

Different Pedunculo-pontine Tegmental Neurons Signal Predicted and Actual Task Rewards

Ken-ichi Okada,¹ Keisuke Toyama,² Yuka Inoue,³ Tadashi Isa,³ and Yasushi Kobayashi^{1,2}

¹Graduate School of Frontier Biosciences, Osaka University, Toyonaka 560-8531, Japan, ²ATR Computational Neuroscience Laboratories, Seika-cho, Kyoto 619-0288, Japan and ³Department of Developmental Physiology, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan

The dopamine system has been implicated in guiding behavior based on rewards. The pedunculo-pontine tegmental nucleus (PPTN) of the brainstem receives afferent inputs from reward-related structures, including the cerebral cortices and the basal ganglia, and in turn provides strong excitatory projections to dopamine neurons. This anatomical evidence predicts that PPTN neurons may carry reward information. To elucidate the functional role of the PPTN in reward-seeking behavior, we recorded single PPTN neurons while monkeys performed a visually guided saccade task in which the predicted reward value was informed by the shape of the fixation target. Two distinct groups of neurons, fixation target (FT) and reward delivery (RD) neurons, carried reward information. The activity of FT neurons persisted between FT onset and reward delivery, with the level of activity associated with the magnitude of the expected reward. RD neurons discharged phasically after reward delivery, with the levels of activity associated with the actual reward. These results suggest that separate populations of PPTN neurons signal predicted and actual reward values, both of which are necessary for the computation of reward prediction error as represented by dopamine neurons.

Introduction

The basic process of reinforcement learning (Houk et al., 1995; Schultz, 2002) involves choosing a behavior for which the maximal reward is predicted and revising this prediction to minimize the reward prediction error (the difference between the predicted and actual reward). Midbrain dopamine neurons encode reward prediction error in tasks in which the animal is cued to predict the reward and revise their predictions based on the reward prediction error (Waelti et al., 2001; Nakahara et al., 2004). The basal ganglia (Hikosaka et al., 2006) and cerebral cortices (Bray and O'Doherty, 2007; Rolls et al., 2008) are implicated in reward prediction. The computation of the reward prediction error requires a temporal memory of the predicted reward (established at cue onset and sustained until reward delivery) and subtraction of the actual reward from the predicted one. The basal ganglia have been implicated in the subtraction process, but the neural mechanisms for the temporal memory of the predicted reward remain elusive.

Anatomical, electrophysiological, and pharmacological studies indicated that the pedunculo-pontine tegmental nucleus (PPTN) of the brainstem receives signals from reward related structures, including the cerebral cortices, amygdala, and basal ganglia (Garcia-Rill, 1991; Semba and Fibiger, 1992; Chiba et al., 2001; Mena-Segovia et al., 2004; Winn, 2006), and provides

strong excitatory inputs (glutamatergic and acetylcholinergic) to dopamine neurons (Scarnati et al., 1984; Blaha and Winn, 1993; Futami et al., 1995; Oakman et al., 1995; Pan and Hyland, 2005; Mena-Segovia et al., 2008). The PPTN has been shown to respond to the sensory and motor task events rather than the task reward (Matsumura et al., 1997; Pan and Hyland, 2005). Electrical stimulation of the PPTN produced burst activity in the dopamine neurons, suggesting that the PPTN conveys the task event information to the dopamine neurons through strong excitatory inputs (Floresco et al., 2003). Furthermore, using a visually guided saccade task (VGST), we found two groups of neurons in the PPTN: one that is tonically active from the onset of the fixation target (FT) until reward delivery (RD), with stronger responses on successful versus failed trials, and a second that responds phasically to reward delivery (Kobayashi et al., 2002b).

In this work, we investigate whether these two groups of neurons might be related to the computation of the reward prediction error. To test this hypothesis, we studied the activity of PPTN neurons in monkeys using a two-valued reward VGST. In this task, the shape of the fixation target (FTS) cues the animal to expect a large or small reward. We also studied PPTN activity when the task reward was withdrawn or delivered in an unexpected manner. Our experimental findings indicate that two subsets of PPTN neurons may provide the neural substrates for the temporal memory of predicted reward magnitude and the actual reward magnitude, both of which are required for the computation of the reward prediction error.

Materials and Methods

Animal preparation. All experimental procedures were performed in accordance with the National Institutes of Health *Guidelines for the Care and Use of Laboratory Animals* and approved by the Committee for Animal Experiment at Okazaki National Institutes and Osaka University.

Received Sept. 2, 2008; revised Feb. 28, 2009; accepted March 12, 2009.

This study was supported by Ministry of Education, Culture, Sports, Science, and Technology Grants 854029, 17022027, 18020019, 20033013, and 20300139. We thank F. A. Miles, M. Kawato, P. Karagiannis, B. Corneil, and I. Ohzawa for helpful comments and C. Sasaki and Y. Takeshima for technical assistance.

Correspondence should be addressed to Dr. Yasushi Kobayashi, Graduate School of Frontier Biosciences, Osaka University, 1-3 Machikaneyama, Toyonaka 560-8531, Japan. E-mail: yasushi@fbs.osaka-u.ac.jp.

DOI:10.1523/JNEUROSCI.4415-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/294858-13\$15.00/0

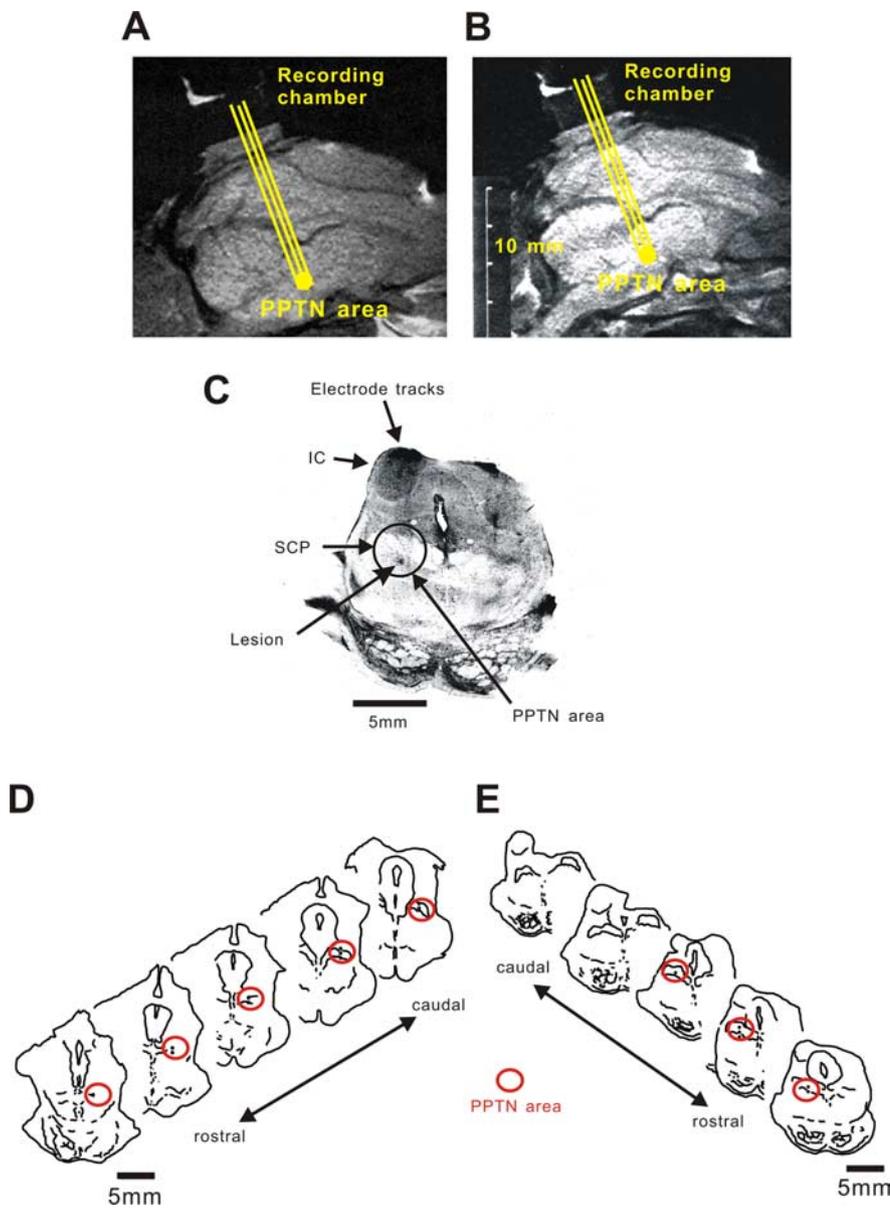


Figure 1. Recording sites. Location of recording sites from MR images of monkey 1 (**A**) and monkey 2 (**B**). Photomicrograph of a histological section cut in the coronal plane, showing electrode tracks and the lesion marking the recording site in the PPTN for monkey 2 (**C**). IC, Inferior colliculus; SCP, superior cerebellar peduncle. Histological drawings are shown for monkey 1 (**D**) and monkey 2 (**E**) with an interval of $\approx 400 \mu\text{m}$. Black dots and red circles indicate reconstructed recording sites and the histologically identified PPTN area, respectively.

The details of the surgical and data acquisition methods were published previously (Kobayashi et al., 2002b). Briefly, two Japanese monkeys [*Macaca fuscata*, one female (monkey 1), 7.0 kg and one male (monkey 2), 13.5 kg] were anesthetized with isoflurane and implanted with scleral search coils (Fuchs and Robinson, 1966), a head holder, and a recording chamber. Three weeks after surgery, the monkeys were trained to perform a VGST rewarded with juice while sitting in a primate chair with their heads restrained. Recording sessions were started after 2 months of training on a normal VGST paradigm, at which point the monkeys' performances were stable (success rate $> 80\%$).

General. All aspects of the behavioral experiment, including presentation of stimuli, monitoring of eye movements, monitoring of neuronal activity, and delivery of reward, were under the control of a personal computer-based real-time data acquisition system (Tempo) with a real-time link to Matlab. Eye position was monitored by means of a scleral search coil system with 1 ms resolution. Stimuli were presented on the screen of a 21 inch cathode ray tube monitor placed 28 cm in front of the

animal. Single-neuron activity was recorded using tungsten microelectrodes (impedance, 1–6 $M\Omega$) positioned through stainless steel guide tubes (23 gauge) using a micromanipulator. The guide tubes were held in position with a delrin grid that was fixed to the recording cylinder (Crist et al., 1988). Eye movements were recorded using a magnetic search coil (Fuchs and Robinson, 1966) (spatial resolution, 0.1° , time resolution, 10 kHz). Single-neuron activity was isolated with a template-matching spike discriminator (time resolution, 20 kHz for waveform matching and spike sampling; MSD). The spike data were registered online by computers running the personal computer-based real-time data acquisition system (Tempo) with a real-time link to Matlab, which also controlled all experimental procedures. Eye position (horizontal and vertical), spike (occurrence of action potential), and task event (visual stimuli and reward on/off) data were sampled with a 1 ms resolution.

Identification of the PPTN. Guide tubes held within the recording chamber were aimed at the PPTN of the two monkeys using magnetic resonance imaging (MRI) (2.2 T). MRI procedures were performed under general anesthesia. The locations of the recorded neurons were reconstructed for each monkey from the readings of the micromanipulator and those of the guide grids of the recording chamber (Fig. 1A,B), referenced to a single marker site selected for each monkey. Correct placement of the recording electrode was confirmed by monitoring the neuronal activity in the surrounding structures, including the auditory responses in the inferior colliculus encountered 3–7 mm before those in the PPTN and high-frequency tonic fiber activity in the cerebellar peduncle close to the PPTN. The PPTN is known to contain both cholinergic and noncholinergic neurons that generate broad and brief action potentials (Matsumura et al., 1997), respectively. Recent studies, however, report that cholinergic PPTN neurons do not always differ significantly from noncholinergic neurons in terms of these electrophysiological features (Kobayashi et al., 2002b). Therefore, rather than choosing neurons with specific electrophysiological properties, we studied all well isolated neurons in the PPTN whose activity changed during the saccade tasks.

At the conclusion of the recording experiments, electrolytic lesions were made at the selected recording sites in the two monkeys (Fig. 1C), the animals were deeply anesthetized with pentobarbital (Nembutal; 200 mg/kg), and the brains were perfused with 10% formaldehyde. Coronal sections of the frozen brain were cut and stained with cresyl violet. Location of the marked site as well as those reconstructed from the micromanipulator readings for the two monkeys were all localized in the PPTN region (126 and 27 neurons for monkeys 1 and 2, respectively) (Fig. 1D,E, red circles).

Two-valued reward VGST. Both monkeys performed the VGST (Fig. 2A,B). Trials began with the presentation of an FT (FTon; a square 0.8° per side or a circle 0.8° in diameter) at the center of the screen. The monkey was required to fixate on the FT within 3000 ms to a precision of $\pm 2^\circ$. If the monkey failed to satisfy these criteria, the trial was regarded as an error trial (fixation failure), and the trial was reinitialized. After fixation on the FT for a variable duration (400–1000 ms), another saccade target (ST) (a circle of 0.8°) appeared 10° left or right of the center of the

screen for 400–600 ms. We occasionally introduced a 200 ms delay (GAP) between the FT disappearance (FToff) and the ST presentation (STon). The monkey was required to saccade to the ST within 80–500 ms to a precision of $\pm 2^\circ$. Successful trials were rewarded with juice presented together with a tone 100 ms after the ST offset (SToff). Intertrial intervals were quasi-randomly varied. Trials in which the monkeys failed to maintain fixation on the FT were regarded as error trials (fixation-hold failure), as were trials in which the animal failed to look at or fixated the ST (saccade failures). The reaction time to fixate the FT (RTft) and saccade to the ST (RTst) was determined as a measure of the motivation to perform the task.

The shape of the FT (square or circle) cued the animal to expect either a large or small reward magnitude for successful completion of the trial. The cue-reward magnitude contingency was switched at quasi-random intervals (20–30 trials). Large and small rewards consisted of the deliver of three or one drops of juice (each drop ~ 0.1 ml). Because the cue-reward contingencies were switched every 20–30 trials, it is possible that the monkeys could anticipate the approximate timings of these switches. For a small number of neurons ($n = 15$), we kept the cue-reward magnitude contingency consistent across three different values (using a square, triangle, and circle to cue rewards of three, two, or one drops of juice, respectively).

Monkeys also performed two other versions of VGST trials. In “temporal reward-omission trials,” we delayed the delivery of the reward for successful trials by a variable delay (500–1000 ms). In “free reward trials,” a reward was suddenly delivered during the intertrial intervals. These two types of trial were randomly inserted in the two-valued reward VGST trials at a rate of 10%. The timing of the task events such as the intertrial intervals from the last reward delivery to the next FT onset (1.5–2 s) and the duration of the FT and ST presentations (400–1000 and 400–600 ms, respectively) were quasi-randomized in a manner that did not impair the monkey’s motivation to perform the task. Such randomization was aimed at minimizing the animal’s prediction of the task event timings and maximizing their efforts to predict reward magnitude for the normal and reversed cues.

Data analysis. The task-event dependency of the neuronal activity was analyzed in three ways: (1) by evaluating the significance ($p < 0.05$, t test) of the difference between the mean spike rates during the pretask and posttask event windows (600 ms for FT and RD and 100 ms for ST and saccade); (2) by constructing peri-task-event spike density functions (SDFs) aligned to each task event; and (3) by estimating the peri-task-event SDFs determined by convolving the registered neuronal spikes with a Gaussian function (σ value = 4 ms).

The precision for neuronal activity to signal the reward magnitude was analyzed in four ways: (1) by evaluating the significance of the difference between the mean spike rates for large and small rewards ($p < 0.05$, t test); (2) by receiver operating characteristic (ROC) analysis (significance level of $p < 0.05$) (Lusted, 1978) for discrimination between the small and large rewards; (3) by mutual information analysis to estimate the information contained in the spike discharges with respect to the magnitude of the reward (Werner and Mountcastle, 1963; Schreiner et al., 1978; Kitazawa et al., 1998); and (4) by regression analysis of the event parameters contributing to neuronal activities (shown below). The second and third analyses were conducted using a sliding time window of 200 ms moved in 1 ms steps. All statistical analyses were conducted using Matlab.

ROC and mutual information analysis. The precision of PPTN neurons to signal the reward magnitude was analyzed in two ways using a 200 ms moving time window across the precue, cue, maintenance, and post-reward delivery periods.

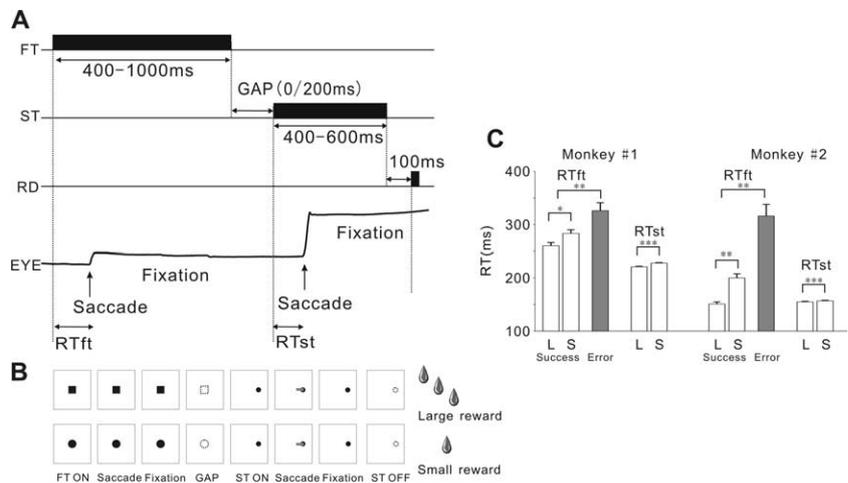


Figure 2. Diagrams for two-valued VGST and behavioral performance. **A**, Time diagram of the VGST. After fixation on the FT for 400–1000 ms, the FT disappeared after a 0 or 200 ms time gap and the ST was presented for 400–600 ms. Monkeys were required to make a saccade to the ST within 500 ms after the ST onset. Rewards for successful trials (RD) were delivered 100 ms after the ST offset. **B**, Diagrams of screen views. Arrows indicate directions of eye movement. The FT shape indicates reward value (square, three drops of juice; circle, one drop). **C**, RTft for success and error trials and RTst for success trials. Error bars indicate SEM. L and S indicate large- and small-reward trials, respectively. * $p = 0.001$, ** $p = 0.0001$, and *** $p = 0.05$, t test.

First, the reliability with which the activity of individual neurons signaled large or small reward was estimated by deriving an ROC value [cumulative probability of the ROC curve (Lusted, 1978)] that measures the accuracy by which an ideal observer could correctly distinguish between large and small reward from the neuronal signal:

$$ROC = \int_0^1 P(Q) dQ,$$

where

$$P(x) = \int_0^x p(x) dx, Q(x) = \int_0^x q(x) dx.$$

x denotes the neuronal activity sampled through the moving window. $p(x)$ and $q(x)$ denote the probability distributions for an ideal observer to answer whether the reward is large or small, respectively; $P(x)$ and $Q(x)$ denote the cumulative probability of these functions. $P(Q)$ represents an ROC curve, the ROC value is the area under the ROC curve evaluated as

$$\int_0^1 P(Q) dQ,$$

and Q is the cumulative probability function for small reward trials that was taken as the reference distribution. In principle, ROC analysis evaluates the reliability with which an ideal observer can tell whether the reward is large or small from the noisy signal in terms of statistical significance of the signal difference between the two rewards compared with the baseline noise. Therefore, an ROC value of 0.5 and >0.56 imply that the answer is 50 and 95% correct, respectively.

Second, the information capacity for the PPTN neuronal ensemble to signal reward magnitude during the three task periods was estimated via mutual information analysis (Werner and Mountcastle, 1963; Schreiner et al., 1978; Kitazawa et al., 1998):

$$I_{\text{reward}} = [- (L/N) \log_2(L/N) - (S/N) \log_2(S/N) - (High/N) \log_2(High/N) - (Low/N) \log_2(Low/N)] - [- \sum_{i=1}^2 (l_i/N) \log_2(l_i/N) - \sum_{i=1}^2 (s_i/N) \log_2(s_i/N)].$$

L , S , and N denote numbers of large- and small-reward and total trials, respectively. High and Low denote the numbers of trials in which the neuronal response was larger and smaller than the median response for all trials, respectively. Therefore, l_1 and l_2 and s_1 and s_2 represent large- and small-reward trials in which the neuronal responses were larger and smaller than the median response, respectively.

Mutual information plots for individual neurons evaluate the information capacity for the neurons to express the reward magnitude in terms of a response correlation with the reward magnitude, and cumulative plots evaluate that for the ensemble neurons for an ideal case in which the individual neuronal responses are perfectly independent. Therefore, the two analyses estimate different aspects of neuronal signal precision, although they are related. Our ROC methods estimate the signal significance compared with the baseline noise, and the mutual information analysis evaluates the signal precision in terms of signal correlation with the reward magnitude.

Multiple regression analysis. We determined the contributions of task event variables to the responses of the 22 FT and 11 RD neurons that showed significant responses to FT or RD and exhibited significantly stronger activity for large- than for small-reward trials. Three time windows [FT/cue period (200–600 ms after FTon), maintenance period of reward prediction (200–600 ms after FToff), and post-RD period (200–600 ms after RD)] were used for the following multiple linear regression analysis, assuming that the responses are a linear sum of the task variables: reward magnitude (REW), cue shape (FTS), saccadic reaction time (RTFT and RTST), and direction of the saccade (DIR). The linear analysis is essentially similar to the one used in a previous study (Leon and Shadlen, 1999) but contains an additional term, RTft, and the reaction times of eye movements were binarized as follows:

$$R^2 = \sum_{i=1}^N \{Z_i - (\alpha \text{REW}_i + b \text{FTS}_i + c \text{RTFT}_i + d \text{RTST}_i + e \text{DIR}_i + f)\}^2,$$

where R^2 is the squared residual between the neuronal response and the modeled response, Z_i the individual neuronal responses in each trial normalized to the peak response of each neuron, REW_i is the binarized variables for reward value (1 and -1 for large and small rewards), FTS_i is the shape of the FT stimulus (1 and -1 for the square and circle shape), RTFT_i is the reaction times to the FT (1 and -1 for reaction times faster and slower than the median values of the individual neuron), RTST_i is the saccadic reaction times to a peripheral saccade target (1 and -1 for saccade reaction times faster and slower than the median values of the individual neuron), and DIR_i is the direction of the saccadic eye movement to the target (1 and -1 for the leftward and rightward saccade). a , b , c , d , e , and f (f is a bias term) are the regression coefficients determined to minimize the R^2 for the trials (N ; 1615 trials for 22 FT neurons and 738 trials for 11 RD neurons) of the activity for each neuron.

Results

Behavioral performance in small- and large-reward trials

The appearance of the FT, which indicated the size of the reward for a successfully completed trial, cued the monkeys to fixate. Regardless of the cue shape–reward contingency, our analysis of the behavioral data that were sampled from 185 neuronal data (156 from monkey 1; 29 from monkey 2; minimum of five trials for each cue shape–reward contingency) showed that the monkeys were more motivated to complete trials associated with larger reward value (see below).

Behavioral performance, such as success/failure, RTft and RTst, was analyzed by ANOVA for a total of 6109 normal (4303 and 1806 for monkeys 1 and 2, respectively) and 5303 reversed cue trials (4041 and 1262 for monkeys 1 and 2, respectively), rejecting the first five trials made after changes in the cue–reward contingency for which a significant reversal effect on a behavioral performance (RTft) was detected (compare Fig. 7D). ANOVA revealed a significant difference in behavioral performance for

the reward magnitude (large or small, $p < 0.05$) but not for the cue shape (square or circle, $p > 0.1$), and therefore, these data were pooled across the two cue shapes for the large and small reward.

The percentage of successful trial was significantly higher for large than for small rewards (88 vs 80% for large vs small rewards, $p < 0.0001$, χ^2 test, correlation coefficient with reward magnitude, 0.15 for monkey 1; and 88 vs 75%, $p < 0.00001$, χ^2 test, the correlation coefficient, 0.13 for monkey 2, respectively).

The types of errors included failures to fixate on the FT (fixation error, 4.9 and 2.9% for monkeys 1 and 2, respectively), to maintain fixation until the appearance of the ST (fixation hold error, 11 and 15% for monkeys 1 and 2, respectively), and to make a saccade toward the ST (1.1 and 0.8%, for monkeys 1 and 2, respectively). The RTft was significantly shorter for the successful than for the error trials (mean \pm SEM; 271 ± 4 vs 326 ± 15 ms, $p < 0.0001$, t test, correlation coefficient of RTft with success/failure, -0.05 , $p < 0.0001$, for monkey 1; and 172 ± 4 vs 316 ± 22 ms, $p < 0.0001$, t test, the correlation coefficient, -0.06 , $p < 0.0001$, for monkey 2, respectively) (Fig. 2C). There was also a systematic difference in RTft for successful trials: those associated with large rewards were significantly shorter than those for small rewards (260 ± 6 vs 151 ± 4 ms, $p < 0.001$, t test; correlation coefficient with the reward magnitude, -0.04 , $p < 0.001$, for monkey 1; and 283 ± 7 vs 200 ± 8 ms, $p < 0.0001$, t test, the correlation coefficient, -0.05 , $p < 0.001$, for monkey 2, respectively) (Fig. 2C). The RTst was shorter than for RTft, and those for large rewards were also significantly shorter than those for small rewards (221 ± 1 vs 155 ± 1 ms, $p < 0.015$, t test, correlation coefficient with the reward magnitude, -0.03 , $p < 0.05$, for monkey 1; and 228 ± 1 ms vs 157 ± 1 ms, $p < 0.01$, t test, the correlation coefficient, -0.01 , $p < 0.02$, for monkey 2, respectively) (Fig. 2C). RTst was shorter in the gap version of the VGST than in the no-gap version (208 ± 1 vs 145 ± 1 ms, $p < 0.02$, t test, for monkey 1; and 241 ± 1 vs 168 ± 1 ms, $p < 0.01$, t test, for monkey 2, respectively).

Neuronal responsiveness to the task events

We analyzed the activity of 185 PPTN neurons during the task for 7961 normal (5534 and 2427 for monkeys 1 and 2, respectively) and 6937 reversed cue trials (5051 and 1886 for monkeys 1 and 2, respectively). For the following analyses, we rejected the first two trials made after changes in the cue–reward contingency, unless otherwise noted (see Fig. 7). ANOVA similar to that for the behavioral performance revealed a significant difference in the reward magnitude ($p < 0.05$) but not in the shape of cue ($p > 0.1$), and the data for the individual neurons were pooled across the cue shape. In agreement with previous studies (Kobayashi et al., 2002b; Pan and Hyland, 2005), PPTN neurons showed sustained or transient responses to various task events. Of the 185 neurons, 153 neurons (126 and 27 from monkeys 1 and 2, respectively) were identified as being significantly modulated around the time of at least one of the following events: FTon, STon, and RD. All analyses conducted by comparing activity in the preevent versus postevent window ($p < 0.05$, t test). Long (600 ms) preevent and postevent time windows were used for the activity after FTon and RD in which the neuronal responses were rather sustained. Short (100 ms) preevent and postevent time windows were used for the other task events (FToff, STon, SToff, and saccade onset), in which the neuronal responses were more transient.

A majority (123) of the 153 neurons responded to either the onset of the FT cue that predicted reward magnitude or the actual reward delivery. FT neurons were defined as those responding to

the onset of the FT cue (86 of 123; 70 and 16 in monkeys 1 and 2, respectively). RD neurons were defined as those significantly more active after reward delivery (35 of 123; 30 and 5 in monkeys 1 and 2, respectively). A small population of FT/RD neurons (2 of 123, 0 and 2 for monkeys 1 and 2, respectively) was equally responsive to both FTon and RD. Approximately half of the FT neurons were selectively responsive to FTon (49 of 86), and the other half (37 of 86) were broadly additionally responsive to succeeding multiple task events, including FToff, STon, saccade onset, and RD. A relatively small number of RD neurons were selective to RD (13 of 35), and many of them were responsive to the preceding multiple task events (22 of 35), including two neurons that very transiently responded to the sound of the solenoid valve for reward feeding (latency <30 ms, duration <10 ms) as well as other environmental sounds.

In agreement with previous studies (Matsumura et al., 1997; Kobayashi et al., 2002b), both FT and RD neurons exhibited a wide range of spike widths [0.17–0.73 ms (mean, 0.34 ms) and 0.20–0.66 ms (mean, 0.38 ms)]; hence, the relationship of these spikes shapes to cholinergic and glutamatergic transmission remains unclear.

Reward dependency of neuronal activity to FT and RD

Neuronal activity after FT/cue onset and RD in the majority of neurons showed a reward-dependent modulation. During the 600 ms after FT/cue onset, 30 of the 86 FT neurons exhibited a significant reward magnitude dependency ($p < 0.05$, t test), all of which showed preference for large rewards. The remaining neurons exhibited no such dependency. The two neuronal groups probably represented separate neuronal populations, because the histogram of the response magnitude correlation with the reward magnitude exhibited a clear bimodal distribution (correlation coefficient peaks at 0.1 and 0.4; Wilcoxon's test, $p < 0.0001$). There was a small population of FT neurons ($n = 6$; 4 and 2 for monkeys 1 and 2, respectively) that showed a weak negative reward magnitude dependency in that the response was smaller in the large-reward trials. These neurons were excluded from the present analysis.

During the 600 ms after RD, 15 of the 35 RD neurons exhibited significantly stronger activity for large- than for small-reward trials ($p < 0.05$, t test). The remaining 15 RD (13 and 2 for monkeys 1 and 2, respectively) neurons exhibited no such dependency on the reward magnitude.

There was a small population of RD neurons ($n = 5$, 4 and 1 for monkeys 1 and 2, respectively) that showed a weak negative reward magnitude dependency in a manner which the response was smaller in the large-reward trials. We primarily focused our analysis on the 30 FT (24 and 6 for monkeys 1 and 2, respectively) and 15 RD (12 and 3 for monkeys 1 and 2, respectively) neurons

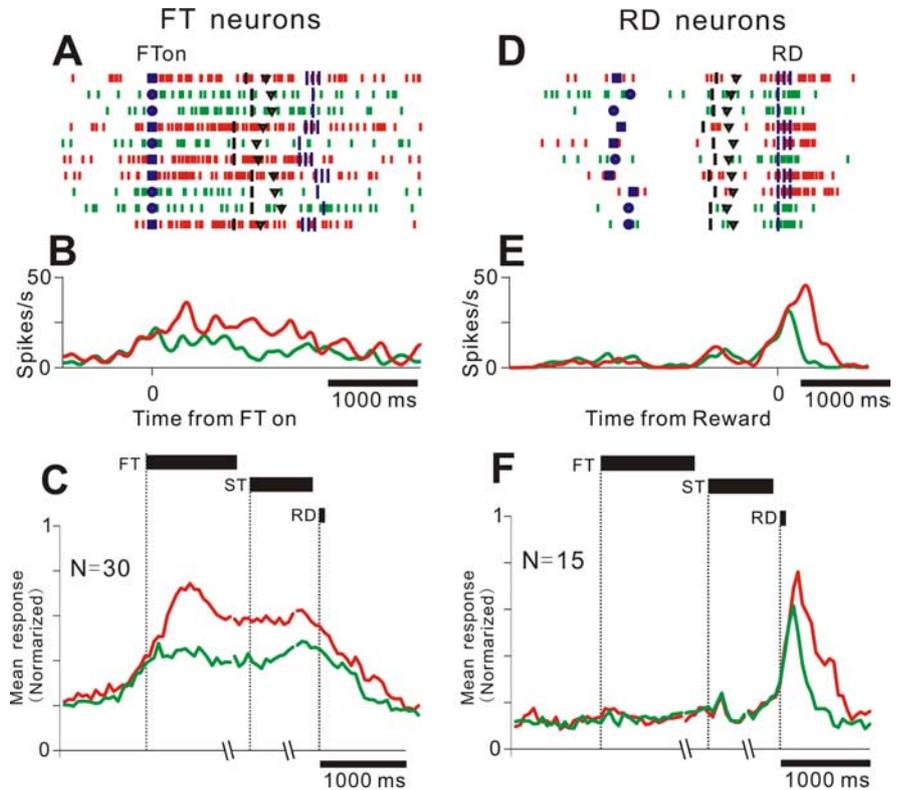


Figure 3. Responses of FT and RD neurons to task events. **A**, Rastergram for activity of a representative FT neuron during 10 successive normal cue trials, aligned to the FT onset. Red and green represent large- and small-reward trials, respectively. **A**, Blue squares and circles, The time of respective FT/cue onsets; black bars, ST onset; black triangles, saccade onset; blue bars, the times for RD, respectively. **B**, Peri-task-event SDF of the activity shown in **A**. Conventions for large- and small-reward trials are the same as for **A**. **C**, Population SDF for 30 FT neurons. Responses are aligned to the FT onset, ST onset, and RD. Population data were averaged for the FT neurons ($n = 30$) sampled in 1504 normal and 1123 reversed cue trials (squares and circles for large and small reward and vice versa), normalized for the peak response of the individual neurons. **D–F**, Similar rastergram and SDF for a representative RD neuron and population SDF for 15 RD neurons. Time axes in **C** and **F** are broken to align the responses to the onset of the FT, ST, and RD. Black bars indicate the periods of FTon, STon, and RD as denoted. Population SDF was averaged for 464 normal and 519 reverse cue trials.

that exhibited significantly more activity for large- than for small-reward trials.

Neuronal activity of reward-magnitude-dependent FT and RD neurons

Figure 3, **A** and **B**, shows raster displays and SDFs for a representative reward-magnitude-dependent FT neuron. This neuron showed elevated firing throughout the trial that was even greater when the cued reward was large. The population SDF plot for the FT neurons ($n = 30$) (Fig. 3C) indicates that these differential responses to large and small rewards generally began to emerge ~ 100 ms after the FT/cue presentation (the first dotted line). The reward-magnitude-dependent differential responses persisted even after the offset of the FT/cue until delivery of the reward (third dotted line) and remained unaffected by other task events such as the onset of the ST (Fig. 3A, black bars; C, second dotted line) and the saccade to the ST (Fig. 3A, triangles). The response differences for large versus small reward and before versus after these task events were estimated as the neuronal firing through the time window of 200 and 100 ms, respectively (t test, $p < 0.05$).

There were nondifferential responses to large and small rewards present even before the FT/cue onset, presumably representing the anticipation of FT onset. We suspect that the animals were able to approximately predict the timing of FT onset, because randomization of the intertrial intervals was rather limited

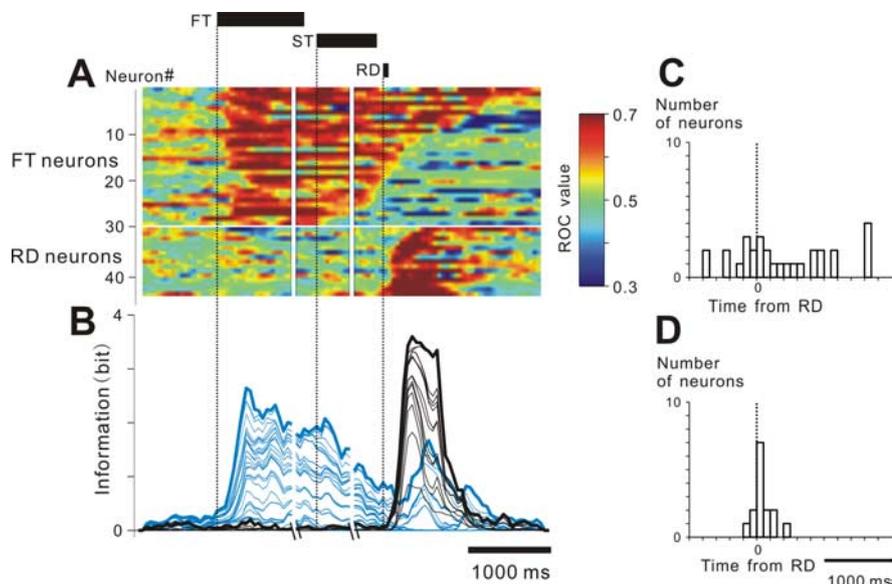


Figure 4. Information analyses of responses in FT and RD neurons. **A**, Pseudocolor plots of the instantaneous ROC values for the large and small rewards for activities in each of the 30 FT and 15 RD neurons. The plots are aligned to the FT and the ST onsets and RD. The horizontal white line separates the FT and RD. **B**, Cumulative plots of mutual information of the reward magnitude encoded by the 30 FT neurons (cyan traces) and 15 RD neurons (black traces). Time axes in **A** and **B** are broken to align the responses to the onset of FT, ST, and RD. **C**, Peri-RD histograms for the ROC value to return to the chance level ($ROC < 0.56$) for the 30 FT neurons. **D**, Peri-RD histograms for the ROC value to exceed the chance level ($ROC > 0.56$) for the 15 RD neurons. Data samples are the same as that for Figure 3, C and F.

so as not to impair the animal's task motivation. In support of this view, the nondifferential response was correlated with the success/failure to fixate FT and reflected the animal's task motivation, estimated as the reaction time to fixate FT (compare Fig. 9A–C). Therefore, the nondifferential response was a reflection of the task motivation and was difficult to abolish.

In contrast, the reward-magnitude-dependent RD neurons discharged rather transiently, reaching a peak shortly after the RD and then rapidly declined back to the baseline (Fig. 3D–F). With the larger reward, the transient response reached a higher peak at a slightly later time and took a few hundred milliseconds longer to decay back to the baseline. Approximately one-half of the RD neurons (8 of 15) also showed small nondifferential responses even before the RD, presumably in anticipation of RD ($p < 0.001$, t test) (compare Fig. 6B).

Additionally, the 30 reward-magnitude-dependent FT neurons (Fig. 3A–C) (also see Fig. 9B) as well as the 52 reward-magnitude-independent FT-responsive neurons (see Fig. 9D) exhibited a similar sustained activity starting around FTon and continuing until or even beyond RD, except for six reward-magnitude-independent FT neurons that responded only transiently (for 100 ms; data not shown) after FT presentation. Conversely, all 15 reward-dependent (Fig. 3D–F) and 15 reward-independent (data not shown) RD neurons responded to RD more transiently than the response of the FT neurons. None of these neurons exhibited a response that lasted longer than 2 s.

ROC and mutual information analyses

Our hypothesis suggests that the reward dependency of the FT neuronal response signals the reward magnitude predicted from the FT/cue stimulus, and the RD neuronal response signals the actual reward. If so, then FT neurons should convey the reward magnitude information from the time of FT onset until that of reward delivery unperturbed by the succeeding task events, and

the RD neurons should maintain the reward magnitude information unperturbed by the preceding task events.

First, we conducted ROC analysis (Lusted, 1978) to estimate the reliability of the individual neuronal responses to encode the difference between the large and small rewards across the cue period lasting from FTon to FToff and the maintenance period of reward prediction lasting from FToff to RD.

Most FT neurons (28 of 30) continued to show significant ROC values (>0.56 , corresponding to significantly higher activity for large- than for small-reward trials in the 200 ms time window; $p < 0.05$, Wilcoxon's signed-rank test) throughout the cue and maintenance period of reward prediction after the FT/cue disappeared (Fig. 4A). Furthermore, the ROC values of more than half of the FT neurons (20 of 30) remained above the chance level even after RD. Conversely, the ROC values for all of the 15 RD neurons generally did not rise above the chance level until after RD, when the ROC values displayed an abrupt and substantial increase often lasting >1 s (Fig. 4A, bottom, D). The ROC values of the FT and RD neurons gradually declined

for a few seconds after the reward delivery. Therefore, individual FT neurons reliably signaled reward magnitude from the time of cue presentation until RD, whereas individual RD neurons did so only after RD.

Second, we conducted mutual information analysis to estimate the precision of the individual and ensemble FT and RD neuronal responses to encode the reward magnitude information across the task events. Cumulative plots of the mutual information of reward magnitude conveyed by individual FT neurons (Fig. 4B, cyan traces) indicated that the information grew rapidly and peaked during the cue period, was maintained during the maintenance period of reward prediction, and then declined after the reward delivery. The 20 FT neurons did not reach the chance level even after the reward was delivered, as shown by the ROC values above. The mutual information of reward magnitude conveyed by individual RD neurons (Fig. 4B, black traces) generally did not rise above the chance level until after RD. The maximum information of reward magnitude conveyed by the populations of FT and RD neurons reached 2.6 (0.09 bits per neuron) and 3.6 bits (0.23 bits per neuron), respectively, and indicated that the FT and RD neuronal ensemble bear the information capacity capable of signaling 6 and 11 levels of the reward magnitudes. Conversely, the six FT and five RD neurons with a negative reward magnitude dependency conveyed significantly smaller reward magnitude information (0.03 and 0.03 bits per neuron, respectively).

We conducted a three-valued reward VGST in a small number of neurons (11 FT and 4 RD neurons). Although the data were insufficient for systematic ROC and mutual information analyses, FT neurons exhibited a significant reward magnitude modulation corresponding to the large, medium, and small rewards across the cue and maintenance periods. RD neurons only displayed a significant reward magnitude modulation during the

post-RD period (statistical significance for large vs medium, medium vs small reward magnitude; all $p < 0.05$, Wilcoxon's test).

Responses in free reward and temporal reward omission trials

The reward-magnitude-dependent responses of the FT neurons to the reward cue and those of the RD neurons to the actual reward delivery may possibly signal the predicted and actual reward magnitudes, respectively. This possibility was tested by comparing the responses to reward delivery for 19 reward-magnitude-dependent FT (15 and 4 for monkeys 1 and 2, respectively) and 9 reward-magnitude-dependent RD (6 and 3 for monkeys 1 and 2, respectively) neurons between the VGST trials in which the large reward was given preceded by the normal cue, and the free reward trials in which the large reward was given unexpectedly during the inter-trial interval.

All of the FT neurons consistently responded to the FT/cue presentation in a reward-magnitude-dependent manner for the two-valued reward VGST trials and remained totally unresponsive to free reward in the free reward trials ($p > 0.5$, t test) (Fig. 5*A,B*). However, all of the RD neurons responded briskly to both task and free reward delivery (Fig. 5*C,D*). The fact that the FT neurons remained totally unresponsive in the free reward trials, in which there was no reward prediction, whereas the RD neurons responded to both the task and free rewards, given in either an expected or unexpected manner, is consistent with the view that the FT and RD neurons encode the predicted and actual reward magnitude, respectively.

This view was further tested in the temporal reward omission trials for 15 FT neurons (12 and 3 for monkeys 1 and 2, respectively), in which the large reward expected for the normal cue was temporally delayed 200–600 ms after the regular RD timing. A majority of FT neurons (10 of 15) maintained their response beyond the expected RD timing until the time of the actual reward delivery (Fig. 6*A,B*, red raster and trace). Importantly, these neurons belonged to the group of 20 FT neurons that maintained reward modulation above chance levels beyond the time of RD (Fig. 4*C*). In a smaller number of FT neurons (5 of 15), the sustained levels of activity continued only until the expected time of RD; all of these neurons belonged to the group of 10 neurons that only maintained reward modulation until RD (Fig. 4*C*). Of the six RD neurons tested in temporal reward omission trials, four produced a small (<20% of the later reward response; $p < 0.00001$, t test) but significant increase in activity from the baseline ($p < 0.005$, t test) at the time of expected RD (Fig. 6*B*, black SDF trace) and produced a second large response to the delivery of the temporally omitted rewards, comparable with those for RD in the reg-

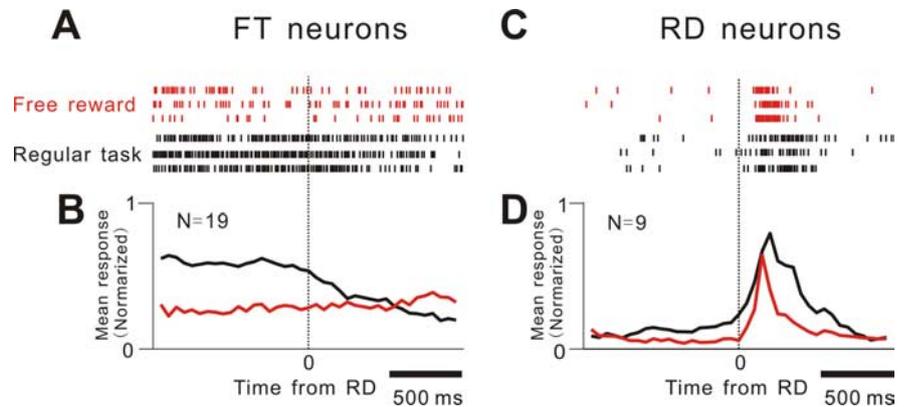


Figure 5. Response of FT and RD neurons to free and task rewards. *A, B*, Rastergram and SDF for representative and ensemble FT neurons ($n = 19$) aligned to delivery of free (red) and task (black) rewards. *B*, For 408 free reward and 1122 normal reward cue trials with a large reward, the responses represent the average firing frequency normalized for the peak responses of the individual neurons (19 FT neurons) whose number of trials was more than five for free reward with a large reward. *C, D*, Similar to *A* and *B* but for representative and ensemble response RD neurons ($n = 9$) sampled in 216 free reward and 459 normal cue trials with large rewards.

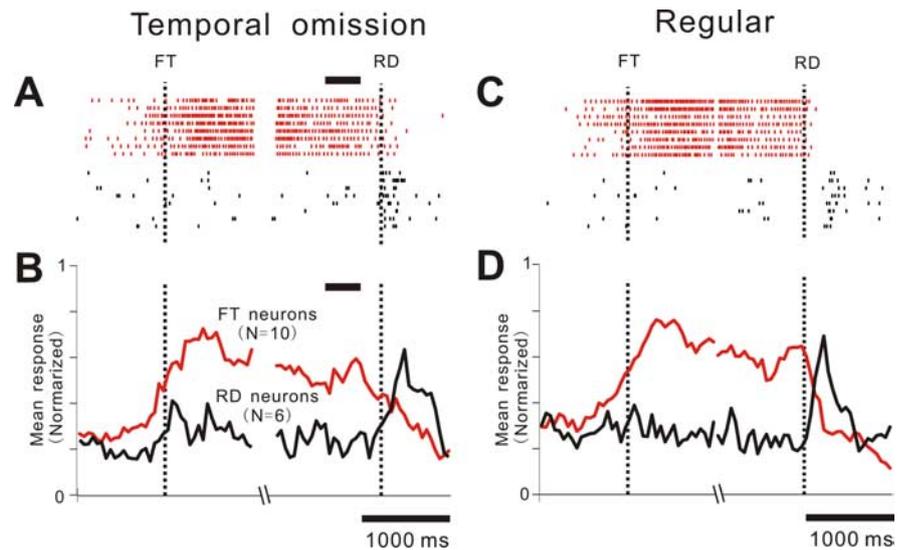


Figure 6. Response of FT and RD neurons to reward omission and task reward. *A*, Rastergram for representative FT (red) and RD (black) neuronal activity aligned to reward omission for a normal cue trial with a large reward. *B*, Ensemble SDF for reward omission in 10 FT (red) and 6 RD (black) neurons from 97 and 52 normal cue trials, respectively. The responses represent the average firing frequency normalized for the peak responses of the individual neurons (10 FT and 6 RD neurons) whose number of trials was more than five for reward omission of a normal cue trial with a large reward. *C, D*, Responses to task reward for the same representative and ensemble FT and RD neurons as in *A* and *B* for 354 and 287 normal cue trials, respectively. Responses are aligned to the time of the FT presentation and the reward delivery (left and right dotted lines, respectively). Black horizontal bars in *A* and *B* indicate the range of the expected time of RD.

ular VGST trials (Fig. 6*D*, black SDF trace). The remaining two RD neurons showed no anticipatory response.

Altogether, the majority of the FT neurons maintained the response of the reward magnitude prediction until the actual RD timing in the temporal reward omission trials, although the minority of neurons terminated the response at the expected RD timing, whereas the RD neurons signaled the magnitude of the reward delivered at the unexpected RD timing.

Responses in the transition phase of cue-reward contingency reversal

We also tested our experimental hypothesis by examining the activity of 22 FT (18 and 4 for monkeys 1 and 2, respectively) and 11 RD (8 and 3 from monkeys 1 and 2, respectively) neurons

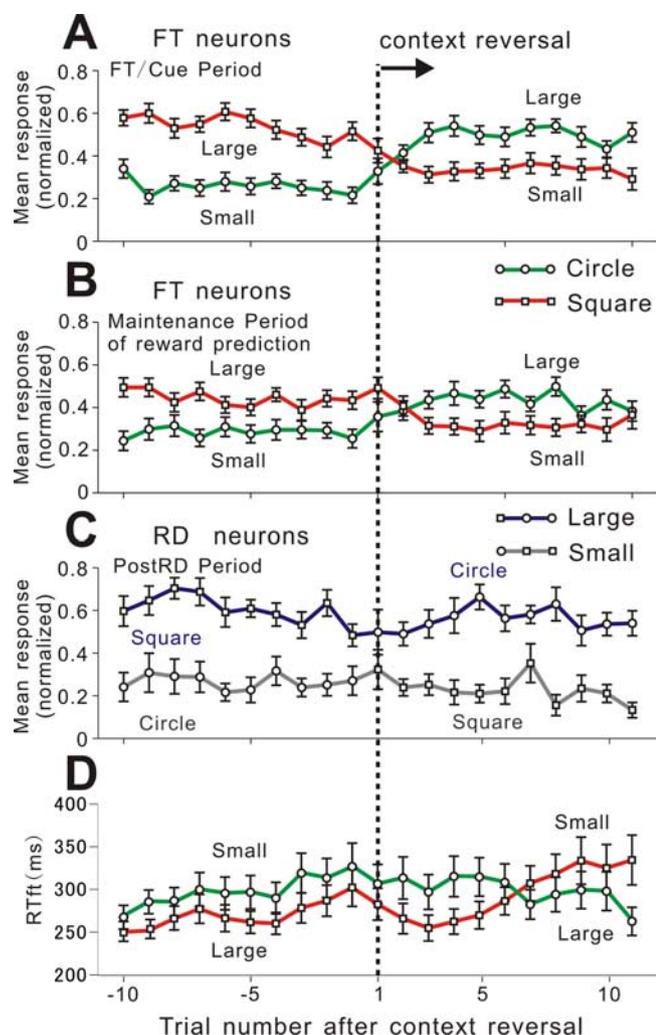


Figure 7. Effects of cue reversal on the responses of FT and RD neurons. **A**, Responses of 22 FT neurons to an FT presentation (200–600 ms after FTon, FT/cue period) before and after reversal of the FT cue (“square-large and circle-small” to “square-small and circle-large”). The response represents the average firing frequency normalized for the peak responses of the individual neurons (22 FT neurons) whose number of trials was > 10 for each of task attributes, including the cue shape and reward magnitude. Red and green lines connect the trials, in which the FT shapes were a square or circle, respectively. **B**, Similar to **A** but during the maintenance period of reward prediction (200–600 ms from FToff). **C**, Similar responses to **A** and **B** but for 11 RD neurons to RD (200–600 ms after RD, post-RD period). Blue and gray lines connect the large- and small-reward trials, respectively. **D**, RTft averaged for 2353 task trials, including 1144 normal (973 and 251 for monkeys 1 and 2, respectively) and 1209 reversed cues (890 and 239) in which the 22 FT and 11 RD neurons were sampled. The convention used to express the cue shape is the same as in **A** and **B**.

sampled across the entire VGST trials using the normal and reversed cues. We specifically focused around the change in cue-reward magnitude contingency from the normal (square, large reward; circle, small reward) to the reverse contingency (square, small reward; circle, large reward). The responses of all FT neurons during both the FT/cue period (Fig. 7A, FTon-FToff) and the subsequent maintenance period of reward prediction (Fig. 7B, FToff-RD) clearly reflected the contingency reversal with a delay of two trials. In the first reversed contingency trial, the animals could not predict the correct reward magnitude because they did not know the contingency reversal yet, and both the FT/cue and reward prediction maintenance period responses did not immediately follow the contingency reversal. As a result, the cue (square) for a large stimulus still produced the smaller re-

sponse and the cue for a small reward (circle) produced the larger response. The false reward prediction in the FT neuronal responses was partially corrected in the next trial (Fig. 7A,B) ($p < 0.0001$, Scheffé test, to compare the response in the trial with the average for the last 10 trials) and perfectly corrected ($p > 0.1$, Scheffé test) in the succeeding 20–30 trials.

These data suggest that the reward prediction of the FT neurons was improved by the reward prediction errors occurring in the two trials after the contingency reversal. Although statistically insignificant ($p > 0.1$, Scheffé test), the responses tended to change even in the first trial, as if the monkey predicted the cue reversal, suggesting that our randomization of the reversal timing (20–30 trials per contingency) was insufficient to completely eliminate the expectation of the reversal timing, although the animal could not anticipate the exact time of contingency reversal. Such anticipation may also explain the rapid revision of the reward prediction after the contingency reversal.

Figure 7C shows the RD neuronal response during the post-RD period before and after the cue reversal. The amplitude of post-RD response remained uninfluenced by the cue reversal (statistical significance for the difference, $p > 0.1$, Scheffé test). The activity of 11 RD neurons all reflected the actual reward value: strong activity if large, weak activity if small (Fig. 7C).

We also observed a significant change ($p < 0.01$, Scheffé test) in the RTft (Fig. 7D). Note that it took seven trials after contingency reversal for the RTft become longer on small- versus large-reward trials. This observation suggests that task motivation determined by RTft followed the reward prediction by the FT neurons with a delay of several trials. There was a similar but earlier (after three trials; significance level, $p < 0.05$, Scheffé test) and smaller change also in the RTst after the contingency reversal (compare Fig. 2C). The quick change of RTst for the contingency reversal is consistent with previous findings (Watanabe and Hikosaka, 2005; Rezvani and Corneil, 2008).

Multiple regression analysis

We used a multiple regression analysis to assess the possible contribution to the activity of the 22 FT and 11 RD neurons of REW, FTS, RTft, RTst, and DIR. This analysis examined the contributions to neural activity after FT/cue presentation, during the maintenance of reward prediction, and after reward delivery (see details in Materials and Methods).

The regression coefficient for REW (Fig. 8A, red bars) contributed most strongly to the FT neuronal response during both the FT/cue period and the maintenance period of reward prediction (the mean regression coefficient and statistical significance, 0.44; $p < 0.001$, t test and 0.26, $p < 0.01$, t test, respectively). RTft (Fig. 8A, purple bars) also contributed slightly to the FT neuronal response during the FT/cue period (the mean regression coefficient and statistical significance, 0.13; $p < 0.01$, t test), corresponding to the response correlation with RTft (compare Fig. 9C,D), whereas other behavioral parameters showed no significant contribution to the FT neuronal response for all of the three task periods ($p > 0.1$, t test). The reward magnitude also contributed most strongly to the activity of RD neurons during the post-RD period (Fig. 8B) (the mean regression coefficient and statistical significance, 0.48; $p < 0.001$, t test) but not for the other two task periods. There was no significant contribution of the other parameters ($p > 0.1$, t test) for all task periods.

This view was further supported by the fact that the residuals of the regression analysis of the ensembles of FT and RD activities for the full model constructed from all task event variables were much smaller ($p < 0.05$, F test) than those of the partial model,

including all variables except REW. Conversely, the residuals for the partial models missing one of the other task event variables, but including REW, were not significantly larger than those of the full model ($p > 0.2$, F test).

These results indicate that the FT and RD neuronal responses encode the predicted and actual reward magnitude information virtually unperturbed by the task events succeeding and preceding the FT/cue and RD, respectively. Altogether, the FT neuronal responses during the FT/cue and working memory periods primarily signal the predicted reward magnitude rather than the motivation to fixate on the FT or execute saccades to the ST.

Correlation of FT response with behavioral performance

Our previous study using single-valued reward VGST reported that PPTN neuronal responses were stronger on trials that were successfully completed (Kobayashi et al., 2002b). Therefore, we questioned whether the reward prediction signaled by the FT neuronal response might also be related to the task motivation. This issue was tested by studying the correlation of the FT response with the task performance for the 30 reward magnitude-dependent FT neurons across 2693 trials. Figure 9 compares representative and ensemble FT neuronal responses to large and small rewards across the fixation error, fixation hold error, and successful trials. This representative neuron showed practically no significant increase in its activity during the entire period of the fixation error trials, in which the animal failed to fixate FT (Fig. 9A, black rasters and SDF traces).

Conversely, for the fixation hold error trials in which the animal did fixate but failed to maintain the fixation (Fig. 9A, cyan rasters and SDF traces), the activity began to increase during the precue period (onset, -100 ms from FT presentation) and declined approximately at the time of the fixation break (200 ms) (Fig. 9A, see an upward arrow in the blue eye movement trace). The precue period response was reward magnitude independent in that the response was almost equal for the large- and small-reward trials (see green and red rasters and SDF traces), whereas the response was magnitude dependent during the cue period, being larger for large- than for small-reward trials (see red and green SDF traces). The FT responses in the successful trials also consisted of the reward-magnitude-independent component during the precue period that almost exactly matched that for the fixation hold error trials (Fig. 9A, red and green SDF traces), and the late reward-magnitude-dependent component that emerged during the cue period became much stronger than that for the fixation hold error trials and was sustained across the maintenance period until the post-RD period. The ensemble response for the FT neurons ($n = 30$) also showed an essentially similar tendency to that for the representative neuron (Fig. 9B). The precue period response was virtually absent in the fixation error trials, but there was a significant precue period response in the fixation hold error and the successful trials. The magnitude-dependent response in the fixation hold error trials was small and transient, whereas that in the successful trials was much larger and sustained until the post-RD period.

The fact that the reward-magnitude-independent precue period response in the FT neurons was absent in the fixation error

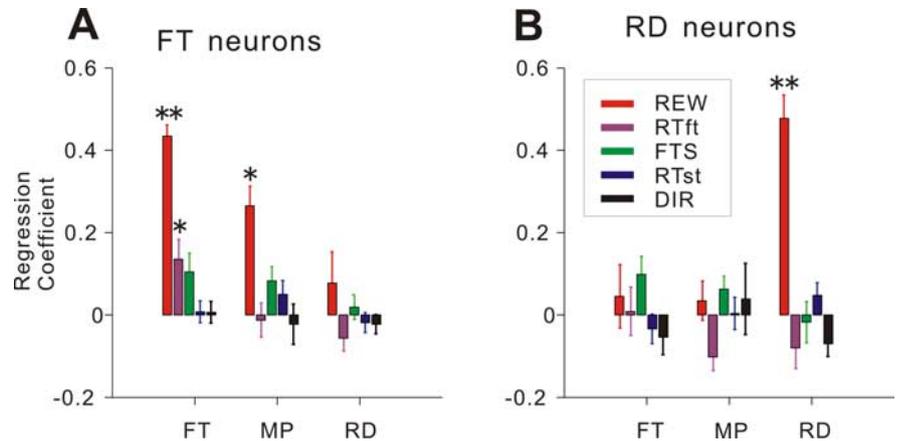


Figure 8. Multiple linear regression analyses of FT and RD neuronal responses. **A**, The regression coefficients of task event variables (REW, red; FTS, green; RTft, purple; RTst, blue; DIR, black) during the three task periods [FT/cue period (FT), maintenance period of reward prediction (MP), and post-RD (RD)] for FT neurons. **B**, The same as in **A** but for RD neurons. Error bars are SEM across neurons ($n = 22$ in **A** and $n = 11$ in **B**). * $p < 0.01$, ** $p < 0.001$, significance level of the regression coefficients.

trials and commonly present in both the fixation hold error and the successful trials indicates that it may reflect the task motivation to fixate FT in anticipation of FT presentation. Although the task intervals were quasi-randomized, the animal might still be able to anticipate FT onset and be motivated to fixate FT in both the fixation hold error and the successful trials before FT onset but probably failed to do so in the fixation error trials. This view was supported by three types of correlation analysis of the 30 FT neuronal response with task performance.

First, the FT neuronal response during the precue period was significantly correlated with the success (including the fixation hold error) and failure of FT fixation ($r = 0.2$; $p < 0.003$). Second, the FT response in the successful trials was correlated with RTft in a time-dependent manner. This correlation became significant ($p < 0.05$) during the precue period, peaked shortly after the FT presentation, and declined back to baseline during the cue period (Fig. 9C, purple trace). These results are consistent with those for the regression analysis, which showed that the FT neuronal response during the FT/cue period significantly contributed to the RTft (Fig. 7A, purple bars).

Second, the RTft in successful trials exhibited a broad unimodal distribution (-500 to 800 ms), a significant fraction of which was much shorter than the normal reaction time for saccades (100 ms) and may represent anticipation of fixation point appearance. We classified trials according to RTft into short (< 100 ms) and long (≥ 100 ms) RTft categories and also according to the reward magnitude (large vs small reward).

The response correlation for the short RTft category was identical between the large- and small-reward categories (compare blue dotted and solid lines), whereas that for long RTft category was significantly ($p < 0.001$) greater for large- than for small-reward categories (Fig. 9D, green dotted and solid lines), whose timing was almost identical with that for the reward magnitude prediction determined as the steep rise of the reward magnitude correlation (Fig. 9C, black dotted line).

Third, the FT response was also correlated with RTst. The response correlation remained at the baseline level during the precue and cue periods and then became significant ($p < 0.05$) during the maintenance period and returned to the baseline during the post-RD period (Fig. 9C, blue trace). The reward magnitude dependency of the response correlation with RTst was undetectable probably because the correlation was much weaker

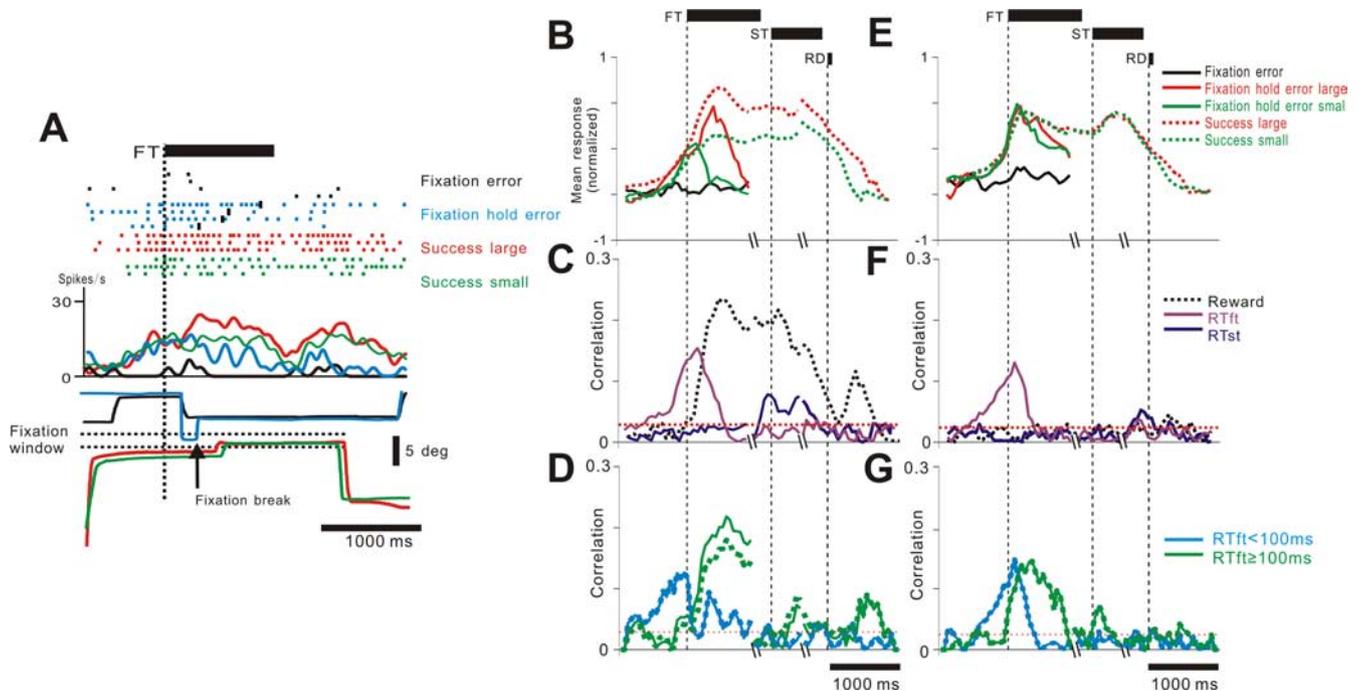


Figure 9. FT neuronal responses in failure and success trials of VGST tasks. **A**, Rastergram and SDF of a representative reward magnitude-dependent FT neuronal response for five successful, fixation hold error, and fixation error trials with normal cue. Black, The fixation error; cyan, fixation hold error; red, successful large-reward trials; green, successful small-reward trials. Bottom, Eye position in single representative case of the four trial categories using the same conventions. Upward arrow indicates the time of the fixation break. Two horizontal dotted lines indicate the fixation window within which the monkey was required to maintain the eye position. **B**, Population SDF of 30 reward-magnitude-dependent FT neurons averaged for 100 fixation error (black solid trace), 236 fixation hold error (solid green and red traces for large and small reward), and 2627 successful trials (dotted red and green dotted traces for large and small reward), aligned to task events including FT, ST, and RD, each of which contained 45, 131, and 1504 normal and 55, 105, and 1123 reversed cue trials, respectively. SDF is the population average normalized for the peaks of the mean individual neuronal responses for each category of success and failure trials, shown synchronized to the times of FT presentation, saccade onset, and RD (left, middle, and right dotted lines), respectively. **C**, Correlation coefficient (absolute value) plot of the FT neuronal responses shown in **B** with RTft (purple), RTst (blue), and reward magnitude (black). The horizontal dotted red line indicates the significance level ($p = 0.05$) of the correlations. **D**, Similar correlation plot to **C** but for four response categories, including those for trials with RTft < 100 ms (cyan) and RTft \geq 100 ms (green) and those for large and small rewards (solid and dotted traces). Number of trials for RTft < 100 ms large- and small-reward and RTft \geq 100 ms and RTft \geq 100 ms large- and small-reward categories, 421, 325, 1005, and 776, respectively. **E–G**, Similar to **B–D** but for 52 reward-magnitude-independent FT responsive neurons (50 FT and 2 FT/RD neurons). The population SDF of the 52 reward-magnitude-independent neurons averaged for 115 fixation error, 325 fixation hold error, and 4470 successful trials, each of which contained 61, 174, and 2573 normal and 54, 151, and 1897 reverse cue trials, respectively. Number of trials for RTft < 100 ms large- and small-reward and RTft \geq 100 ms and RTft \geq 100 ms large- and small-reward categories, 683, 747, 1275, and 1765, respectively.

than that for RTft. These findings indicate that the motivational drive, determined as the FT neuronal response correlation with RTft, includes an early and late component based on the prediction of timing for FT onset and that of the reward magnitude cued by FT, respectively.

The responses of the reward-magnitude-independent FT neurons ($n = 52$; 50 FT and 2 FT/RD neurons) during the precue period was identical to those of the reward-magnitude-dependent FT neurons (Fig. 9B,E, red and green traces), those for the fixation error trials being virtually absent, and those for the fixation hold error and successful trials built up during the precue period (Fig. 9B,E, black, blue, red and green traces). The responses during the cue and maintenance periods in the fixation hold error and successful trials also resembled those for the reward-magnitude-dependent FT neurons, except that they remained the same for large and small rewards. Therefore, we questioned whether the sustained response of the reward-magnitude-independent FT neurons was also correlated with the task performance.

The response of the reward-magnitude-independent neurons during the precue period was correlated with the success/failure of task initiation in an almost the same manner as that for the reward-magnitude-dependent neurons ($r = 0.2$; $p < 0.004$). The reward-magnitude-independent neurons also showed a response correlation to RTft, as did the reward-magnitude-dependent neurons (Fig. 9F, purple and blue trace), but lacked the response

correlation to RTst. Additionally, we conducted similar response correlation analyses using the four response categories as shown in Figure 9E. The RTft (<100 and \geq 100 ms groups) and RTst were not significantly different across the reward magnitude-dependent and -independent neuronal groups ($p > 0.1$, t test). This result indicates that the correlation lacked the reward magnitude dependency, the correlation being identical between the large versus small reward categories for both the short and long RTft categories (Fig. 9G). Altogether, the reward magnitude-independent neurons shared the component of the response correlation related to the prediction of cue onset with the reward-magnitude-dependent neurons but not that related to the cue implication.

These findings indicate that the reward magnitude-independent neurons signal the early component of the motivational drive to fixate FT in an almost equal manner to that for the reward-magnitude-dependent FT neurons but not the late component nor for the saccade to the ST. Similar analysis to that for the FT neurons revealed no response correlation for RD neurons with task performance, including success/failure, RTft or RTst ($p > 0.5$, t test). Additionally, only one of the six reward-magnitude-independent FT neurons that responded transiently to the FT/cue stimulus showed a significant correlation with the success/failure in the two-valued reward VGST trials ($p < 0.05$, t test), which is consistent with the results of our previous study (Kobayashi et al., 2002b).

Discussion

We demonstrated previously that PPTN activity in the fixation period of a simple visually guided saccade task predicted task outcome (Kobayashi et al., 2002b). Using a two-valued reward VGST, we have revealed new functional aspects of PPTN activity. The profile of the activity of FT and RD neurons in this task indicated that these functional neuronal classes may encode the predicted and actual reward magnitude, respectively.

FT and RD neurons responded to the FT/cue stimulus and reward delivery in a manner signaling the predicted and actual reward magnitude.

ROC analysis of the magnitude-dependent FT and RD neuronal responses in our task revealed that most FT and RD neurons reliably signaled whether reward is large or small. Mutual information analysis further showed that FT and RD neuronal signaled reward magnitude with a high precision (the maximum information capacity of 2.6 and 3.5 bits and 0.04 and 0.25 bits per neuron), comparable with those reported for the sensory [0.2 bits/neuron (Gochin et al., 1994)] and motor systems [0.05 bits/neuron (Kitazawa et al., 1998)]. The information capacities of FT and RD neurons imply that they are potentially capable of differentiating 6 and 11 levels of reward magnitude, respectively. Mutual information analysis also showed that FT neurons conveyed information about predicted reward magnitude throughout the cue and maintenance periods, with no significant attenuation until RD neurons signaled actual reward magnitude.

Regression analysis indicated that the reward magnitude contributed most strongly to FT and RD neuronal responses, and the reaction times (both to the fixation and saccade targets, measures of task motivation) contributed only slightly.

Finally, responses of FT neurons responded to changes in the cue–reward contingency within two trials. This finding indicates that FT neurons rapidly revised their prediction of reward magnitude across changes in cue shape. Such rapid revision of reward prediction may be partly attributable to insufficient randomization of the reversal timing, allowing our animals to anticipate contingency reversals. FT neurons remained totally unresponsive for unexpected rewards, and most FT neurons sustained their responses until reward delivery, even when the reward was delayed. These results are consistent with a role of FT neurons in reward prediction.

Conversely, RD neurons responded more whenever the larger rewards were delivered, regardless of cue shape, or whether the reward was delayed or delivered unexpectedly. These results are consistent with RD neurons signaling the magnitude of the delivered reward.

These observations support the view that the FT and RD neurons signal the predicted and actual reward magnitude, respectively. The continuation of the FT neuronal response after the disappearance of the cue until reward delivery indicates that the FT neurons may maintain the signals of the predicted reward from cue presentation until the RD neurons signal the actual reward magnitude. This study revealed that the strong excitatory inputs exerted by the PPTN on midbrain dopamine neurons (Mena-Segovia et al., 2004; Pan and Hyland, 2005; Winn, 2006) convey the memory of the predicted reward and the signals of the actual reward, two essential elements needed for computing the reward prediction error. The high information capacity of the FT and RD neurons to signal the reward magnitude may help accurate computation of the reward prediction error and the efficient execution of reinforcement learning.

Interestingly, the predictive and actual reward responses of

the FT and RD neurons follow comparable time courses with those supposed for the value function and the actual reward signals in the temporal difference (TD) model of reinforcement learning, respectively (Houk et al., 1995; Schultz et al., 1997; Sutton and Barto, 1998; Doya, 2000; Suri, 2002). Therefore, the reward prediction error may be computed in the dopamine neurons from the FT and RD signals, using the TD algorithm (Doya, 2000). It is known from the classical conditioning paradigm of reinforcement learning that dopaminergic neurons show transient excitatory responses to cue presentation but not to reward delivery and inhibitory responses to reward omission at the expected RD timing (Brown et al., 1999; Contreras-Vidal and Schultz, 1999; Doya, 2000; Fiorillo et al., 2008). The FT neuronal response that slowly rises at FT/cue presentation may be conveyed to the dopamine neurons, transformed by temporal differentiation of the TD mechanism as transient excitatory and inhibitory signals timed at FT presentation and reward delivery, respectively, and summed with the actual reward signals of the RD neurons, for computation of reward prediction errors. Those excitatory transients impinge on the dopamine neurons in the absence of RD neuronal signals, producing a sharp cue response, whereas during reward delivery, the inhibitory transients are summed with the excitatory actual reward signals for computation of the reward prediction error, producing no response when the reward prediction matches with the actual one (Tobler et al., 2003; Fiorillo et al., 2008).

The FT responses do not primarily explain the inhibitory omission response of the dopamine neurons, because the response of the majority of FT neurons was shut down at the actual, rather than the expected, RD timing in the temporal reward omission experiments. Therefore, they would feed the inhibitory transients to the dopamine neurons through the TD mechanism, at the actual rather than the expected reward timing. However, the minority of FT neurons whose responses were terminated at the expected RD timing could convey the inhibitory transients to the dopamine neurons, producing the inhibitory omission response. It is possible that the former and latter FT neurons whose response is shut down at the actual and expected reward timing represent the value functions $V(t)$ and $V(t + 1)$ for the current and predicted task events (Houk et al., 1995; Sutton and Barto, 1998; Doya, 2000).

From where do the FT and RD neurons receive the predictive and actual reward signals? The FT neurons may receive the signals of reward prediction from the orbitofrontal cortex (OFC) (Tremblay and Schultz, 1999; Hikosaka and Watanabe, 2000; Roesch and Olson, 2004; Simmons and Richmond, 2008), prefrontal cortex (Leon and Shadlen, 1999; Kobayashi et al., 2002a; Roesch and Olson, 2003), or the striatum (Mena-Segovia et al., 2004; Hikosaka et al., 2006; Winn, 2006). These structures may learn the cue–reward magnitude contingency during the training and task periods as a synaptic memory and recall that memory as the signals of the predicted reward magnitude at the time of cue presentation. These signals would be transferred to the FT neurons and stored as the working memory of the reward prediction until the time of reward delivery. OFC neurons signal the relative, rather than the absolute, value of the cue stimulus (Tremblay and Schultz, 1999). However, this possibility is unlikely for the FT neurons, because in the three-reward valued VGST, we observed that single FT neurons were capable of coding the three-reward magnitudes. The high information capacity of the FT neuronal response in the two-reward valued VGST is also consistent with this view. The RD neurons may receive the actual reward signals

from the lateral hypothalamus (Rolls et al., 1980; Fukuda et al., 1986).

The FT neurons may also signal task motivation, because the reaction times to the fixation and saccade targets were significantly correlated with the success/failure of the task, and the FT neuronal responses for fixation error trials were significantly smaller than those for the successful trials. Additionally, the FT neuronal response during the precue and cue periods was correlated with the RTft and, during the maintenance period of reward prediction, with RTst. Therefore, the reward prediction signaled by the FT neurons may be transformed to the motivational drive to fixate FT and saccade to ST. Supporting this hypothesis, in the cue reversal experiments RTft followed the FT response change with a delay of a few trials, which may imply the time required for transformation of reward prediction to task motivation.

Conversely, correlation analysis of the FT neuronal response with the RTft demonstrated the existence of the early motivational drive to fixate FT that emerged during the precue period and was probably attributable to anticipation of the reward cue. The late motivational drive that emerged during the cue period in parallel with the response correlation with the reward magnitude was probably attributable to the prediction of the reward magnitude. These findings suggest the existence of two mechanisms for the conversion of reward prediction to task motivation: the quick one, whereby the reward prediction is almost immediately converted to the task motivation, i.e., the response correlation with RTft during the precue and cue periods; and the slow one, whereby the reward prediction is gradually converted to the task motivation in several trials, i.e., those found in the cue reversal experiments. The slow mechanism may represent the belief state of the reward prediction controlling the gain whereby the reward prediction is converted to motivation by the early mechanism. At the time of the cue reversal, the reward prediction is revised by the reward prediction error in a few trials and converted to the task motivation through the early mechanism, whereas the belief state remains low and gradually builds up during the next few trials.

We also found significant number of reward magnitude-independent FT neurons that exhibited a significant response correlation with the success/failure of the task as well as RTft during the cue period. The functional implication of the reward-magnitude-independent FT neurons remains unclear, but they may represent the timestamp of the reward expectation (Pan and Hyland, 2005). The neurons responsive to RD also included reward-magnitude-dependent and -independent groups. However, none of these RD neurons showed a response correlation with the success/failure of the task, which is consistent with the view that they monitor the time and magnitude of the actual task reward. Finally, the responses of the reward-magnitude-dependent FT and RD neurons do not purely signal the reward magnitude but partially signal the timestamps of the reward expectation like those found in the reward-magnitude-independent FT and RD neurons, which were reflected by the anticipatory responses preceding the onset of FT and RD.

References

- Blaug CD, Winn P (1993) Modulation of dopamine efflux in the striatum following cholinergic stimulation of the substantia nigra in intact and pedunculopontine tegmental nucleus-lesioned rats. *J Neurosci* 13:1035–1044.
- Bray S, O'Doherty J (2007) Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *J Neurophysiol* 97:3036–3045.
- Brown J, Bullock D, Grossberg S (1999) How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J Neurosci* 19:10502–10511.
- Chiba T, Kayahara T, Nakano K (2001) Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, *Macaca fuscata*. *Brain Res* 888:83–101.
- Contreras-Vidal JL, Schultz W (1999) A predictive reinforcement model of dopamine neurons for learning approach behavior. *J Comput Neurosci* 6:191–214.
- Crist CF, Yamasaki DS, Komatsu H, Wurtz RH (1988) A grid system and a microsyringe for single cell recording. *J Neurosci Methods* 26:117–122.
- Doya K (2000) Reinforcement learning in continuous time and space. *Neural Comput* 12:219–245.
- Fiorillo CD, Newsome WT, Schultz W (2008) The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci* 11:966–973.
- Floresco SB, West AR, Ash B, Moore H, Grace AA (2003) Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci* 6:968–973.
- Fuchs AF, Robinson DA (1966) A method for measuring horizontal and vertical eye movement chronically in the monkey. *J Appl Physiol* 21:1068–1070.
- Fukuda M, Ono T, Nishino H, Nakamura K (1986) Neuronal responses in monkey lateral hypothalamus during operant feeding behavior. *Brain Res Bull* 17:879–883.
- Futami T, Takakusaki K, Kitai ST (1995) Glutamatergic and cholinergic inputs from the pedunculopontine tegmental nucleus to dopamine neurons in the substantia nigra pars compacta. *Neurosci Res* 21:331–342.
- Garcia-Rill E (1991) The pedunculopontine nucleus. *Prog Neurobiol* 36:363–389.
- Gochin PM, Colombo M, Dorfman GA, Gerstein GL, Gross CG (1994) Neural ensemble coding in inferior temporal cortex. *J Neurophysiol* 71:2325–2337.
- Hikosaka K, Watanabe M (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb Cortex* 10:263–271.
- Hikosaka O, Nakamura K, Nakahara H (2006) Basal ganglia orient eyes to reward. *J Neurophysiol* 95:567–584.
- Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia*, pp 249–270. New York: MIT.
- Kitazawa S, Kimura T, Yin PB (1998) Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature* 392:494–497.
- Kobayashi S, Lauwereyns J, Koizumi M, Sakagami M, Hikosaka O (2002a) Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *J Neurophysiol* 87:1488–1498.
- Kobayashi Y, Inoue Y, Yamamoto M, Isa T, Aizawa H (2002b) Contribution of pedunculopontine tegmental nucleus neurons to performance of visually guided saccade tasks in monkeys. *J Neurophysiol* 88:715–731.
- Leon MI, Shadlen MN (1999) Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24:415–425.
- Lusted LB (1978) General problems in medical decision making with comments on ROC analysis. *Semin Nucl Med* 8:299–306.
- Matsumura M, Watanabe K, Ohye C (1997) Single-unit activity in the primate nucleus tegmenti pedunculopontinus related to voluntary arm movement. *Neurosci Res* 28:155–165.
- Mena-Segovia J, Bolam JP, Magill PJ (2004) Pedunculopontine nucleus and basal ganglia: distant relatives or part of the same family? *Trends Neurosci* 27:585–588.
- Mena-Segovia J, Winn P, Bolam JP (2008) Cholinergic modulation of mid-brain dopaminergic systems. *Brain Res Rev* 58:265–271.
- Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopamine neurons can represent context-dependent prediction error. *Neuron* 41:269–280.
- Oakman SA, Faris PL, Kerr PE, Cozzari C, Hartman BK (1995) Distribution of pontomesencephalic cholinergic neurons projecting to substantia nigra differs significantly from those projecting to ventral tegmental area. *J Neurosci* 15:5859–5869.
- Pan WX, Hyland BI (2005) Pedunculopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. *J Neurosci* 25:4725–4732.
- Rezvani S, Corneil BD (2008) Recruitment of a head-turning synergy by

- low-frequency activity in the primate superior colliculus. *J Neurophysiol* 100:397–411.
- 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.
- Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304:307–310.
- Rolls ET, Burton MJ, Mora F (1980) Neurophysiological analysis of brain-stimulation reward in the monkey. *Brain Res* 194:339–357.
- Rolls ET, McCabe C, Redoute J (2008) Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb Cortex* 18:652–663.
- Scarnati E, Campana E, Pacitti C (1984) Pedunculopontine-evoked excitation of substantia nigra neurons in the rat. *Brain Res* 304:351–361.
- Schreiner RC, Essick GK, Whitsel BL (1978) Variability in somatosensory cortical neuron discharge: effects on capacity to signal different stimulus conditions using a mean rate code. *J Neurophysiol* 41:338–349.
- Schultz W (2002) Getting formal with dopamine and reward. *Neuron* 36:241–263.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Semba K, Fibiger HC (1992) Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: a retro- and anterograde transport and immunohistochemical study. *J Comp Neurol* 323:387–410.
- Simmons JM, Richmond BJ (2008) Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. *Cereb Cortex* 18:93–103.
- Suri RE (2002) TD models of reward predictive responses in dopamine neurons. *Neural Netw* 15:523–533.
- Sutton RS, Barto AG (1998) Reinforcement learning. New York: MIT.
- Tobler PN, Dickinson A, Schultz W (2003) Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J Neurosci* 23:10402–10410.
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.
- Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48.
- Watanabe K, Hikosaka O (2005) Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. *J Neurophysiol* 94:1879–1887.
- Werner G, Mountcastle VB (1963) The variability of central neural activity in a sensory system, and its implications for the central reflection of sensory events. *J Neurophysiol* 26:958–977.
- Winn P (2006) How best to consider the structure and function of the pedunculopontine tegmental nucleus: evidence from animal studies. *J Neurol Sci* 248:234–250.