Operant Conditioning of Primate Prefrontal Neurons

Shunsuke Kobayashi,1,2 Wolfram Schultz,2 and Masamichi Sakagami1
1Brain Science Institute, Tamagawa University, Tokyo, Japan; and 2Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, United Kingdom

Submitted 11 March 2009; accepted in final form 24 January 2010

Kobayashi S, Schultz W, Sakagami M. Operant conditioning of primate prefrontal neurons. J Neurophysiol 103: 1843–1855, 2010. First published January 27, 2010; doi:10.1152/jn.00173.2009. An operant is a behavioral act that has an impact on the environment to produce an outcome, constituting an important component of voluntary behavior. Because the environment can be volatile, the same action may cause different outcomes. Thus to obtain an optimal outcome, it is crucial to detect action–outcome relationships and adapt the behavior accordingly. Although prefrontal neurons are known to change activity depending on expected reward, it remains unknown whether prefrontal activity contributes to obtaining reward. We investigated this issue by setting variable relationships between levels of single-neuron activity and rewarding outcomes. Lateral prefrontal neurons changed their spiking activity according to the specific requirements for gaining reward, without the animals making a motor response. Thus spiking activity constituted an operant response. Data from a control task suggested that these changes were unlikely to reflect simple reward predictions. These data demonstrate a remarkable capacity of prefrontal neurons to adapt to specific operant requirements at the single-neuron level.

INTRODUCTION

Behavior is influenced by its consequences. For example, we reward and punish people so that they will behave in different ways. In a laboratory, animals take an action (e.g., lever pressing) more often if it is followed by a positive outcome (e.g., a drop of juice). A behavioral act that influences the outcome is called an operant response and operant conditioning involves the association between an operant response and its outcome (Balleine and Dickinson 1998; Skinner 1938; Thorndike 1911).

Various parts of the prefrontal cortex and basal ganglia have been implicated in the processing of action–outcome relationships. For example, neurons in the primate medial prefrontal cortex change the level of spiking activity depending on the combinations of action and outcome (Matsumoto et al. 2003). Striatal projection neurons track the changes in action–outcome contingency and predict animals’ choice of actions (Lau and Glimcher 2008; Samejima et al. 2005). Several lines of evidence suggest that the primate lateral prefrontal cortex (LPFC) contributes to operant behavior at a level higher than the control of specific motor responses. Single-neuron activity in the LPFC changes by behavioral demand (Asaad et al. 2000; Pan et al. 2008; Rainer et al. 1998; Sakagami and Niki 1994) and correlates with performance rates in operant tasks (Kobayashi et al. 2006). Patients with lesions in the LPFC lose intentional control of behavioral responses, making automatic responses that blindly follow given instructions (Lhermitte 1986; Luria 1973; Vendrell et al. 1995). These observations led to the commonly held notion that the LPFC guides purposeful flexible behavior (Fuster 2000; Miller and Cohen 2001; Rossi et al. 2009; Sakagami and Pan 2007; Tanji and Hoshi 2008; Wise 2008).

A fundamental aspect of purposeful and efficient behavior consists of the capability to interact with the environment to maximize desirable outcomes and minimize undesirable outcomes. If the environment is changeable, the same behavior may result in different consequences. Thus the environmental changes have to be detected and behavior needs to be updated. This process might be possible if the brain can assess the consequences of its own computation and modify the responses accordingly.

To investigate this issue, we provided reward to animals as a direct consequence of their neuronal activity and tested various types of activity–reward relations. Animals increased their chance to obtain juice reward by increasing or decreasing neuronal activity when the activity–reward relation was positive or negative, respectively. To measure baseline activity without operant requirement, we used a control schedule in which reward was given independent of neuronal discharges.

In the present study, we took neuronal activity as an operant response without requiring specific behavioral responses. This technique has been applied to control external devices using motor-related cortical activity (Chapin et al. 1999; Fetz 1969; Fetz and Baker 1973; Moritz et al. 2008; Musallam et al. 2003; Taylor et al. 2002). For example, spatially tuned cortical activity can be used to guide spatial cursor movement on a computer monitor (Musallam et al. 2004; Taylor et al. 2002). Prefrontal neurons may not carry kinematic information that can guide spatial response, but may change activity in the direction that optimizes its consequences.

METHODS

Subjects and surgery

We used two adult male Japanese monkeys (Macaca fuscata), monkeys A and B (weight: 9 and 8 kg, respectively). Preparation for the recording experiments included implanting a head holder, a chamber for unit recording, and a scleral search coil under general anesthesia. Details on surgical procedures are available elsewhere (Kobayashi et al. 2002). All surgical experimental protocols were approved by the Tamagawa University Animal Care and Use Committee and were in accordance with the National Institutes of Health Guide for the Care and Use of Animals.

Experimental design

Monkeys were mildly fluid deprived and sat in a primate chair inside a completely enclosed sound-attenuated room with their head
fixed. Visual stimuli were presented on a 21-in. computer monitor (GDM-F500, Sony, Tokyo or FlexScan T966, EIZO, Ishikawa, Japan) placed 70.5 cm in front of the monkey. A trial started with the onset of a central fixation spot (visual angle of 0.21°). After the monkey gazed at the fixation spot for 0.5 s (fixation period), a fractal picture (visual angle of 6.8°) was presented at the center for 1 s (cue period). After the picture was removed, the fixation spot remained for 0.7 s (delay period) until the time of reinforcement. Monkeys were required to fixate on the central spot during the fixation, cue, and delay periods. Breaking fixation during any of these periods resulted in immediate termination of the trial and a long time-out period (10 s) before the protocol was repeated.

To provide feedback about the operant response, a horizontal green bar was superimposed on the fractal picture during the cue period (Fig. 1A). The bar stretched dynamically in both directions along the horizontal axis, indicating proximity to the target, which was the edge of the fractal picture. The bar was invisible (width = 0) at the onset of the visual cue and it increased its width given any impulses generated in every refresh rate of the monitor (14.3 ms) (see Eqs. 1 and 2 in the following text). If the bar reached that target during the cue period, the color of the bar changed from green to red and a drop of liquid reward was given at the end of the delay period. If the target was not acquired, the bar remained green and predicted the absence of reward. In either case, the bar remained on the monitor during the delay period with a constant size and color. Activity of the same neuron in the two schedules was compared in a block design (see following text for details on the schedules).

A fractal picture was presented to indicate the time window for neuronal conditioning (cue period), to indicate the target of the bar (the edges of the picture), and to evoke visual response efficiently so that the effects of neuronal conditioning could be measured as changes in visually evoked activity. The same fractal picture was used in both control and operant schedules. To keep visual stimulation identical, there was no explicit cue that instructed the relevant schedule. A short break (~30 s) between adjacent blocks indicated a possible schedule change and it was possible for the animals to learn the relevant schedule within a block of trials. The block design was used because we used the mean activity in the control schedule as a criterion in the following operant schedule. The block design also allowed us to study the speed of neuronal adaptation to the activity–reward relation within each block of trials.

1. Control schedule (with a visual indicator). During the control schedule, the reward was given randomly at 50% probability and a bar indicator predicted the outcome. To indicate the proximity to the reward and to mimic the visual presentation in the operant schedule (see following text), the bar indicator stretched horizontally at a random pace. A computer generated virtual spike trains by a Poisson process at a given mean discharge rate in every trial. Every virtual spike stretched a green bar horizontally during the cue period. The incremental bar stretch was determined such that the bar reached the edges of the fractal picture randomly in 50% of the trials. The size of the bar indicated the proximity to the goal and the color of the bar predicted either the presence or the absence of reward outcome and thus the bar served as a Pavlovian reward predictor.

2. Operant schedule (with a visual indicator and fixed criterion). During the operant schedule, we rewarded animals based on the spiking activity of the particular neuron under investigation at that time. The task timeline and visual display were identical to the control schedule. Unlike in the control schedule, real spike trains sampled from prefrontal neurons controlled the bar in real time and determined reward outcomes. In the standard operant schedule, neurons were required to generate spikes beyond the criterion during the cue period (1 s). Once the criterion was reached, the color of the bar changed from green to red and remained on the monitor until the end of the delay period. If the criterion was not reached, the bar remained in green. The criterion in this schedule was set at the mean cue-period activity of the same neuron in the previous control schedule. Note that number of spikes during the cue period was generally normally distributed and there was negligible difference between mean and median values of spike counts (assumptions of normal distribution were accepted for all neurons by both Jarque–Bera and Lilliefors tests, P > 0.05). Therefore if the mean cue-period activity increased in the operant schedule, compared with the control schedule, more than half of the trials were rewarded (reward rate >50%). In other words, the growth rate of the reward-indicator bar was set so that the median spike count in the control schedule was just enough to reach the target.

In some of the experiments, neurons were further tested with the negative activity–reward relation (see following text). Reward was omitted for discharge rates above a specific criterion but was delivered for discharge rates below the criterion.

3. Activity–reward relation. In the operant schedule, reward outcome was conditional on the levels of cue-period activity (cf. Gibbon et al. 1974; Hammond 1980). We tested neurons in the positive activity–reward relation by giving reward when activity was above the criterion, i.e., p (reward | activity > criterion) = 1, and omitting reward otherwise, i.e., p (reward | activity < criterion) = 0. We also tested the negative activity–reward relation by giving reward only when activity was below the criterion, p (reward | activity < criterion) = 1, and omitting reward otherwise, i.e., p (reward | activity > criterion) = 0. In the control schedule, reward was given at 50% independent of the activity level.

4. Visual indicator. During the experiments with a visual indicator, the horizontal size of the bar (y) was controlled during the cue period according to

\[ y = \frac{w \cdot s(t)}{c} \quad \text{while} \quad s(t) \leq c \text{ (green bar)} \quad (1) \]

\[ y = w \quad \text{after} \quad s(t) > c \text{ (red bar)} \quad (2) \]

where w is the horizontal length of the fractal picture (8.3 cm), c is the number of the spikes set as the criterion, and s(t) is the cumulative number of virtual or real spikes in the control or operant schedule, respectively, at time t as measured from the onset of picture presentation (0 ≤ t ≤ 1,000 ms). In both schedules, the bar size indicated the

Fig. 1. The experimental design. A: behavioral task. After the fixation period (0.5 s), a picture with a bar indicator was presented for 1 s. The bar stretched either randomly (control schedule) or proportionally to the neuronal activity (operant schedule). If the bar reached that target (the edges of the fractal picture) during the cue period, the color of the bar changed from green to red and a drop of liquid reward was given at the end of the delay period (rewarded trial, top). If the target was not acquired, the bar remained in green and predicted the absence of reward (nonrewarded trial, bottom). In either case, the bar remained on the monitor during the delay period (0.7 s) with a constant size and color. B: a magnetic resonance imaging showing the lateral surface of the brain. The recording area is indicated by shade.
percentage of the criterion achieved. The step of bar stretch caused by every single spike (a virtual spike in the control schedule and a real spike in the operant schedule) was \( w/c \), as defined in Eq. 1. Since the horizontal length of the fractal picture \( (w) \) was fixed throughout the experiment, \( c \) determined the width of a single bar step. To control for the width of a single bar step in both schedules, \( c \) in the control schedule (the mean discharge rate of the virtual neurons) matched that of the real neurons. During the delay period, the bar was displayed on the monitor with constant size and color. The bar was presented in the same way by Eqs. 1 and 2 in the operant schedule with both positive and negative relations.

5. Experiments without a visual indicator. In separate experiments, we recorded single-neuron activity without visual feedback. The task was the same as described earlier, except that the bar was not displayed in either the control or the operant schedule, so the reward outcome could not be predicted based on feedback. The fractal picture was presented at the center of the monitor for the same duration as that in the experiments using the moving bar stimulus.

6. Operant schedule (with a visual indicator and a variable criterion). We devised a variable-criterion algorithm in which the criterion was updated in every trial depending on the outcomes of previous three trials. If the reward rate in the previous three trials was \( 3/3, 2/3, 1/3, \) or \( 0/3 \), the criterion in the next trial changed by \( +25.0 \), \( +6.25, -6.25, \) or \( -12.5\% \), respectively. The aim of this algorithm was to give reward in half of the trials overall by adjusting the criterion to the trial-by-trial change of neuronal activity, while keeping the schedule to reward activity higher than criterion and not to reward activity below criterion. The variable-criterion algorithm also matched the speed of bar stretching between operant and control schedules. As described in Eq. 1, the bar speed depended on \( w/c \). If neuronal activity in the operant schedule tended to increase, the criterion \( c \) was calibrated by the variable-criterion algorithm to larger values. Thus the width of the single bar step became smaller and the speed of bar stretch was kept similar even with higher activity.

The variability of the criterion has conceivably impaired the activity–reward contingency. We quantified the level of contingency by using point-biserial correlation coefficient. The point-biserial correlation is mathematically equivalent to the Pearson’s product moment correlation when one variable is continuously measured (e.g., activity [spikes/s]) and the other variable is dichotomous (e.g., reward [presence or absence]).

**History of training/recording**

The animals were first trained in a fixation task for about 10 days. Then they were trained in the control schedule with visual feedback in which they received a reward randomly in 50% of the trials if they maintained fixation. After this training for about 3 wk, recording was started using the operant schedule. Initially, neurons were recorded with visual feedback (monkey A, 45 neurons; monkey B, 28 neurons; Fig. 4A presents neurophysiological data in chronological order). We then tested 18 neurons from monkey A in the absence of visual cues and data sampling.

**Recording procedures**

Conventional extracellular single-unit recording was conducted while monkeys performed the tasks described earlier. We used the Tempo System (Reflective Computing, St. Louis, MO) and Matlab (The MathWorks, Natick, MA) for task control, behavioral monitoring, and data sampling.

An electrode (diameter, 0.25 mm; 1–5 MΩ; measured at 1 kHz; FHC, Bowdoinham, ME) was inserted into the brain through a stainless steel guide tube (diameter, 0.8 mm) that penetrated the pachymeninx and held above the cortex. A micromanipulator (MO-95-S, Narishige, Tokyo) advanced an electrode vertically to the cortical surface. Action potentials were amplified, filtered (500 Hz to 2 kHz), and matched in real time with a unique waveform to select out a single unit (MSD 3,21; Alpha Omega Engineering, Nazareth, Israel). The amplified electrode signal was placed in a buffer containing the last 100 samples and compared continuously with a template, which was constructed from eight equally spaced points separated by 0.1 ms. Each template was defined by the experimenter when a single neuron was first isolated. The sum of squares of the differences between eight points in the buffer (starting 0.4 ms from the beginning of the buffer and equally spaced at 0.1 ms) and the template was calculated. The hardware reported when this sum reached a minimum that was below an experimenter-defined threshold. A personal computer generated raster displays of neuronal activity on-line. The times of spike occurrences were stored with a time resolution of 1 ms. Eye positions were recorded at a sampling rate of 500 Hz using a magnetic search-coil technique (MEL-25, Enzanshi-Kogyo, Tokyo; Judge et al. 1980; Robinson 1963). The licking response was sensed by interruption of the beam of an infrared photodiode positioned in front of the spout through which liquid was delivered (E3X-NH, Omron, Kyoto, Japan). We monitored hand movement by using a video camera under infrared lighting (DCR-RP220, Sony).

**Recording sites**

The cylinder position was verified using magnetic resonance imaging (1.5 T, Sonata, Siemens, Munich, Germany) and postmortem histology.

**Data analysis**

After recording 30–40 trials in the control schedule, we analyzed statistically whether the neuron showed a significant cue response. If the cue-period activity was significantly higher than the baseline activity \( (P < 0.05, \) one-tailed paired \( t\)-test), the neuron was further tested in the operant schedule. We analyzed trials in which monkeys correctly performed a fixation task. To evaluate the effect of reward prediction, we compared the mean discharge rate during the cue period in the control schedule between rewarded \( (R_{\text{RWD}}) \) and nonrewarded \( (R_{\text{NRW}}) \) trials and quantified the effect by Pavlovian Index, defined as \( (R_{\text{RWD}} - R_{\text{NRW}})/(R_{\text{RWD}} + R_{\text{NRW}}) \). The main analyses compared the cue-period activity between operant and control schedules. Operant Index was defined as \( (R_{\text{OPE}} - R_{\text{CON}})/(R_{\text{OPE}} + R_{\text{CON}}) \), where \( R_{\text{OPE}} \) and \( R_{\text{CON}} \) indicate the mean cue-period activity in the operant and control schedules, respectively.

For our main tests with the fixed criterion, we used the following block design: i) control schedule, 2) positive-relation operant schedule, and 3) negative-relation operant schedule. Each schedule was tested once in a block of 30–40 correct trials in the preceding order. In neurophysiological experiments that compare activity across blocks of trials, it is important to ensure that any neural effects are not artifacts of the experimental design, such as electrode drift over time. We used several techniques to ensure that such artifacts did not influence our data. First, we tracked the activity of a single neuron by using a strict template-matching algorithm (see earlier text). Neural recording was discontinued when the waveform of sampled activity changed (~25% of neurons dropped out due to drift of the electrode or cell injury). Second, we tested whether baseline activity changed across blocks. The activity of the reported neuron did not change during intertrial intervals across blocks \( (P > 0.05, \) two-tailed \( t\)-test). Third, we tested neurons across multiple blocks of trials with different activity–reward relations. An activity change that followed different activity–reward relations across blocks suggested that the change was not due to uncontrolled influences.

We identified microsaccades with a protocol modified from Martinez-Conde et al. (2000) and Engbert and Kliegl (2003). We first smoothed the horizontal and vertical eye position data by a Gaussian kernel \( (\text{SD} = 25 \text{ ms}) \). We prepared an array of vectors representing the instantaneous motion of the eye \( (\sqrt{d^2x + d^2y}) \) at each 2 ms
interval. We determined the onset and offset of each microsaccade by velocity thresholds (2-5°/s) adjusted to the signal-to-noise level of each recording. Detected microsaccades were supervised by visual inspection on a computer monitor. Detection of microsaccade onset and offset determined peak velocity and duration. We also calculated microsaccade frequency during the cue period. To quantify the difference in microsaccade frequency between the two task schedules, we introduced the Microsaccade Index, defined as \((F_{OPE} - F_{CON}) / (F_{OPE} + F_{CON})\), where \(F_{OPE}\) and \(F_{CON}\) indicate the frequency of microsaccade during the cue period in the operant and control schedules, respectively. To examine the relationships between microsaccade and neuronal activity, we took neuronal activity before and after (0 to 250 ms) microsaccades during the cue period and compared with baseline activity without microsaccades (activity < −500 ms before and >500 ms after any microsaccade). Neurons with significant excitation or suppression in either the pre- or the postmicrosaccade period were labeled as microsaccade-related neurons (two-tailed t-test, \(P < 0.05\)).

RESULTS

This study investigated how neuronal activity in the LPFC could serve as an operant response. To this end, our primary experiment compared neuronal activity between the following two schedules. In the operant schedule, the animal obtained a drop of juice reward if neuronal activity was higher than a specific criterion and received no reward otherwise. The animal received visual feedback by a green bar that stretched in proportion to the neuronal activity and turned red when reaching the criterion (Fig. 1A). By contrast, the control schedule produced reward randomly and independently of neuronal activity. It served to measure the base rate of neuronal responding in the absence of operant conditioning. Thus the difference in activity between the two schedules reflected the induced level of operant neuronal response, as measured by the Operant Index (see METHODS).

We also used the control schedule to assess the potential role of reward prediction in increasing prefrontal activity. In this schedule, the bar stretched at a random pace unlinked to neuronal activity. The animal obtained a reward if the bar stretched fast and acquired the goal within the cue period and received no reward otherwise. The difference in activity between rewarded and nonrewarded trials would reflect the neuron’s sensitivity to reward prediction in the absence of operant control, as measured by the Pavlovian Index (see METHODS).

Prefrontal neuronal activity in operant and control schedules

An example of prefrontal neuronal activity in the two task schedules is illustrated in Fig. 2. In the control schedule, the bar stretched randomly, acquiring the goal in 50% of trials independent of the neuronal activity. In this situation, the neuron showed only a mild visual response to the cue (Fig. 2A). The cue response failed to differentiate the reward outcomes predicted by the bar (Fig. 2A, C). In the operant schedule, every spike generated by the tested neuron served to stretch the bar during the cue period. If the bar reached the edge of the picture, its color changed from green to red and the animal received a drop of liquid reward. This neuron showed significantly more spikes in this operant situation (Fig. 2B, blue line in Fig. 2D) compared with the control schedule (Fig. 2A, gray line in Fig. 2D). To reach the target and produce a reward in the operant schedule, the neuron was required to generate >20 spikes during the cue period, which was the mean cue-period activity in the control schedule. This target was acquired in 90% of trials (Fig. 2B, top) and the animal missed out on reward in only 10% of trials as a result of insufficient spikes generated during the cue period (Fig. 2B, bottom).

We sampled single-neuron activity widely from the LPFC (Fig. 1B). A total of 98 of 147 isolated neurons (67%) showed excitatory visual responses to the fractal picture. Of these, 73 neurons were fully tested in both control and operant schedules. The majority of these neurons showed significantly higher activity in the operant schedule than that in the control schedule (40/73 neurons, red symbols in Fig. 3A; \(P < 0.05\), two-tailed t-test). Although it was the activity during the cue period that determined the reward outcome, activity already started to increase during the initial fixation period in the operant schedule (Fig. 3B). This might be due to the block design that separated the two schedules, which already allowed the animals during the fixation period to expect the cue and prepare for the operant control.

A Control schedule

- Rewarded trials
  - Fix on
  - Cue on
  - Reinforcement

- Nonrewarded trials

B Operant schedule

- Rewarded trials
  - Fix on
  - Cue on
  - Reinforcement

- Nonrewarded trials

C Activity in rewarded and nonrewarded trials in control schedule

D Activity in control and operant schedules

\[ \text{Microsaccade Index} = \frac{F_{OPE} - F_{CON}}{F_{OPE} + F_{CON}} \]
We quantified the difference of neuronal activity between the operant and control schedules by using the Operant Index, which was positive in the majority of tested neurons, with its mean significantly greater than zero (0.10 ± 0.18, \(P < 0.001\), \(t\)-test). The Pavlovian Index quantified the effects of simple reward prediction by contrasting the activity between rewarded and nonrewarded trials in the control schedule. The mean Pavlovian Index was not significantly different from zero (-0.01 ± 0.07, \(P = 0.52\), \(t\)-test; Fig. 3C). These data suggest that prefrontal neurons increased activity when the activity could directly change the reward outcome.

As a consequence of the task design, the operant control of neuronal activity influenced the reward rate. If neurons were insensitive to activity–reward relations, reward would be expected at 50% in both control and operant schedules. However, the animals often obtained reward at higher rates in the operant schedule as a direct consequence of increased cue-period activity (53/73 blocks, \(P < 0.001\), paired \(t\)-test; Fig. 3D).

During the control schedule, reward probability was 50% in each block of trials. However, reward probability fluctuated in shorter ranges of time due to the randomization procedure. For example, by taking the past three trials into account, local reward probability ranged from 0 to 100% (reward rate of 0/3, 1/3, 2/3, 3/3 trials). We examined the influence of local reward probability on neuronal activity. Normalized cue-period activity is displayed as a function of reward rate in the past 3 trials. The activities that showed positive operant effect were averaged (40 neurons) and normalized by the mean activity in the operant schedule relative to the control schedule. Figure 3 depicts the averaged discharge rates from 40 neurons whose cue-period activity was significantly higher in the operant schedule (red line) than that in the control schedule (blue line). Error bar, SD.
circles with a dotted line; $P = 0.66$). In sum, neuronal activity was enhanced by local reward probability only during the operant schedule. The results suggest that operant neuronal conditioning was more efficient with successive positive reinforcements.

**Task experience and anatomical location**

To examine the possibility that the operant neuronal response developed through the animals’ experience in the task, we plotted the Operant and Pavlovian Indices in the chronological order of recording (Fig. 4A; filled squares and solid line, Operant Index; open diamonds and dotted line, Pavlovian Index). The variability of the indices did not correlate with the number of sessions that the animals had experienced ($P > 0.1$, Spearman’s rank-correlation test).

Figure 4B shows anatomical distributions of the tested neurons. Neurons that showed significant activity increases during the operant schedule (filled circles) were intermingled with neurons that did not show this property (open squares) within the LPFC. In sum, the degree of activity change during the operant schedule was largely independent of the animals’ experience on the task and anatomical location within the LPFC.

**Correlation between neuronal activity and movements**

Our task required animals to gaze at the central fixation spot, but did not require overt behavioral responses to produce reward. However, some of the prefrontal activities may have correlated with movements and animals may have been able to use such movements as operant response means for increasing the neuronal activity. To assess this possibility, we analyzed licking and eye movements.

First, we analyzed licking responses. Licking responses were minimal during the cue period (Fig. 5, A and B) and the duration of licking did not differ between the operant and control schedules (Fig. 5C; $P > 0.05$, two-tailed paired $t$-test). To be entirely free from the potential confound of licking, we reanalyzed the neuronal data by excluding trials in which animals made anticipatory licking. However, even when considering only lick-free trials, the majority of recorded neurons still showed the operant-related activity enhancement (34/61 neurons, red symbols; $P < 0.05$, two-tailed $t$-test, Fig. 5D). Thus licking did not seem to provide a simple explanation for the neuronal operant effect.

Next, we examined eye movements. Figure 6 shows the temporal profile of the eye position (Fig. 6, A and C) and angular velocity (Fig. 6, B and D) during the control and operant schedules averaged across all the recording sessions. The animals made occasional microsaccades (Fig. 6, A–D). The small eye movements after cue onset could be reflexive microsaccade or vergence. If the microsaccades were directly linked to neuronal excitation, the animals could gain reward by making frequent microsaccades during the operant schedule, but not during the control schedule. Interestingly, microsaccades were less frequent during the operant schedule compared with that during the control schedule (Fig. 7A; $P < 0.05$, two-tailed $t$-test). We quantified the difference in microsaccade frequency between the two schedules by using a Microsaccade Index (see METHODS). The Microsaccade Index correlated only insignificantly with the degree of operant-related neuronal activity changes measured by Operant Index (Fig. 7B; $P > 0.1$, Pearson’s correlation coefficient test). This result indicates that the neuronal operant effect was not explained by the frequency of microsaccades.

Further analysis on microsaccade kinematics revealed small schedule-dependent variations in peak velocity, amplitude, and duration (Table 1). To be strictly free from microsaccade confounds, we reanalyzed neuronal data based on microsaccade-free trials. Data from 49 neurons had a sufficient number of trials without microsaccades (>3 trials in both operant and control schedules). Of these neurons, 12 showed activity increases (red symbols in Fig. 7C) and 3 showed activity decreases (blue symbols) during the operant schedule compared with that during the control schedule ($P < 0.05$, two-tailed $t$-test). Thus removing trials with microsaccades failed to affect the main finding of neuronal operant conditioning.
To positively identify neurons with microsaccade-related activity, we constructed microsaccade-triggered spike histograms (Fig. 7, D–F). Only a limited fraction of our database (37 neurons) exhibited microsaccade-related activity (excitation, 5 neurons [red lines]; suppression, 12 neurons [blue lines]; two-tailed t-test, \( P < 0.05 \); see METHODS). Most neurons that showed a positive operant effect lacked microsaccade-related responses (30/40 neurons; black lines in Fig. 7D). Conversely, only a small fraction of microsaccade-related neurons increased activity during the operant schedule (2/17 neurons; red lines in Fig. 7D). These results argue against the possibility that the activity enhancement during the operant schedule was mediated by uncontrolled microsaccade responses.

**Controls for reward rate and bar speed**

We found that the bar stretched faster and the reward rate was higher in the operant schedule compared with those in the control schedule (Figs. 8A and 3D). Although these were direct consequences of higher neuronal activity in the operant schedule compared with that in the control schedule, it is possible that the faster bar movement and the higher reward rate contributed to the increase in neuronal activity.

To control for these potential confounds, we conducted separate experiments in which the criterion of the operant schedule was adjusted in every trial so that the bar movement and reward rate were similar to those in the control schedule but the activity was still significantly higher during the operant schedule. This method successfully matched the bar speed and reward rate between the control and operant schedules (Fig. 8D). This was mainly due to the difficulty to stretch the bar and to obtain reward (Fig. 8D). This method successfully matched the bar speed and reward rate between the control and operant schedules (Fig. 8D), but the activity was still significantly higher during the operant schedule compared with that during the control schedule in 12
of 20 tested prefrontal neurons (Fig. 8E, red circles, increases; blue circle, decrease; P < 0.05, two-tailed t-test). This result is also reflected in the increase of the criterion across operant trials (Fig. 8, D and F). Taken together, neuronal activity increased in the operant schedule even when the visual stimulation of bar movement and the reward rate were controlled.

Although neuronal activity increased systematically, we observed a limited effect of neuronal operant conditioning when compared with the main result shown in Fig. 3A. A possible explanation is the impaired activity–reward contingency due to the variation of criterion. For example, the same level of activity was rewarded when the criterion was low but not rewarded when the criterion was higher. We quantified the level of activity–reward contingency by using the point-biserial correlation coefficient (see METHODS). This contingency measure would produce coefficients of 1.0 and 0 for perfectly positive and perfectly random relations, respectively. As expected, this coefficient was smaller with the variable criterion (0.60 ± 0.10, mean ± SD; closed bars in top histogram of Fig. 8G) compared with the fixed criterion (0.71 ± 0.24; open bars). Thus the variability of the criterion reduced the level of activity–reward contingency. Accordingly, the induced activity changes were smaller with the variable criterion (Operant Index of 0.066 ± 0.079; large squares) compared with the fixed criterion (0.108 ± 0.195; small circles). The present result is consistent with the idea that activity in the tested neurons was sensitive to the level of operant contingency.

**Operant schedule with negative activity–reward relation**

The data presented so far were obtained in the operant schedule in which neuronal activity was positively associated with reward. In the next experiment, we tested whether prefrontal neurons could decrease activity to gain access to reward

### Table 1. Microsaccade parameters

<table>
<thead>
<tr>
<th>Schedule</th>
<th>Frequency, N/s</th>
<th>Peak Velocity, deg/s</th>
<th>Amplitude, deg</th>
<th>Duration, ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operant</td>
<td>0.63 ± 0.48*</td>
<td>37.90 ± 10.69*</td>
<td>0.89 ± 0.31*</td>
<td>16.7 ± 4.4*</td>
</tr>
<tr>
<td>A: 0.52 ± 0.46*</td>
<td>A: 38.47 ± 11.30*</td>
<td>A: 0.95 ± 0.33*</td>
<td>A: 18.0 ± 4.4*</td>
<td></td>
</tr>
<tr>
<td>B: 0.81 ± 0.47</td>
<td>B: 37.27 ± 9.94*</td>
<td>B: 0.84 ± 0.28*</td>
<td>B: 15.4 ± 3.9</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>0.72 ± 0.54</td>
<td>36.90 ± 11.36</td>
<td>0.85 ± 0.32</td>
<td>16.2 ± 4.3</td>
</tr>
<tr>
<td>A: 0.63 ± 0.54</td>
<td>A: 34.84 ± 12.18</td>
<td>A: 0.84 ± 0.35</td>
<td>A: 16.9 ± 4.5</td>
<td></td>
</tr>
<tr>
<td>B: 0.85 ± 0.50</td>
<td>B: 39.10 ± 9.97</td>
<td>B: 0.87 ± 0.28</td>
<td>B: 15.6 ± 4.1</td>
<td></td>
</tr>
</tbody>
</table>

Values are means ± SD. * and † indicate significantly higher and lower values, respectively, in the operant schedule compared with the control schedule (P < 0.05, two-tailed t-test). A and B refer to the two animals used.
(negative activity–reward relation; see METHODS). As before, the bar stretched in proportion to the level of neuronal discharges.

Out of 73 neurons tested in the standard operant schedule, we further investigated 45 neurons with the negative-relation schedule. Although we used the same cue picture in all schedules, short breaks of about 30 s between trial blocks served to indicate the schedule change to the animals.

Figure 9A shows the trial-by-trial change of population activity in three schedule types. When we rewarded the animals independent of the neuronal activity (block 1, control schedule), the activity was initially high but decreased quickly in the first few trials and remained nearly constant thereafter. When activity higher than the criterion was rewarded (block 2, positive-relation schedule), activity remained high throughout the block. Subsequently, we shifted the activity–reward relation from positive to negative by rewarding only low-activity trials (block 3, negative-relation schedule) and average activity decreased slowly toward the baseline level. Figure 9B shows normalized activity of individual neurons in the positive (left) and negative (right) relation schedules after activity reached asymptote (mean activity after 20 trials). Compared with the baseline in the control schedule, cue responses increased by 26.9% on average in the positive-relation schedule (block 2). The increases were statistically significant in 55.6% of neurons (25/45 neurons, closed circles in Fig. 9B, left, P < 0.05, one-tailed t-test). By contrast, cue responses decreased by 17.5% from the elevated level when the operant schedule
required lower than control activity (block 3). The decreases were statistically significant in 22 neurons (48.9%; \( P < 0.05 \), one-tailed \( t \)-test). Furthermore, the asymptotic activity decreased significantly from the control activity in 31.1% of neurons (14/45 neurons, closed circles in Fig. 9B, right; \( P < 0.05 \), two-tailed \( t \)-test). The average asymptotic activity in block 3 stayed above the baseline level (normalized activity = 1.10 ± 0.54 [mean ± SD]) and activity of 33.3% of neurons remained significantly above the baseline (15/45 neurons, \( P < 0.05 \), two-tailed \( t \)-test).

Figure 9C illustrates the relationships between asymptotic, normalized neuronal activity, and reward rate in each schedule (>20th trial). The solid and dotted regression lines are based on data in the positive (open squares) and negative (closed squares) relation schedules, respectively. As expected, depending on the given schedule, reward rate increased with either increase (positive slope = 0.38, \( r^2 = 0.51 \), \( P < 0.001 \)) or decrease (negative slope = −0.60, \( r^2 = 0.68 \), \( P < 0.001 \)) of neuronal activity. We designed the control schedule to produce reward independent of spiking activity and indeed found an insignificant relation between activity and reward rate (\( r^2 = 0.01 \), \( P = 0.61 \); data not plotted). The result indicates that neuronal activity (abscissa) determined reward rate (ordinate) based on the given activity–reward relationship. However, the inverse was not the case; reward rate did not predict neuronal activity without taking into account the activity–reward relationship (\( r^2 = 0.00 \), \( P = 0.81 \)). These data support the notion that neuronal activity reflected the operant schedule but not simply the prediction of reward.

Next, we examined whether the changes of activity in the upward and downward directions were quantitatively correlated. Figure 9D shows the percentage activity changes of each neuron from block 1 to block 2 (abscissa) and from block 2 to block 3 (ordinate). The magnitude of activity changes were significantly correlated between upward and downward directions (correlation coefficient, −0.39, \( P < 0.01 \), permutation test); neurons that showed a large activity increase in block 2 generally showed a large activity decrease in block 3.

**Operant schedule without visual feedback**

As with many other studies (Fetz 1969; Fetz and Baker 1973; Musallam et al. 2004; Taylor et al. 2002), our experiment involved continuous feedback that provided visual information of ongoing neuronal activity, together with the juice delivery at trial end. Thus successful neuronal conditioning by juice reward might have been aided by the visual feedback. To examine the contribution of this feedback to neuronal conditioning, we repeated the experiment by removing the visual bar on the ceiling (schedules 4-6 in Table 1). The data from negative activity–reward relations are plotted with closed squares and the regression line is drawn with a dotted line (slope = −0.60, \( r^2 = 0.68 \)). Differences in mean cue-period activity between the 1st and 2nd blocks (abscissa) and between the 2nd and 3rd blocks (ordinate) are plotted with closed circles and open squares, respectively. The result indicates that neuronal activity reward relations (closed circles, left; open squares, right) strongly reflected the rewarded operant schedule.

**FIG. 9.** Directional flexibility of operant neuronal conditioning. A: trial-by-trial change of population activity (\( n = 45 \)) in 3 schedule types. The cue-period activity is normalized to the control activity (mean activity in the control schedule) and plotted as a function of trial count in each block. Error bar, SE. B: normalized cue-period activity of each neuron tested in the positive (left) and negative (right) activity reward relations (\( n = 45 \)). In all, 25 neurons (55.6%) increased activity significantly above 1.0 with positive activity–reward relations (closed circles, left); 14 neurons (31.1%) decreased activity significantly below 1.0 with negative activity–reward relations (closed circles, right). C: relations between the normalized activity and reward rate in each block of trials for each neuron. Data from positive activity–reward relations are plotted with open squares and the regression line is drawn with a solid line (slope = 0.38, \( r^2 = 0.51 \)). Data from negative activity–reward relations are plotted with closed squares and the regression line is drawn with a dotted line (slope = −0.60, \( r^2 = 0.68 \)). D: differences in mean cue-period activity between the 1st and 2nd blocks (abscissa) and between the 2nd and 3rd blocks (ordinate), with corresponding bar graphs (black bars, \( P < 0.05 \); white bars \( P > 0.05 \), \( t \)-test). Activity generally increased from 1st to 2nd block and decreased from 2nd to 3rd block. The degrees of activity changes from 1st to 2nd and 2nd to 3rd blocks were significantly correlated (correlation coefficient, −0.39, \( P < 0.01 \), permutation test). The line shows the first principal component vector.

**FIG. 10.** Activity in the control and operant schedules without visual feedback. A: comparison of cue-period activity in the 2 schedules (\( n = 18 \)). White circles, \( P > 0.05 \); red circles, increases, \( P < 0.05 \); blue circles, decreases, \( P < 0.05 \) (\( n = 2 \)), 2-tailed \( t \)-test. Error bar, SE. B: time course of the population activity in the control (blue) and operant (red) schedules. Error bar, SE.
**Discussion**

This study shows that neuronal activity in the primate LPFC can serve as an operant response for obtaining reward. Prefrontal activity showed remarkable changes according to the given activity–reward relation. The changes in discharge rate were not explained by simple visual stimulation because prefrontal activity failed to change when reward delivery was not under operant control. Our data suggest that the primate LPFC neurons can influence and optimize the occurrence of outcomes meeting the given operant requirements. Neuronal sensitivity to the operant schedule may possibly reflect the animals’ intentional processes; the animals may have learned to modify their neuronal activity to manipulate the bar indicator. The following arguments aim to assess the potential mechanisms underlying the observed neuronal activity changes, including reward prediction, attention, arousal, and mediating behavior.

**Dissociating neural processes of operant and simple reward prediction**

Our operant schedule involved both the control of reward outcome and the prediction of reward. To dissociate between the two processes, we used the control schedule, which included simple reward prediction without involving operant control. We tested the effect of reward prediction by contrasting neuronal activity between rewarded and nonrewarded trials and found little impact of reward prediction on neuronal activity, except for a few neurons (see delay period activity in Fig. 2A and horizontal outliers in Fig. 3C). This result suggests that presently tested prefrontal neurons were insensitive to the prediction of immediate reward outcomes in the absence of operant control.

One may argue that the high level of activation during the operant schedule could be explained by high reward probability. If this were true, neuronal activity should increase when animals experience frequent rewards even in the absence of operant control. Analysis of local reward probability suggests that this is not the case (Fig. 3E). The processes of reward prediction and operant control were also dissociable in the tests with positive and negative rewards; neuronal activity changed in the direction of the operant requirement. Thus activity did not correlate unconditionally with the probability of reward (Fig. 9C).

It should be noted that, compared with the main result with the fixed criterion, the operant enhancement effect was considerably reduced when we used the variable criterion. Although the level of operant contingency may explain the difference between the two experiments (Fig. 8G), overall reward rate may also have influenced the level of neuronal activations. Thus the present study does not exclude the influence of Pavlovian reward prediction on prefrontal activity.

**Operant control versus general processes of attention and arousal**

General attention and arousal are likely to play an important role in operant responses. However, the animal psychology literature presents a more complex picture; attention and arousal may increase with larger amounts of reward but do not consistently increase operant performance (e.g., Hendry 1962; Killeen 1985; Reed and Wright 1988; see Reed 1991). Because prefrontal neurons appear to be sensitive to operant requirements, it would be necessary to assess a potential contribution of attention and arousal in the observed neuronal operant control. Three arguments may be worth considering.

First, the experiment required eye fixation within a given tolerance window. The animals showed similar gaze accuracy between the operant and control schedules (Fig. 6, A and B), possibly reflecting comparable levels of visual attention. Therefore the effects of operant conditioning may not be explained by gaze accuracy. Second, the activity changed in a specific temporal pattern during the operant schedule: activity started to build up after the fixation onset, peaked during the cue period, and came back to the baseline well before reinforcement (Fig. 3B). By contrast, general attention and arousal would be sustained until reinforcement occurred. Third, if the operant control of prefrontal neurons simply reflected attention to the visual stimuli, this might indeed have optimized reward outcomes during the operant schedule when the activity–reward relation was positive, but not when it was negative. Taken together, an attention account cannot be completely ruled out as explanation for neuronal operant control but is unlikely to explain all observed neuronal modulations.

**Direction of activity change in neuronal conditioning**

We tested different operant requirements in separate blocks of trials interrupted by short breaks. At the beginning of each block, the cue response was higher than control activity and subsequently adapted to each specific schedule (Fig. 9A). In fact, increasing the activity at the beginning of a new schedule was a reasonable strategy because we required high activity more often (73 blocks) than low activity (45 blocks). Furthermore, the animal would immediately gain reward if the new schedule required high activity and would not miss out any reward in the control schedule. Finally, many fewer trials would be required to identify the current schedule at high activity compared with baseline because activity at the baseline level was associated with 50% reward in any of the given schedules.

Neurons changed activity flexibly in both upward and downward directions. However, more neurons achieved significant upward shifts than downward shifts from baseline (closed circles in Fig. 9B). A previous study on the primary motor area also showed that changes below baseline were more difficult to obtain than upward changes (Fetz and Baker 1973). The electrically evoked spinal stretch reflex (H-reflex) showed bidirectional changes, but different underlying mechanisms are suggested for the changes in different directions (Carp et al. 2006). The asymmetry observed in the previous and present studies might imply that it is more difficult to suppress than to enhance neuronal activity.

**Potential confounds of movement**

In principle, any movement that influenced prefrontal activity could constitute a confound for the neuronal operant response. However, the regularly monitored licking response did not show systematic differences during the cue period between the operant and control schedules (Fig. 5). Only a limited number of the sampled neurons showed correlation to the eye movement and the neuronal operant effect was observed even after excluding trials with microsaccades (Fig. 7). Our monitoring of the animals’ limbs
and trunk by video cameras revealed very few movements. In addition, only 1.4% of the task-related activity in the LPFC is related to arm movement (Hoshi et al. 2000). The limited motor representation in the LPFC would make it difficult for the animals to use movements for controlling prefrontal activity in the present study.

To make use of movements for changing neuronal activity in the present task, animals would have needed to find out what movement might be linked to the specific activity of the neuron being recorded at a given moment. The effective movement must have varied from neuron to neuron, if there was any. The different movement–reward associations would conceivably take a long time to learn. Indeed, in similar tests with the primary motor cortex, naïve animals usually needed substantial training sessions before they could reliably use the activity of each neuron and the ability improved with experience (Fetz 1969; Fetz and Baker 1973; Moritz et al. 2008). In contrast, our animals did not improve the efficiency of activity control by their experience within a session (Fig. 9A) or across sessions (Fig. 4A). In conclusion, the use of movements does not seem to be an efficient strategy to change neuronal activity and thus may not constitute a major confound in the present study.

Effects of visual stimuli

We studied visually sensitive neurons in the LPFC and found that their visual response could be modified in both directions according to the operant requirement (Fig. 9). Because the bar size changed unidirectionally in proportion to the neuronal activity, the bidirectional modulation of neuronal response could not be explained by simple visual stimulation. Also, the operant neuronal responses occurred even when the speed of the bar movement strictly matched between the two schedules (Fig. 8).

Although visual stimulation did not explain the observed operant neuronal control, continuous visual feedback played an essential role in the animals’ capacity for neuronal conditioning (Fig. 10). Perhaps the visual signal provided secondary reinforcement during the delay before the primary juice reinforcer occurred. Previous studies also demonstrated successful neuronal conditioning by giving real-time feedback (Chapin et al. 1999; Fetz and Baker 1973; Musallam et al. 2003; Taylor et al. 2002). The present data confirm that the operant conditioning in the LPFC greatly benefits from real-time feedback, but that the conditioning itself is not explained by visual stimulation.

Operant conditioning of neuronal activity and behavioral control

Behavioral adaptation to the changing environment would be a key advantage for animals in their quest for survival. Thus evolutionary pressure might have contributed to the development of flexible behavior that allows more efficient foraging than stereotyped behavior. In the present study, we took the measured neuronal activity as an operant response by using a brain–computer interface. It remains an open question how the observed prefrontal activity contributes to animal behavior in natural situations.

Previous single-unit studies that required operant behavioral responses showed that reward prediction boosts movement preparatory activity in the superior colliculus (Ikeda and Hikosaka 2007) and premotor area (Roesch and Olson 2004). Since we found that the operant signals in the LPFC did not depend on movement, the movement-related areas would need to receive additional action signals to execute movements. Thus the role of the LPFC might be to enhance the action signals to ensure the motor execution when operant control is required.

Further investigation is required to clarify whether the LPFC constitutes a neural substrate for operant behavior. Testing operant behavior after lesions in the LPFC would provide important evidence. It would be also important to examine operant neuronal conditioning in other brain structures. For example, the orbitofrontal cortex and striatum, which are also suggested to be involved in goal-directed behavior, are interesting candidates for neuronal conditioning experiments. Indeed, recent imaging studies showed that hemodynamic responses in these structures change in parallel with operant contingency in behavioral tasks (Tanaka et al. 2008; Tricomi et al. 2004). Together with simultaneous recordings in downstream movement-related areas, neuronal conditioning experiments may be useful to clarify the top-down control mechanisms underlying voluntary, intentional behavior.

Acknowledgments

We thank M. Koizumi, K. Nomoto, A. Noritake, and X. Pan for technical assistance; M. Watanabe and A. Dickinson for helpful discussion; and K. Sawa, P. Tobler, M. O’Neill, W. Stauffer, and E. Fetz for critical reading of the manuscript.

Grants

This work was supported by a Japan Society for the Promotion of Science grant to S. Kobayashi, a Wellcome Trust grant to S. Kobayashi and W. Schultz, a Human Frontier Science Program grant to W. Schultz and M. Sakagami, a Precursory Research for Embryonic Science and Technology Organization/ Japan Science and Technology grant, a Grant-in-Aid for Scientific Research on Priority Areas, and a Tamagawa University Center of Excellence grant from the Ministry of Education, Culture, Sports, Science and Technology (Japan) to M. Sakagami.

References


