

# 博士論文

Primate midbrain dopamine neurons  
monitor reward values changing from  
moment to moment

(ドーパミンニューロンの持続的活動は連  
続的に変動する報酬価値をコードする)

令和3年度

筑波大学大学院人間総合科学研究科  
感性認知脳科学専攻

王亜偉

筑波大学

# Table of contents

1. Abstract .....	1
2. Introduction.....	2
2.1. Reward system .....	2
2.2. Midbrain dopamine neurons and their phasic activity .....	5
2.3. Gap between previous studies on dopamine neuron signals and real-life situation .....	6
2.4. Objective.....	8
3. Materials and Methods .....	9
3.1. Animals .....	9
3.2. Behavioral tasks.....	9
3.3. Single-unit recording .....	14
3.4. Localization and identification of dopamine neurons .....	15
3.5. Data analysis.....	17
4. Results.....	22
4.1. Pavlovian procedure in which the value of a cued reward gradually changed . .....	22
4.2. Tonic activity of dopamine neurons tracking gradually changing reward values .....	23
4.3. Phasic activity of dopamine neurons evoked when the reward value stopped changing.....	40
4.4. Relationship between the tonic and phasic dopamine neuron activities.....	44
4.5. Burst and non-burst spike firing modes of dopamine neurons .....	48
5. Discussion.....	52
5.1. What variable does the tonic activity of dopamine neurons represent?.....	52
5.2. Comparison between the present findings and previous reports on tonic dopamine signals .....	55

5.3. What roles does the tonic activity of dopamine neurons play in behavior?...	57
5.4. Conclusion .....	58
6. Acknowledgments .....	59
7. References.....	61

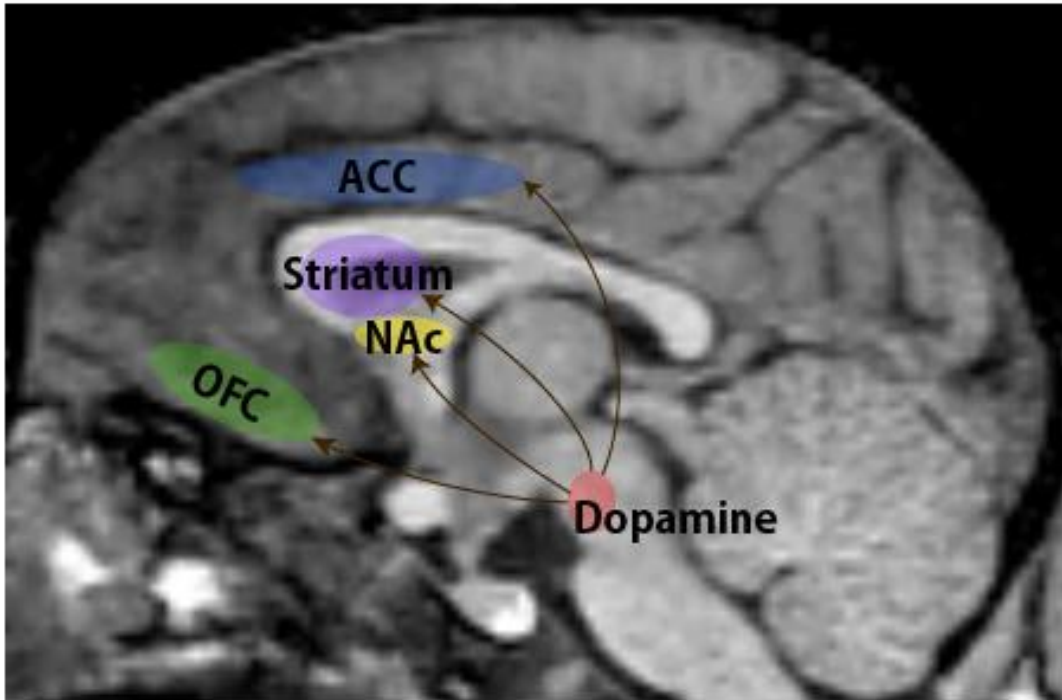
## 1. Abstract

Animal behavior is regulated based on the values of future rewards. The phasic activity of midbrain dopamine neurons signals these values. Because reward values often change over time, even on a subsecond-by-subsecond basis, appropriate behavioral regulation requires continuous value monitoring. However, the phasic dopamine activity, which is sporadic and has a short duration, likely fails continuous monitoring. Here, I demonstrate a tonic firing mode of dopamine neurons that effectively tracks changing reward values. I recorded dopamine neuron activity in monkeys during a Pavlovian procedure in which the value of a cued reward gradually increased or decreased. Dopamine neurons tonically increased and decreased their activity as the reward value changed. This tonic activity was evoked more strongly by non-burst spikes than burst spikes producing a conventional phasic activity. My findings suggest that dopamine neurons change their firing mode to effectively signal reward values in a given situation.

## 2. Introduction

### 2.1. Reward system

Rewards are vital for survival and strongly influence animal behavior such as learning and motivation. How the brain processes reward information has been a fundamental issue in systems neuroscience. For example, intracranial self-stimulation experiments have identified brain regions of which electrical stimulation evokes approach behavior (Heath, 1963; Olds and Milner, 1954). In these studies, electrodes were implanted into given brain regions, and electrical stimulation was delivered when animals took a particular action (e.g., pressing a lever). The animals repeatedly took the actions if the electrical stimulation was delivered into particular brain regions. These brain regions are regarded as being included in the brain's "reward system", which extends across cortical and subcortical regions such as the orbitofrontal cortex (OFC), anterior cingulate cortex (ACC), striatum, nucleus accumbens (NAc), midbrain dopamine regions (substantia nigra pars compacta (SNc) and ventral tegmental area (VTA)) and so on.



**Figure 1. Brain reward system.**

The reward system includes the orbitofrontal cortex (OFC), anterior cingulate cortex (ACC), striatum, nucleus accumbens (NAc), midbrain dopamine regions (substantia nigra pars compacta (SNc) and ventral tegmental area (VTA)) and so on.

The OFC is a ventral part of the prefrontal cortex, and neurons in this region are known to encode the value of rewards (Hikosaka and Watanabe, 2000; 2004; Morrison and Salzman, 2009; Tremblay and Schultz, 1999). This region is also known to play key roles in choosing one among multiple options based on their values (i.e., value-based decision-making) by signaling the values of options and updating

stimulus–reward associations (Padoa-Schioppa and Assad, 2006; 2008; Padoa-Schioppa, 2013).

The ACC is another key prefrontal subdivision of the reward system, which lies on the medial surface of the prefrontal cortex. This region is known as a crucial hub between the reward system and other systems including the executive and motor systems. Neurons in the ACC, as well as OFC neurons, encode the value of rewards. A critical feature of the ACC value signal is that this signal not only represents the value of upcoming and ongoing rewards, but also reflects the value of rewards that have been obtained in the past (Kawai et al., 2015; Kolling et al., 2016; Seo and Lee, 2007; Shima and Tanji, 1998). Therefore, the ACC is thought to regulate animal behavior based on past experiences.

The striatum is a subcortical structure, which is an essential component of the cortical-basal ganglia loop that regulates multiple behavioral processes such as executive, motivational, motor and oculomotor functions (Alexander et al., 1986). Neurons in the striatum encode the value of rewards (Apicella et al. 1991; Bowman et al. 1996), and are thought to modulate the behavioral processes by signaling the reward value information (Hikosaka, 2007; Lau and Glimcher, 2008; Nakamura et

al., 2012; Samejima et al., 2005).

Midbrain dopamine neurons are a key component of the reward system. These neurons widely project to other regions of the reward system including cortical (Bjorklund and Dunnett, 2007; Takada and Hattori, 1987; Williams and Goldman-Rakic, 1998) and subcortical structures (Bjorklund and Dunnett, 2007; Francois et al., 1999; Gauthier et al., 1999; Lynd-Balta and Haber, 1994), and are well known for their striking responses to rewards and reward predicting cues. The next section summarizes studies on unique roles of dopamine neurons in processing reward information.

## **2.2. Midbrain dopamine neurons and their phasic activity**

Midbrain dopamine neurons respond to rewards and reward predicting cues with a phasic increase or decrease in activity (Bayer and Glimcher, 2005; Cohen et al., 2012; Kawagoe et al., 2004; Morris et al., 2004; Satoh et al., 2003; Schultz, 1998). These neurons exhibit a phasic activity increase if the rewarding events are better than predicted, while they exhibit a phasic activity decrease if the rewarding events are worse than predicted. An influential theory proposed that such phasic activity of



dopamine neurons encode “reward prediction errors”, which indicate discrepancies between obtained and predicted reward values, and that the dopamine prediction error signals play a crucial role in reinforcement learning as a teaching signal (Doya, 2002; Montague et al., 1996; Schultz et al., 1997). Thus, positive dopamine prediction error signals (i.e., phasic dopamine activity increase) are thought to reinforce actions that lead to better-than-predicted rewards, while negative dopamine prediction error signals (i.e., phasic dopamine activity decrease) are thought to suppress actions that lead to worse-than-predicted rewards. Indeed, the reward-evoked phasic activity of dopamine neurons has been shown to regulate this type of learning (Chang et al., 2016; Stauffer et al., 2016; Steinberg et al., 2013; Tsai et al., 2009). In addition, the cue-evoked phasic activity of dopamine neurons has been reported to reflect the value of cued rewards (Matsumoto and Hikosaka, 2009; Roesch et al., 2007; Tobler et al., 2005), and has recently been demonstrated to be involved in reward-seeking behavior associated with cued rewards (Maes et al., 2020; Morrens et al., 2020).

### **2.3. Gap between previous studies on dopamine neuron signals and real-life situation**

Although, as described above, a wealth of studies have revealed roles of the

phasic dopamine neuron activity in reward processing (i.e., value coding and reward prediction error signaling), it is unclear whether the same principle is applicable to real-life situations. You often experience situations in which the value of future rewards changes on a moment-by-moment basis. For example, as a bartender pours a cocktail into your glass, the amount (i.e., the value) of the cocktail that you will receive gradually increases. The phasic activity of dopamine neurons, which is sporadic and has a short duration (around a couple of hundred milliseconds), likely fails to continuously monitor these changing values. Therefore, although dopamine neurons have been shown to regulate reward-seeking behavior by signaling reward values (Maes et al., 2020; Morrens et al., 2020), the phasic activity seems to be unable to properly regulate reward-seeking behavior in changing environments in which reward values fluctuate.

I presume that a possible way for dopamine neurons to signal changing reward values is to monitor changing values by tonically increasing and decreasing their activity in parallel with the value change. Indeed, several previous studies have reported that dopamine neurons exhibit tonic activity changes in multiple behavioral situations (Bromberg-Martin et al., 2010; Engelhard et al., 2019; Howard et al., 2017; Hughes, et al., 2020). For example, the tonic activity of nigrostriatal dopamine

neurons is associated with internal action selection and necessary for appropriate choice behavior (Howard et al., 2017). A tonic increase in dopamine activity is also observed while animals move towards long-term goals, which provides sustained motivational drive (Engelhard et al., 2019). However, these studies did not show whether the tonic activity of dopamine neurons can encode the value of rewards as their phasic activity does.

#### **2.4. Objective**

To address whether and how dopamine neurons signal changing reward values, I designed a Pavlovian procedure in which a visual cue indicated the value (i.e., the amount) of a liquid reward that gradually increased or decreased moment-by-moment. I recorded single-unit activity from dopamine neurons while macaque monkeys were conditioned with this Pavlovian procedure. Here, I demonstrate a tonic firing mode of dopamine neurons that effectively tracks changing reward values.

### **3. Materials and Methods**

#### **3.1. Animals**

Two adult macaque monkeys (*Macaca fuscata*; monkey H, female, 6.6 kg, 8 years old; monkey P, female, 6.8 kg, 7 years old) were used for the experiments. All procedures for animal care and experimentation were approved by the University of Tsukuba Animal Experiment Committee (permission number, 14-137), and were carried out in accordance with the guidelines described in Guide for the Care and Use of Laboratory Animals published by the Institute for Laboratory Animal Research.

#### **3.2. Behavioral tasks**

Behavioral tasks were controlled using MATLAB (Mathworks, MA, USA) with Psychtoolbox, a freely available toolbox for MATLAB. The monkeys sat in a primate chair facing a computer monitor in a sound-attenuated and electrically shielded room. Eye movements were monitored using an infrared eye-tracking system (EYE-TRAC 6; Applied Science Laboratories, MA, USA) with sampling at 240 Hz. A liquid reward was delivered through a spout that was positioned in front of the monkey's mouth. Licking of the monkeys was monitored during the recording of 68 of the 99

recorded dopamine neurons. To monitor licking, a strain gauge was attached to the spout and measured the strain of the spout caused by licking.

I designed a Pavlovian procedure with three different conditions (Figure 2A-C) and applied it to the two monkeys. In each condition, trials started with the presentation of a central fixation point ( $0.8^\circ \times 0.8^\circ$ ). The monkey was required to fixate this point. After a 400-ms fixation period, a bar stimulus ( $2.9^\circ \times 9.4^\circ$ ) with red and green areas was presented as a conditioned stimulus (CS). The size of the green area indicated the amount (i.e., the value) of a liquid reward that the monkey would receive. The larger the green area became, the larger the associated reward value. In the first condition (value-increase condition, Figure 2A), the green area was minimal at the beginning and gradually increased ( $3.8^\circ/\text{s}$ ), i.e., the reward value gradually increased ( $0.082 \text{ ml/s}$ ). The gradual increase randomly stopped within 2450 ms after the onset of the bar stimulus (uniform distribution from 0 to 2450 ms) so that the monkey was unable to precisely predict the reward value until the green area stopped increasing. In the second condition (value-decrease condition, Figure 2B), the green area was maximal at the beginning and gradually decreased ( $3.8^\circ/\text{s}$ ), i.e., the reward value gradually decreased ( $0.082 \text{ ml/s}$ ). The gradual decrease randomly stopped within 2450 ms after the bar onset (uniform distribution from 0

to 2450 ms). In the third condition (value-fixed condition, Figure 2C), the size of the green area did not change and was instead fixed at the minimum, half, or maximum that predicted a 0.1-ml, 0.2-ml, or 0.3-ml reward, respectively (these reward amounts were the same as those indicated by the green area in the value-increase and value-decrease conditions). In addition to these CSs, a CS constituting of vertical red and green bars was included, and this was randomly followed by a 0.1- to 0.3-ml reward (i.e., the reward value was uncertain). These four CSs were presented with an equal probability (25% each). In each condition, the total time during which the CS was presented was fixed at 2850 ms, and the monkey was required to maintain the central fixation during this period. Immediately after the offset of the bar stimulus, the reward indicated by the green area was delivered simultaneously with a tone (1000 Hz). The delay between the CS onset and the reward delivery did not change across trials. Each condition consisted of a block of 50 trials, and I collected data by repeating the three conditions once or more for each neuron (Figure 2D). The order of the three conditions, 1) the value-fixed, 2) value-increase, and 3) value-decrease conditions, was fixed across recording sessions. The total amount of reward was the same (10 ml) among blocks. Trials were aborted immediately if the monkey 1) did not start the central fixation within 4000 ms after the onset of the fixation point, or

2) broke the central fixation during the initial 400-ms fixation period or the 2850-ms CS period (i.e., a continuous 3250-ms fixation was required). These error trials were signaled by a beep tone (100 Hz) and excluded from analyses.

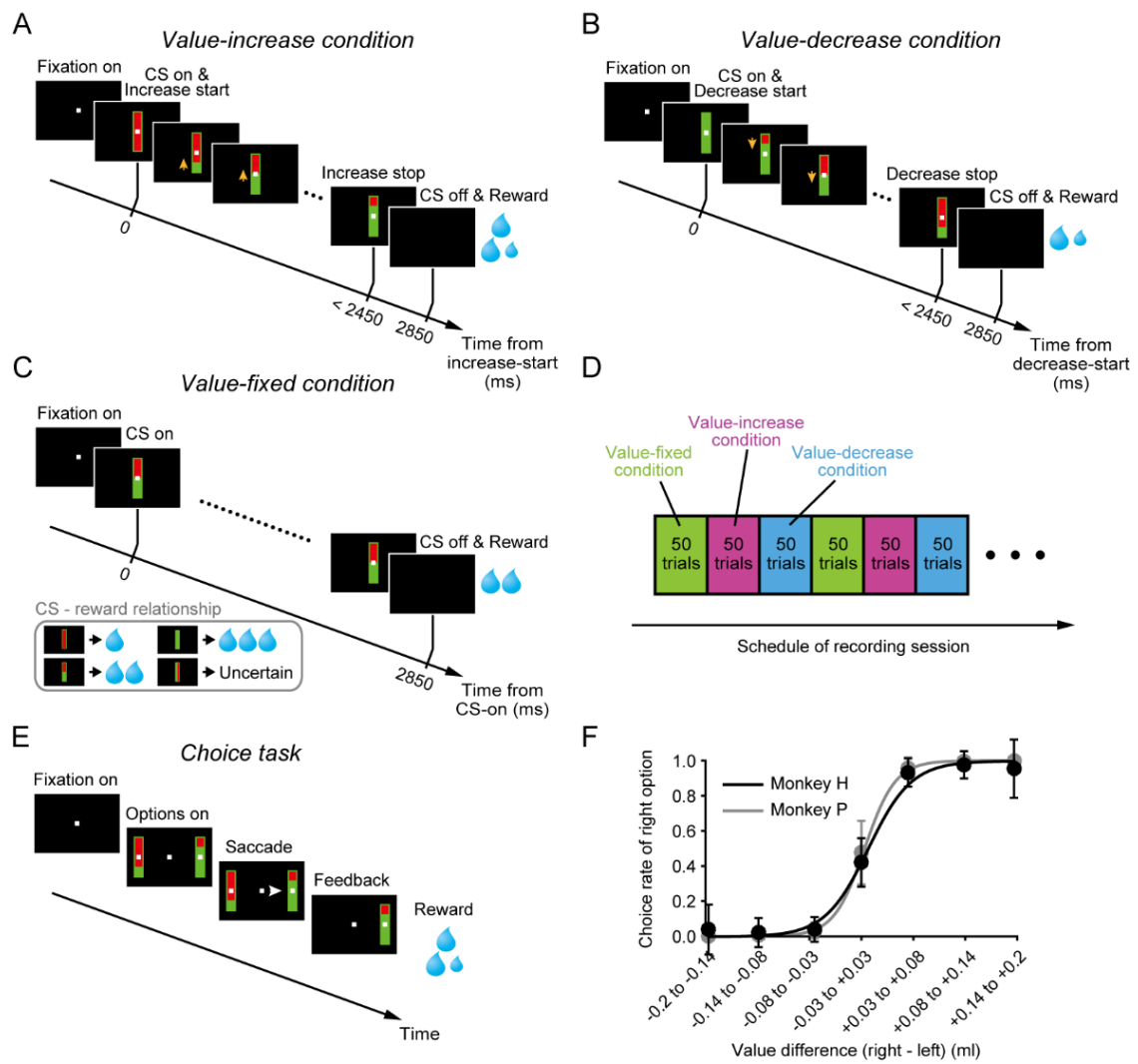


Figure 2. Pavlovian procedure in which the value of a cued reward gradually changed.

(A-C) Value-increase (A), value-decrease (B) and value-fixed (C) conditions in the Pavlovian procedure. (D) Schedule of recording session. Each condition consisted of a block of 50 trials and was repeated once or more. (E) Choice task. (F) Choice rate of the right bar stimulus as a function of the value difference between the right and left bar stimuli in monkey H (n = 85 sessions) (black) and monkey P (n = 68 sessions) (gray). Error bars indicate SEM.

I also used a choice task (Figure 2E). Each trial started with the presentation of a central fixation point ( $0.8^\circ \times 0.8^\circ$ ), and the monkey was required to fixate this point. After a 400-ms fixation period, two bar stimuli with red and green areas ( $2.9^\circ \times 9.4^\circ$ ) were presented on the right and left sides of the fixation point (eccentricity:  $8.8^\circ$ ). The monkey was required to choose one of the bar stimuli by making a saccade immediately after the presentation of the bar stimuli. These bar stimuli were identical to those used in the Pavlovian procedure except that the green area did not increase or decrease over time. The size of the green area indicated the value of the reward that the monkey would obtain by choosing that bar stimulus, and was randomly assigned to each bar stimulus (uniform distribution from 0.1 to 0.3 ml). Immediately after the monkey chose a bar stimulus, the other bar stimulus disappeared. Then, the reward indicated by the green area of the chosen bar stimulus was delivered simultaneously with a tone (1000 Hz). Trials were aborted immediately if the monkey 1) did not start the central fixation within 4000 ms after the onset of the fixation point, or 2) broke the central fixation during the 400-ms



fixation period. These error trials were signaled by a beep tone (100 Hz) and excluded from analyses.

### **3.3. Single-unit recording**

A plastic head holder and recording chamber were fixed to the skull under general anesthesia and sterile surgical conditions. The recording chamber was placed over the midline of the frontoparietal lobes, and aimed at the SNc and VTA in both hemispheres. The head holder and recording chamber were embedded in dental acrylic that covered the top of the skull and were firmly anchored to the skull by plastic screws.

Single-unit recordings were performed using tungsten electrodes with impedances of approximately 3.0 M $\Omega$  (Frederick Haer, ME, USA) that were introduced into the brain through a stainless-steel guide tube using an oil-driven micromanipulator (MO-97-S; Narishige, Tokyo, Japan). Recording sites were determined using a grid system, which allowed recordings at every 1 mm between penetrations. For a finer mapping of neurons, I also used a complementary grid that allowed electrode penetrations between the holes of the original grid.

Electrophysiological signals were amplified, band-pass filtered (200 Hz to 3 kHz; RZ5D, Tucker-Davis Technologies, FL, USA), and stored in a computer at the sampling rate of 24.4 kHz. Single-unit potentials were isolated using a window discrimination software (OpenEx, Tucker-Davis Technologies, FL, USA).

### 3.4. Localization and identification of dopamine neurons

I recorded single-unit activity from putative dopamine neurons in the SNc and VTA. To localize the recording regions, the monkeys underwent a magnetic resonance imaging (MRI) scan to determine the positions of the SNc and VTA (Figure 3).

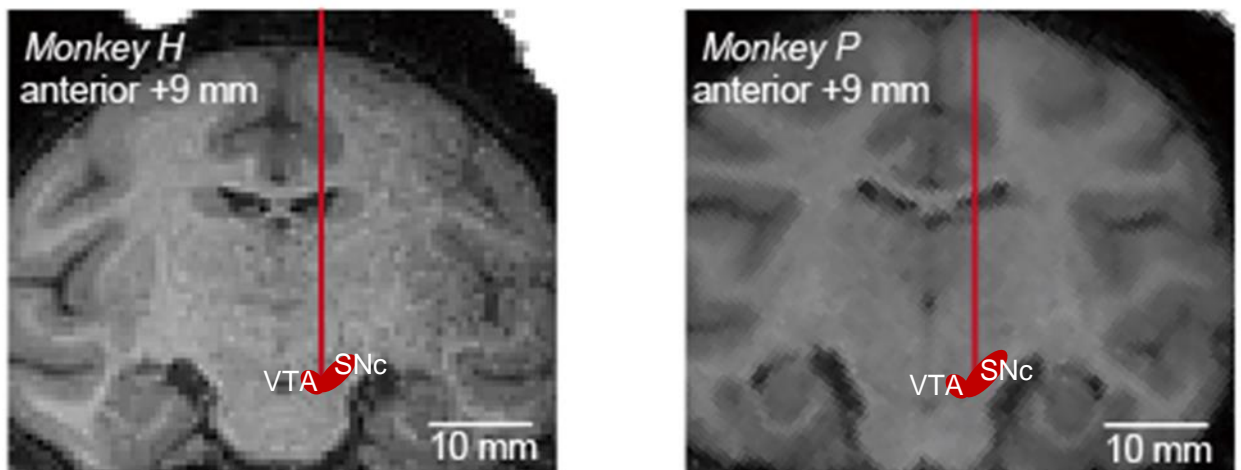


Figure 3. Localization of recording sites

Images obtained by MRI scans for monkey H (left) and monkey P (right) are shown. Vertical red lines indicate representative recording tracks. The approximate anteroposterior distance (millimeters) from the interaural line is shown in the top-left corner of each panel.

Putative dopamine neurons were identified based on their well-established electrophysiological signatures: a low background firing rate at around 5 Hz, a broad spike potential in clear contrast to neighboring neurons with a high background firing rate in the substantia nigra pars reticulata (Figure 4), and a phasic excitation in response to free reward.

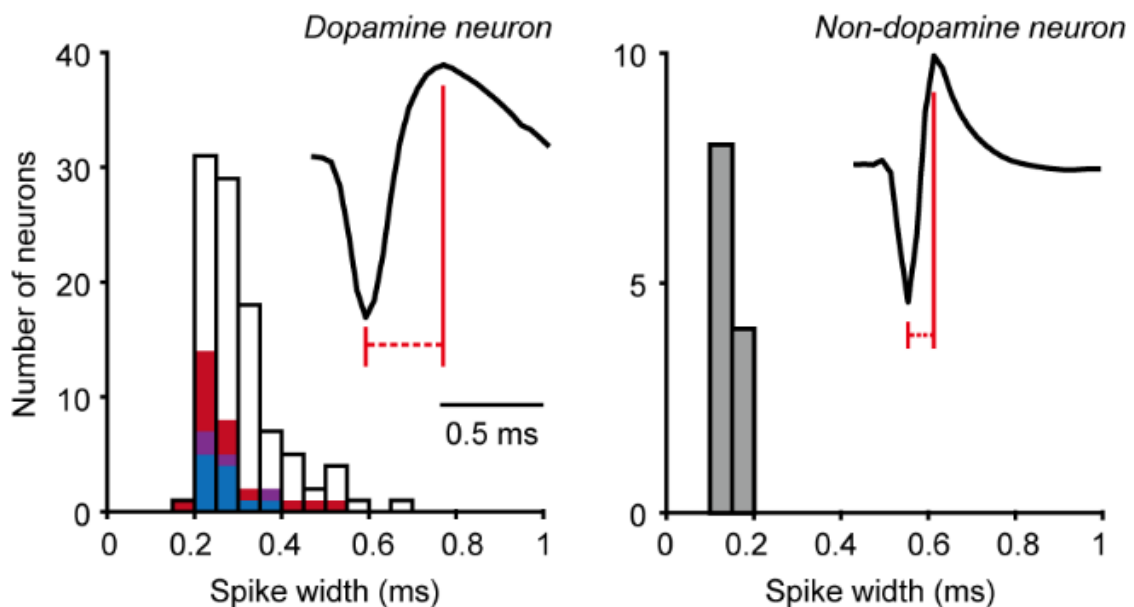


Figure 4. Spike waveforms of dopamine neurons.

Distributions of the spike widths of dopamine neurons (left,  $n = 99$ ) and non-dopamine neurons in the substantia nigra pars reticulata (right,  $n = 12$ ). Red bars indicate dopamine neurons with a significantly positive regression slope between firing rate and time in the value-increase condition, while blue bars indicate dopamine neurons with a significantly negative regression slope in the value-decrease condition ( $p < 0.05$ ; linear regression analysis). Purple bars indicate dopamine neurons with both of them, and open bars indicate dopamine neurons with neither of them. The spike widths of the dopamine neurons with a significantly positive regression slope in the value-increase condition, those with a significantly negative regression slope in the value-decrease condition, and those with neither of them were not significantly different (positive versus negative:  $p = 0.77$ ; positive versus no significance:  $p = 0.051$ ; negative versus no significance:  $p = 0.058$ ; Wilcoxon rank-sum test). Spike shapes of example dopamine and non-dopamine neurons are shown in the top-right corner of each panel. The two vertical lines show how the spike duration was measured.

### 3.5. Data analysis

For null hypothesis testing, 95% confidence intervals ( $p < 0.05$ ) were used to define statistical significance in all analyses.

To examine whether the monkeys accurately predicted the reward values according to the size of the green area (Figure 2F), the choice rate of the right bar stimulus was fitted by the following logistic function:

$$P = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 \times (V_{\text{right}} - V_{\text{left}})))}$$

where  $P$  indicates the choice rate of the right bar stimulus,  $V_{\text{right}}$  and  $V_{\text{left}}$  indicate the reward values obtained by choosing the right and left bar stimuli, respectively, and  $\beta_0$  and  $\beta_1$  indicate the coefficients determined by logistic regression.

To calculate spike density functions (SDFs), each spike was replaced by a Gaussian curve ( $\sigma = 15$  ms).

To calculate SDFs aligned by the CS onset in the value-increase and value-decrease conditions (Figure 5A, B, E, F, and Figure 13A), spikes that occurred before the reward value stopped increasing or decreasing were used.

To calculate averaged SDFs across neurons (Figure 5E, F, Figure 11C and Figure 13A, B), the baseline firing rate of each neuron was measured using a window from 500 to 0 ms before the fixation point onset, and was subtracted from the original firing rate of each neuron.

To quantify tonic activity changes in the value-increase and value-decrease conditions (Figure 5C, D), the slope of the regression line between firing rate and time was calculated for each neuron. First, a window from 650 to 2450 ms after the

CS onset was divided into 200-ms bins, and the firing rate in each bin was measured. Then, the regression line between the firing rate in each bin and the time at the center of each bin was calculated. The calculation window from 650 to 2450 ms after the CS onset was determined to exclude the effect of phasic responses evoked immediately after the CS onset.

To examine phasic activity modulations evoked by the reward value in the value-fixed condition, the regression coefficient between firing rate and reward value was calculated for each neuron. The firing rate was measured using a window from 100 to 400 ms after the CS onset. The calculation window was determined such that the window includes the major part of the neural modulation in the averaged activity.

To calculate SDFs aligned by the onset of each saccade (Figure 10B, F), the onset was determined as the time when angular eye velocity exceeded 40 °/s.

To examine phasic activity modulation evoked when the reward value stopped increasing or decreasing in the value-increase and value-decrease conditions (Figure 11B, D), the regression coefficient between firing rate and reward value was calculated for each neuron. First, all trials were divided into three groups based on the reward value (large: 0.23 to 0.3 ml, medium: 0.16 to 0.23 ml, small: 0.1 to 0.16

ml), and the firing rate was measured using a window from 150 to 500 ms after the stop onset for each trial group. Then, the regression coefficient between the firing rate and the mean reward value in each trial group was calculated. The calculation window was determined such that the window includes the major part of the neural modulation in the averaged activity.

To statistically test the difference in firing rate (original firing rate – baseline firing rate) between burst and non-burst spike firing (left in Figure 13A), a bootstrap procedure was applied. I first divided the calculation time window (650 to 2450 ms after the CS onset), which I used to detect the significantly positive regression slope of the 19 dopamine neurons (horizontal gray bar in the left column of Figure 5E), into initial (650 to 1250 ms), middle (1250 to 1850 ms), and late (1850 to 2450 ms) periods. The 19 dopamine neurons were randomly resampled with replacements to form a new bootstrap dataset that had the same number of neurons as the original dataset. Using the new dataset, I compared the firing rate (original firing rate – baseline firing rate) between burst and non-burst spike firing for the initial, middle, and late periods. This random resampling and comparison process was repeated 1000 times. If the firing rate was larger for burst spike firing than for non-burst spike firing or vice versa in more than 975 repetitions, the difference in firing rate

was regarded as significant ( $p < 0.05$ ; bootstrap test with 1000 repetitions).



## 4. Results

### 4.1. Pavlovian procedure in which the value of a cued reward gradually changed

I designed a Pavlovian procedure with three different conditions (Figure 2A-C) and applied it to two macaque monkeys (monkey H and monkey P). In each condition, a bar stimulus with red and green areas was presented as a conditioned stimulus (CS) after a fixation period. The size of the green area indicated the amount (i.e., the value) of a liquid reward that the monkey would receive. In the first condition (value-increase condition, Figure 2A), the green area (i.e., the reward value) gradually increased (0.082 ml/s). In the second condition (value-decrease condition, Figure 2B), the green area (i.e., the reward value) gradually decreased (0.082 ml/s). In the third condition (value-fixed condition, Figure 2C), the size of the green area did not change and was instead fixed. Each condition consisted of a block of 50 trials and was repeated once or more for each recording session (Figure 2D).

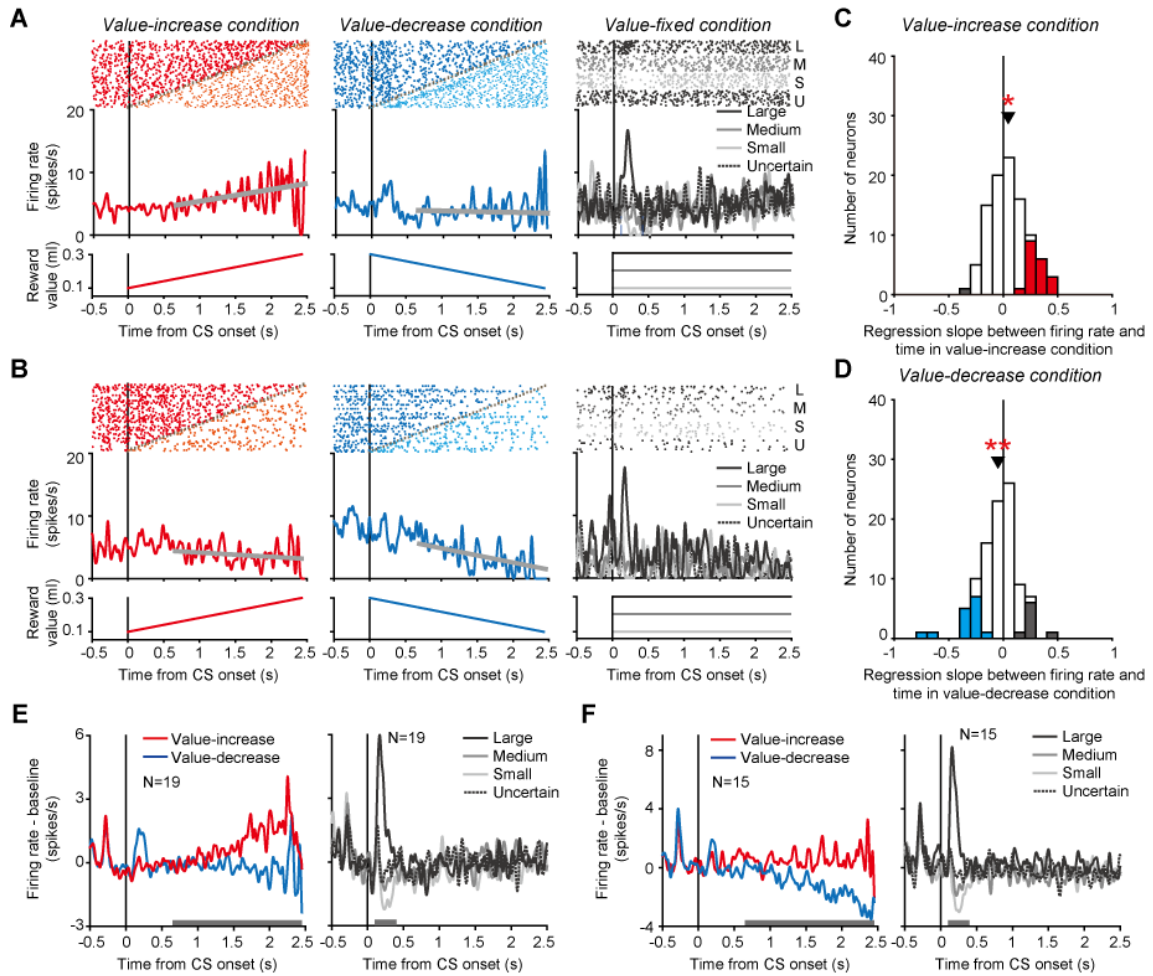
To examine whether the monkeys accurately predicted the reward values according to the size of the green area, I applied a choice task in which two bar stimuli with red and green areas were presented on the right and left sides of a fixation point (Figure 2E). The monkeys were required to choose one of the bar stimuli by making a saccade. I observed that the monkeys were more likely to choose

the bar stimulus with the larger green area, i.e., the logistic regression slope between the choice rate of the right bar stimulus and the value difference (right bar's value – left bar's value) was significantly larger than zero (monkey H:  $n = 85$  sessions, mean  $\pm$  SD =  $0.94 \pm 1.64$ ,  $p = 1.17 \times 10^{-15}$ ; monkey P:  $n = 68$  sessions, mean  $\pm$  SD =  $0.98 \pm 1.65$ ,  $p = 7.64 \times 10^{-13}$ ; Wilcoxon signed-rank test) (Figure 2F). This suggests that the monkeys correctly predicted the reward values according to the size of the green area.

#### **4.2. Tonic activity of dopamine neurons tracking gradually changing reward values**

While the monkeys participated in the Pavlovian procedure, I recorded single-unit activity from 99 dopamine neurons (monkey H,  $n = 46$ ; monkey P,  $n = 53$ ) in and around the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA) (Figure 3, also see Figure 4 for the characteristic spike waveforms of the recorded dopamine neurons). Here, I presumed that a possible way for dopamine neurons to signal changing reward values is to track the changing values by tonically increasing and decreasing their activity in parallel with the value change. Indeed, an example dopamine neuron tonically increased its activity as the reward value gradually increased in the value-increase condition (left in Figure 5A). This tonic

increase in activity was clearly in contrast with the phasic responses to the CSs in the value-fixed condition, in which the reward value did not change (right in Figure 5A). This neuron did not exhibit clear activity modulation as the reward value gradually decreased in the value-decrease condition (middle in Figure 5A). Another example dopamine neuron tonically decreased its activity in the value-decrease condition (middle in Figure 5B), but it did not exhibit clear activity modulation in the value-increase condition (left in Figure 5B). This neuron also displayed phasic responses to the CSs in the value-fixed condition (right in Figure 5B).



**Figure 5. Tonic activity of dopamine neurons tracking gradually changing reward values.**

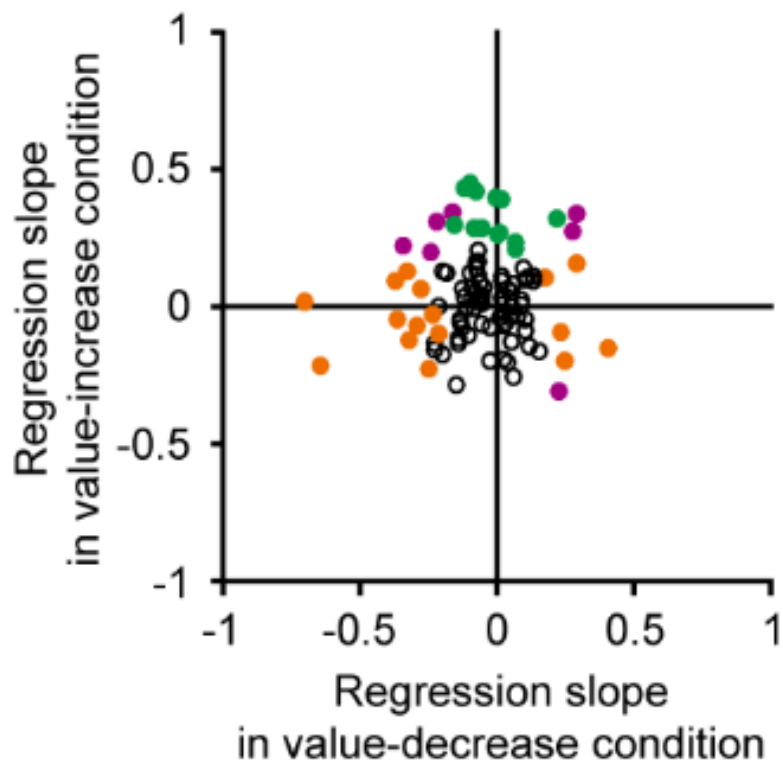
(A and B) Activity of two example dopamine neurons in the value-increase (left), value-decrease (middle) and value-fixed (right) conditions. Rasters and spike density functions (SDFs) are aligned by the CS onset. Left: Red and orange rasters indicate spikes occurring before and after, respectively, the reward value stopped increasing. Gray plots indicate the time at which the reward value stopped changing. The spikes occurring before the stop were used to calculate the SDF. The thick gray line indicates the regression line between firing rate

and time. The bottom graph indicates the reward value as a function of time. Middle: Blue and cyan rasters indicate spikes occurring before and after, respectively, the reward value stopped decreasing. Other conventions are as in the left panel. Right: Black (top), dark gray, light gray, and black (bottom) rasters indicate spikes in trials in which the reward value was large, medium, small, and uncertain, respectively. Black (solid), dark gray, light gray, and black (dotted) curves indicate the SDFs in these trials. (C and D) Regression slope between firing rate and time in the value-increase (C) and value-decrease (D) conditions ( $n = 99$  neurons). Red and gray bars indicate neurons with significantly positive and negative slopes, respectively, in the value-increase condition (C), while blue and gray bars indicate neurons with a significantly negative and positive slope, respectively, in the value-decrease condition (D) ( $p < 0.05$ ; linear regression analysis). Arrowheads indicate the mean regression slopes. Single and double asterisks indicate a significant deviation from zero ( $p < 0.05$  and  $p < 0.01$ , respectively; Wilcoxon signed-rank test). (E and F) Averaged SDFs of the 19 dopamine neurons with a significantly positive regression slope in the value-increase condition (E) and those of the 15 dopamine neurons with a significantly negative regression slope in the value-decrease condition (F). Left: Red and blue curves indicate the SDFs in the value-increase and value-decrease conditions, respectively. Right: Black (solid), dark gray, light gray, and black (dotted) curves indicate the SDFs in trials in which the reward value was large, medium, small, and uncertain, respectively. Horizontal gray bars indicate the time windows used for analyses.

To quantify tonic activity changes in the value-increase and value-decrease conditions, I calculated the slope of the regression line between firing rate and time

for each neuron (thick gray lines in the left and middle columns of Figure 5A, B). In the value-increase condition, 19 of the 99 dopamine neurons exhibited a significantly positive slope ( $p < 0.05$ ; linear regression analysis), and, on average, the regression slope was significantly larger than zero ( $n = 99$  neurons,  $\text{mean} \pm \text{SD} = 0.046 \pm 0.17$ ,  $p = 0.029$ ; Wilcoxon signed-rank test) (Figure 5C). The averaged activity of the 19 dopamine neurons exhibited a clear tonic increase as the reward value gradually increased (left in Figure 5E), and displayed value-representing phasic responses to the CSs in the value-fixed condition (significantly positive regression coefficient between firing rate and value:  $n = 19$  neurons,  $\text{mean} \pm \text{SD} = 1.61 \pm 1.42$ ,  $p = 6.25 \times 10^{-4}$ ; Wilcoxon signed-rank test) (right in Figure 5E). In the value-decrease condition, 15 of the 99 dopamine neurons exhibited a significantly negative slope ( $p < 0.05$ ; linear regression analysis), and, on average, the regression slope was significantly smaller than zero ( $n = 99$  neurons,  $\text{mean} \pm \text{SD} = -0.048 \pm 0.18$ ,  $p = 0.01$ ; Wilcoxon signed-rank test) (Figure 5D). The averaged activity of the 15 dopamine neurons exhibited a clear tonic decrease as the reward value gradually decreased (left in Figure 5F), and displayed value-representing phasic responses to the CSs in the value-fixed condition (significantly positive regression coefficient between firing rate and value:  $n = 15$  neurons,  $\text{mean} \pm \text{SD} = 2.10 \pm 0.97$ ,  $p = 6.10 \times 10^{-5}$ ; Wilcoxon signed-

rank test) (right in Figure 5F). Notably, single dopamine neurons rarely exhibited both tonic increase and decrease in activity in the value-increase and value-decrease conditions, respectively (no significant correlation between the regression slopes in the value-increase and value-decrease conditions:  $n = 99$  neurons,  $r = 0.048$ ,  $p = 0.64$ ; correlation analysis) (Figure 6). These results suggest that a subset of dopamine neurons monitored the gradually increasing and decreasing reward values by tonically changing their activity, although the neurons did not signal these values in a unified manner.



**Figure 6. Relationship between the regression slopes in the value-increase and value-decrease conditions.**

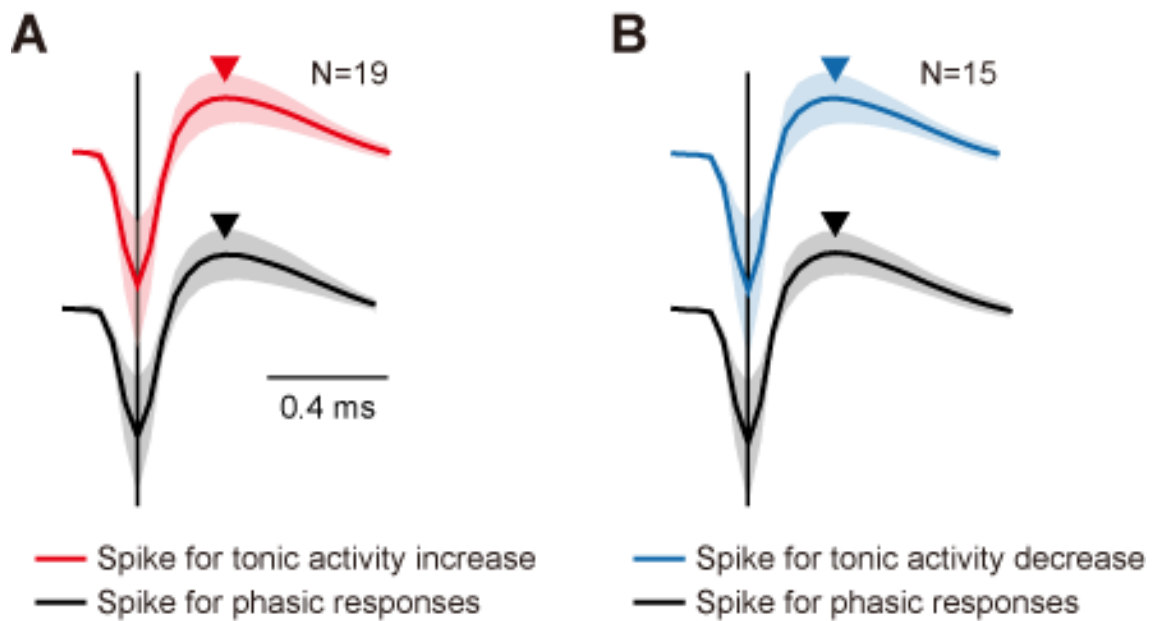
Regression slope between firing rate and time in the value-increase condition (ordinate) and that in the value-decrease condition (abscissa) ( $n = 99$  neurons). Green and orange dots indicate dopamine neurons with a significant regression slope in the value-increase and value-decrease conditions, respectively ( $p < 0.05$ ; linear regression analysis). Purple dots indicate neurons with a significant regression slope in both conditions.

To exclude the possibility that the tonic increase and decrease in activity were caused by spikes of non-dopamine neurons contaminating single-unit recording, I



compared spike waveform (i.e., the width of spikes: a characteristic electrophysiological feature of dopamine neurons) between the time window during which the tonic activity changes occurred in the value-increase and value-decrease conditions and the time window during which the conventional, CS-aligned phasic dopamine responses occurred in the value-fixed condition (Figure 7A, B). There was no significant difference in spike width between the time windows for the 19 dopamine neurons with a tonic increase in activity in the value-increase condition ( $n = 19$ , spike width during the time window to detect the tonic activity increase in the value-increase condition: mean  $\pm$  SD =  $0.32 \pm 0.08$  ms, spike width during the time window to detect the CS-aligned phasic responses in the value-fixed condition: mean  $\pm$  SD =  $0.33 \pm 0.10$  ms,  $p = 0.12$ ; Wilcoxon signed-rank test) (Figure 7A) or for the 15 dopamine neurons with a tonic decrease in activity in the value-decrease condition ( $n = 15$ , spike width during the time window to detect the tonic activity decrease in the value-decrease condition: mean  $\pm$  SD =  $0.31 \pm 0.07$  ms, spike width during the time window to detect the CS-aligned phasic responses in the value-fixed condition: mean  $\pm$  SD =  $0.31 \pm 0.06$  ms,  $p = 0.53$ ; Wilcoxon signed-rank test) (Figure 7B). These results suggest that the tonic increase and decrease in activity were not caused by spikes of non-dopamine neurons contaminating single-unit recording. Furthermore,

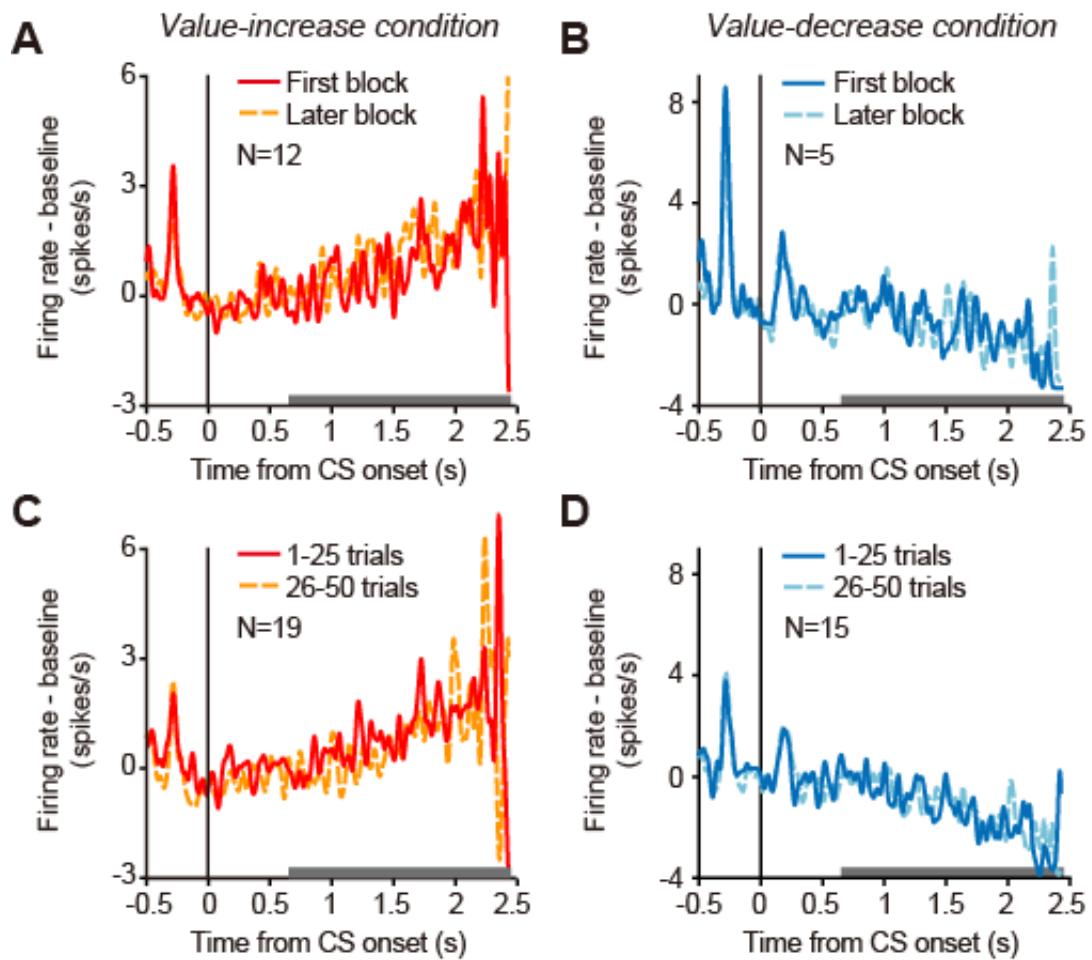
to exclude the possibility that the difference in activity patterns among the value-increase, value-decrease, and value-fixed conditions (i.e., tonic increase, tonic decrease, and phasic changes, respectively) was due to the loss of spikes or emergence of spikes of other neurons during single-unit recording, I compared the baseline firing rate (i.e., firing rate from 500 to 0 ms before the fixation point onset) of the 19 and 15 dopamine neurons among the three conditions. On average, there was no significant difference in baseline firing rate among the conditions for the 19 dopamine neurons ( $n = 19$ , value-increase condition: mean  $\pm$  SD =  $3.4 \pm 1.2$  spikes/s, value-decrease condition: mean  $\pm$  SD =  $3.2 \pm 1.3$  spikes/s, value-fixed condition: mean  $\pm$  SD =  $3.6 \pm 1.2$  spikes/s,  $p = 0.50$ ; Kruskal-Wallis test) or the 15 dopamine neurons ( $n = 15$ , value-increase condition: mean  $\pm$  SD =  $3.7 \pm 1.0$  spikes/s, value-decrease condition: mean  $\pm$  SD =  $3.8 \pm 1.3$  spikes/s, value-fixed condition: mean  $\pm$  SD =  $3.8 \pm 1.0$  spikes/s,  $p = 0.94$ ; Kruskal-Wallis test). Even at the single-neuron level, only a few neurons (4 of the 19 neurons, and 4 of the 15 neurons) exhibited a significant difference in baseline firing among the three conditions ( $p < 0.05$ ; Kruskal-Wallis test). These results suggest that the three different activity patterns were not due to the loss of spikes or emergence of other spikes during single-unit recording.



**Figure 7. Waveforms of spikes causing the tonic activity changes and the CS-aligned phasic responses.**

(A and B) Averaged spike waveforms of the 19 dopamine neurons with a tonic increase in activity in the value-increase condition (A) and the 15 dopamine neurons with a tonic decrease in activity in the value-decrease condition (B). The waveforms are shown for the spikes that occurred during the time window from 650 to 2450 ms after the CS onset in the value-increase (red curve) and value-decrease (blue curve) conditions and the spikes that occurred during the time window from 100 to 400 ms after the CS onset in the value-fixed condition (black curves). Vertical lines indicate the first negative peaks of the averaged spikes, and arrowheads indicate the following positive peaks of the spikes. Shaded areas indicate SD.

Although each condition consisted of a block of 50 trials and was repeated once or more for each recording session (Figure 2D), the tonic increase and decrease in activity were consistently observed between the first and later blocks of each condition (Figure 8A, B). No neuron exhibited a significantly different regression slope between the first and later blocks ( $p > 0.05$ ; comparison of two regression slopes). Furthermore, the tonic increase and decrease in activity were also consistently observed within each block (Figure 8C, D). No neuron exhibited a significantly different regression slope between the first half and latter half trials of blocks ( $p > 0.05$ ; comparison of two regression slopes). These results suggest that the tonic activity changes occurred in consistent manners between and within blocks.



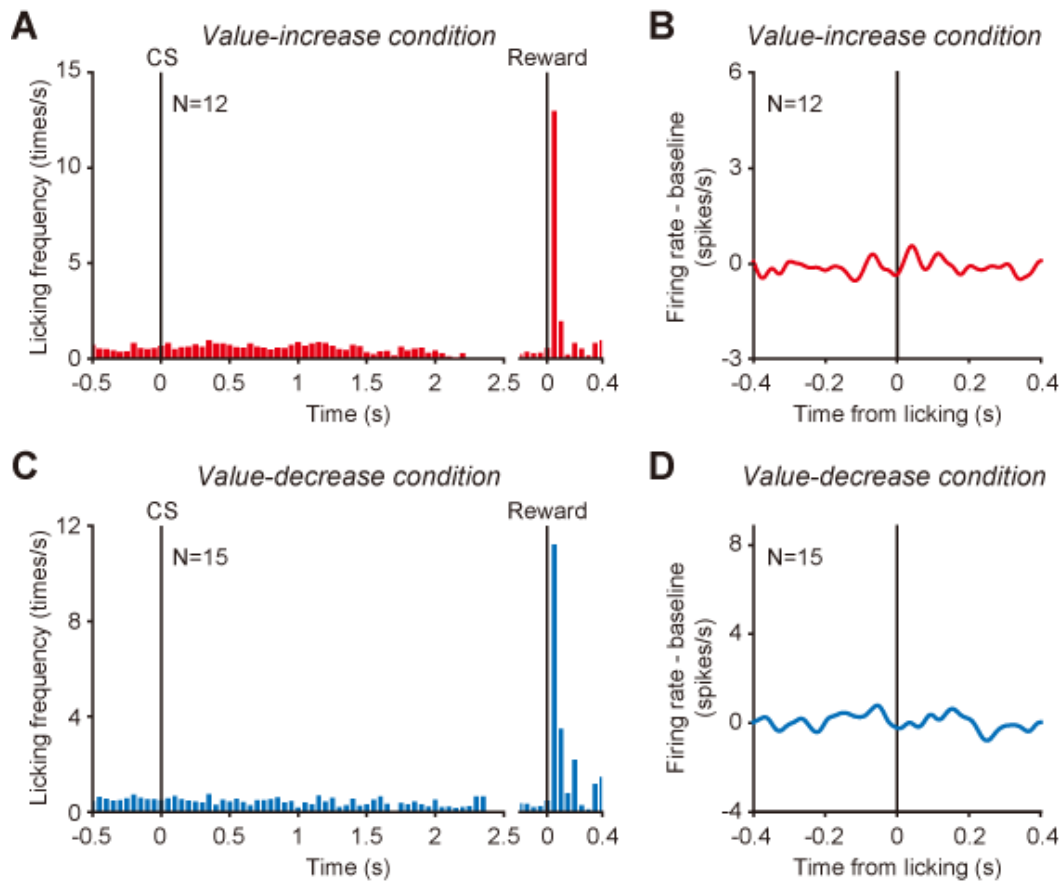
**Figure 8.** Changes in the tonic activity of dopamine neurons between and within blocks.

(A) Averaged SDFs of dopamine neurons with a tonic increase in activity in the value-increase condition. The SDFs are shown for the first (red solid curve) and later (orange dashed curve) blocks. 12 of the 19 dopamine neurons with a tonic increase in activity could be used for this analysis because recording from these neurons was repeated for two or more blocks. (B) Averaged SDFs of dopamine neurons with a tonic decrease in activity in the value-decrease condition. The SDFs are shown for the first (blue solid curve) and later (cyan dashed curve) blocks. 5 of the 15 dopamine neurons with a tonic

decrease in activity could be used for this analysis because recording from these neurons was repeated for two or more blocks. (C) Averaged SDFs of the 19 dopamine neurons with a tonic increase in activity in the value-increase condition. The SDFs are shown for the first half (red solid curve) and latter half (orange dashed curve) trials of blocks. (D) Averaged SDFs of the 15 dopamine neurons with a tonic decrease in activity in the value-decrease condition. The SDFs are shown for the first half (blue solid curve) and latter half (cyan dashed curve) trials of blocks.

A previous study reported that dopamine neurons exhibited a tonic increase in activity when future rewards were uncertain (Fiorillo et al., 2003). Thus, the tonic increase in activity, which I observed in the value-increase condition, might be accounted for by an uncertainty-evoked activity increase, because the reward value was uncertain until the green area stopped increasing in this condition. Contrary to this assumption, the 19 dopamine neurons, which exhibited a tonic increase in activity in the value-increase condition, did not exhibit a tonic increase in activity when the CS did not predict the future reward with certainty in the value-fixed condition (dotted curves in the right column of Figure 5E). On average, the regression slope between firing rate and time was not significantly different from zero ( $n = 19$  neurons,  $\text{mean} \pm \text{SD} = 0.015 \pm 0.098$ ,  $p = 0.20$ ; Wilcoxon signed-rank

test). This result suggests that the tonic increase in activity in the value-increase condition was not evoked by uncertainty.



**Figure 9. Effect of licking on dopamine neuron activity.**

(A) Frequency of licking of the monkeys aligned by the CS onset (left) and the reward onset (right) in the value-increase condition. The data used for this analysis was measured during the recording of 12 of the 19 dopamine neurons with a tonic increase in activity in the value-increase condition. During the recording of the other dopamine neurons, I did not measure licking. (B) Averaged SDF of the 12 dopamine neurons aligned by the onset of each licking

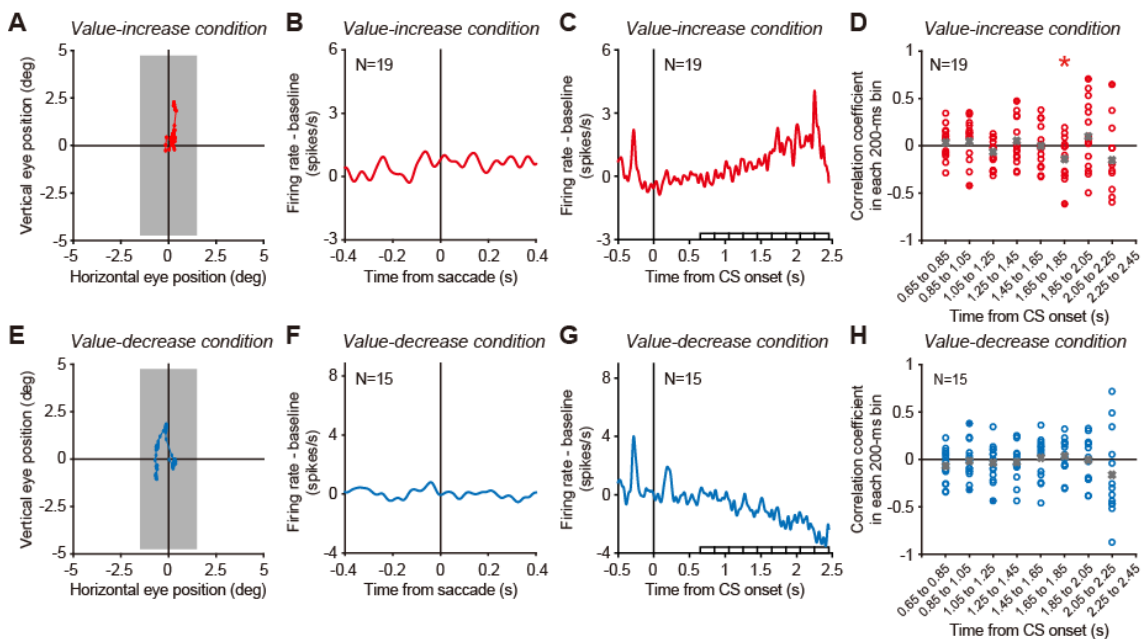
event in the value-increase condition. The spikes that occurred during value change was used to calculate the SDF. (C) Frequency of licking of the monkeys in the value-decrease condition. The data used for this analysis was measured during the recording of the 15 dopamine neurons with a tonic decrease in activity in the value-decrease condition. (D) Averaged SDF of the 15 dopamine neurons aligned by the onset of each licking event in the value-decrease condition.

Recent studies using rodents have shown that a subgroup of dopamine neurons increases their activity when animals simply initiate a body movement (da Silva et al., 2018; Howe and Dombeck, 2016), although such movement-related dopamine neuron activation has not been reported in primates. In the present study, I observed that the monkeys sporadically licked the spout for reward delivery while the reward value gradually changed (see Figure 9A and C for the value-increase and value-decrease conditions, respectively). To examine whether the tonic increase and decrease in dopamine neuron activity were evoked by licking, I aligned the activity of the dopamine neurons exhibiting a tonic increase or decrease in activity by the onset of each licking event. These dopamine neurons did not exhibit clear activity modulation around the onset (see Figure 9B for the dopamine neurons with a tonic increase in activity in the value-increase condition, and see Figure 9D for the



dopamine neurons with a tonic decrease in activity in the value-decrease condition), suggesting that the tonic increase and decrease in dopamine neuron activity were not caused by licking. I also observed that the monkeys often made eye movements along the vertical bar stimulus (i.e., CS) while the reward value gradually changed (see Figure 10A and E for example trials in the value-increase and value-decrease conditions, respectively). To examine whether the tonic increase and decrease in dopamine neuron activity were evoked by eye movements, I aligned the activities of the dopamine neurons exhibiting a tonic increase and decrease in activity by the onset of each saccade. These dopamine neurons did not exhibit clear activity modulation around the onset (see Figure 10B for the dopamine neurons with a tonic increase in activity in the value-increase condition, and see Figure 10F for the dopamine neurons with a tonic decrease in activity in the value-decrease condition). I further analyzed the relationship between firing rate and vertical gaze position. In this analysis, I used the time window from 650 to 2450 ms (divided into 200-ms bins) after the CS onset (see the horizontal white bars in Figure 10C, G), which I used to calculate the regression slope between firing rate and time. In each 200-ms bin, I calculated the correlation coefficient between firing rate and vertical gaze position for each of the dopamine neurons with a tonic increase in activity in the value

increase condition (Figure 10D) and each of the dopamine neurons exhibiting a tonic decrease in activity in the value-decrease condition (Figure 10H). On average, these dopamine neurons did not exhibit a significant correlation coefficient between firing rate and vertical gaze position in either bin except the bin from 1650 to 1850 ms in the value-increase condition (n = 19 neurons, mean  $\pm$  SD =  $-0.14 \pm 0.20$  p = 0.011; Wilcoxon signed-rank test). These results suggest that the tonic increase and decrease in dopamine neuron activity were not caused by eye movements.



**Figure 10. Effect of eye movements on dopamine neuron activity.**

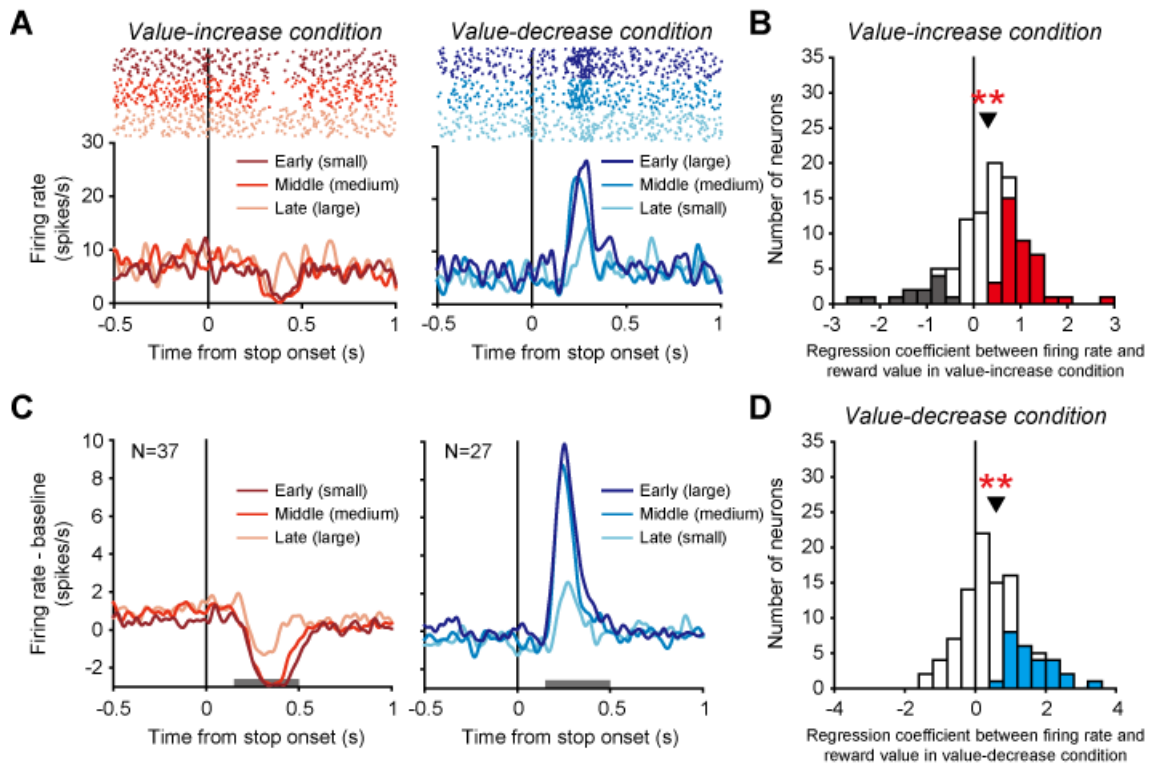
(A and E) Gaze positions of example trials during value change in the value-increase (A) and

value-decrease (E) conditions. The position was plotted for each 25 ms. Gray areas indicate the vertical bar stimulus (i.e., CS). (B and F) Averaged SDFs aligned by the onset of each saccade. The averaged SDF of the 19 dopamine neurons with a tonic increase in activity in the value-increase condition (B) and that of the 15 dopamine neurons with a tonic decrease in activity in the value-decrease condition (F) are shown. The spikes that occurred during value change was used to calculate the SDFs. (C and G) Averaged SDFs aligned by the CS onset and calculation time window. The averaged SDF of the 19 dopamine neurons (C) and that of the 15 dopamine neurons (G) are shown with the calculation time window divided into 200-ms bins (horizontal white bars). (D and H) Correlation coefficients between firing rate and vertical gaze position in each 200-ms bin in the value-increase (D) and value-decrease (H) conditions. Filled plots indicate dopamine neurons with a significant regression coefficient ( $p < 0.05$ ; correlation analysis), and open plots indicate dopamine neurons with no significance. Gray crosses indicate the mean of the correlation coefficients in each bin. Single asterisk above plots indicates a significant deviation from zero in a bin ( $p < 0.05$ ; Wilcoxon signed-rank test). The data in the bin from 2250 to 2450 ms were excluded from the correlation analysis because this bin included less than 5 trials for each neuron and, consequently, I was unable to conduct a statistically valid analysis.

#### **4.3. Phasic activity of dopamine neurons evoked when the reward value stopped changing**

I have observed the tonic activity of dopamine neurons that gradually increased and decreased as the reward value increased and decreased, respectively. I next

found unique modulations of dopamine neuron activity that were evoked when the reward value stopped increasing and decreasing. An example dopamine neuron phasically decreased its activity when the value stopped increasing in the value-increase condition (left in Figure 11A), and phasically increased its activity when the value stopped decreasing in the value-decrease condition (right in Figure 11A). Such phasic decrease and increase in activity were generally observed across the recorded dopamine neurons, and seem to be accounted for by conventional phasic dopamine inhibitory and excitatory responses to undesirable and desirable events. That is, the stopping of the value increase was undesirable and evoked the phasic decrease in dopamine neuron activity, whereas the stopping of the value decrease was desirable and evoked the phasic increase in dopamine neuron activity.



**Figure 11. Phasic activity of dopamine neurons evoked when the reward value stopped changing.**

(A) Activity of an example dopamine neuron. Rasters and SDFs are aligned by the time at which the reward value stopped increasing or decreasing in the value-increase (left) and value-decrease (right) conditions, respectively. Left: Dark red, red, and light red indicate the activity in trials in which the stopping of the value increase occurred at early (small reward: 0.1 to 0.16 ml), middle (medium reward: 0.16 to 0.23 ml) and late (large reward: 0.23 to 0.3 ml) times, respectively. Right: Dark blue, blue, and light blue indicate the activity in trials in which the stopping of the value decrease occurred at early (large reward: 0.23 to 0.3 ml), middle (medium reward: 0.16 to 0.23 ml) and late (small reward: 0.1 to 0.16 ml) times, respectively. (B and D) Regression coefficient between firing rate and reward value in the

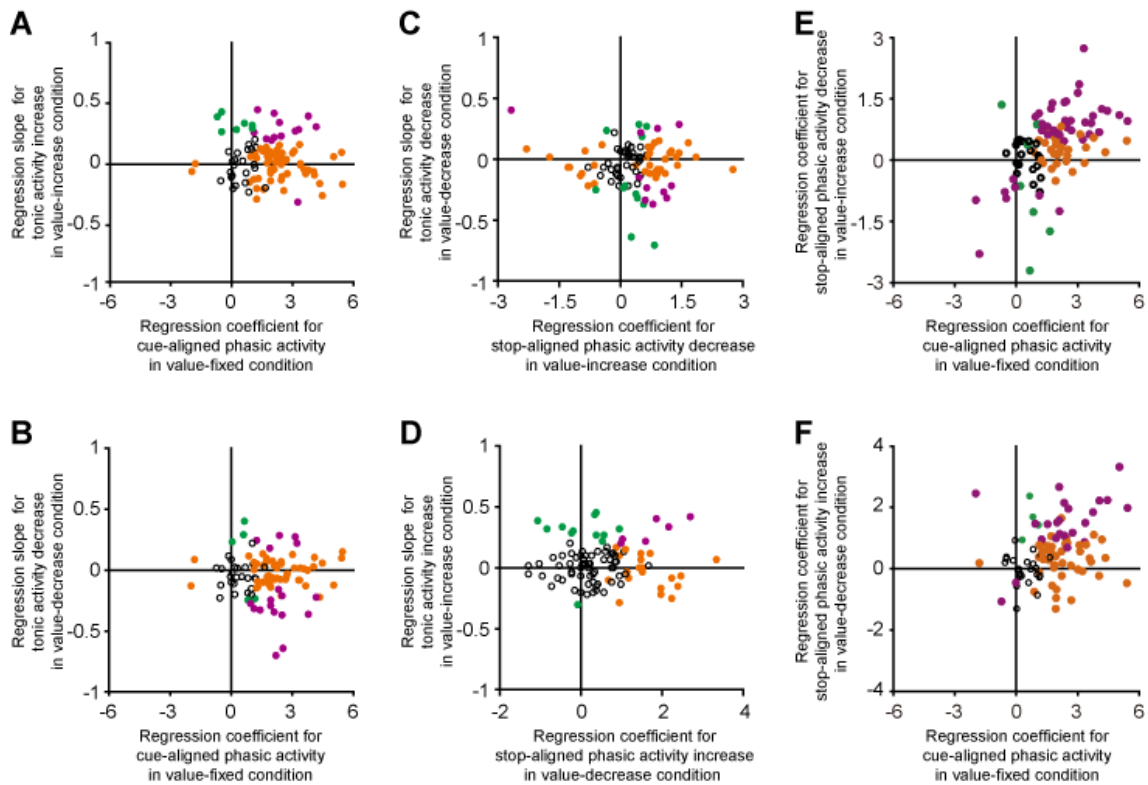
value-increase (B) and value-decrease (D) conditions (n = 99 neurons). Conventions are as in Figure 5C, D. (C) Averaged SDFs of the 37 dopamine neurons with a significantly positive regression coefficient in the value-increase condition (left) and those of the 27 dopamine neurons with a significantly positive regression coefficient in the value-decrease condition (right). Horizontal gray bars indicate the time windows used for analyses. Conventions are as in A.

Notably, the earlier the reward value stopped increasing (i.e., the smaller the reward value), the more strongly the phasic activity decreased (significantly positive regression coefficient between firing rate and reward value: n = 99 neurons, mean  $\pm$  SD =  $0.31 \pm 0.81$ ,  $p = 1.97 \times 10^{-5}$ ; Wilcoxon signed-rank test) (Figure 11B, also see left in Figure 11C for the averaged activity of 37 dopamine neurons with a significantly positive regression coefficient). Furthermore, the earlier the reward value stopped decreasing (i.e., the larger the reward value), the more strongly the phasic activity increased (significantly positive regression coefficient between firing rate and reward value: n = 99 neurons, mean  $\pm$  SD =  $0.59 \pm 1.01$ ,  $p = 7.66 \times 10^{-8}$ ; Wilcoxon signed-rank test) (Figure 11D, also see right in Figure 11C for the averaged activity of 27 dopamine neurons with a significantly positive regression coefficient). Thus, the phasic decrease and increase in dopamine neuron activity likely reflected the

reward value that was fixed when the green area stopped changing.

#### **4.4. Relationship between the tonic and phasic dopamine neuron activities**

So far, I have reported that dopamine neurons exhibited not only the phasic activity but also the tonic one to effectively signal reward values. I next examined the relationship between the tonic and phasic activities. Even if dopamine neurons exhibited a cue-aligned phasic activity that represented the value of cued reward in the value-fixed condition, these neurons did not necessarily display a tonic increase in activity (i.e., a positive regression slope between firing rate and time) in the value-increase condition (no significant correlation between the regression coefficient and the regression slope:  $n = 99$  neurons,  $r = -0.083$ ,  $p = 0.42$ ; correlation analysis) (Figure 12A) or a tonic decrease in activity (i.e., a negative regression slope between firing rate and time) in the value-decrease condition (no significant correlation between the regression coefficient and the regression slope:  $n = 99$  neurons,  $r = -0.003$ ,  $p = 0.97$ ; correlation analysis) (Figure 12B).



**Figure 12. Relationship between the tonic and phasic dopamine neuron activities.**

- (A) Regression slope between firing rate and time for the tonic activity increase in the value-increase condition (ordinate) and regression coefficient between firing rate and reward value for the cue-aligned phasic activity in the value-fixed condition (abscissa) ( $n = 99$  neurons).
- (B) Regression slope between firing rate and time for the tonic activity decrease in the value-decrease condition (ordinate) and regression coefficient between firing rate and reward value for the cue-aligned phasic activity in the value-fixed condition (abscissa) ( $n = 99$  neurons).
- (C) Regression slope between firing rate and time for the tonic activity decrease in the value-decrease condition (ordinate) and regression coefficient between firing rate and reward value for the stop-aligned phasic activity decrease in the value-increase condition (abscissa) ( $n = 99$  neurons).
- (D) Regression slope between firing rate and time for the tonic activity increase



in the value-increase condition (ordinate) and regression coefficient between firing rate and reward value for the stop-aligned phasic activity increase in the value-decrease condition (abscissa) (n = 99 neurons). (E) Regression slope between firing rate and reward value for the stop-aligned phasic activity decrease in the value-increase condition (ordinate) and regression coefficient between firing rate and reward value for the cue-aligned phasic activity in the value-fixed condition (abscissa) (n = 99 neurons). (F) Regression slope between firing rate and reward value for the stop-aligned phasic activity increase in the value-decrease condition (ordinate) and regression coefficient between firing rate and reward value for the cue-aligned phasic activity in the value-fixed condition (abscissa) (n = 99 neurons). Green and orange dots indicate neurons with a significant difference from zero in the ordinate and abscissa, respectively ( $p < 0.05$ ; linear regression analysis). Purple dots indicate neurons with significance for both.

Furthermore, even if dopamine neurons exhibited a stop-aligned phasic activity decrease that reflected the reward value in the value-increase condition, these neurons did not necessarily display a tonic decrease in activity (i.e., a negative regression slope between firing rate and time) in the value-decrease condition (no significant correlation between the regression coefficient and the regression slope: n = 99 neurons,  $r = -0.069$ ,  $p = 0.50$ ; correlation analysis) (Figure 12C). Thus, there was no clear relationship between the phasic and tonic activity decreases. In addition,

even if dopamine neurons exhibited a stop-aligned phasic activity increase that reflected the reward value in the value-decrease condition, these neurons did not necessarily display a tonic increase in activity (i.e., a positive regression slope between firing rate and time) in the value-increase (no significant correlation between the regression coefficient and the regression slope:  $n = 99$  neurons,  $r = 0.061$ ,  $p = 0.55$ ; correlation analysis) (Figure 12D). Thus, there was no clear relationship between the phasic and tonic activity increases. Taken together, the above results suggest that the tonic activity tracking the changing reward values was evoked independently of the phasic activity that represent the static, unchanged reward values.

I also examined the relationship between the cue-aligned phasic activity in the value-fixed condition and the stop-aligned phasic activity in the value-increase or value-decrease condition (Figure 12E, F). These phasic responses were significantly correlated with each other (the regression coefficient for the cue-aligned phasic activity in the value-fixed condition versus the regression coefficient for the stop-aligned phasic activity decrease in the value-increase condition: Figure 12E,  $n = 99$  neurons,  $r = 0.48$ ,  $p = 5.6 \times 10^{-7}$ ; the regression coefficient for the cue-aligned phasic activity in the value-fixed condition versus the regression coefficient for the stop-

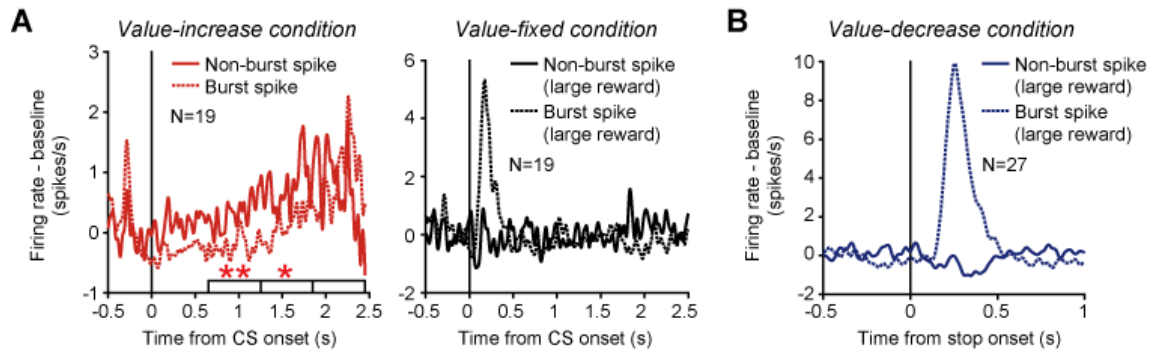
aligned phasic activity increase in the value-decrease condition: Figure 12F,  $n = 99$  neurons,  $r = 0.29$ ,  $p = 0.003$ ; correlation analysis). Thus, the same dopamine neurons tended to exhibit both of the cue-aligned and stop-aligned phasic activities.

#### **4.5. Burst and non-burst spike firing modes of dopamine neurons**

Dopamine neurons are known to exhibit two different firing modes: burst spike firing and non-burst spike firing (Grace and Bunney, 1984a, b). The phasic activity increase of these neurons, which has been shown to reflect reward prediction errors, is produced by burst spike firing (Bayer et al., 2007). Here, I attempted to determine which firing mode produced the tonic activity increase of dopamine neurons. The burst and non-burst spike firing modes were identified based on interspike intervals (ISIs) using previously established criteria (Grace and Bunney, 1984a). These criteria were 1) an ISI of 80 ms or less to identify the onset of the burst spike firing mode, and 2) an ISI of more than 160 ms to identify the end of the firing mode. I calculated the averaged activity of the 19 dopamine neurons that exhibited a tonic increase in activity in the value-increase condition separately for burst and non-burst spikes, and found that non-burst spike firing in these neurons increased

tonically as the reward value gradually increased (left in Figure 13A). On the other hand, burst spike firing in these neurons increased when the reward value came close to the maximum value near the end of the value-increase period. To statistically test this trend, I divided the calculation time window, which I used to detect the significantly positive regression slope of the 19 dopamine neurons, into initial, middle, and late periods (horizontal white bars in the left column of Figure 13A). I found that, while non-burst spike firing exhibited a significantly larger firing rate (original firing rate – baseline firing rate) compared with burst spike firing in the initial and middle periods, there was no significant difference in the late period (initial period:  $p = 0.006$ ; middle period:  $p = 0.018$ ; late period:  $p = 0.43$ ; bootstrap test with 1000 repetitions). Notably, these neurons exhibited a phasic increase in activity only with burst spike firing and not with non-burst spike firing in response to the CS predicting the large reward in the value-fixed condition (right in Figure 13A). In addition, the phasic increase in activity evoked when the reward value stopped decreasing in the value-decrease condition was also produced only by burst spike firing (Figure 13B). Taken together, these data indicate that, while the conventional phasic activity of dopamine neurons signaling the static, unchanged reward values was evoked by burst spike firing, the tonic activity tracking the

changing reward values was produced more largely by non-burst spike firing.



**Figure 13. Phasic and tonic activities of dopamine neurons produced by burst and non-burst spikes.**

(A) Averaged SDFs of the 19 dopamine neurons with a tonic increase in activity in the value-increase condition. Left: SDFs in the value-increase condition. Right: SDFs in the value-fixed condition. The SDFs are aligned by the CS onset, and shown for burst (dotted curve) and non-burst (solid curve) spikes. Horizontal white bars indicate the initial, middle, and late periods to calculate the neural modulation between burst and non-burst spike firing. Single and double asterisks above the horizontal bars indicate periods during which the firing rate (original firing rate – baseline firing rate) was significantly larger for non-burst spike firing than for burst spike firing ( $p < 0.05$  and  $p < 0.01$ , respectively; bootstrap test with 1000 repetitions). The SDFs in the value-fixed condition were calculated using large reward trials. (B) Averaged SDFs of the 27 dopamine neurons with a phasic activity that was modulated by reward values when the values stopped decreasing in the value-decrease condition. The SDFs in large reward trials are shown, and aligned by the time at which the reward value stopped

decreasing. Other conventions are as in Figure 5A.

## 5. Discussion

Midbrain dopamine neurons respond to reward-predicting cues with phasic activity, and this phasic activity has been shown to represent the value of cued rewards (Matsumoto and Hikosaka, 2009; Roesch et al., 2007; Tobler et al., 2005). On the other hand, using a Pavlovian procedure in which a reward value gradually increased or decreased moment-by-moment, I found that a subset of dopamine neurons tracked the changing reward value by a tonic increase and decrease in activity rather than by the phasic activity. These neurons exhibited a conventional value-representing phasic activity when the reward value was static and when the value stopped changing. While the phasic activity was evoked by burst spike firing, the tonic activity tracking the changing reward value was produced not only by burst spike firing but also by non-burst spike firing. Particularly, the tonic activity was more largely produced by non-burst spike firing than burst spike firing.

### 5.1. What variable does the tonic activity of dopamine neurons represent?

Although the phasic activity of dopamine neurons has attracted attention for its crucial roles in reward processing, recent studies using freely-moving rodents reported that dopamine release in the striatum exhibited a tonic increase (or ramp)

as animals approached rewards (Engelhard et al., 2019; Hamid et al., 2016; Howe et al., 2013; Mohebi et al., 2019). Such tonic increase has not been observed in dopamine neuron spike activity, at least in the VTA (Mohebi et al., 2019). Although the phasic activity of dopamine neurons has long been thought to represent reward prediction errors, the tonic increase in dopamine release has been proposed to represent “state values” that gradually increase as animals spatially and temporally approach rewards (Hamid et al., 2016; Mohebi et al., 2019). However, it is not immediately clear whether the tonic increase and decrease in dopamine neuron spike activity, which I observed in head-fixed monkeys, represent state values or reward prediction errors. Because both the state values and reward prediction errors gradually changed in the same fashion in my Pavlovian procedure, I was unable to definitely identify which variable the tonic activity of dopamine neurons represented. On the other hand, I observed that, even if dopamine neurons exhibited a strong phasic activity, these neurons did not necessarily display a tonic increase or decrease in activity. That is, the tonic activity occurred independently of the phasic activity that has been thought to represent reward prediction errors. This suggests that the tonic activity might not represent reward prediction errors, and appears to be in accord with a recent perspective that dopamine neurons represent not only reward



prediction errors but also multiple signals related to various events and behaviors, including punishments, movements, subjective sensory experience, response inhibition, decision-making and working memory (Brischoux et al., 2009; da Silva et al., 2018; de Lafuente and Romo, 2011; Engelhard et al., 2019; Howard et al., 2017; Howe and Dombeck, 2016; Matsumoto and Hikosaka, 2009; Matsumoto and Takada, 2013; Menegas et al., 2018; Ogasawara et al., 2018; Yun et al., 2020).

An attribute that might cause the tonic increase and decrease in dopamine activity is body movements. Recent studies using rodents have shown that a subgroup of dopamine neurons increases their activity when animals simply initiate a body movement (da Silva et al., 2018; Howe and Dombeck, 2016), although such movement-related dopamine neuron activation has not been reported in primates. Contrary this assumption, I found that two types of monkeys' movements (licks of the spout and eye movements), which were often observed during my Pavlovian procedure, did not influence the tonic increase or decrease in dopamine neuron activity. Another attribute that might cause these tonic activity changes is the visual motion of the green area. Since the green area dynamically increased and decreased in the value-increase and value-decrease conditions, respectively, in which I observed the tonic increase and decrease in dopamine neuron activity, it might be possible that

these motions of the green area caused the tonic activity changes irrespective of the changing reward value. On the other hand, a previous study in monkeys observed that dopamine neurons exhibited a phasic activation but not tonic activity change in response to moving-dot images (Nomoto et al., 2010). According to this observation, the motion itself was unlikely to cause the tonic increase and decrease in dopamine neuron activity. The time distance to the reward delivery might also cause the tonic activity changes. This time distance became shorter as the green area increased and decreased. However, dopamine neurons exhibited the opposite modulations (i.e., the tonic increase and decrease in activity) in the value-increase and value-decrease conditions in which the time distance to the reward delivery equally became shorter as the green area changed. Therefore, the time distance was also unlikely to cause the tonic increase and decrease in dopamine neuron activity.

## **5.2. Comparison between the present findings and previous reports on tonic dopamine signals**

Although I found that some dopamine neurons exhibited tonic activities signaling reward values, some other studies have also reported tonic dopamine signals. Howe et al. (2013) reported that dopamine released in the striatum

gradually increased as animals moved towards a goal to obtain a reward, which seemed to represent the distance to the goal. Hughes et al. (2020) reported that a subpopulation of dopamine neurons exhibited a tonic increase in activity while animals moved forward and a tonic decrease in activity while animals moved backward, which might represent the direction of locomotion. Consistent with this view, these dopamine neurons did not exhibit a clear response to a reward. Howard et al. (2017) reported that, when mice dynamically switched an action selected at different time points, dopamine neuron activity as well as dopamine released in the striatum exhibited gradual changes associated with action selection. Finally, Bromberg-Martin et al. (2010) reported that dopamine neurons exhibited tonic changes in activity during inter-trial-intervals during which animals waited for the start of the next trial. The tonic activity reflected the expected timing distribution of the next trial start.

As described above, some previous studies have observed tonic changes in dopamine neuron activity as well as dopamine release while animals were performing given tasks. However, the findings of these studies were fundamentally different from what I observed in the present study. The tonic dopamine signals reported in these studies were not directly associated with value change. Therefore,

I believe that the tonic dopamine neuron activity, which I observed in the present study, is the first experimental evidence that dopamine neurons represent reward values by gradually changing their activity in particular situation.

### **5.3. What roles does the tonic activity of dopamine neurons play in behavior?**

While the phasic dopamine activity, which is presumed to represent reward prediction errors, has been considered to provide teaching signals in reinforcement learning (Doya, 2002; Montague et al., 1996; Schultz et al., 1997), the tonic dopamine activity has been proposed to regulate motivation (Cagniard et al., 2006; Niv, 2007). A theory called “incentive salience theory” proposes that dopamine signals assign incentive values to goals or actions of behavioral processes, and motivate actions aimed at acquiring rewards (Berridge and Robinson, 1998). According to this theory, the tonic dopamine neuron activity, which I observed in the present study, seems to be a good candidate for a neural substrate that regulates the motivational vigor of actions in changing environments in which reward values fluctuate. That is, by tracking the changing values of future rewards, the tonic activity might enhance and suppress the motivational vigor of actions to obtain the rewards on a moment-by-

moment basis.

#### **5.4. Conclusion**

In conclusion, I demonstrated that, instead of phasic firing, tonic firing of dopamine neurons continuously tracked reward values changing moment-by-moment. This tonic activity was produced more largely by non-burst spikes than burst spikes that evoked the phasic dopamine neuron activity. My findings expand the current knowledge on dopamine neuron signals by highlighting the alternating firing modes of these neurons, which effectively signal reward values in a given situation. Further investigations are called for to definitely determine whether the tonic activity of dopamine neurons represents state values, reward prediction errors or other reward-related variables.

## 6. Acknowledgments

First, I would like to express my sincere gratitude to my supervisor Professor Masayuki Matsumoto, the Faculty of Medicine at the University of Tsukuba. I am grateful to Professor Matsumoto for his willingness to accept me, a person without any professional foundation in this field to start my master's program in neuroscience and his extreme patience throughout my study. His insightful comments and thorough revisions helped me bring this research to fruition. As the Chinese saying goes, it would be better to teach someone how to fish than to give him a fish. His guidance enabled me to think critically and complete my research independently, which has undoubtedly laid a strong foundation for my future research career. I am fortunate and honored to have conducted my master's and Ph.D. programs at Matsumoto Laboratory.

Also, I would like to send my gratitude to Dr. H. Yamada and Dr. J. Kunimatsu for their kind and earnest supervision in completing this thesis. And I would like to acknowledge Dr. M. Nejime and other members of Matsumoto Laboratory for their support and helpful suggestions. I also appreciate K. Bunzui, T. Okano, and Y. Narita for their help in animal care and Dr. T. Kawai for his support in data analyses.

Moreover, I owe my family members my great gratitude for their financial and mental support over these years.

In the end, I would like to show special thanks to the Chinese drama "Word Of Honor." As my academic career was coming to an end, the two main actors, Zhang Zhehan and Gong Jun, gave wonderful performances that made me have the immersive experience of deep following the drama for the first time in my life. It is an extraordinary and precious life experience for me. Like the rest of the audience of this drama, I am very glad to have encountered this drama in March, which is an experience we will all never forget. And I believe that this unforgettable experience will help me in my future academic research.

## 7. References

- Alexander, G. E., DeLong, M.R. & Strick, P.L. (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.*, **9**, 357-381.
- Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. (1991) Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res.*, **85**, 491-500.
- Bayer, H. M., Lau, B. & Glimcher, P.W. (2007) Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.*, **98**, 1428-1439.
- Bayer, H. M. & Glimcher, P.W. (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, **47**, 129-141.
- Berridge, K. C. & Robinson, T.E. (1998) What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.*, **28**, 309-369.
- Bjorklund, A. & Dunnett, S.B. (2007) Dopamine neuron systems in the brain: An update. *Trends Neurosci.*, **30**, 194-202.
- Bowman, E. M., Aigner, T.G. & Richmond, B.J. (1996) Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol.*, **75**, 1061-1073.
- Brischoux, F., Chakraborty, S., Brierley, D.I. & Ungless, M.A. (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 4894-4899.



- Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. (2010) Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron*, **67**, 144-155.
- Cagniard, B., Balsam, P.D., Brunner, D. & Zhuang, X. (2006) Mice with chronically elevated dopamine exhibit enhanced motivation, but not learning, for a food reward. *Neuropsychopharmacology*, **31**, 1362-1370.
- Chang, C. Y., Esber, G.R., Marrero-Garcia, Y., Yau, H.J., Bonci, A. & Schoenbaum, G. (2016) Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nat. Neurosci.*, **19**, 111-116.
- Cohen, J. Y., Haesler, S., Vong, L., Lowell, B.B. & Uchida, N. (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, **482**, 85-88.
- da Silva, J. A., Tecuapetla, F., Paixao, V. & Costa, R.M. (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature*, **554**, 244-248.
- de Lafuente, V. & Romo, R. (2011) Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 19767-19771.
- Doya, K. (2002) Metalearning and neuromodulation. *Neural Netw.*, **15**, 495-506.
- Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W. & Witten, I.B. (2019)

- Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*, **570**, 509-513.
- Fiorillo, C. D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1898-1902.
- Francois, C., Yelnik, J., Tande, D., Agid, Y. & Hirsch, E.C. (1999) Dopaminergic cell group A8 in the monkey: Anatomical organization and projections to the striatum. *J. Comp. Neurol.*, **414**, 334-347.
- Gauthier, J., Parent, M., Levesque, M. & Parent, A. (1999) The axonal arborization of single nigrostriatal neurons in rats. *Brain Res.*, **834**, 228-232.
- Grace, A. A. & Bunney, B.S. (1984a) The control of firing pattern in nigral dopamine neurons: Burst firing. *J. Neurosci.*, **4**, 2877-2890.
- Grace, A. A. & Bunney, B.S. (1984b) The control of firing pattern in nigral dopamine neurons: Single spike firing. *J. Neurosci.*, **4**, 2866-2876.
- Hamid, A. A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J. & Berke, J.D. (2016) Mesolimbic dopamine signals the value of work. *Nat. Neurosci.*, **19**, 117-126.
- Heath, R. G. (1963) Electrical self-stimulation of the brain in man. *Am. J. Psychiatry*, **120**, 571-577.
- Hikosaka, K. & Watanabe, M. (2004) Long- and short-range reward expectancy in the primate orbitofrontal cortex. *Eur. J. Neurosci.*, **19**, 1046-1054.

- 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. (2007) Basal ganglia mechanisms of reward-oriented eye movement. *Ann. N. Y. Acad. Sci.*, **1104**, 229-249.
- Howard, C. D., Li, H., Geddes, C.E. & Jin, X. (2017) Dynamic nigrostriatal dopamine biases action selection. *Neuron*, **93**, 1436-1450.e8.
- Howe, M. W. & Dombeck, D.A. (2016) Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, **535**, 505-510.
- Howe, M. W., Tierney, P.L., Sandberg, S.G., Phillips, P.E. & Graybiel, A.M. (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, **500**, 575-579.
- Hughes, R. N., Bakhurin, K.I., Petter, E.A., Watson, G.D.R., Kim, N., Friedman, A.D. & Yin, H.H. (2020) Ventral tegmental dopamine neurons control the impulse vector during motivated behavior. *Curr. Biol.*, **30**, 2681-2694.e5.
- Kawagoe, R., Takikawa, Y. & Hikosaka, O. (2004) Reward-predicting activity of dopamine and caudate neurons--a possible mechanism of motivational control of saccadic eye movement. *J. Neurophysiol.*, **91**, 1013-1024.
- Kawai, T., Yamada, H., Sato, N., Takada, M. & Matsumoto, M. (2015) Roles of the lateral habenula and anterior cingulate cortex in negative outcome monitoring and behavioral adjustment in nonhuman primates. *Neuron*, **88**, 792-804.

- Kolling, N., Wittmann, M.K., Behrens, T.E., Boorman, E.D., Mars, R.B. & Rushworth, M.F. (2016) Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.*, **19**, 1280-1285.
- Lau, B. & Glimcher, P.W. (2008) Value representations in the primate striatum during matching behavior. *Neuron*, **58**, 451-463.
- Lynd-Balta, E. & Haber, S.N. (1994) The organization of midbrain projections to the striatum in the primate: Sensorimotor-related striatum versus ventral striatum. *Neuroscience*, **59**, 625-640.
- Maes, E. J. P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G. & Iordanova, M.D. (2020) Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. *Nat. Neurosci.*, **23**, 176-178.
- Matsumoto, M. & Takada, M. (2013) Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. *Neuron*, **79**, 1011-1024.
- Matsumoto, M. & Hikosaka, O. (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, **459**, 837-841.
- Menegas, W., Akiti, K., Amo, R., Uchida, N. & Watabe-Uchida, M. (2018) Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat. Neurosci.*, **21**, 1421-1430.

- Mohebi, A., Pettibone, J.R., Hamid, A.A., Wong, J.T., Vinson, L.T., Patriarchi, T., Tian, L., Kennedy, R.T. & Berke, J.D. (2019) Dissociable dopamine dynamics for learning and motivation. *Nature*, **570**, 65-70.
- Montague, P. R., Dayan, P. & Sejnowski, T.J. (1996) A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.*, **16**, 1936-1947.
- Morreens, J., Aydin, C., Janse van Rensburg, A., Esquivelzeta Rabell, J. & Haesler, S. (2020) Cue-evoked dopamine promotes conditioned responding during learning. *Neuron*, **106**, 142-153.e7.
- Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, **43**, 133-143.
- Morrison, S. E. & Salzman, C.D. (2009) The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.*, **29**, 11471-11483.
- Nakamura, K., Santos, G.S., Matsuzaki, R. & Nakahara, H. (2012) Differential reward coding in the subdivisions of the primate caudate during an oculomotor task. *J. Neurosci.*, **32**, 15963-15982.
- Niv, Y. (2007) Cost, benefit, tonic, phasic: What do response rates tell us about dopamine and motivation? *Ann. N. Y. Acad. Sci.*, **1104**, 357-376.

- Nomoto, K., Schultz, W., Watanabe, T. & Sakagami, M. (2010) Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.*, **30**, 10692-10702.
- Ogasawara, T., Nejime, M., Takada, M. & Matsumoto, M. (2018) Primate nigrostriatal dopamine system regulates saccadic response inhibition. *Neuron*, **100**, 1513-1526.e4.
- Olds, J. & Milner, P. (1954) Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.*, **47**, 419-427.
- Padoa-Schioppa, C. (2013) Neuronal origins of choice variability in economic decisions. *Neuron*, **80**, 1322-1336.
- Padoa-Schioppa, C. & Assad, J.A. (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.*, **11**, 95-102.
- Padoa-Schioppa, C. & Assad, J.A. (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature*, **441**, 223-226.
- Roesch, M. R., Calu, D.J. & Schoenbaum, G. (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.*, **10**, 1615-1624.
- Samejima, K., Ueda, Y., Doya, K. & Kimura, M. (2005) Representation of action-specific reward values in the striatum. *Science*, **310**, 1337-1340.

- Satoh, T., Nakai, S., Sato, T. & Kimura, M. (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.*, **23**, 9913-9923.
- Schultz, W. (1998) Predictive reward signal of dopamine neurons. *J. Neurophysiol.*, **80**, 1-27.
- Schultz, W., Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and reward. *Science*, **275**, 1593-1599.
- Seo, H. & Lee, D. (2007) Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.*, **27**, 8366-8377.
- Shima, K. & Tanji, J. (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, **282**, 1335-1338.
- Stauffer, W. R., Lak, A., Yang, A., Borel, M., Paulsen, O., Boyden, E.S. & Schultz, W. (2016) Dopamine neuron-specific optogenetic stimulation in rhesus macaques. *Cell*, **166**, 1564-1571.e6.
- Steinberg, E. E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K. & Janak, P.H. (2013) A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.*, **16**, 966-973.
- Takada, M. & Hattori, T. (1987) Organization of ventral tegmental area cells projecting to the occipital cortex and forebrain in the rat. *Brain Res.*, **418**, 27-33.
- Tobler, P. N., Fiorillo, C.D. & Schultz, W. (2005) Adaptive coding of reward value by dopamine neurons. *Science*, **307**, 1642-1645.

- Tremblay, L. & Schultz, W. (1999) Relative reward preference in primate orbitofrontal cortex. *Nature*, **398**, 704-708.
- Tsai, H. C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L. & Deisseroth, K. (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*, **324**, 1080-1084.
- Williams, S. M. & Goldman-Rakic, P.S. (1998) Widespread origin of the primate mesofrontal dopamine system. *Cereb. Cortex*, **8**, 321-345.
- Yun, M., Kawai, T., Nejime, M., Yamada, H. & Matsumoto, M. (2020) Signal dynamics of midbrain dopamine neurons during economic decision-making in monkeys. *Sci. Adv.*, **6**, eaba4962.