*: population reward responses. The reward responses of all recorded neurons are shown, separated by schedule. The scattered points compare the firing rate (mean ± SE) to small vs. large rewards in constant (*left*)- and variable (*right*)-reward schedules for all recorded neurons. Filled data points indicate a significant difference between the activities for small vs. large rewards. For the constant schedule, 59% (80/135) of recorded cells showed a significant difference in firing rate between reward magnitudes ($P < 0.05$, ANOVA). Only 17% (24/135) of neurons exhibited a difference in the variable schedule ($P < 0.05$, ANOVA). This result was consistent across *monkeys H*, *F*, and *M* and for both recording procedures. At the population level, we find that neural activity in the medial intraparietal area (MIP) is modulated differently for the same reward between the constant and variable schedules.

large-reward trials during the constant schedule. During the variable schedule, this number decreased to 24 [22% of reward- and schedule-sensitive neurons, or 18% of all recorded neurons (24/135)]. We observed a similar difference in the number of reward-sensitive cells across schedules when memory period firing rates were normalized by pretrial baseline activity (70/135 in the constant schedule vs. 29/135 in the variable schedule). Increased firing rate was not always associated with the promise of large rewards. In the constant schedule, a large reward increased the firing rate of 61% (51/83) of cells selective for reward. In the variable schedule, 58% of neurons selective for reward had an analogous response.

We used MI to quantify the discriminability between the two reward magnitudes. For individual neurons, the MI during constant schedules was significantly greater than the MI during variable schedules ($P < 0.05$, 2-sample *t*-test on bootstrapped distributions) for 87% (95/109) of reward-responding neurons (Fig. 4*A*). At the population level, the MI during the variable schedule was significantly lower than the MI during the constant schedule ($P = 6.8 \times 10^{-7}$, paired *t*-test; see Fig. 4*B*).

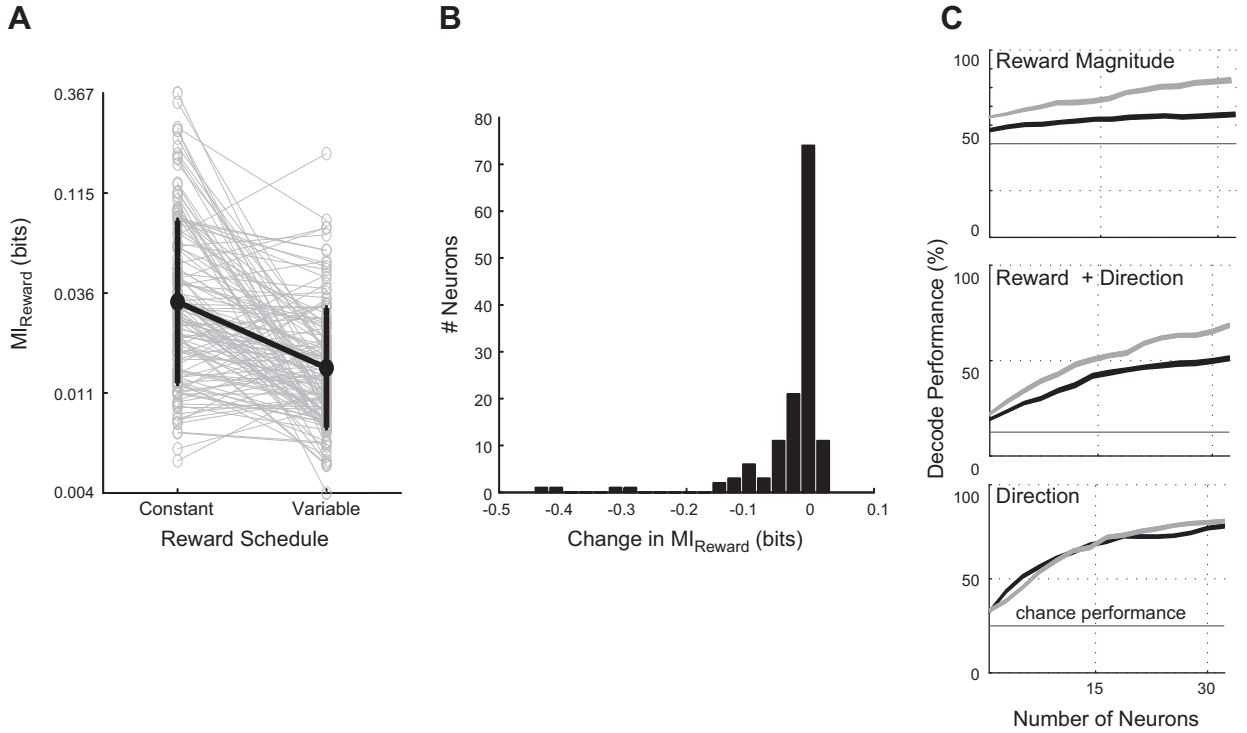


Fig. 4. *A* and *B*: difference in reward information between schedules. We used the metric of mutual information (MI, measured in bits) to measure a neuron's ability to discriminate between the 2 reward magnitudes. The MI between memory period firing rate (plotted in log scale) and reward for each recorded neuron is shown in *A*, separated by reward schedule. For individual neurons, we find a significant decrease in MI between constant and variable schedules ($P < 0.05$, 2-sample *t*-test on bootstrapped distributions) for 87% (95/109) of reward-modulated neurons. The change in MI at the population level is shown in *B*; we found a significant decrease in MI between constant and variable schedules ($P = 6.8 \times 10^{-7}$, paired *t*-test), indicating that neurons discriminate more between small and large rewards in the constant schedule than in the variable schedule. *C*: performance of a discriminant classifier attempting to predict reward magnitude (chance = 50%), reward and direction (chance = 12.5%), and direction (chance = 25%) with the memory period mean firing rate for each schedule. The constant and variable schedules are depicted by gray and black traces, respectively. The firing rate was obtained from neurons simultaneously recorded from an array implanted in MIP. We used up to 32 neurons that were tuned to any of direction, reward, or schedule. The thickness of the classification curves is the 95% confidence interval around the mean.

Decode Performance of Identical Rewards Varies with Schedule

We next examined the influence of reward schedule on our ability to decode reward magnitude. During the constant schedule, reward magnitude was classified with a maximum success rate of 83.9 ± 5.3% (Fig. 4*C*, *top*). This result is significantly greater than the decode performance during the variable schedule (VS: 65.0 ± 6%, chance = 50%; paired *t*-test between best performance, $P << 0.05$, $n = 50$). In contrast, our ability to decode reach direction was similar between the schedules (CS: 80.5 ± 4.3%, VS: 78.7 ± 4.5%, chance = 25%; paired *t*-test, $P = 0.04$, $n = 50$; Fig. 4*C*, *bottom*). Decoding reach direction and reward magnitude simultaneously was significantly better during the constant schedule (CS: 68.3 ± 5.7%, VS: 51.6. ± 6.4%, chance = 12.5%; paired *t*-test, $P << 0$, $n = 50$; Fig. 4*C*, *middle*). Thus decoding the same reward magnitude (with or without direction) was unreliable, as performance was dependent on reward schedule. As expected, the decode performance improved as the number of neurons increased.

Reward Schedule Does Not Affect Behavioral Metrics

We explored the contribution of reach error, reaction time, and motion time to the dependence of reward sensitivity on schedule. For both schedules, the reach error, reaction time, and motion time decreased when large rewards were expected (example of 1 behavioral session in Fig. 5*A*). For the population, the mean reach error, mean reaction time, and mean motion time for small- versus large-reward trials, separated by schedule, are scattered in Fig. 5*B*. Unlike the neural activity, however, there was no significant difference in MI across schedules for all behavioral variables (*t*-test, $P = 0.29$, $P = 0.44$, and $P = 0.11$ for error, motion time, and reaction time, respectively; see Fig. 5*C*). There was a significant change in MI of reaction times and motion times with reward magnitude for all (46/46) recording sessions ($P < 0.05$, permutation test) for both schedules. Half of the recording sessions (23/46) also showed significant change in MI between reach error and reward magnitude. Thus changes in reward magnitude led to changes in behavior (measured by MI), but this occurred equally for both schedules. Thus variables that define the metrics of reaching cannot be exclusively responsible for the observed dependence on schedule.

Effect of Time Between Schedules on Neural Activity

We next investigated whether the greater separation in time (or "temporal distance") between trials of each stimulus in the constant schedule produced the schedule dependence. During the constant schedule, animals spend considerable time receiving a single reward before the reward magnitude changes. The time difference between small and large rewards could capture nonspecific temporal or nonsta-

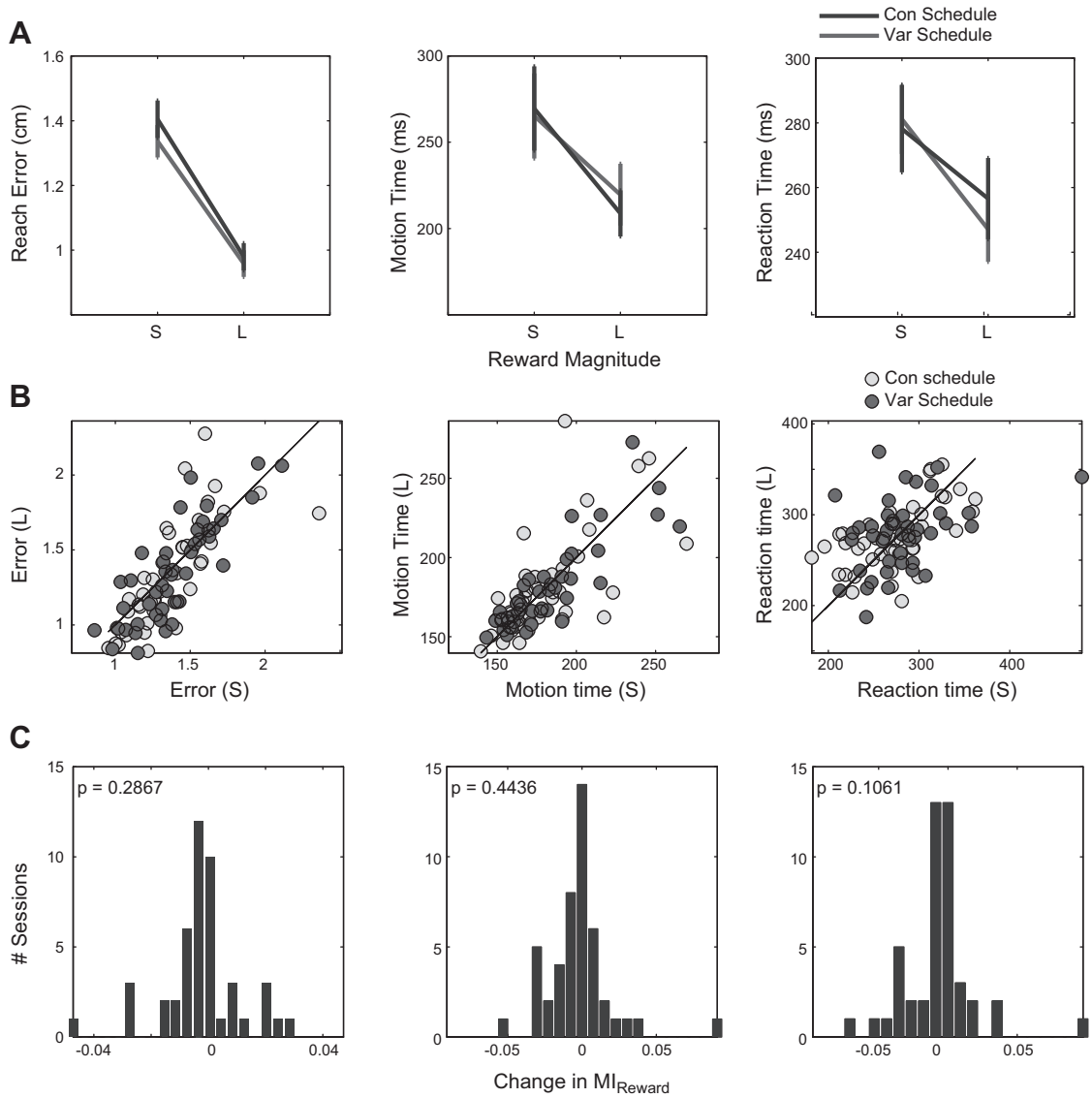
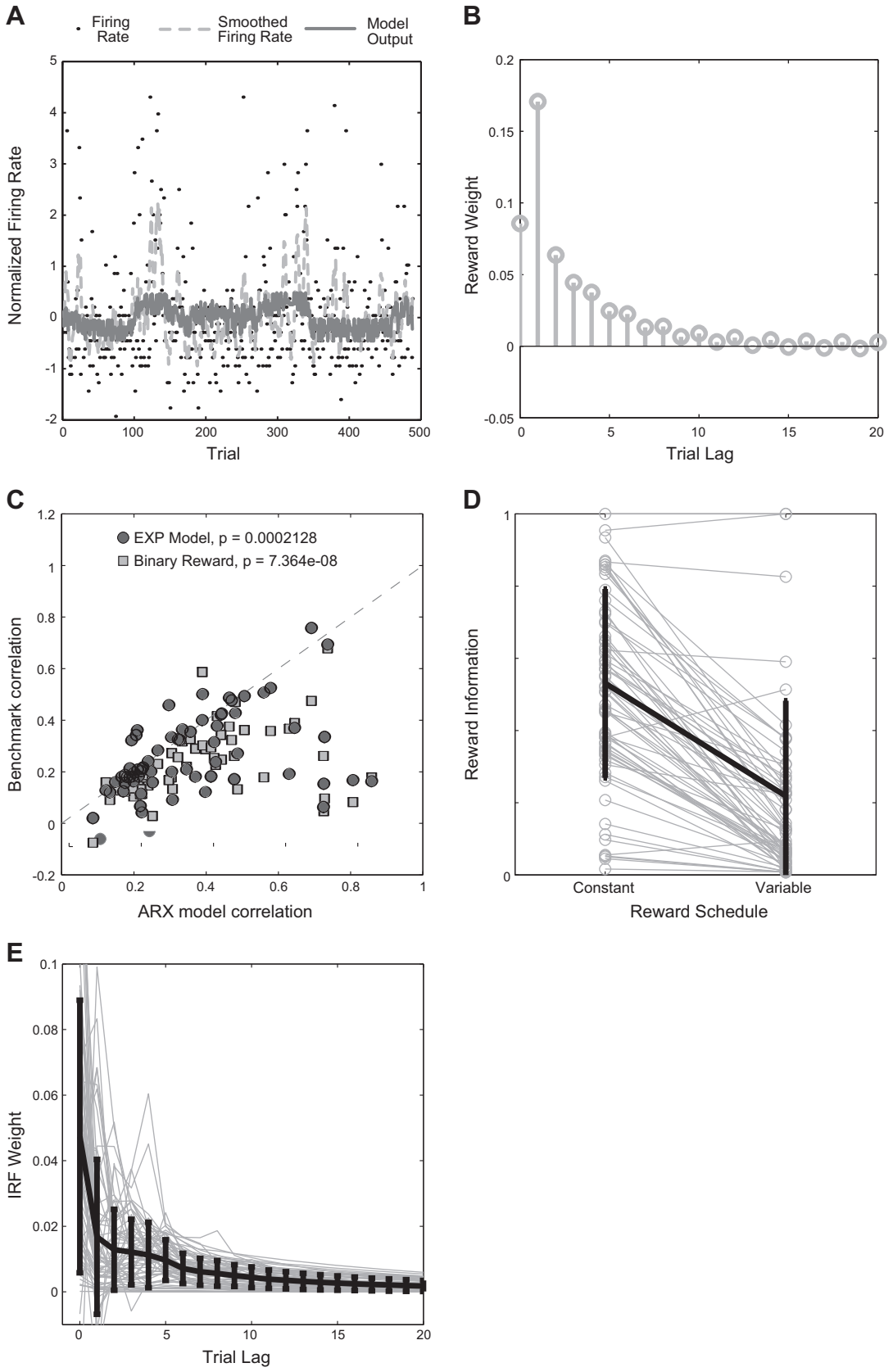


Fig. 5. *A*: behavioral response for example session. For an example recording session, the reach error, motion time, and reaction time (mean ± SE) for both reward magnitudes are shown for constant and variable schedules. We observe decreased error, reaction time, and motion time for large reward (L) compared with small reward (S), for the constant and variable schedules, but no interaction between reward magnitude and schedule for this session. In general, animals understood cue-reward associations, as evidenced by significant MI between reward magnitude and behavioral variables. *B* and *C*: comparison of behavior between schedules. The results of Fig. 3, *C* and *D*, are replicated for behavioral variables. We characterize behavior with reach error, reaction time, and motion time. In *B*, the mean behavioral responses to small vs. large rewards are shown for the constant and variable reward schedules for all recording sessions. Behavior was significantly different between reward magnitudes ($P < 0.05$, permutation test on MI), suggesting that the animals indeed discriminated between the 2 reward magnitudes. In contrast, scattered data seem to overlap for the reward schedules. In *C*, we show the change in reward discrimination between the reward schedules, using the metric of MI, in the behavioral variables. Unlike the neural activity, there is no significant difference in the reward responses ($P = 0.287$, $P = 0.444$, $P = 0.107$ for reach error, motion time, and reaction time, respectively). Thus this trend is not an epiphenomenon of behavior but a neural response to reward.

tionary artifacts. To test for this possibility, we generated 100 random sets of trials, selecting equal numbers of small- and large-reward trials from all neurons. We then computed the normalized MI between reward and the firing rates. For each of these sets, we computed the mean distance (number of trials) between small-reward trials and the mean distance between large-reward trials. We then computed the correlation coefficient between the mean trial distances and the reward discrimination within a set. If indeed there were a distance effect, then a significant positive correlation will exist. We did not find a significant correlation in 94.5% (103/109) of cells, which coincides with our statistical criterion for chance ($P < 0.05$).

ARX Model of Neural Activity

We next explored the origin of the influence of reward schedule. Constant- and variable-schedule trials only differ by the sequence of rewards, suggesting a correlation between firing rate and reward history. To test this possibility, we used the ARX model (linear autoregressive with exogenous inputs) to converge onto a vector of weights that best explained current neural activity when these weights were convolved with past rewards. This analysis applies to neurons that are selective to reward magnitude and are sensitive to reward schedule ($P < 0.05$, 3-way ANOVA, main or interaction effect). There were 68 such neurons, and we refer to them as *population I*. Figure 6*A* shows the firing rate of an



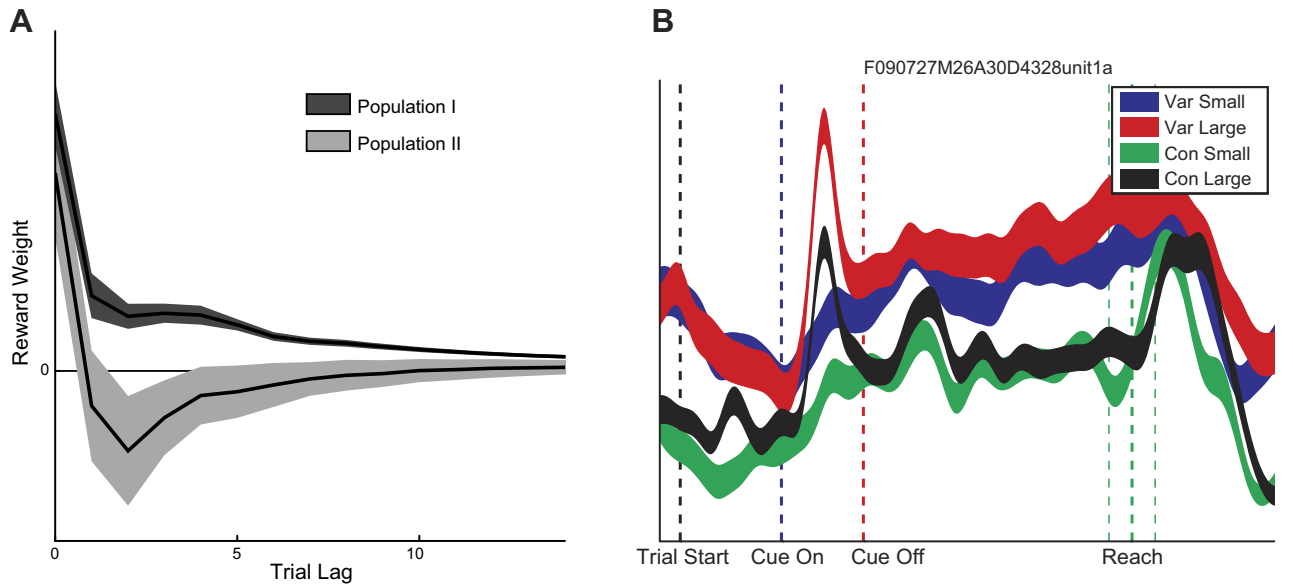


Fig. 7. *A*: ARX IRFs for 2 populations. Two populations are identified in the set of reward-responding neurons, and the corresponding ARX modeling results are shown. Solid line shows the average IRF across neurons of each population. Shaded area corresponds to the SE. In general, there is a weighted sum of past rewards, with weights decreasing with trial lag for *population I*, whereas a difference between current trial and past trial rewards best explains the firing rate of *population II* neurons. Intuitively, *population II* cells seem to exhibit a sensitization of reward, by increasing their firing activity for one schedule over the other. *B*: example PSTH: PSTH (mean ± SE) for an example neuron belonging to *population II*. Axes and formatting are identical to Fig. 3*A*, where an example neuron belonging to *population I* is shown. This *population II* cell shows increased activity for the variable schedule compared with the constant schedule during the memory period.

example neuron plotted with the model estimate and the smoothed firing rate (obtained by convolving the firing rate with a boxcar function of width 5). The model output is significantly correlated to the firing rate of this neuron ($r = 0.21$, $P < 10^{-6}$). IRF describes the weight of past trial rewards (Fig. 6*B*). The model output is a weighted sum of past rewards, with weights increasing and then decreasing with trial lag. At the population level, the model correlated significantly with the firing rate for most *population I* neurons ($P < 0.01$ for 63/68 neurons, mean correlation $r = 0.36$).

To validate the ARX model, we compared it to the parametric EXP model for all *population I* neurons (Fig. 6*C*). The ARX model was better correlated with neural firing rate for most neurons. At the population level, there is a significant increase in explained variance for the ARX model compared with the EXP model (paired *t*-test, $P = 2.1 \times 10^{-4}$). The ARX model was also better correlated with firing rate than the raw, binary reward signal (paired *t*-test, $P = 7.4 \times 10^{-8}$) (Fig. 6*C*). The ARX model replicated the observed decrease in reward discrimination in the variable schedule compared with the constant schedule (Fig. 6*D*). Thus the ARX model both fits the data well and exhibits similar response dynamics. The effective median duration estimated from the IRF of all *population I* neurons (Fig. 6*E*) was 9.8 trials (4.5, 18.2) (25th and 75th percentiles) (Orfanidis 1996), suggesting that neurons in MIP have a median memory of 10 trials.

The ARX model revealed a subset of 35 neurons best fit by weights that differentiated the input. These 35 neurons (*population II*) were modulated by reward schedule without having any selectivity for reward magnitude. The modeling analysis on *population II* cells showed that a difference between current trial and past trial rewards best explained the firing rate of these neurons, with negative model weights for low lags. The model was significantly correlated with firing rate for just over half of the cells ($P < 0.01$ for 17/35 neurons, mean correlation of $r = 0.2031$). Figure 7*A* depicts the IRFs (mean ± SE) for both neural populations. The series of negative weights (light gray curve, Fig. 7*A*) suggests that these neurons differentiate their

Fig. 6. *A* and *B*: ARX model for an example neuron. We used the ARX approach to systematically select the function of reward that best explains the neural activity. The firing rate of an example neuron (black dots) with the ARX model estimate (solid curve) is shown in *A*. For comparison, the smoothed firing rate, obtained by convolution with a boxcar function of width 5, is overlaid on the model output (dashed curve). For visualization only, the modulation by reach direction was reinstated through multiplication by the mean firing rate per direction. The model output was significantly correlated to the neural response ($r = 0.21$, $P < 10^{-6}$). For this example, the impulse response function (IRF), which describes the weight of past trial rewards, is shown in *B*. Firing rate is approximated as a weighted sum of past rewards, with weights decreasing with trial lag. *C*: comparison of ARX and benchmark models. To validate the ARX model, we compare it to the parametric EXP model and see that the ARX model has comparable or better correlation with neuronal firing rate (circles). At the population level, there is a significant increase in performance with the ARX model (paired *t*-test, $P = 22.1 \times 10^{-4}$) relative to the EXP model. Additionally, the scattered squares compare the correlations of ARX model and the raw, binary reward signal with firing rate, showing that the ARX model also has greater correlation with neuronal firing rate than the reward signal (paired *t*-test, $P = 7.4 \times 10^{-8}$). *D*: difference in reward information between schedules. The predicted change in reward discrimination, measured as MI with reward, for both reward schedules is shown for all modeled neurons. The dark solid line shows the mean MI for each schedule. This result replicates the observed phenomena of decreased reward discrimination in neural data (see Fig. 3*D*) in the variable schedule compared with the constant schedule. Thus the ARX model exhibits similar response dynamics. *E*: ARX IRFs. To characterize the resulting ARX models, the IRFs for all models are shown. The dark solid line shows the average IRF across neurons. In general, there is a weighted sum of past rewards, with weights decreasing with trial lag. Although our models have infinite impulse responses, we estimated the effective duration of the IRFs to be 9.8 trials (4.5, 18.2) (25th and 75th percentiles), suggesting that neurons in MIP have a "memory" of ~10 trials in the past.

input, using previous trials to uncover the schedule. The effective median duration for this population was 6.5 trials (3.3, 13.6) (25th and 75th percentiles). Figure 7*B* shows the peristimulus time histogram (mean ± SE) for an example neuron belonging to *population II*. This cell shows increased activity for the variable schedule compared with the constant schedule during the memory period.

In addition to the ARX model, we used CMI, a complementary information theoretic approach, to verify that past trials do indeed modulate current trial neural responses. We estimated the amount of information about the current trial contained in previous trials by calculating $I(Fe_{rn};R_{n-1}|R_n)$ and $I(FR_n;FR_{n-1}|R_{n-1})$; 84% (57/68) and 75% (26/35) of cells from *populations I* and *II* contained information about the current trial in either past trial firing rates or past trial rewards ($P < 0.05$, permutation test). For the majority of neurons, the current trial reward, past trial reward, and past trial firing rate all contain independent information regarding current trial firing rate.

Decode Using Haar Wavelet Coefficients

We used Haar wavelet decomposition of the spike train to increase the dimensions of our model. Wavelet analysis is well suited for the complexities introduced by the reward schedule (Mallat 1999). We repeated the classifications of direction, reward and Direction*Reward, using a model built from a family of Haar wavelet coefficients (Fig. 8). Overall, decode accuracy was significantly enhanced for all variables and for both schedules, although trials from the constant-reward schedule still yielded better results. Decode accuracy for reach direction was comparable for both schedules (chance = 25%: CS: 97.4 ± 1.7; VS: 95.2 ± 2.5%). Decode performance of reward and Direction*Reward were also significantly improved, especially for the variable schedule trials (black curve, Fig. 8) (reward, chance = 50%: CS: 93.1 ± 5.1%, VS: 79.2 ± 5.3; Direction*Reward, chance = 12.5%: CS: 92.6 ± 0.5%, VS: 84.7 ± 1.3%).

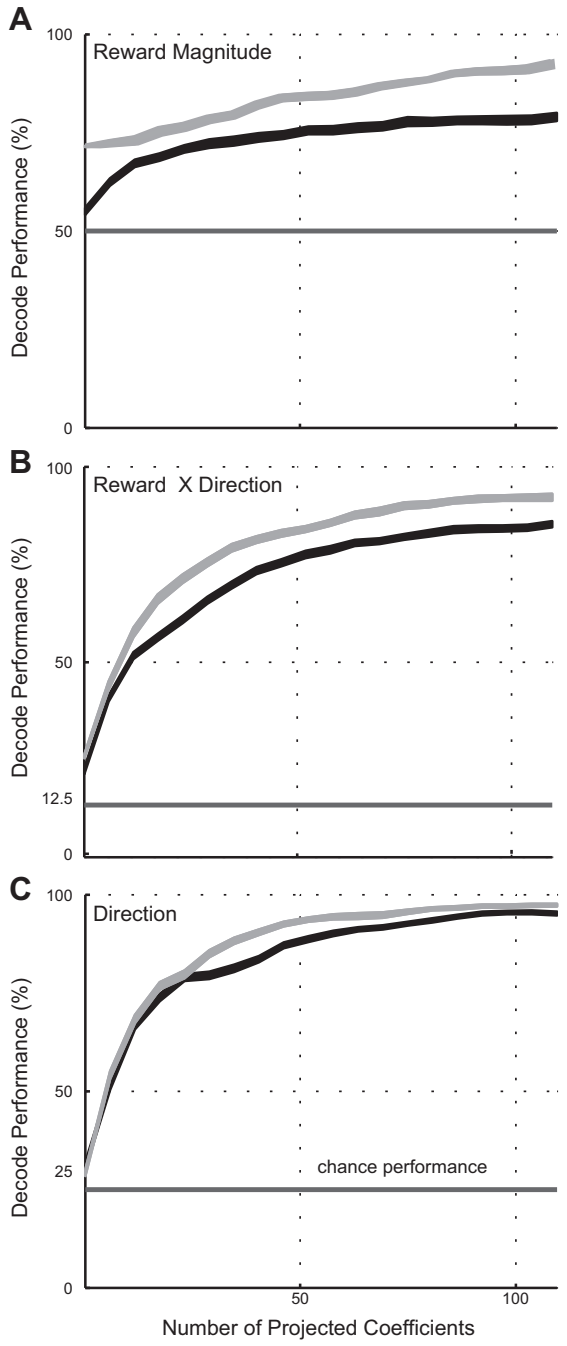


Fig. 8. *A–C*: classification results when decode included Haar wavelet coefficients. *x*-Axis is the number of coefficients used in the classification; *y*-axis is the decode accuracy during the constant (gray) and variable (black) schedules, as a function of the number of coefficients. Direction×Reward refers to the simultaneous classification of reward and direction. Horizontal lines indicate chance performance. The dependence of decoding accuracy on the number of wavelet coefficients is shown. The number of coefficients used to classify was randomly selected from a maximum of 120 coefficients. The thickness of the classification curves is the 95% confidence interval around the mean.

DISCUSSION

Previously, we found neurons in PRR sensitive to the reward associated with the reach (Musallam et al. 2004). Reward value was successfully decoded with reach neurons and improved our ability to decode reach information. This strategy is consistent with one of the key challenges in neural prosthetic research: to increase the information that can be extracted from cortical areas (Tehovnik et al. 2013). Reward can enhance the functionality of prosthetic systems by providing information about preferences, decisions, or the state of the patient (Andersen et al. 2010). Other strategies can enhance learning or extinguish bad habits by feeding back the decoded reward signal. It was unclear, however, whether neural sensitivity to reward recorded from reach neurons reflects complex decision-making signals dependent on reward history, available choices, or other contexts. Previous studies on reach neurons had not addressed this issue, as trial-by-trial randomization of reward value can mask these context dependencies of reward. To validate the utility of reward for prosthetic systems, we investigated the response of reach neurons to identical rewards by presenting reward value in schedules. Reward schedules introduced contextual dependence by varying the history of reward for each trial. The memory period firing rate of MIP neurons modulated by reward was found to vary with reward history (represented here as a reward schedule). The difference in firing rate for reach trials associated with high and low rewards was greater during the constant-reward blocks than during the variable-reward blocks. We did not find schedule-dependent behavioral changes, suggesting that reward modulates neural activity in MIP.

Reward-Related Signals in MIP

Reward-related modulation of neurons from other motor areas during decision-making paradigms has been attributed to behavioral confounds such as motivation, covarying behavioral variables, salience, or attention (Desimone and Duncan 1995; Eason et al. 1969; Hare et al. 2008, 2011; Leathers and Olson 2012; Maunsell 2004; Padoa-Schioppa and Assad 2006). Our findings do not conclusively rule out these hypotheses for neurons in MIP, but the observed changes in firing rate responses between constant and variable schedules (see Fig. 4*B*) are unlikely to be caused by behavioral confounds, since we failed to observe the changes in behavior across schedules (see Fig. 5*C*). Our results support the hypothesis that neurons in MIP are sensitive to reward rather than confounding behavioral variables.

Our results further imply that neurons in MIP are sensitive to the history of reward. The constant and variable schedules are simply block and shuffled trial configurations, composed of identical large- and small-reward trials. A difference in the firing rate to identical stimuli that appeared in block and shuffled trials implies that neural activity is dependent on reward history. We used a systems identification approach to model firing rate as a function of past trials. We found that the firing rate computed by a weighted summation of past rewards reliably fits the mean firing data for a significant population of neurons (*population I*). A smaller population of neurons (*population II*) was better described by a differential of past rewards, suggesting that the rate of change of past rewards is an important modulator of neural activity for these neurons.

A subject's confusion about the meaning of the cue during the variable schedule may have contributed to the difference in neural response between the schedules. If true, then behavioral metrics between large- and small-reward trials in the variable schedule should be similar. The significant change in MI between reward magnitude and behavioral variables during the variable schedule shows that animals understood cue-reward associations (Fig. 5). We also eliminated the possibility that our results were confounded by nonstationary temporal distance between reward trials caused by a lack of pseudorandom reward presentation in the constant schedule. We measured the occurrence of nonstationary artifacts in our data and failed to find distance effects in 103 of 109 neurons. We also safeguarded against such confounds by randomizing the order of different blocks in different recording sessions.

The dependence of reward sensitivity on reward history in reach neurons is consistent with experiments probing risk sensitivity (Buchkremer and Reinhold 2010; Kacelnik and Bateson 1997). The striatum utilizes information about reward history during decision-making and action selection (Muranishi et al. 2011). Similarly, learning to use a prosthetic device using M1 signals also relies on the striatum (Koralek et al. 2012). The functional link between premotor neurons and the reward system may allow for reward manipulations that improve learning and performance of prosthetic devices. Similarly, different reward schedules can reinforce advantageous behavior or extinguish undesirable behavior (Pipkin and Vollmer 2009). Human ACC (Behrens et al. 2007), lateral habenula and dopamine neurons (Bromberg-Martin et al. 2010), neurons in ACCd, prefrontal cortex, and parietal cortex have reward memories that range over multiple timescales that can be used to assess multiple strategies before making a choice (Bernacchia et al. 2011). The dependence of reward history in MIP may afford similar computations for decision-making by selecting one of many potential actions (Cisek and Kalaska 2010).

For each neuron, the information about reward was dependent on schedule. By summing past rewards, a neuron low-pass filters the reward sequence, filtering unpredictable or volatile reward sources. Neural sensitivity to the variance of reward computed from reward history is useful for decision-making agents and foraging activity (Arnauld and Nicole 1964; Herrnstein 1961; Sutton and Barto 1998). Many animals prefer a low-variance reward source when deciding between a constant reward source (low variance) or a variable reward source (high variance), provided the time delay between rewards is not variable (Kacelnik and Bateson 1997). Similarly, models of decision-making that incorporate weighted past rewards accurately predict variance sensitivity in animal behavior (Buchkremer and Reinhold 2010). The different sensitivities to identical rewards found between the two schedules may be due to the small and large variances associated with the constant and variable schedules.

Consequences for Neural Prosthetic Applications

Using the mean firing rate in the memory period, we were unable to achieve high decode performance of Direction*Reward when the reward history was not controlled. Initially, this was problematic, as schedule degraded reward decoding (Fig. 4). However, we successfully decoded reward, and simultaneously decoded reward and reach goal, using Haar wavelet coefficients computed from the firing rates. The improved decode accuracy achieved with wavelets suggests that additional information may be contained in dynamics of the spike trains and at multiple timescales. This information may be less sensitive to context, as shown by the high classification rates of behavioral variables. Adding local field potentials to the decode paradigm may further enhance the success rate (Hwang and Andersen 2013). We conclude that neural activity in MIP is sensitive to reach goals and the value of reward at multiple timescales. Furthermore, MIP is a reliable source of motor and cognitive information that can enhance prosthetic function.

ACKNOWLEDGMENTS

We thank Dr. Erik Cook, Dr. Christopher Pack, and Dr. Edward Tehovnik for their comments on a previous version of this manuscript. We also thank Stephen Nuara for animal help and Walter Kucharski for technical support.

GRANTS

Financial support was received from the Canadian Institute of Health Research, the National Science and Engineering Research Council, the Canadian Foundation for Innovation, the Canada Research Chair Program, and FQRNT.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: R.R. and S.M. analyzed data; R.R. and S.M. interpreted results of experiments; R.R. and S.M. prepared figures; R.R. and S.M. drafted manuscript; R.R., R.G.S., and S.M. edited and revised manuscript;

R.G.S., G.T., and S.M. performed experiments; S.M. conception and design of research; S.M. approved final version of manuscript.

REFERENCES

Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 19: 716–723, 1974.

Andersen RA, Hwang EJ, Mulliken GH. Cognitive neural prosthetics. *Annu Rev Psychol* 61: 169–190, 2010.

Andersen RA, Musallam S, Pesaran B. Selecting the signals for a brain-machine interface. *Curr Opin Neurobiol* 14: 720–726, 2004.

Arnauld A, Nicole P. *The Art of Thinking; Port-Royal Logic*. Indianapolis, IN: Bobbs-Merrill, 1964.

Barraclough DJ, Conroy ML, Lee D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7: 404–410, 2004.

Batista AP, Buneo CA, Snyder LH, Andersen RA. Reach plans in eye-centered coordinates. *Science* 285: 257–260, 1999.

Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47: 129–141, 2005.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci* 10: 1214–1221, 2007.

Bernacchia A, Seo H, Lee D, Wang XJ. A reservoir of time constants for memory traces in cortical neurons. *Nat Neurosci* 14: 366–372, 2011.

Birbaumer N. Breaking the silence: brain-computer interfaces (BCI) for communication and motor control. *Psychophysiology* 43: 517–532, 2006.

Bisley JW, Goldberg ME. Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299: 81–86, 2003.

Boucher L, Palmeri TJ, Logan GD, Schall JD. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol Rev* 114: 376–397, 2007.

Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O. Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* 67: 499–510, 2010.

Buchkremer EM, Reinhold K. The emergence of variance-sensitivity with successful decision rules. *Behav Ecol* 21: 576–583, 2010.

Cao S. *Spike Train Characterization and Decoding for Neural Prosthetic Devices* (PhD thesis). Pasadena, CA: California Inst. of Technology, 2003.

Cisek P, Kalaska JF. Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci* 33: 269–298, 2010.

Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18: 193–222, 1995.

Dorris MC, Glimcher PW. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44: 365–378, 2004.

Doya K. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, MA: MIT Press, 2007.

Duda RO, Hart PE, Stork DG. *Pattern Classification*. New York: Wiley, 2001.

Eason RG, Harter MR, White C. Effects of attention and arousal on visually evoked cortical potentials and reaction time in man. *Physiol Behav* 4: 283–289, 1969.

Estepp JR, Klosterman SL, Christensen JC. An assessment of non-stationarity in physiological cognitive state assessment using artificial neural networks. *Conf Proc IEEE Eng Med Biol Soc* 2011: 6552–6555, 2011.

Fisher R. Statistical methods and scientific induction. *J R Stat Soc B* 22: 69–78, 1955.

Gilja V, Chestek CA, Diester I, Henderson JM, Deisseroth K, Shenoy KV. Challenges and opportunities for next-generation intracortically based neural prostheses. *IEEE Trans Biomed Eng* 58: 1891–1899, 2011.

Gold JI, Shadlen MN. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci* 5: 10–16, 2001.

Gold JI, Shadlen MN. The neural basis of decision making. *Annu Rev Neurosci* 22: 535–574, 2007.

Gottlieb JP, Kusunoki M, Goldberg ME. The representation of visual salience in monkey parietal cortex. *Nature* 391: 481–484, 1998.

Grunewald A, Linden JF, Andersen RA. Responses to auditory stimuli in macaque lateral intraparietal area. I. Effects of training. *J Neurophysiol* 82: 330–342, 1999.

Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28: 5623–5630, 2008.

Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A. Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci USA* 108: 18120–18125, 2011.

Hatsopoulos NG, Donoghue JP. The science of neural interface systems. *Annu Rev Neurosci* 32: 249–266, 2009.

Herrnstein RJ. Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav* 4: 267, 1961.

Hwang EJ, Andersen RA. The utility of multichannel local field potentials for brain-machine interfaces. *J Neural Eng* 10: 046005, 2013.

Ikeda T, Hikosaka O. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39: 693–700, 2003.

Izawa J, Rane T, Donchin O, Shadmehr R. Motor adaptation as a process of reoptimization. *J Neurosci* 28: 2883–2891, 2008.

Kacelnik A. Normative and descriptive models of decision making: time discounting and risk sensitivity. *Ciba Found Symp* 208: 51–67, 1997.

Kacelnik A, Bateson M. Risk-sensitivity: crossroads for theories of decision-making. *Trends Cogn Sci* 1: 304–309, 1997.

Kiani R, Hanks TD, Shadlen MN. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci* 28: 3017–3029, 2008.

Kiani R, Shadlen MN. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324: 759–764, 2009.

Koralek AC, Jin X, Ii JD, Costa RM, Carmena JM. Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nature* 483: 331–335, 2012.

Krusienski DJ, Grosse-Wentrup M, Galan F, Coyle D, Miller KJ, Forney E, Anderson CW. Critical issues in state-of-the-art brain-computer interface signal processing. *J Neural Eng* 8: 025002, 2011.

Lau B, Glimcher PW. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84: 555, 2005.

Leathers ML, Olson CR. In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science* 338: 132–135, 2012.

Lee D, Seo H. Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Ann NY Acad Sci* 1104: 108–122, 2007.

Liu D, Todorov E. Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J Neurosci* 27: 9354–9368, 2007.

Mahmoudi B, Principe JC, Sanchez JC. Symbiotic brain-machine interface decoding using simultaneous motor and reward neural representation. *5th Int IEEE/EMBS Conf Neural Eng* 2011: 597–600, 2011.

Mallat SG. *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1999.

Maunsell JH. Neuronal representations of cognitive state: reward or attention? *Trends Cogn Sci* 8: 261–265, 2004.

Minamimoto T, Hori Y, Kimura M. Complementary process to response bias in the centromedian nucleus of the thalamus. *Science* 308: 1798–1801, 2005.

Muranishi M, Inokawa H, Yamada H, Ueda Y, Matsumoto N, Nakagawa M, Kimura M. Inactivation of the putamen selectively impairs reward history-based action selection. *Exp Brain Res* 209: 235–246, 2011.

Musallam S, Corneil BD, Greger B, Scherberger H, Andersen RA. Cognitive control signals for neural prosthetics. *Science* 305: 258–262, 2004.

O'Doherty JP. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14: 769–776, 2004.

O'Sullivan I, Burdet E, Diedrichsen J. Dissociating variability and effort as determinants of coordination. *Plos Comput Biol* 5: e1000345, 2009.

Orfanidis S. *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1996.

Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. *Annu Rev Neurosci* 34: 333, 2011.

Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441: 223–226, 2006.

Peck CJ, Jangraw DC, Suzuki M, Efem R, Gottlieb J. Reward modulates attention independently of action value in posterior parietal cortex. *J Neurosci* 29: 11182–11191, 2009.

Pipkin CS, Vollmer TR. Applied implications of reinforcement history effects. *J Appl Behav Anal* 42: 83–103, 2009.

Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature* 400: 233–238, 1999.

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

Schall JD, Thompson KG. Neural selection and control of visually guided eye movements. *Annu Rev Neurosci* 22: 241–259, 1999.

Scheidt RA, Dingwell JB, Mussa-Ivaldi FA. Learning to move amid uncertainty. *J Neurophysiol* 86: 971–985, 2001.

Schultz W. Multiple reward signals in the brain. *Nat Rev Neurosci* 1: 199–207, 2000.

Schultz W. Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57: 87–115, 2006.

Scott SH. Optimal feedback control and the neural basis of volitional motor control. *Nat Rev Neurosci* 5: 534–546, 2004.

Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci* 27: 8366–8377, 2007.

Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86: 1916–1936, 2001.

Shadmehr R, de Xivry JJ, Xu-Wilson M, Shih TY. Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30: 10507–10516, 2010.

Snyder LH, Batista AP, Andersen RA. Coding of intention in the posterior parietal cortex. *Nature* 386: 167–170, 1997.

Snyder LH, Batista AP, Andersen RA. Intention-related activity in the posterior parietal cortex: a review. *Vision Res* 40: 1433–1441, 2000.

Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science* 304: 1782–1787, 2004.

Sutton RS, Barto AG. *Reinforcement Learning: an Introduction*. Cambridge, MA: MIT Press, 1998.

Tehovnik EJ, Woods LC, Slocum WM. Transfer of information by BMI. *Neuroscience* 255: 134–146, 2013.

Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol* 93: 1074–1089, 2005.

Wallis JD, Miller EK. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18: 2069–2081, 2003.

Zander TO, Jatzev S. Detecting affective covert user states with passive brain-computer interfaces. *3rd Int Conf Affective Comput Intell Interaction Workshops* 2009: 1–9, 2009.