Behavioral/Cognitive

# Representation of Functional Category in the Monkey Prefrontal Cortex and Its Rule-Dependent Use for Behavioral Selection

Ken-Ichiro Tsutsui,* Takayuki Hosokawa,* Munekazu Yamada, and Toshio Iijima

Laboratory of Systems Neuroscience, Tohoku University Graduate School of Life Sciences, Aoba, Sendai 980-8577, Japan

Humans, monkeys, and other animals are considered to have the cognitive ability to use functional categories—that is, stimulus groups based on functional equivalence independent of physical properties. To investigate the underlying neural mechanisms of the use of functional categories, we recorded single-unit activity in the prefrontal cortex of monkeys performing a behavioral task in which the rule-dependent usage of functional category was needed to select an appropriate response. We found a neural correlate of functional categories on the single-neuron level and found that category information is coded independently of other task-relevant information such as rule and contingency information. Analysis of the time course of the information activation suggested that contingency information used for action selection is derived by integrating incoming category information with rule information maintained throughout a session. Such neural computation can be considered as the neural background of flexible behavioral control based on category and rule.

*Key words:* monkey; single-unit recording; prefrontal cortex; functional category

---

**Significance Statement**

Perceptual categories are based on perceptual similarity, whereas functional categories are based on functional equivalence independently of their physical properties. The neural background of perceptual categories has been investigated in a number of studies, but that of functional categories has been investigated in only a few. In the present study, we found neural correlates of functional categories on the single-neuron level. Further analysis suggested that functional category information and rule information is integrated within the prefrontal cortex to derive contingency information necessary for action selection. This study has shown how the brain uses the category to select an action appropriate to the behavioral context.

---

## Introduction

Humans, monkeys, and other species of higher vertebrates have the ability to group similar items into a category (Bruner et al., 1966; Sidman and Tailby, 1982; Lazar et al., 1984; Vaughan, 1988;

Dube et al., 1993; Schusterman and Kastak, 1998; Ward-Robinson and Hall, 1999; Jitsumori et al., 2002; Urcuioli et al., 2006; Truppa et al., 2010). Major theories of cognitive psychology suggest that there are two types of categories: perceptual and functional (Bruner et al., 1966). A perceptual category is based on perceptual similarity among stimuli. In contrast, a functional category, or concept, is a set of stimuli that are regarded as equivalent independently of their physical properties and elicit the same behavioral response. Perceptual and functional categories may be based on different neural mechanisms because the perceptual categories are based on sensory nature of stimuli, whereas the functional categories are based on experience and learning. Neural representation of perceptual categories (i.e., common response to perceptually similar stimuli) has been found in both the prefrontal cortex (PFC) (Freedman et al., 2001, 2003; Cromer et al., 2010; Roy et al., 2010; Antzoulatos and Miller, 2011) and in the posterior association cortices such as the inferotemporal cortex (Vogels, 1999; Freedman et al., 2003; Freedman and Assad, 2006; Kiani et al., 2007) and the posterior parietal cortex (Freedman and Assad, 2006; Goodwin et al., 2012). Meanwhile, con-

cerning functional categories, neurons in the PFC (Brincat and Miller, 2015), as well as the inferotemporal (Tomita et al., 1999; Mogami and Tanaka, 2006) and perirhinal cortex (Mogami and Tanaka, 2006), have been reported to respond commonly to arbitrary visual stimuli that are associated with a specific stimulus or outcome. Those studies, however, did not examine whether the recorded neurons coded the functional category itself or if they were simply being activated in relation to the retrieval of the representation of a specific stimulus or outcome that the stimuli were associated with. Indeed, many prefrontal neurons (Watanabe, 1990; Asaad et al., 1998; Pasupathy and Miller, 2005; Pan et al., 2008) and perirhinal cortex neurons (Ohyama et al., 2012) have been reported to reverse their stimulus selectivity when stimulus–outcome or stimulus–response relations were reversed, suggesting that those neurons are coding the predicted outcome or response. Therefore, it is still unclear whether functional categories are coded explicitly on the single-neuron level.

The behavioral paradigm that we used to investigate the neural correlate of functional categories was a "group reversal" task, a modification of the task developed by Vaughan (1988). Four arbitrary selected stimuli (group A) were associated with fruit juice and the other four (group B) were associated with saline (the A+/B− condition), so that monkeys had to learn to make acquisition and avoidance reactions, respectively, to the stimuli of groups A and B, when each of these eight stimuli was presented in a different random order in a trial block. After several blocks, the stimulus–outcome contingencies were reversed for all stimuli: the stimuli in subset A were associated with saline and those in subset B were associated with juice (the A−/B+ condition). The contingencies were repeatedly reversed after being fixed for six to 12 blocks. With this behavioral task, we dissociated the category (stimulus group) from the contingency (outcome or response). Neurons that respond to the same subset of stimuli in both the A+/B− and A−/B+ conditions can be regarded as category-coding neurons and neurons that respond to one subset in the A+/B− condition and to the other subset in the A−/B+ condition can be regarded as contingency-coding neurons. In our previous study analyzing the same dataset of prefrontal neuron activity (Yamada et al., 2010), we found rule-dependent activity, such as being active in one condition and inactive in the other. In the present study, we analyzed the data regarding category, rule, and contingency as three major factors for neuron activity variance and looked for an explicit and independent coding of category information and its integration with rule information to retrieve contingency information for future behavioral selection.

## Materials and Methods

*Subjects.* Two male Japanese monkeys (*Macaca fuscata*) were used as experimental subjects. Throughout the experiments, they were treated in accordance with the National Institutes of Health's *Guide for the Care and Use of Laboratory Animals* and the Tohoku University's *Guidelines for Animal Care and Use*. This project was approved by the Center for Laboratory Animal Research of Tohoku University. Monkeys were housed individually in a cage in a room with natural lighting.

*Apparatus.* In the laboratory, a monkey sat on a home-made primate chair to which a touch-key sensor (Supertech) was attached. A liquid-crystal display (LCD) (Prolite E431S; Iiyama) on which the visual stimuli were presented was placed 35 cm in front of the monkey at eye level. A double-spout device through which two types of liquid were delivered as reward and punishment was placed in front of the monkey's mouth. An infrared sensor (Supertech) was attached to the spout device to monitor the monkey's spout-licking behavior. All of the abstract figures used as visual stimuli had been generated as computer graphics files (bitmap format, 142 × 142 pixels) and stored on a hard disk. When a visual

stimulus was presented on the LCD, its size extended to 6° × 6° in visual angle. Orange juice (Toris Conc; Suntory) was used as a reward and concentrated saline (7%) was used as a punishment. Visual stimulus presentation was controlled by graphics presentation software (Presentation; Neurobehavioral Systems) on a personal computer (xSeries100; IBM Japan) that was synchronized with and controlled by a home-made host computer. Juice/saline delivery was controlled by the opening of solenoid valves (CKD) remotely controlled by the host computer.

*Behavioral task.* Monkeys performed a task in which they had to adapt their behavior to a repeated stimulus–outcome reversal with a set of eight distinct visual stimuli and two distinct outcomes (Fig. 1). In every trial, a visual stimulus was presented to the subject and then its associated outcome (juice or saline) was delivered through the spout in front of the mouth. Therefore, the stimulus served as a cue to predict the outcome of the trial. The correct response at the time of liquid delivery was to lick if the preceding stimulus had been associated with juice (reward acquisition by "go" response) and to withhold licking if it had been associated with saline (punishment avoidance by "no-go" response). The precise time sequence of the task events was as follows. When a red fixation spot (4 mm in diameter) appeared at the center of the LCD, the monkey touched the key and fixated on the spot. After a pre-cue period (either variable between 1.0 and 1.5 s or fixed to 1.25 s), a visual stimulus was presented at the center of the LCD for 0.75 s. After cue offset, a delay period lasted for either a variable duration between 0.75 and 1.5 s or a fixed duration of 1.25 s. When the fixation spot turned from red to green, the monkey released the key and the fixation spot disappeared. After the key release, another delay period lasted for 0.5 s (or 0 s in some training sessions) and then either juice or saline was delivered. For the pre-cue and first delay period, variable timing was used during the training sessions and initial 3 months of the recording sessions. Later, fixed timing was used. We counted a trial as correct if a monkey licked the spout within a certain time window including the juice delivery time (from 200 ms before juice onset to 500 ms after juice onset) in juice trials and did not lick the spout within the time window in saline trials. We repeated the same trial (correction trial) if a monkey terminated a trial erroneously by fixation break or early key release so that the monkeys learned to complete a trial even if the predicted outcome was saline. In each trial, a stimulus was chosen pseudorandomly in such a way that each of the eight stimuli in a set was used once in a block of eight trials. The rule of stimulus–outcome associations, four of eight stimuli being associated with juice and the rest being associated with saline, was constant within a session of six to 12 blocks (48–96 trials) and, between sessions, they were reversed without any explicit cue. The first trial after the reversal was always a saline trial so that a monkey would notice the reversal. Therefore, the punishment, an unexpected delivery of saline in the first trial after the rule reversal, served as a cue for the rule reversal. Eye movements were monitored using an infrared eye movement recording system (ETL-200; I-scan). The trial was canceled immediately if the eye position exceeded the limit of 1° from the fixation spot. The eye movements during each trial were also examined offline to confirm eye fixation. Before the single-unit recording, the monkeys were trained with three different regular stimulus sets.

*Single-unit recording.* Before single-unit recording, a stereotaxic MRI scan of the brain was taken for each monkey. Then, a head-fixation device and a recording chamber were implanted in a standard surgical procedure using pentobarbital (Nembutal; Dainippon Sumitomo Pharma) for general anesthesia. The recording chamber was cylindrical in shape with an inner diameter of 18 mm. The skull over the lateral PFC (LPFC) was removed, giving a skull opening size 18 mm in diameter, and the cylindrical recording chamber was implanted at 45 degree of angle over the opening of the skull. After the monkey had recovered from surgery, extracellular single-unit recording was performed in the PFC during the performance of the task by using an Epoxylite- or glass-coated tungsten microelectrode (impedance: 1.5 M at 1 kHz; FHC) attached to a hydraulic x–y stage microdrive (MO-97S; Narishige) to penetrate and advance into the brain. Electrophysiological signals were amplified (10,000 times) and band-pass filtered (low cut: 100 Hz; high cut: 10,000 Hz) with a standard biophysical amplifier (BioAmp A2-v6; Supertech) and displayed on an

oscilloscope (CS-4125A; Kenwood). The amplified electrophysiological signals were also audibilized and presented to the experimenter through a speakerphone. The action potentials of isolated neurons were sorted by a window-discriminator (DDIS-1; Bak Electronics) and displayed on a digital storage oscilloscope (DCS-7040; Kenwood). The recorded electrophysiological signals were digitized at 25 kHz by an analog-to-digital conversion interface (Power 1401; CED) and then stored on the hard disk of the personal computer. The times of the detected action potentials were also stored on that hard disk, together with those of key touch/release, licking responses, eye movements, and task events. Rastergrams and histograms showing the neuronal activity in each cue condition were displayed online on an LCD video screen. In daily recording sessions, we randomly selected one of the three stimulus sets at the beginning of the session. We searched for a neuron while a monkey performed the task with the stimulus set. After isolating a neuron, we recorded the neuronal activity with the same stimulus set, during which the rule was normally reversed from three to five times. Between the reversals, the rule was fixed for 6–12 blocks (48–96 trials). After finishing the recording of neuronal activity, we usually changed the stimulus set to another one by random selection and then began to search for another neuron.

*Offline analysis of recorded neuronal activity.* The recorded single-unit activity was analyzed offline using both a home-made data analysis program and a commercial one (Matlab; The MathWorks). We analyzed neuronal activity using the following multiple regression equation with four variables:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2 + b_4 x_3$$

where $y$ denotes the spike rate during the cue period, $b_0$ the offset, $b_1$–$b_4$ the coefficients for the explanatory variables $x_1$–$x_3$. $x_1$: rule (1: A+/B−, −1: A−/B+), $x_2$: category group (1: stimulus in group A, −1: stimulus in group B), $x_1 x_2$: interaction term between rule and category factors, which also corresponds to contingency (1: juice, −1: saline), and $x_3$: cumulative trial number (to account for a gradual activity change over trials). We excluded error trials and the following correction trials from the analysis of neuronal activity. To minimize the influence of neuronal activity change due to the reversal, we excluded from the analyses one block (eight trials) of data after the reversal.

To determine whether the proportion of neurons of a given type is higher than the chance level, we conducted a randomization test ($n = 1000$) in which the spike rates were shuffled between trials while keeping the relationship of the explanatory variables the same. We calculated the expected percentage of neurons for each type from the randomized data to create the distribution of percentages for each type of neurons (1000 points for each type).

To determine whether the stimulus selectivity of the category-coding neurons reflected the functional category or a noncategorical combination of four stimuli, we conducted another randomization test in which the relationship between the stimulus and group were shuffled 1000 times. For example, we randomly assigned A1 stimulus to group A or B and did the same for the remaining seven stimuli, with the constraint that both groups consist of four stimuli, and then conducted regression anal-
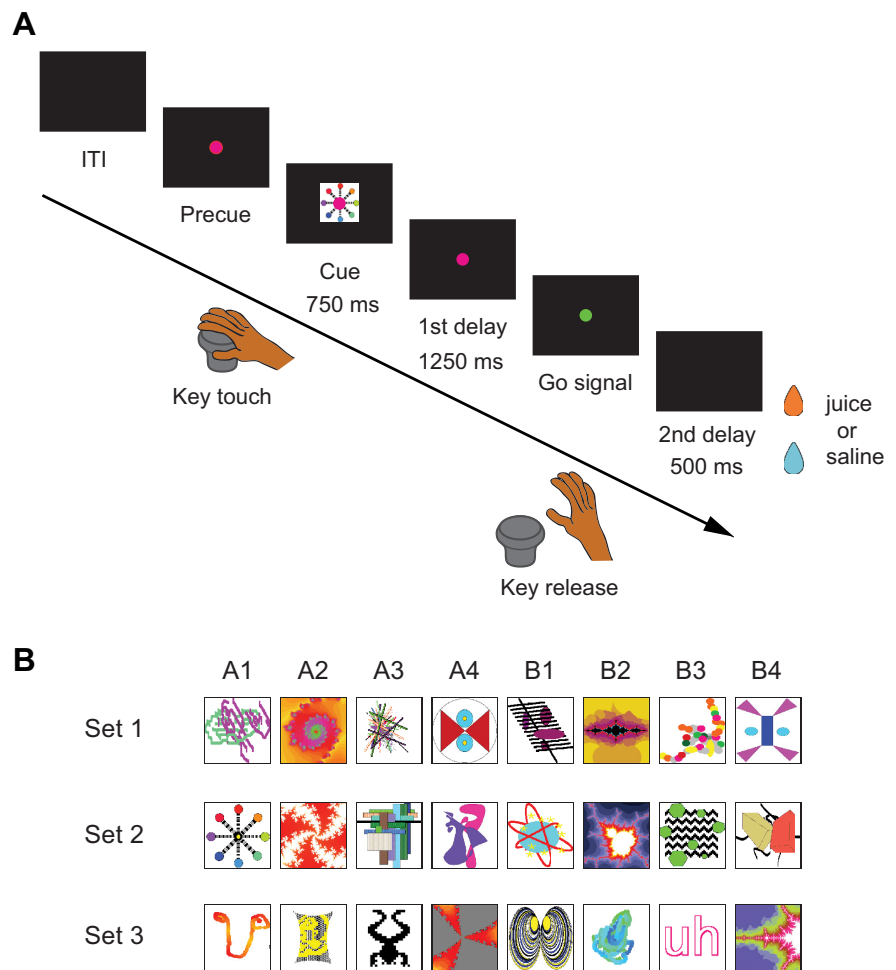


**Figure 1.** Group reversal task. **A**, Time course of the group reversal task. In every trial, a visual stimulus predicted whether the appetitive or aversive type of liquid (fruit juice or saline) was going to be delivered at the end of the trial. **B**, Stimuli used in the group reversal task. Three stimulus sets were used repeatedly during the training and the unit recording. Each set consisted of eight abstract figures. In each trial, a stimulus was selected from the stimulus set pseudorandomly in such a way that each stimulus in a set was presented once in a block of eight trials. Four of the stimuli in a set were associated with juice and the other four were associated with saline and the association relationships were occasionally reversed without an explicit cue (rule reversal). We call one of the four-stimulus groups A and the other B. Even when the stimulus–outcome relationship was reversed at the rule reversal, the functional equivalence of each of the four stimuli in a group was always preserved (i.e., they predicted the same outcome) and thus constituted a functional category.

ysis to see whether the category factor was significant. Because the relationship between firing rate and stimulus was preserved in this procedure, we could evaluate the category effects independently of stimulus effects on neuronal activity. We repeated this procedure 1000 times for all neurons to calculate the expected percentage of neurons with selectivity to the shuffled combinations of stimuli and created a distribution of the percentages. We also calculated the coefficient of partial determination (CPD) for the category factor in the regression model and compared the CPD calculated from the original data with that calculated from the randomized data. The CPD is defined as follows:

$$CPD = (SSEr - SSEf)/SSEr$$

where *SSEr* is the sum of squared errors in the regression model that excludes the category factor and *SSEf* is the sum of squared errors in the regression model with all factors. A CPD ranges from 0 to 1.

## Results

Two male Japanese monkeys were trained to perform a group reversal task (Fig. 1). Through the training, the monkeys learned to respond correctly to all of remaining seven stimuli after being
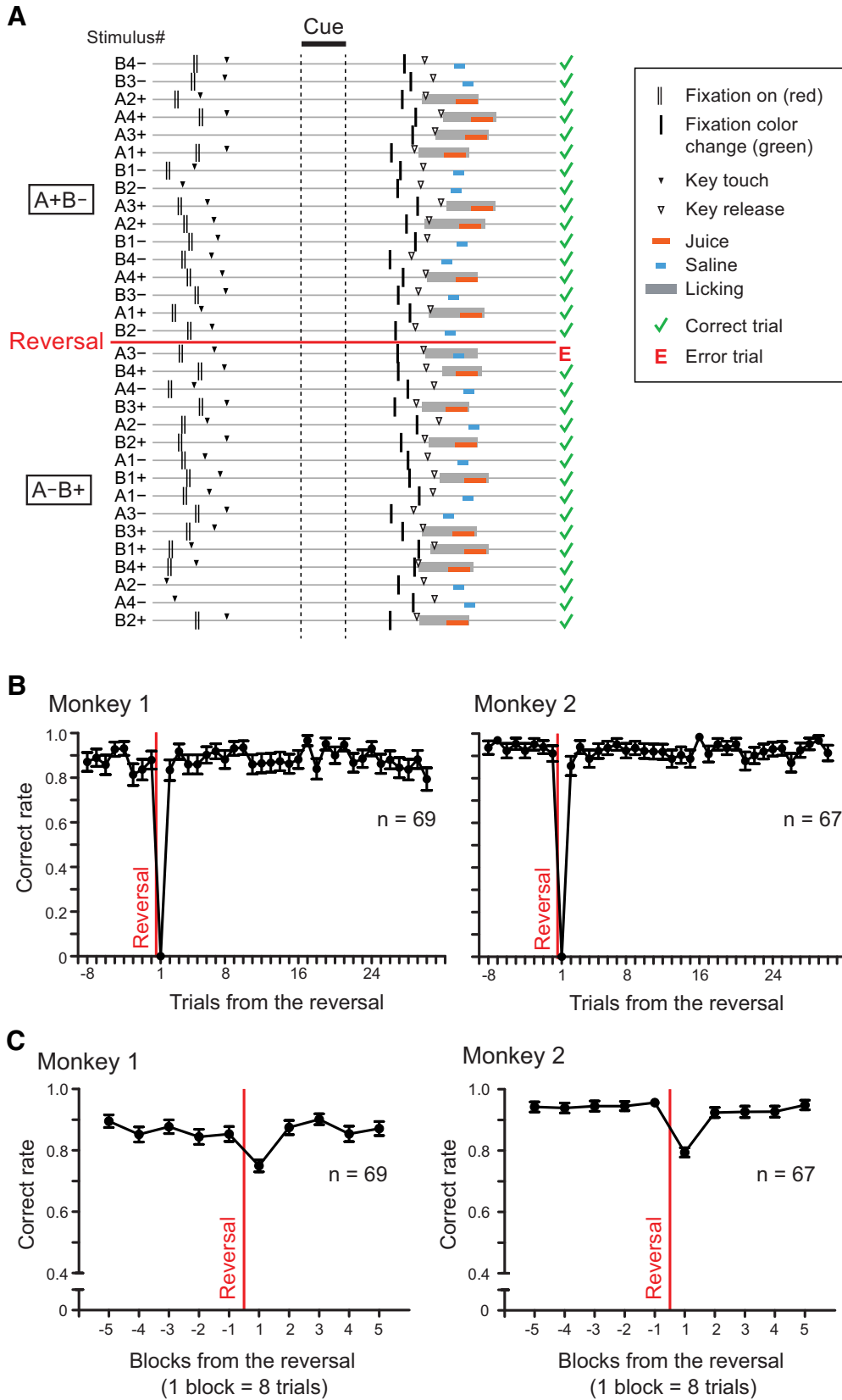
**Figure 2.** Immediate behavioral adaptation after the reversal. **A**, Behavioral performance around the reversal in a session. Each horizontal line indicates a trial. Trials progressed from top to bottom. The monkey made an error response (licking saline) on the first trial after the reversal, but made correct responses from the second trial after the reversal. Note that the monkey had not experienced the stimulus–outcome relation change for the cue on the second to eighth trial after the reversal (stimulus B4, A4 . . . B1). **B**, Trialwise mean behavioral performance in all recording sessions. Mean correct rates of eight trials before the reversal and 30 trials after the reversal are shown (mean ± SEM). **C**, Blockwise mean behavioral performance in all recording sessions. Mean correct rates of five blocks before and after the reversal are shown (mean ± SEM).
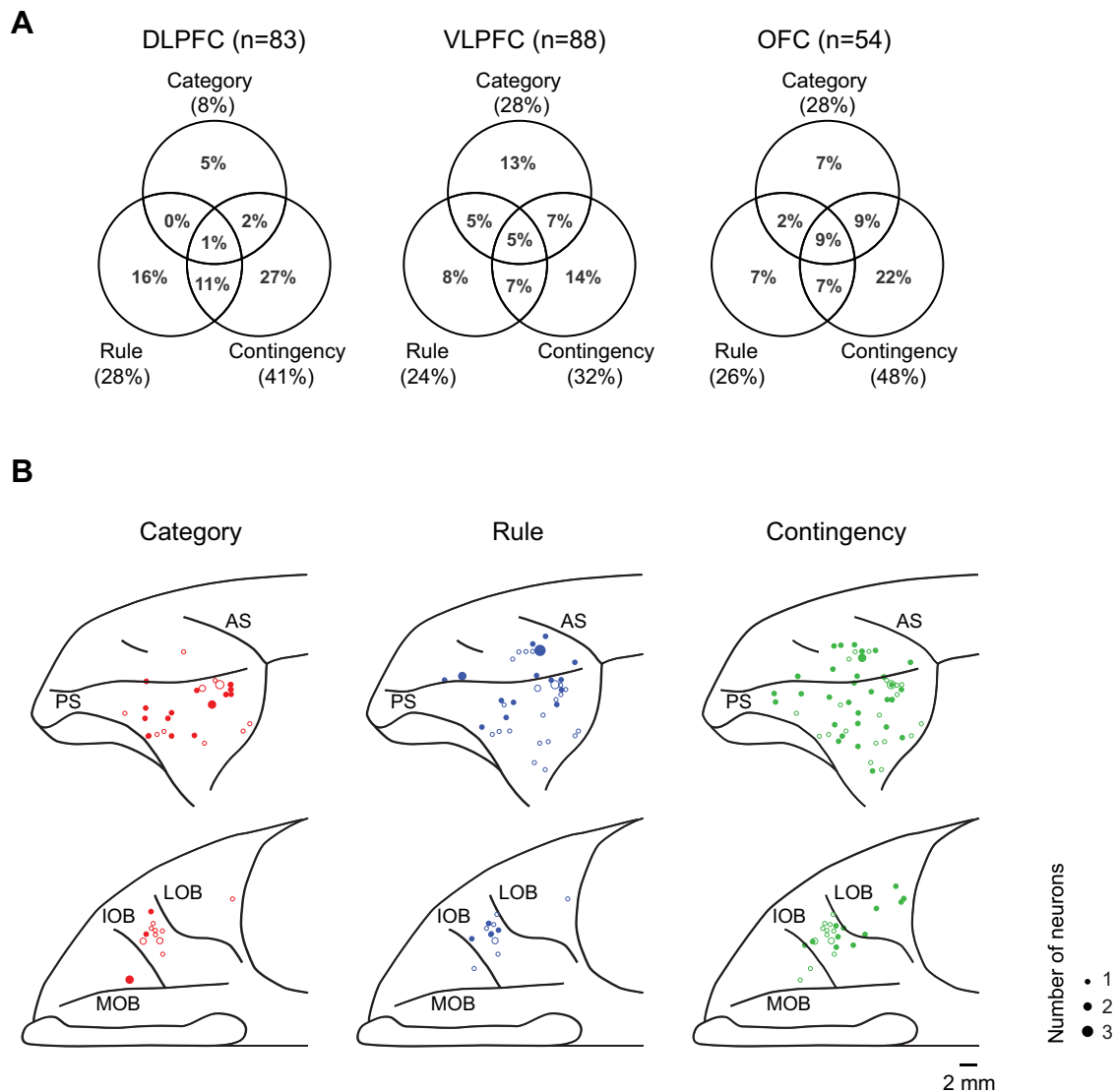
**A**



**B**



**Figure 3.** Proportion of neurons in each brain region and locations of each type of neurons. **A**, Proportion of each type of neuron in the cue period. **B**, Recording locations for each type of neuron. Filled circles indicate the locations of neurons with activity that was dependent on a single factor; open circles indicate neurons with activity that was dependent on multiple factors. Top row shows the lateral view of the PFC; bottom row shows the bottom view of the PFC. AS, Arcuate sulcus; PS, principal sulcus; LPB, lateral orbital sulcus; IOB, intermediate orbital sulcus; MOB, medial orbital sulcus.

exposed to the reversed contingency for the first stimuli in the first reversed block (Fig. 2). Such quick adaptation to the contingency reversal indicates that monkeys can form and use functional categories, rather than learning the contingency of each stimulus independently, to perform efficiently in the group reversal task.

**Coding of category, rule, and contingency information during the cue period**

During the performance of the task, we recorded the single-unit activity of 225 neurons in the PFC: 83 in the dorsal LPFC (DLPFC), 88 in the ventral LPFC (VLPFC), and 54 in the orbitofrontal cortex (OFC). In this study, we focused on the neuronal activity during the presentation of the visual stimulus (cue period). We used a multiple linear regression to determine whether the variance of neuronal activity across trials was related to the category (A or B), the rule (A+/B− or A−/B+), or the contingency [juice (+) or saline (−)].

The results of the multiple linear regression analysis are summarized in Figure 3A. In the DLPFC, 8(5)%, 28*(16*)%, and 41*(27*)% of the recorded neurons changed their activity in relation to category, rule, and contingency, respectively (numbers in parentheses indicate the percentages of neurons that were dependent on the single factor alone; asterisks indicate that the proportion turned out to be above chance level by the randomization test; $p < 0.05$; see Materials and Methods for the details of the randomization test.) In the VLPFC, 28*(13*)%, 24*(8*)%, and 32*(14*)% of the recorded neurons changed their activity in relation to category, rule, and contingency, respectively. In the OFC, 28*(7*)%, 26*(7*)%, and 48*(22*)% of the recorded neurons changed their activity in relation to category, rule, and contingency, respectively. Neurons coding category were more abundant in the VLPFC (28%) and OFC (28%) than in the DLPFC (8%) (Bonferroni-corrected $\chi^2$ test, $p < 0.05$) and the proportion of neurons coding category alone was above the chance level only in the VLPFC. Neurons coding rule were distributed widely over the recorded areas, but the proportion of neurons coding rule alone was above the chance level only in the DLPFC. Neurons coding contingency were also distributed widely over the
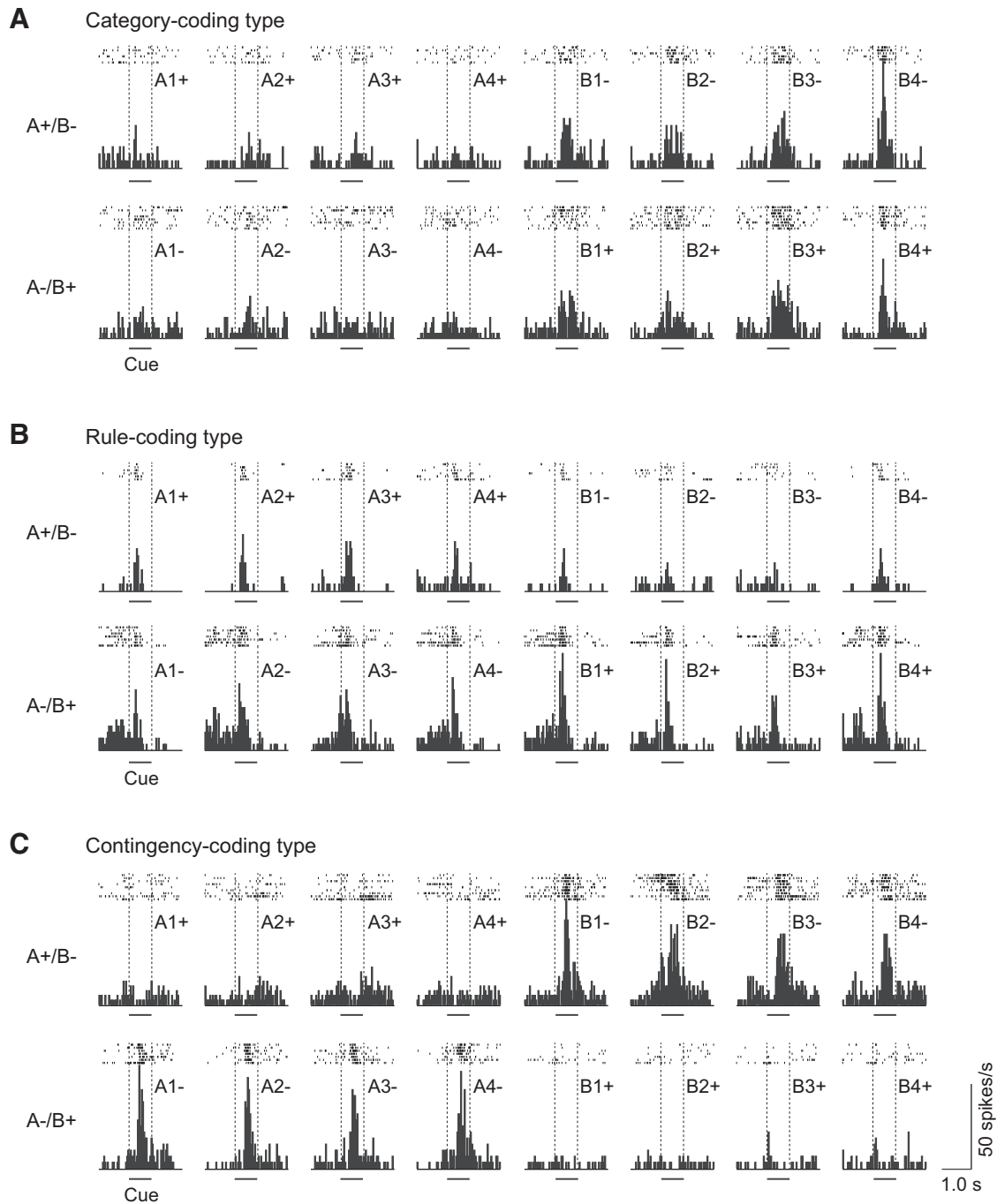
**Figure 4.** Examples of the activity of prefrontal neurons coding category, rule, or contingency information. *A*, *B*, and *C* show the activity of three different types of neurons in the PFC, all of which show the rastergram and histogram for each stimulus (*A1–A4* and *B1–B4*) under different rules. Top row shows the neuronal activity under the A+/B− rule, where group A stimuli predicted juice and group B stimuli predicted saline. Bottom row shows the neuronal activity under the A−/B+ rule, where group A stimuli predicted saline and group B stimuli predicted juice. The left vertical dashed line in each panel represents the cue onset and the right one represents the cue offset. *A*, Activity of a representative category-coding neuron that responded to group B stimuli regardless of which outcome they were associated with. *B*, Activity of a representative rule-coding neuron that showed higher activity under the A−/B+ rule. *C*, Activity of a representative contingency-coding neuron that responded to saline-predicting cue stimuli.

recorded areas, but their proportion tended to be higher in the DLPFC (41%) and OFC (48%) than in the VLPFC (32%). These trends in neuron distribution are clearly visualized in the surface plots of the neuron-recording sites (Fig. 3B).

Figure 4A shows the activity of a typical neuron that had significant variance of activity dependent on the category information during the cue presentation (category-coding neuron). This particular neuron consistently responded to group B stimuli re-

gardless of whether they predicted juice (+) or saline (−), suggesting that the activity of this neuron is based on the functional equivalence of the stimuli in group B. Thus, we found a clear neuronal coding of functional category information within the PFC. Figure 4B shows the activity of a typical neuron that had a significant variance of activity dependent on rule information (rule-coding neuron). The activity of this particular neuron was higher under the A−/B+ rule than under the A+/B− rule. We
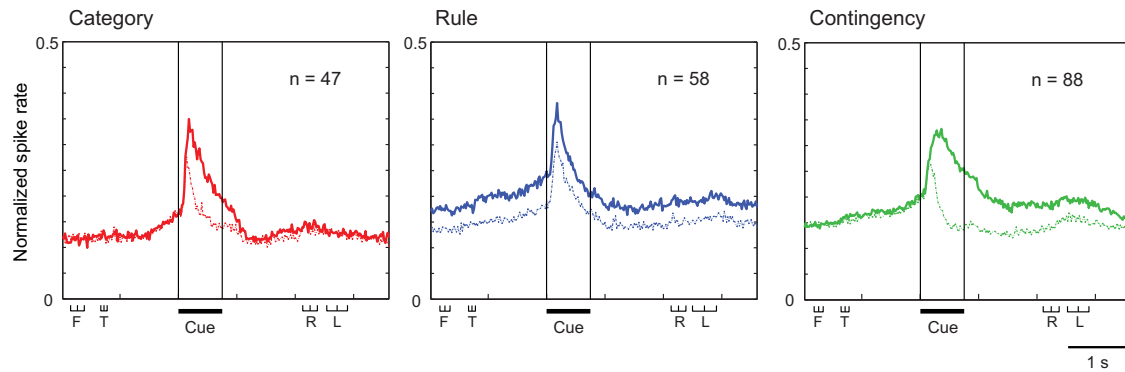
**Figure 5.** Population histograms of the activity of the category-, rule-, and contingency-coding neurons. Neuronal activity was normalized to the maximum activity of each neuron before averaging across the neurons. Solid lines show the activity in the preferred condition (e.g., preferred category for category-coding neurons); dashed lines show activity in the nonpreferred condition. Labels underneath each panel indicate the duration of cue presentation (Cue) and the average timings and SD for fixation spot onset (F), key touch (T), key release (R), and liquid delivery (L).
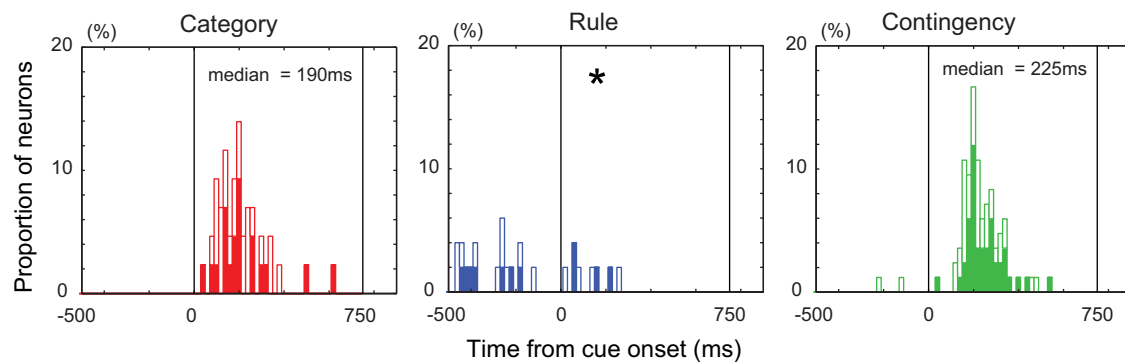


**Figure 6.** Frequency distribution histograms of the differential latencies of the category-, rule-, and contingency-coding neurons. Filled bars indicate the percentage of neurons that were dependent on the single factor alone; open bars indicate the percentage of neurons that were dependent on multiple factors. Asterisk indicates that >60% of rule-coding neurons were out of the time range of the histogram.

also found a considerable number of neurons that had a significant variance of activity dependent on the predicted contingency information (contingency-coding neurons). Figure 4C shows the activity of a typical contingency-coding neuron. The neuron in Figure 4C responded to group B stimuli under the A+/B− rule, where group A stimuli predicted juice and group B stimuli predicted saline, and to group A stimuli under the A−/B+ rule, where group A stimuli predicted saline and group B stimuli predicted juice. Therefore, this particular neuron responded to saline-predicting stimuli. This type of neurons may code the outcome (juice or saline) that the visual cue predicted, or the required response (go or no-go), but not the identity of the stimuli in a group.

To examine the possibility that the selectivity of category-coding neurons might reflect the sensitivity to random subsets of stimuli, we compared the number of category-coding neurons in the original data against the number detected after randomly reassigning the stimuli to surrogate categories that did not bear a relation to functional significance in the task (see Materials and Methods). In all areas, the percentage of neurons coding the category factor was significantly higher than that of neurons showing selectivity to the randomized stimulus combinations (DLPFC: original data = 8% vs randomized data = 3%; VLPFC: original data = 28% vs randomized data = 17%; OFC: original data = 28% vs randomized data = 22%; $p < 0.05$). We also computed the CPD for the category factor as a category-selectivity index for each neuron and compared the mean CPD calculated from the original data with that calculated from the

randomized data. We found that, in all areas, the mean CPD calculated from the original data was significantly higher than that calculated from randomized data (DLPFC: original CPD = $0.10 \pm 0.02$ vs randomized CPD = $0.03 \pm 0.00$; VLPFC: original CPD = $0.16 \pm 0.02$ vs randomized CPD = $0.05 \pm 0.00$; OFC: original CPD = $0.12 \pm 0.01$ vs randomized CPD = $0.04 \pm 0.00$; mean $\pm$ SEM, $t$ test, $p < 0.05$). These results indicate that the category-coding neurons code the functional category rather than random subsets of stimuli.

### Time course of the activation of category, rule, and contingency information

Figure 5 shows the population histogram of the activity of category-, rule-, and contingency-coding neurons. All three types of neurons were activated at the visual cue presentation. In rule-coding neurons, the differentiation of activity between preferred and nonpreferred conditions was observed throughout the trial. In category- and contingency-coding neurons, the differentiation of activity between preferred and nonpreferred conditions started after the cue onset. The activity of category-coding neurons became nondifferential again soon after the disappearance of the cue, but that of contingency-coding neurons continued to be differential until the end of the trial.

Figure 6 shows the distributions of the onset of differential activity between preferred and nonpreferred conditions in category-, rule-, and contingency-coding neurons. The onset of differential activity was determined as the time of occurrence of the first of 5 consecutive significantly differential bins ($t$ test, $p <$
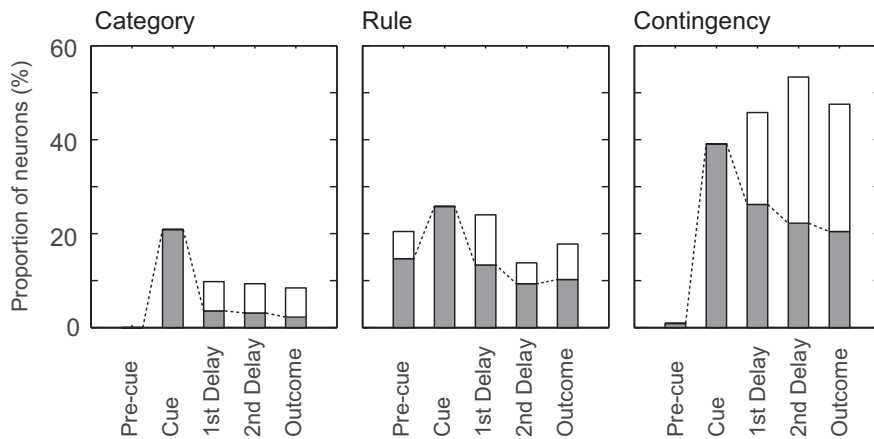
**Figure 7.** Proportion of each type of neuron found in each task period. Filled portions of the bar indicate the proportion of neurons that were selective to the corresponding factor (i.e., category, rule, or contingency, respectively) in the cue period.

0.05). In rule-coding neurons, most neurons started to show differential activity before the cue onset (>60% of neurons showed the onset of the differential activity more than 500 ms before the cue onset or even already showed differential activity when the computer started storing the data in the intertrial interval before the trial onset). In category- and contingency-coding neurons, the onset of the differential activity was distributed within the range of the cue presentation. The onset of the differential activity of category-coding neurons (median 190 ms) was significantly earlier than that of contingency-coding neurons (median 225 ms) (Kolmogorov–Smirnov test, $p < 0.05$).

So far, we had conducted data analysis on neuronal activity during the cue period. Here, we have also conducted regression analysis for the neuronal data in the pre-cue, first delay, second delay, and outcome periods with the same regression model (Fig. 7). The proportion of category-coding neurons was highest in the cue period and much lower in other periods, suggesting that the category information was mainly encoded during the cue period. The proportion of rule-coding neurons stayed relatively high across the periods, suggesting that the rule information was coded throughout a session. The proportion of contingency-coding neurons became high during the cue period, even more neurons were recruited during the delay period, and then many held the information active up to the outcome period. Therefore, although the rule information was kept active throughout the trial, category information was quickly activated after the cue onset and faded out soon after the disappearance of the cue. Contingency information was activated immediately after the activation of category information and was maintained active until the end of the trial.

**Behavioral adaptation to the reversal and neuronal activity change**

Figure 2, *B* and *C*, shows the trialwise and blockwise behavioral performances before and after the reversal. The trialwise analysis shows that, in both monkeys, the correct rate of the first trial after the reversal was 0%, but the correct rate of the second trial after the reversal was >80% (Fig. 2B). The blockwise analysis shows that the behavioral performance was considerably high (>70%), normally one or two errors, in the first block after the reversal and recovered to the plateau level in the second block (Fig. 2C). These results suggest that the monkeys used the category information for behavioral adaptation after the reversal. Figure 8 shows the normalized population activity of each type of

neuron before and after the reversal. The activity of category-coding neurons did not change before and after the reversal. Conversely, the activity of rule-coding and contingency-coding neurons changed drastically in the first block after the reversal, indicating quick neuronal adaptation after the reversal in those neurons.

## Discussion

We recorded single-unit activity in the monkey PFC during the performance of a group reversal task designed to dissociate clearly functional categories from perceptual categories. The task also dissociated the category information from the rule or outcome/response contingency information. We found a group of neurons that maintain their response selectivity to functionally equivalent stimuli over contingency changes and the activity of those neurons was less affected by visual features. Those category-coding neurons were a discrete functional group that was dissociable from those coding rule or contingency information, whereas there was some overlap in coding those different types of information. We also found the specificity of localization of the category-coding neurons that large numbers of which were distributed in the VLPFC and OFC, with neurons coding category information alone (independent of rule or contingency information) being especially abundant in the VLPFC (Fig. 3).

Humans, monkeys, and other species of higher vertebrates are considered to have the ability to use functional categories (Bruner et al., 1966; Sidman and Tailby, 1982; Lazar et al., 1984; Vaughan, 1988; Dube et al., 1993; Schusterman and Kastak, 1998; Ward-Robinson and Hall, 1999; Jitsumori et al., 2002; Urcuioli et al., 2006; Truppa et al., 2010). Categories play an important role in compressing the enormous amount of redundant information that we get from the environment into an abstract form of information stored for future use, which is generally referred to as "knowledge." In human neuropsychology, it is known that prefrontal damage leads to deficits in abstraction ability, including the use of category, together with various other cognitive functions related to executive control and working memory (Stuss and Benson, 1986; Fuster, 2008). Human neuroimaging studies have shown that the inferior frontal cortex, which may correspond to the monkey VLPFC, exhibits higher activity when subjects are required to select among categorically related alternatives (Thompson-Schill et al., 1997) or the retrieval of conceptual knowledge from memory (Wagner et al., 2001; Gold and Buckner, 2002).

Functional category is generally thought to be a product of learning and to be stored in long-term memory. There arises the question of whether the category-coding neurons that we found in the PFC are themselves involved in long-term storage of the functional category or if they receive the functional category information stored and activated in some other brain region. Accumulating information from human studies suggests that knowledge and concepts are stored in the temporal cortex (Tranel et al., 1997; Vandenbulcke et al., 2006). Indeed, the VLPFC, where most of the category-coding neurons were recorded (Fig. 3), is the main target of fiber projections from the inferior temporal cortex (Webster et al., 1994; Borra et al., 2011; Saleem et al., 2014), which is known to be dedicated to visual
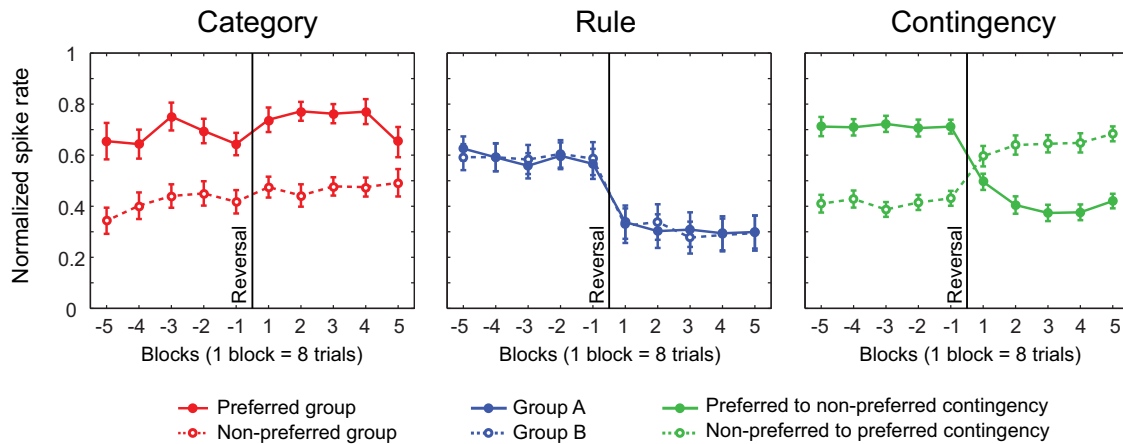
**Figure 8.** Normalized population activity of each type of neuron (category, rule, and contingency coding) before and after reversal. For category-coding neurons, filled circles (connected by solid lines) represent the mean activity in response to the cue stimuli of the preferred group, whereas open circles (connected by dotted lines) represent that in response to the cue stimuli of the nonpreferred group. For rule-coding neurons, filled and open circles represent the mean activity in response to the cue stimuli of groups A and B, respectively. For contingency-coding neurons, filled circles represent the mean activity in response to the cue stimuli of the group that were associated with the preferred contingency before the reversal (and the nonpreferred contingency after the reversal), whereas open circles represent that in response to the cue stimuli of the group that were associated with the nonpreferred contingency before the reversal (and the preferred contingency after the reversal). For category and contingency-coding neurons, we averaged the data before and after the first reversal after stimulus set change; for rule-coding neurons, we averaged the data before and after the reversal when the rule was changed from preferred to nonpreferred one for the first time.

long-term object memory in monkeys (Miyashita, 1988; Messinger et al., 2001; Naya et al., 2001). It may be reasonable to assume that, in every trial, the functional category information stored in the temporal cortex is activated by the visual cue presentation and then forwarded to the category-coding neurons in the VLPFC.

A previous study reported that monkeys with lesions to the LPFC could still learn a category (Minamimoto et al., 2010), suggesting that the LPFC is not necessary for the formation and storage of categories. The study found no significant difference between the task performances of monkeys with and without LPFC lesions in the task in which the category was based on the associated reward alone and the monkeys were not required to select different operant responses based on the category. It is probable that the PFC is not necessary unless the activated category information is needed for a behavioral selection, and therefore, the effect of LPFC lesions on the behavioral performance was not found in a task without a behavioral selection.

The rule-coding neurons were distributed widely in the recording areas, but most of those coding only rule information were found in the DLPFC, especially around the principal sulcus. Many of the rule-coding neurons were actively holding the rule information throughout the trial and a large proportion of them became most active at the time of the cue onset. The result that many DLPFC neurons coded rule information is consistent with previous lesion (Buckley et al., 2009) and single-unit studies (White and Wise, 1999; Wallis et al., 2001; Eiselt and Nieder, 2013; Zhang et al., 2013), which suggests that the DLPFC plays an important role in actively maintaining the behavioral context information.

Using the group reversal task, we were able to dissociate category coding from outcome or response contingency coding, which has been well known to be prevalent in the PFC (Watanabe, 1986; Quintana et al., 1988; Watanabe, 1990; Sakagami and Niki, 1994; Rolls et al., 1996; Watanabe, 1996; Asaad et al., 1998; Pasupathy and Miller, 2005; Pan et al., 2008). We did not intend to distinguish between outcome and response contingency. In this task, the prediction of a certain outcome always led to the selection of a certain action: when juice was predicted, the

appropriate response was "go (lick)," and when saline was predicted, the appropriate response was "no-go (no lick)." Therefore, the neurons classified as coding contingency may include both outcome and response contingency coding types. Because the DLPFC is anatomically strongly connected to motor-related areas (Dum and Strick, 1991; Bates and Goldman-Rakic, 1993; He et al., 1993; Galea and Darian-Smith, 1994; Lu et al., 1994; Rushworth, 2000), it is likely that the contingency-coding neurons in the DLPFC are coding response contingency and are involved in planning and preparing the impending action. Conversely, the VLPFC and OFC are connected with limbic areas rather than motor areas (Morecraft et al., 1993; Carmichael and Price, 1995; Cavada et al., 2000), so the contingency-coding neurons in the VLPFC and OFC may reflect the outcome contingency.

We propose a simplified model based on the results of the analysis of the temporal change of information coding (Figs. 5, 6) explaining how category and rule information are integrated within the PFC to produce the contingency information needed for action selection in this task. During the task performance, including intertrial intervals, rule information (i.e., which category would be rewarded and which would be punished) is maintained by the rule-coding neurons and is activated as a trial begins. Then, at the cue presentation, temporal cortex neurons that store functional category information are activated by the visual input and the information is forwarded to the category-coding neurons in the PFC. Rule and category information is progressively integrated in the neurons that code multiple factors and the product of the information processing is coded by the contingency-coding neurons. Contingency-coding neurons then send this information to motor-related areas, where it is used in the preparation of appropriate actions. The actual neural computation, however, may not be simple as that and we may need to combine a computational approach and knowledge of the prefrontal local circuit if we are to construct a detailed information-processing model explaining how the contingency information is derived by integrating category and rule information in the prefrontal neural circuit.

In this study, monkeys were required to apply the category information to the rule information to foresee the outcome and select an appropriate action. During the task performance, we found neurons that coded functional category information, as well as neurons that coded rule or contingency information and combinations of these kinds of information. In general terms, a functional category can be regarded as a form of "knowledge" and a rule can be regarded as a "behavioral context." Knowledge is constructed by elaborating the information obtained through experience by connecting pieces of information into a symbolized and structured form. We apply knowledge to practical cases depending on the behavioral context to foresee their outcomes and select appropriate actions. The results of this study are consistent with previous studies in human neuropsychology indicating the role of the PFC in the behavior control based on abstract thinking involving categories and rules (Stuss and Benson, 1986; Fuster, 2008) and they suggest possible neural mechanisms underlying that role.

## References

Antzoulatos EG, Miller EK (2011) Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. Neuron 71:243–249. CrossRef Medline

Asaad WF, Rainer G, Miller EK (1998) Neural activity in the primate prefrontal cortex during associative learning. Neuron 21:1399–1407. CrossRef Medline

Bates JF, Goldman-Rakic PS (1993) Prefrontal connections of medial motor areas in the rhesus monkey. J Comp Neurol 336:211–228. CrossRef Medline

Borra E, Gerbella M, Rozzi S, Luppino G (2011) Anatomical evidence for the involvement of the macaque ventrolateral prefrontal area 12r in controlling goal-directed actions. J Neurosci 31:12351–12363. CrossRef Medline

Brincat SL, Miller EK (2015) Frequency-specific hippocampal-prefrontal interactions during associative learning. Nat Neurosci 18:576–581. CrossRef Medline

Bruner JS, Olver RR, Greenfield PM (1966) Studies in cognitive growth: a collaboration at the Center for Cognitive Studies. New York: Wiley.

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, Tanaka K (2009) Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science 325:52–58. CrossRef Medline

Carmichael ST, Price JL (1995) Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 363:642–664. CrossRef Medline

Cavada C, Compañy T, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suárez F (2000) The anatomical connections of the macaque monkey orbitofrontal cortex: a review. Cereb Cortex 10:220–242. CrossRef Medline

Cromer JA, Roy JE, Miller EK (2010) Representation of multiple, independent categories in the primate prefrontal cortex. Neuron 66:796–807. CrossRef Medline

Dube WV, Green G, Serna RW (1993) Auditory successive conditional discrimination and auditory stimulus equivalence classes. Journal of the Experimental Analysis of Behavior 59:103–114. CrossRef Medline

Dum RP, Strick PL (1991) The origin of corticospinal projections from the premotor areas in the frontal lobe. J Neurosci 11:667–689. Medline

Eiselt AK, Nieder A (2013) Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. J Neurosci 33:7526–7534. CrossRef Medline

Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. Nature 443:85–88. CrossRef Medline

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. Science 291:312–316. CrossRef Medline

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23:5235–5246. Medline

Fuster JM (2008) The prefrontal cortex, Ed 4. San Diego: Academic.

Galea MP, Darian-Smith I (1994) Multiple corticospinal neuron populations in the macaque monkey are specified by their unique cortical ori-

gins, spinal terminations, and connections. Cereb Cortex 4:166–194. CrossRef Medline

Gold BT, Buckner RL (2002) Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron 35:803–812. CrossRef Medline

Goodwin SJ, Blackman RK, Sakellaridi S, Chafee MV (2012) Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex. J Neurosci 32:3499–3515. CrossRef Medline

He SQ, Dum RP, Strick PL (1993) Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. J Neurosci 13:952–980. Medline

Jitsumori M, Siemann M, Lehr M, Delius JD (2002) A new approach to the formation of equivalence classes in pigeons. Journal of the Experimental Analysis of Behavior 78:397–408. CrossRef Medline

Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296–4309. CrossRef Medline

Lazar RM, Davis-Lang D, Sanchez L (1984) The formation of visual stimulus equivalences in children. J Exp Anal Behav 41:251–266. CrossRef Medline

Lu MT, Preston JB, Strick PL (1994) Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. J Comp Neurol 341:375–392. CrossRef Medline

Messinger A, Squire LR, Zola SM, Albright TD (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A 98:12239–12244. CrossRef Medline

Minamimoto T, Saunders RC, Richmond BJ (2010) Monkeys quickly learn and generalize visual categories without lateral prefrontal cortex. Neuron 66:501–507. CrossRef Medline

Miyashita Y (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. Nature 335:817–820. CrossRef Medline

Mogami T, Tanaka K (2006) Reward association affects neuronal responses to visual stimuli in macaque te and perirhinal cortices. J Neurosci 26:6761–6770. CrossRef Medline

Morecraft RJ, Geula C, Mesulam MM (1993) Architecture of connectivity within a cingulo-fronto-parietal neurocognitive network for directed attention. Arch Neurol 50:279–284. CrossRef Medline

Naya Y, Yoshida M, Miyashita Y (2001) Backward spreading of memory-retrieval signal in the primate temporal cortex. Science 291:661–664. CrossRef Medline

Ohyama K, Sugase-Miyamoto Y, Matsumoto N, Shidara M, Sato C (2012) Stimulus-related activity during conditional associations in monkey perirhinal cortex neurons depends on upcoming reward outcome. J Neurosci 32:17407–17419. CrossRef Medline

Pan X, Sawa K, Tsuda I, Tsukada M, Sakagami M (2008) Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. Nat Neurosci 11:703–712. CrossRef Medline

Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433:873–876. CrossRef Medline

Quintana J, Yajeya J, Fuster JM (1988) Prefrontal representation of stimulus attributes during delay tasks. I. Unit activity in cross-temporal integration of sensory and sensory-motor information. Brain Res 474:211–221. CrossRef Medline

Rolls ET, Critchley HD, Mason R, Wakeman EA (1996) Orbitofrontal cortex neurons: role in olfactory and visual association learning. J Neurophysiol 75:1970–1981. Medline

Roy JE, Riesenhuber M, Poggio T, Miller EK (2010) Prefrontal cortex activity during flexible categorization. J Neurosci 30:8519–8528. CrossRef Medline

Rushworth MS (2000) Anatomical and functional subdivision within the primate lateral prefrontal cortex. Psychobiology 28:187–196.

Sakagami M, Niki H (1994) Encoding of behavioral significance of visual stimuli by primate prefrontal neurons: relation to relevant task conditions. Exp Brain Res 97:423–436. Medline

Saleem KS, Miller B, Price JL (2014) Subdivisions and connectional networks of the lateral prefrontal cortex in the macaque monkey. J Comp Neurol 522:1641–1690. CrossRef Medline

Schusterman RJ, Kastak D (1998) Functional equivalence in a California sea

lion: Relevance to animal social and communicative interactions. Animal Behaviour 55:1087–1095. CrossRef Medline

Sidman M, Tailby W (1982) Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. J Exp Anal Behav 37:5–22. CrossRef Medline

Stuss DT, Benson DF (1986) The frontal lobes. Baltimore: Raven.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ (1997) Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc Natl Acad Sci U S A 94:14792–14797. CrossRef Medline

Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y (1999) Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature 401:699–703. CrossRef Medline

Tranel D, Damasio H, Damasio AR (1997) A neural basis for the retrieval of conceptual knowledge. Neuropsychologia 35:1319–1327. CrossRef Medline

Truppa V, Garofoli D, Castorina G, Piano Mortari E, Natale F, Visalberghi E (2010) Identity concept learning in matching-to-sample tasks by tufted capuchin monkeys (*Cebus apella*). Animal Cognition 13:835–848. CrossRef Medline

Urcuioli PJ, Lionello-DeNolf K, Michalek S, Vasconcelos M (2006) Some tests of response membership in acquired equivalence classes. Journal of the Experimental Analysis of Behavior 86:81–107. CrossRef Medline

Vandenbulcke M, Peeters R, Fannes K, Vandenberghe R (2006) Knowledge of visual attributes in the right hemisphere. Nat Neurosci 9:964–970. CrossRef Medline

Vaughan W (1988) Formation of equivalence sets in pigeons. Journal of Experimental Psychology-Animal Behavior Processes 14:36–42. CrossRef

Vogels R (1999) Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. Eur J Neurosci 11:1239–1255. CrossRef Medline

Wagner AD, Paré-Blagoev EJ, Clark J, Poldrack RA (2001) Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31:329–338. CrossRef Medline

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953–956. CrossRef Medline

Ward-Robinson J, Hall G (1999) The role of mediated conditioning in acquired equivalence. Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology 52:335–350.

Watanabe M (1986) Prefrontal unit activity during delayed conditional Go/No-Go discrimination in the monkey. II. Relation to Go and No-Go responses. Brain Res 382:15–27. CrossRef Medline

Watanabe M (1990) Prefrontal unit activity during associative learning in the monkey. Exp Brain Res 80:296–309. Medline

Watanabe M (1996) Reward expectancy in primate prefrontal neurons. Nature 382:629–632. CrossRef Medline

Webster MJ, Bachevalier J, Ungerleider LG (1994) Connections of inferior temporal areas TEO and TE with parietal and frontal-cortex in macaque monkeys. Cereb Cortex 4:470–483. CrossRef Medline

White IM, Wise SP (1999) Rule-dependent neuronal activity in the prefrontal cortex. Exp Brain Res 126:315–335. CrossRef Medline

Yamada M, Pita MC, Iijima T, Tsutsui K (2010) Rule-dependent anticipatory activity in prefrontal neurons. Neurosci Res 67:162–171. CrossRef Medline

Zhang J, Kriegeskorte N, Carlin JD, Rowe JB (2013) Choosing the rules: distinct and overlapping frontoparietal representations of task rules for perceptual decisions. J Neurosci 33:11852–11862. CrossRef Medline