

A Possible Role of Midbrain Dopamine Neurons in Short- and Long-Term Adaptation of Saccades to Position-Reward Mapping

Yoriko Takikawa,¹ Reiko Kawagoe,¹ and Okihide Hikosaka^{1,2}

¹Department of Physiology, Juntendo University, School of Medicine, Tokyo 113-8421, Japan; and ²Laboratory of Sensorimotor Research, National Eye Institute, National Institute of Health, Bethesda, Maryland 20892-4435

Submitted 9 March 2004; accepted in final form 18 May 2004

Takikawa, Yoriko, Reiko Kawagoe, and Okihide Hikosaka. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J Neurophysiol* 92: 2520–2529, 2004. First published May 26, 2004; 10.1152/jn.00238.2004. Dopamine (DA) neurons respond to sensory stimuli that predict reward. To understand how DA neurons acquire such ability, we trained monkeys on a one-direction-rewarded version of memory-guided saccade task (1DR) only when we recorded from single DA neurons. In 1DR, position-reward mapping was changed across blocks of trials. In the early stage of training of 1DR, DA neurons responded to reward delivery; in the later stages, they responded predominantly to the visual cue that predicted reward or no reward (reward predictor) differentially. We found that such a shift of activity from reward to reward predictor also occurred within a block of trials after position-reward mapping was altered. A main effect of long-term training was to accelerate the within-block reward-to-predictor shift of DA neuronal responses. The within-block shift appeared first in the intermediate stage, but was slow, and DA neurons often responded to the cue that indicated reward in the preceding block. In the advanced stage, the reward-to-predictor shift occurred quickly such that the DA neurons' responses to visual cues faithfully matched the current position-reward mapping. Changes in the DA neuronal responses co-varied with the reward-predictive differentiation of saccade latency both in short-term (within-block) and long-term adaptation. DA neurons' response to the fixation point also underwent long-term changes until it occurred predominantly in the first trial within a block. This might trigger a switch between the learned sets. These results suggest that midbrain DA neurons play an essential role in adapting oculomotor behavior to frequent switches in position-reward mapping.

INTRODUCTION

The ability to predict future reward is important for goal-directed behavior (Dickinson and Balleine 1994). This ability needs to be flexible since the context and location of reward may change unpredictably. Midbrain dopamine (DA) neurons comprise a prominent neuronal system that convey and process reward signals in the brain (Schultz 2002). DA neurons in monkeys respond to reward and to sensory stimuli that predict reward (Schultz 1998). More importantly, DA neurons respond to unexpected reward and show depression of activity when reward is unexpectedly omitted (Schultz et al. 1997) or shifted (Hollerman and Schultz 1998). These observations led Schultz and colleagues to suggest that DA neurons encode "reward prediction error" (Schultz 1998; Schultz and Dickinson 2000).

Address for reprint requests and other correspondence: O. Hikosaka, Laboratory of Sensorimotor Research, National Eye Institute, National Institute of Health, 49 Convent Dr., Bldg. 49, Rm. 2A50, Bethesda, MD 20892-4435 (E-mail: oh@lsr.nei.nih.gov).

Consistent with this hypothesis, we previously found that DA neurons increased or decreased their activity in a predictive manner when a cue, not reward, was presented (Kawagoe et al. 2004). In our experiments, the subject (a macaque monkey) performed a memory-guided saccade task while a reward was given after a saccade to one particular direction out of four [one-direction-rewarded (1DR) version of a memory-guided saccade task] (Kawagoe et al. 1998). A visual cue stimulus indicated the saccade direction and the presence or absence of reward. We found that DA neurons responded to the cue stimulus in a selective manner: a phasic excitation in response to a reward-indicating cue and a phasic suppression to a non-reward-indicating cue (Kawagoe et al. 2004). These responses were interpreted as reflecting reward prediction error because the likelihood that the trial would be rewarded was 25% (1 of 4) before the cue came on but changed to either 100% (after a reward-indicating cue) or 0% (after a non-reward-indicating cue).

A prominent feature of our 1DR task was that the reward contingency was changed systematically across blocks of trials. There were four sets of position-reward mapping, all of which the monkey had experienced extensively, and the question was to choose one particular set of position-reward mapping in one block and switch to another in the next block. We found that, as a new block of trials started, DA neurons changed their cue responses quickly to match the altered position-reward mapping (Kawagoe et al. 2004). That these changes in DA neuronal activity may have significant behavioral consequences has been suggested by a series of studies in the basal ganglia from our laboratory. Activity of presumed projection neurons in the caudate nucleus, to which DA neurons may project, responded to the cues depending on the current position-reward mapping (Kawagoe et al. 1998). Neurons in the substantia nigra pars reticulata (Sato and Hikosaka 2002) and those in the superior colliculus (Ikeda and Hikosaka 2003), which may receive inputs from caudate projection neurons, exhibited cue responses depending on the position-reward mapping. Finally, saccadic eye movements were strongly influenced by the position-reward mapping such that the saccades to the rewarded direction had shorter latencies and higher velocities than those to the unrewarded directions (Takikawa et al. 2002). Thus the neural circuits inside and emanating from the basal ganglia seem to play an important role in quick adaptation of goal-oriented behavior.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Having suggested the scheme for position-reward mapping, we were still puzzled by the quickness of the neural and behavioral adaptation. An obvious explanation would be long-term learning. In all of the preceding studies, we examined neural activity and behavior (saccadic eye movement) after the monkey had been trained to perform IDR extensively. The ability to quickly adapt to the altered position-reward mapping may be acquired through long-term training on IDR. By training monkeys on operant and classical conditioning tasks, Ljungberg et al. (1992), Schultz et al. (1993), and Mirenowicz and Schultz (1994) showed that DA neurons in naive monkeys responded to a reward, whereas DA neurons in well-trained monkeys responded to the earliest stimulus that indicated the reward but not the reward. Their studies suggest that the pattern of DA neuronal activity changes with long-term training. However, there is a critical difference between the tasks employed by Schultz's group and our task, IDR. In the tasks of Schultz et al., the reward contingency was fixed throughout the long-term training. In our IDR task, the reward contingency was altered frequently; it required switches between four learned sets of position-reward mapping.

In this study, we found that overall responses of DA neurons shifted from reward to reward-predictor during a long-term training period, consistent with the findings by Ljungberg et al. (1992), Schultz et al. (1993), and Mirenowicz and Schultz (1994). We further found that the shift from reward to reward-predictor occurred every time the position-reward mapping was altered. Long-term training changed the speed and accuracy of the adaptation to the altered position-reward mapping. This appeared as changes in DA neuronal activity as well as changes in the reward-contingent bias of saccadic eye movements.

METHODS

General

We used two male Japanese monkeys (*Macaca fuscata*). The monkeys were kept in individual primate cages in an air-conditioned room where food was always available. At the beginning of each experimental session, they were moved to the experimental room in a primate chair. The monkeys were given restricted amounts of fluid during periods of training and recording. Their body weight and appetite were checked daily. Supplementary water and fruit were provided daily. All surgical and experimental protocols were approved by the Juntendo University Animal Care and Use Committee and are in accordance with the National Institutes of Health Guide for the Care and Use of Animals.

The experiments were carried out while the monkey's head was fixed and his eye movements were recorded. For this purpose, a head holder, a chamber for unit recording, and an eye coil were implanted under surgical procedures. The monkey was sedated by intramuscular injections of ketamine (4.0–5.0 mg/kg) and xylazine (1.0–2.0 mg/kg). General anesthesia was induced by intravenous injection of pentobarbital sodium (5 mg/kg/h). Surgical procedures were conducted in aseptic conditions. After exposing the skull, 15–20 acrylic screws were bolted into it and fixed with dental acrylic resin. The screws served as anchors by which a head holder and a recording chamber, both made of delrin, were fixed to the skull. A scleral eye coil was implanted in one eye for monitoring eye position (Robinson 1963; Judge et al. 1980). The recording chamber, which was rectangular (antero-posterior: 42 mm; lateral: 30 mm; depth: 10 mm), was placed over the fronto-parietal cortices, tilted laterally by 35°. The monkey

received antibiotics (sodium ampicillin 25–40 mg/kg im each day) after the operation.

Behavioral tasks

We used the memory-guided saccade task (Hikosaka and Wurtz 1983) in two different reward conditions: all-directions-rewarded condition (ADR) and IDR (Kawagoe et al. 1998) (Fig. 1). In both conditions, a task trial started with onset of a central fixation point on which the monkeys had to fixate. A cue stimulus (spot of light) came on 1 s after onset of the fixation point (duration: 100 ms), and the monkeys had to remember its location. After 1–1.5 s, the fixation point turned off, and the monkeys were required to make a saccade to the previously cued location. The target came on 400 ms later for 150 ms at the cued location. The saccade was judged to be correct if the eye position was within a window around the target (usually within $\pm 3^\circ$) when the target turned off. The monkeys made the saccade before target onset based on memory, because, otherwise, the eyes could not reach the target window within the 150-ms target-on period; the target was presented only to give the monkeys the accurate feedback information. The next trial started after an intertrial interval of 3.5–4 s. In ADR, every correct saccade was rewarded with the liquid reward together with the tone stimulus. In IDR, an asymmetric

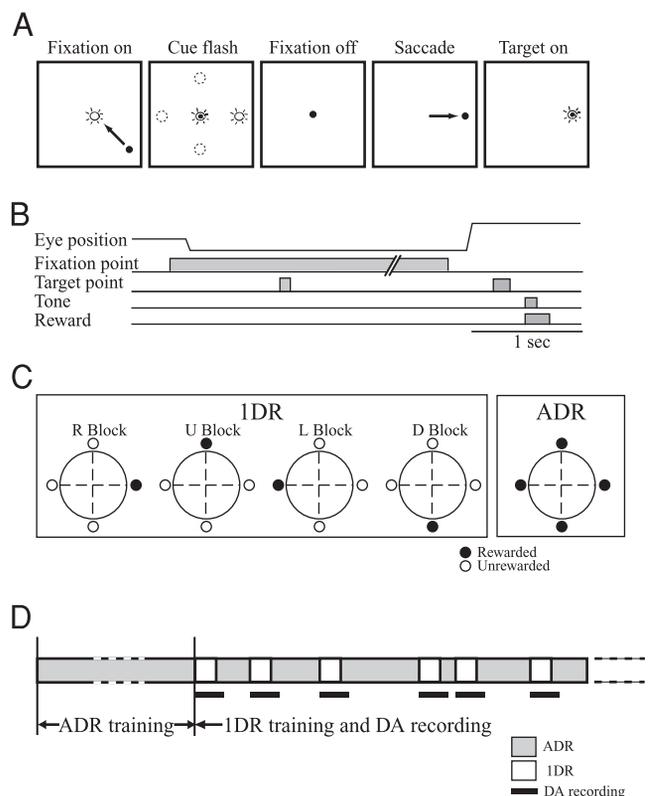


FIG. 1. Behavioral paradigm and experimental schedule. *A*: schematic display of visual stimuli and eye movements in a memory-guided saccade task. Arrows indicate saccadic eye movements. In this case, the monkey was required to saccade to the right direction. *B*: timing of stimulus presentation and eye movements. *C*: schedule for a single experiment while recording from a dopamine (DA) neuron. The monkey performed the memory-guided saccade task, which consisted of 4 blocks of 1-direction-rewarded (IDR) and 1 block of all-directions-rewarded (ADR). In IDR, only 1 of 4 directions was rewarded throughout a block of 60 trials, and the rewarded direction was changed across blocks. Order of the blocks was randomized. *D*: long-term training schedule to examine changes in activity of DA neurons. After the monkey mastered ADR with the initial training of 4 days, we started recording DA neurons. IDR was done only when a DA neuron was being recorded; otherwise, the monkey continued to perform ADR.

reward schedule was used such that only one of the four directions was rewarded, while the other directions were not rewarded. The rewarded direction was fixed in a block of experiments, which included 60 successful trials. Even for the unrewarded direction, the monkeys had to make a correct saccade, because the same trial was repeated if the saccade was incorrect. The amount of reward per trial was set approximately the same between IDR and ADR. The cue was chosen pseudorandomly such that the four directions were randomized in every sub-block of four trials; thus one block of experiment (60 trials) contained 15 trials for each direction (15 sub-blocks). Other than the actual reward, no indication was given to the monkeys as to which direction was currently rewarded. IDR was performed in four blocks, in each of which a different direction was rewarded. The order of the four blocks of IDR was randomized. To change the task schedule (IDR or ADR) or the reward direction in IDR, there was an interblock interval (about 30 s).

Experimental procedures

Each monkey was first trained to perform ADR for 4 days (about 100 blocks every each monkey). After the monkey mastered ADR, we started recordings from DA neurons (Fig. 1D). IDR was performed only when a DA neuron was being recorded; otherwise, the monkey continued to perform ADR. The target was chosen randomly out of four locations of equal eccentricity (10 or 20°), arranged in either normal or oblique angles. Once a candidate DA neuron was isolated, we examined whether it responded to the delivery of reward (free reward). A drop of water as a reward was given with a random time interval (4–9 s) while the monkey was sitting in the dim experimental room. The reward consisted of a tone followed by actual delivery of water to the monkey from a spout. The water delivery was delayed by about 150 ms from the tone, largely due to the compliance of the connecting silicon tube. Only when the neuron responded to reward did we ask the monkey to perform IDR. This was done in four blocks, in each of which one out of four directions was rewarded; in addition, we performed one block of ADR (Fig. 1C). We repeated some of the IDR blocks to confirm the stability of recording. During this standard set of experiments, we kept recording from the same DA neuron. This procedure was repeated every time a DA neuron was recorded.

Recording procedures

Eye movements were recorded using the search coil method (Enzanshi Kogyo MEL-20U) (Robinson 1963; Judge et al. 1980; Matsumura et al. 1992). Eye positions were sampled at 500 Hz. The behavioral tasks and storage and display of data were controlled by a computer (PC 9801RA, NEC, Tokyo, Japan). The unitary action potentials were passed through a window discriminator (model DDIS-1, Bak), and the times of their occurrences were stored with a resolution of 1 ms.

Before the single unit recording experiment, we obtained MR images (AIRIS, 0.3 T, Hitachi) such that they were perpendicular to the recording chamber. We then determined the recording sites in the substantia nigra based on the chamber-based coordinates (Kawagoe et al. 1998). The recording sites were further verified by MRI of a plastic guide tube through which the electrodes were inserted.

Single unit recordings were performed using tungsten electrodes (0.25 mm diam, 1–5 M Ω , measured at 1 KHz; Frederick Haer). To introduce the electrode into the brain, we first inserted a stainless steel guide tube (OD, 0.85 mm; ID, 0.60 mm) containing the electrode. A hydraulic microdrive (MO95-S, Narishige) was used both to insert the guide tube and subsequently to advance the electrode into the brain. We sometimes implanted a plastic guide tube (OD, 1.1 mm; ID, 0.8 mm) semi-permanently at the location where DA neurons were concentrated. The location of the guide tube was visualized on MRIs and was confirmed to be directed to the substantia nigra pars compacta (SNc).

DA neurons were identified by their irregular and tonic firing around 5 spikes/s with broad spike potentials. Extracellular spikes may have an initial positive component or may be followed by prolonged positive component (Schultz and Romo 1987). A neuron with these features was thus determined to be a DA neuron candidate. Near the end of a long-term experimental session, we made electrolytic microlesions at the recording sites of DA neurons for later histological analysis. Later histological examination showed that the presumed DA neurons were located among tyrosinehydroxylase (TH) positive neurons. They were usually in the SNc (A9) and sometimes in the area medio-dorsal to the SNc (A8) (Kawagoe et al. 2004).

Definition of the learning stage

We divided the long-term training period of two monkeys into three stages based on the differentiation of saccade latencies between the rewarded and unrewarded conditions. First, for every sub-block, we obtained the saccade latency for the one rewarded trial and the mean saccade latency for the three unrewarded trials. We compared pair-wise the saccade latencies in the rewarded trials and the saccade latencies in the unrewarded trials using Wilcoxon signed-rank test ($P < 0.05$). We chose the pair-wise test because it was more tolerant to slow change in saccade latency that occurred within a block of 60 trials whether or not the saccade was rewarded (see Fig. 7). The statistical test was performed for the four blocks of IDR while one DA neuron was recorded.

For defining the learning stages, we divided the data in a block of 60 trials into the first and second halves. This was because, in an earlier learning stage, any difference between the rewarded and unrewarded trials in one block tended to be carried over to the next block (see Fig. 7, stage 2); this feature would not be captured by the data taken from the whole IDR block. If the statistically significant difference in saccade latency was observed in the second half for 2 consecutive experimental days, we defined that stage 2 started. If it was observed in the first half, we defined that stage 3 started.

Data analysis

NEURONAL ACTIVITY. For each DA neuron, we calculated the spontaneous activity and its responses to the fixation point, to the cue stimulus, and to the reward as the spike frequency. The spontaneous activity (–500 to 0 ms before fixation onset) was calculated for each neuron, and these were averaged for each stage. The fixation response (100–200 ms after fixation onset) was calculated for each neuron for the four blocks of IDR. The cue (100–300 ms for *monkey G* and 150–350 ms for *monkey C*) and reward (200–400 ms after reward onset) responses were calculated for each neuron separately for rewarded and unrewarded conditions.

STATISTICAL ANALYSIS FOR THE CUE AND REWARD RESPONSES OF DA NEURON. To determine the presence of cue and reward responses, we compared the spike frequency in the control period and the spike frequency in the test period using Wilcoxon signed-rank tests ($P < 0.01$) separately for rewarded and unrewarded trials in IDR for each neuron. The test period was 100–300 ms for *monkey G* and 150–350 ms for *monkey C* after cue onset for cue response and 200–400 ms after reward onset for reward response. The control period was set just before the onset of the fixation point, and its duration was the same as that of the test period. If the cue and reward responses were determined statistically, the responses were classified into two groups: increase and decrease.

SACCADE LATENCY. We calculated the mean saccade latency separately for the rewarded and unrewarded conditions in two ways. To examine long-term changes of saccade latency, we averaged saccade latencies for all trials in each experiment that consists of four blocks

of 1DR. To examine short-term (within a block) changes, we averaged saccade latencies for every sub-block in a block.

RESULTS

Overview

The data presented in this paper are based on the long-term training of two monkeys. During the training period of 1DR (3 mo in each monkey), we recorded 264 and 169 DA neurons in *monkey C* and *monkey G*, respectively. Among these neurons, 164 of 264 (62%) and 104 of 169 (62%) in each monkey responded to free reward by phasically increasing its discharge rate. In this experiment, we focused on neurons that responded to the delivery of free reward. We examined 53 of 164 neurons and 50 of 104 neurons in each monkey by performing four blocks of 1DR and one block of ADR. *Monkey C* and *monkey G* performed 626 and 431 blocks of 1DR, respectively, in addition to many blocks of ADR. In this paper, we will show only the data of 1DR.

Long-term changes in saccade parameters

The monkey's behavior changed during the long-term training period of the 1DR task. As shown in Fig. 2, saccade latency became shorter initially and then became longer gradually. Another clear change was the differentiation between the rewarded trials and the unrewarded trials. Saccade latency was initially not different between the two conditions, but the difference became clearer as the monkey experienced more blocks of 1DR such that saccade latencies in the rewarded condition were shorter than those in the unrewarded condition.

We defined the learning stages based on the reward-dependent differentiation of saccade latency. In learning stage 1, there was no differentiation in either the first half or the second half of the 1DR block. In learning stage 2, a statistical difference appeared only in the second half of the 1DR block (Wilcoxon signed-rank test, $P < 0.05$). In learning stage 3, there was a statistical difference already in the first half. The rationale for the definition came from the observation that the reward-dependent differentiation appeared earlier within a block as learning progressed. This phenomenon will be one of the major results described here.

The learning stages thus defined are indicated in Fig. 2. The reward-dependent differentiation of saccade latency occurred more quickly in *monkey G* than in *monkey C*. For example, stage 1 was much shorter in *monkey G*. Nonetheless, there were similarities between them. In both monkeys, learning stage 1 roughly corresponded to the period in which the saccade latency became shorter in general. Saccade latency appears to be shortest in stage 2. In stage 3, the saccade latency for the rewarded trials stayed relatively short, whereas that for the unrewarded trials became longer gradually.

Long-term shift in DA neuronal activity

Along with the long-term changes in saccade latency, the responses of DA neurons changed its pattern. Figure 3 shows the population activity of DA neurons of two monkeys at each learning stage separately for rewarded (black lines) and unrewarded (gray lines) trials. DA neurons showed three types of response: the response to the onset of the fixation point (fixation response), the response to the cue stimulus that indicated

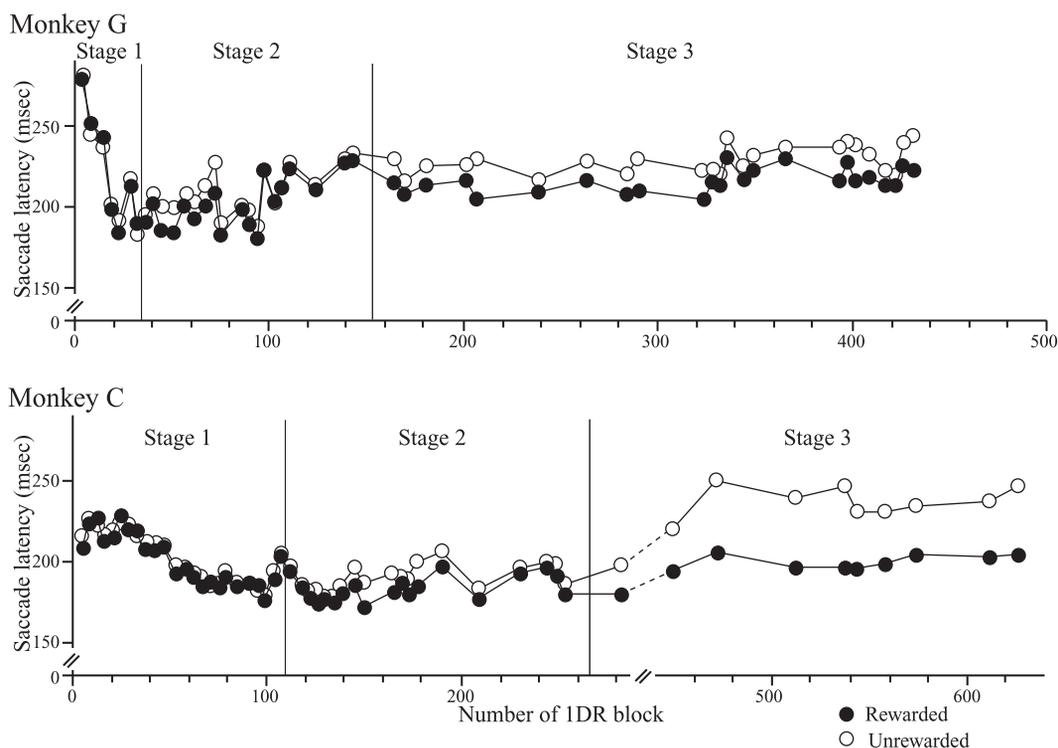


FIG. 2. Reward-contingent changes in saccade latency with long-term training of 1DR for *monkey G* (top) and *monkey C* (bottom). Mean saccade latencies for each experiment (i.e., 4 blocks of 1DR) are plotted, separately for rewarded trials (●) and unrewarded trials (○), against the cumulative number of 1DR blocks that the monkey experienced. We divided the long-term training period into 3 stages based on the differentiation of saccade latencies between the rewarded and unrewarded trials (see METHODS).

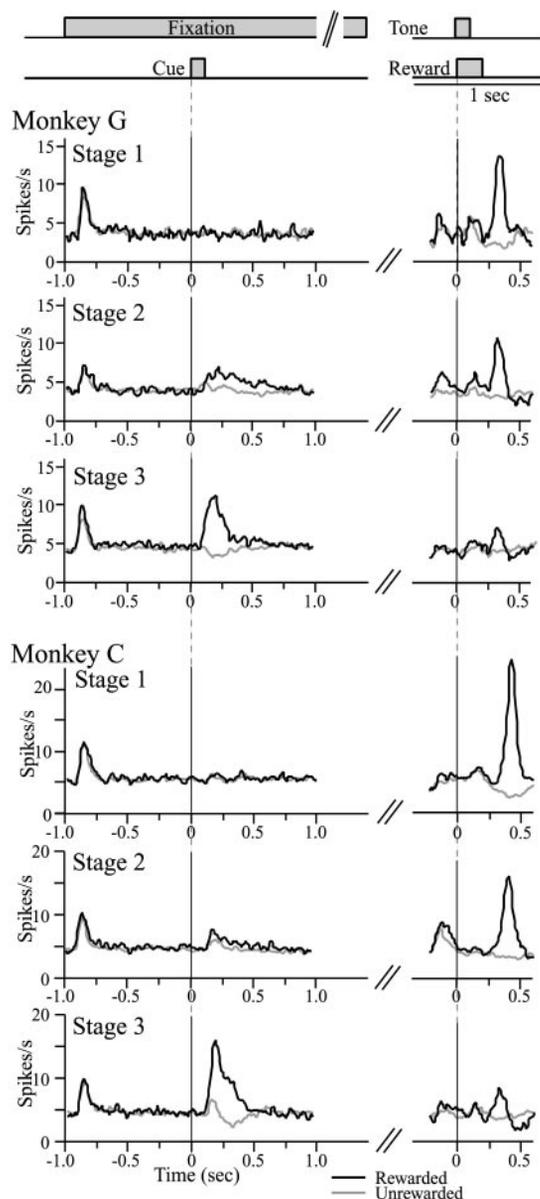


FIG. 3. Changes in population activity of DA neurons across 3 training stages for monkey *G* (top) and monkey *C* (bottom). Population activities are shown separately for rewarded trials (black lines) and unrewarded trials (gray lines) of IDR. They are aligned on fixation (left vertical axis), cue (left line), and reward onset (right line). The numbers of experiments were 7 and 24 in stage 1, 19 and 19 in stage 2, and 24 and 10 in stage 3 in monkey *G* and monkey *C*, respectively.

an upcoming reward (cue response), and the response to reward (reward response).

The cue and reward responses appeared differently depending on the reward conditions and the learning stages. In the rewarded trials (black lines in Fig. 3), all cue and reward responses were positive responses (increase in activity). The reward response was strong in stage 1, but became weaker in stages 2 and 3. In contrast, the cue response was absent in stage 1, weak in stage 2, and strong in stage 3. In the unrewarded trials (gray lines in Fig. 3), the responses were relatively weak and predominantly negative (decrease in activity). A negative response at the time of reward was clear in stage 1, whereas a negative response after the non-reward-indicating cue was clear in stage 3.

Long-term changes were also evident in the proportion of DA neurons that showed these responses. In general, progressively fewer neurons responded to the reward (Table 1), whereas progressively more neurons responded to the cue (Table 2). Cue responses were initially nonselective, mostly positive for both the reward-indicating and non-reward-indicating cues, but became selective, positive for the reward-indicating cue and negative for the non-reward-indicating cue (Table 2).

Short-term shift in cue and reward responses of DA neurons

The data in Fig. 3 show that the bi-directional reward responses of DA neuron (positive and negative responses in rewarded and unrewarded trials, respectively) were replaced with the bi-directional cue responses during the long-term training of IDR (long-term shift). However, in stages 2 and 3, the replacement of cue and reward responses was also observed within a block that took only ~ 7 min to perform (short-term shift). Such within-block changes are shown in Fig. 4. We plotted the mean magnitude of cue and reward responses against "sub-block number" (from 1 to 15). Since the four cue directions were chosen randomly in every block of four trials (which we call "sub-block"), each cue appeared once for each sub-block and 15 times in the entire block of 60 trials. The cue and reward responses are shown for two monkeys as population averages at each learning stage separately for rewarded (black) and unrewarded (dark gray) trials. In stage 1, DA neurons continued to respond to reward with bi-directional manner, but not to the cue. In stage 2, the reward response decreased gradually, whereas the response to the reward-indicating cue increased gradually. In stage 3, the shift of response from reward to the cue occurred quickly. This was true for both rewarded trials and unrewarded trials.

Short-term and long-term changes in cue responses of DA neurons

As indicated in METHODS, the task IDR was done in four blocks in which four different directions were rewarded in a random order. The rewarded direction was indicated only by reward itself when the monkey experienced the particular

TABLE 1. Response at the time of reward of DA neurons across three training stages in two monkeys

Stage	<i>N</i>	Rewarded	Nonrewarded
<i>Monkey G</i>			
1	7	6 (86%)	5 (71%)
2	19	7 (37%)	3 (16%)
3	24	3 (13%)	0 (—)
<i>Monkey C</i>			
1	24	22 (92%)	18 (75%)
2	19	11 (58%)	9 (47%)
3	10	2 (20%)	1 (10%)

For each neuron, the firing rates at the time of reward were compared with the firing rates in the control period separately for rewarded and unrewarded trials (Wilcoxon signed-rank test $P < 0.01$). All statistically significant responses were positive (increase in firing rate) in rewarded trials and negative (decrease in firing rate) in unrewarded trials. *N*, number of neurons.

TABLE 2. Cue response of DA neurons across three training stages in two monkeys

Stage	N	Reward-indicating cue		Non-reward-indicating cue	
		Positive	Negative	Positive	Negative
<i>Monkey G</i>					
1	7	1 (14%)	0 (—)	1 (14%)	0 (—)
2	19	10 (53%)	0 (—)	6 (32%)	1 (5%)
3	24	19 (79%)	0 (—)	0 (—)	6 (25%)
<i>Monkey C</i>					
1	24	2 (8%)	0 (—)	3 (13%)	0 (—)
2	19	6 (32%)	0 (—)	5 (26%)	2 (11%)
3	10	9 (90%)	0 (—)	0 (—)	3 (30%)

For each neuron, the firing rates after cue onset were compared with the firing rates in the control period separately for reward, and non-reward-indicating cues (Wilcoxon signed rank test $P < 0.01$). The statistically significant cue responses were classified into positive (increase in firing rate) or negative (decrease in firing rate) types. *N*, number of neurons.

block of 1DR. Therefore it was interesting to examine the transition of DA neuronal activity from one 1DR block to another. This is shown in Fig. 5 for two typical neurons recorded in stages 2 and 3. In stage 2, the DA neuron appeared to respond to the reward-indicating cue (indicated by a dotted circle), but the response was evident only in the later trials of a particular block. In the early trials, the neuron continued to respond to a non-reward-indicating cue that had been the reward-indicating cue in the preceding block. For example, in the left-up (LU) block (2nd column), the neuron responded to the right-down (RD) cue stimulus in the early trials (rewarded direction in the preceding block) and responded to the LU cue stimulus in the later trials (current rewarded direction). In the next block [left-down (LD) block, 3rd column], the neuron continued to respond to the LU cue (rewarded direction in the preceding block). Such transitional effects were much weaker in stage 3. The DA neuron's response to the reward-indicating cue was established quickly in one 1DR block. There was no indication that the neuron continued to respond to the same cue in the next block when it no longer indicated reward.

We found that the transitional effects shown in Fig. 5 were common among DA neurons (Fig. 6). Here the within-block changes of population cue responses of DA neurons are shown separately for three kinds of cues: the reward-indicating cue (R-cue), the non-reward-indicating cue that indicated reward in the preceding block (RN-cue), and the non-reward-indicating cue that indicated no reward in the preceding block (NN-cue). In stage 1, there was virtually no response to either of the cues. In stage 2, the response to R-cue started from a low value (close to the spontaneous activity) and increased gradually toward the end of the block (Spearman $P = 0.0010$ for *monkey G*, $P = 0.0023$ for *monkey C*). In contrast, the response to RN-cue started from a high value and decreased gradually; this was significant in *monkey G* (Spearman $P = 0.0037$ for *monkey G*, $P = 0.1115$ for *monkey C*). This contrasted with the response to NN-cue, which stayed close to the spontaneous activity (Spearman $P = 0.0249$ for *monkey G*, $P = 0.3419$ for *monkey C*). These results indicate that, in stage 2, the reward-predictive signals

in DA neurons were carried over to the next block of experiment, and it took many trials to be adapted to the new condition. In stage 3, the responses to R-cue grew faster and reached a high plateau level (Spearman $P = 0.0042$ for *monkey G*, $P = 0.0102$ for *monkey C*), whereas the response to RN- and NN-cues started from the spontaneous level and gradually deviated negatively from the spontaneous level (Spearman $P = 0.0294$ for *monkey G*, $P = 0.0012$ for *monkey C*). There was no obvious difference between the responses to RN- and NN-cues. The carry-over effect was minimal in stage 3.

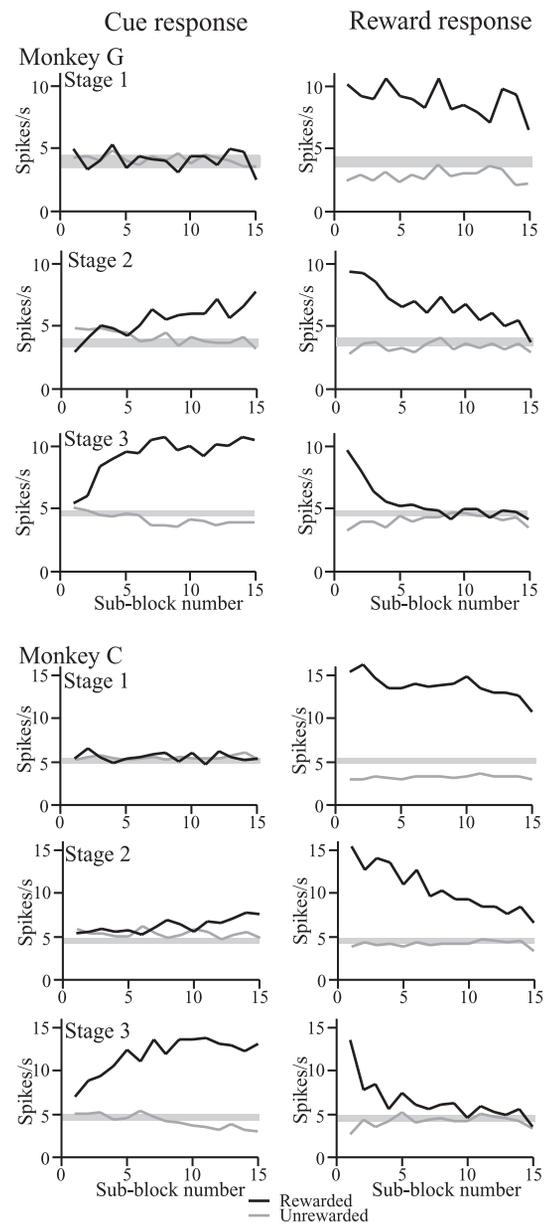


FIG. 4. Within-block changes of population responses of DA neurons to cue (left) and reward (right) across 3 training stages for *monkey G* (top) and *monkey C* (bottom). Data are shown separately for rewarded (black) and unrewarded (dark gray) trials. Abscissa indicates the sub-block number in 1 block of 1DR (maximum: 15). Light gray line indicates the mean \pm SE of spontaneous activity. The same database as in Fig. 3.

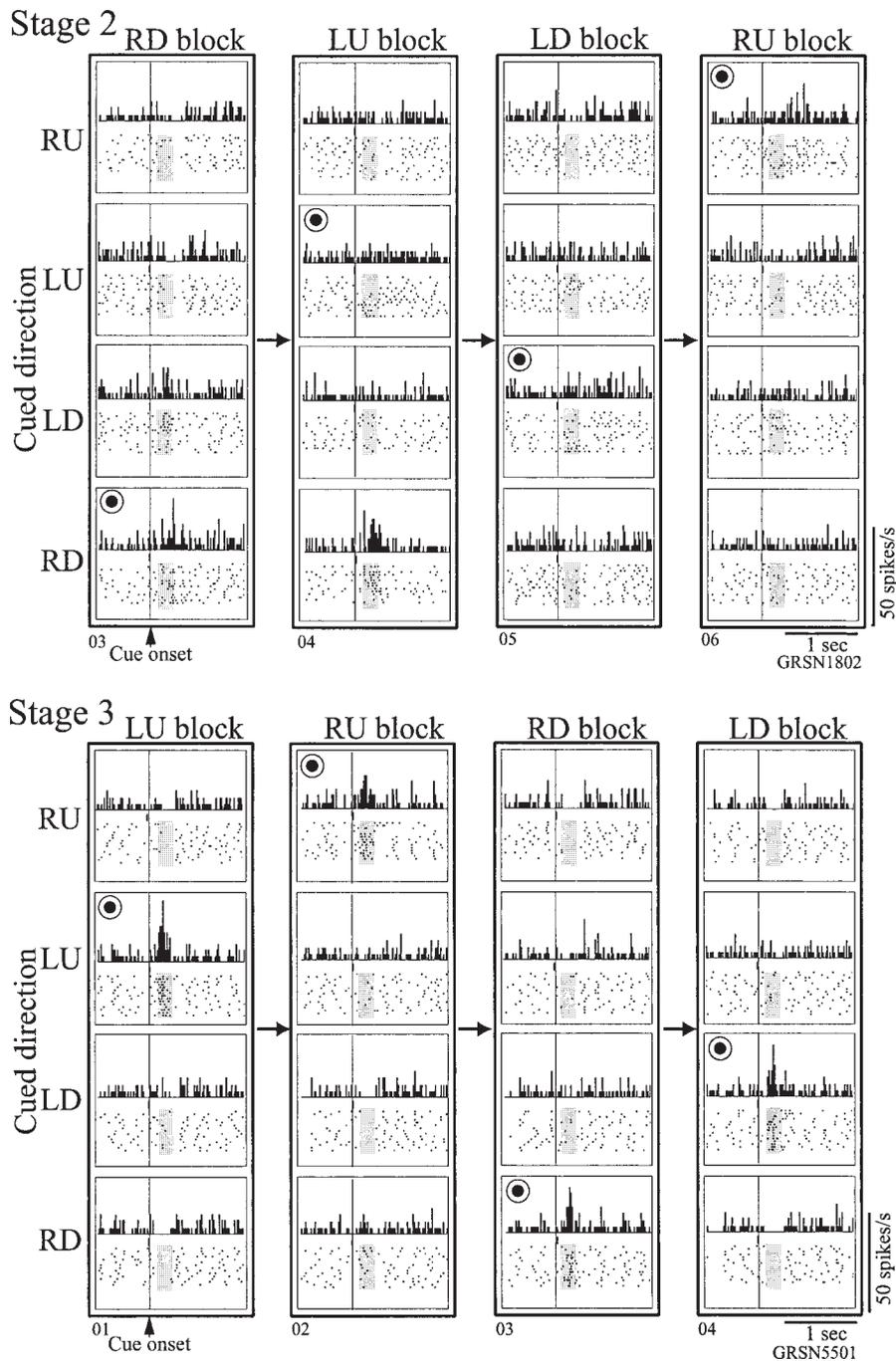


FIG. 5. Sample DA neuronal responses to the cue stimulus in stage 2 (*top*) and stage 3 (*bottom*) in monkey G. Data in each column include 1 block of 60 trials of IDR. In the histogram/raster display (bin width: 20 ms), spike activity aligned on cue onset is shown separately for different cue directions (RU, right-up; LU, left-up; LD, left-down; RD, right-down). Rewarded direction is indicated by dotted circle. Data of 4 blocks of IDR are arranged according to the actual order in the experiment from *left to right*. Target eccentricity was 20°. Gray bars indicate postcue activity period (100–300 ms after cue onset).

Long-term co-variation of DA neuronal activity and saccadic eye movements

We have shown that, while the monkey experienced IDR daily, both saccade latency and DA neuronal responses became differentiated proactively between the rewarded and unrewarded trials. These data raise the possibility that DA neuronal activity influences saccadic motor outputs. To examine this possibility further, we compared the within-block changes of DA neuronal responses and saccade latency.

Figure 7 shows the within-block changes of saccade latency in each stage, using the same format as used for DA neuronal responses (Fig. 6). In stage 1, saccade latency was not clearly differentiated among the three kinds of cue (i.e., R-, RN-,

NN-cues). In stage 2, saccade latency became shorter gradually for the R-cue, became longer gradually for the RN-cue, and remained long for the NN-cue. In stage 3, saccade latency started from similar levels for these cues, but became differentiated such that latencies for RN- and NN-cues were longer than for R-cue. These changes in saccade latency, especially the within-block changes in stage 2, are reminiscent of those in DA neuronal activity shown in Fig. 6.

Long-term change in response to fixation onset of DA neurons

As shown in Fig. 3, DA neurons responded to the onset of the fixation point, and the response was present throughout the

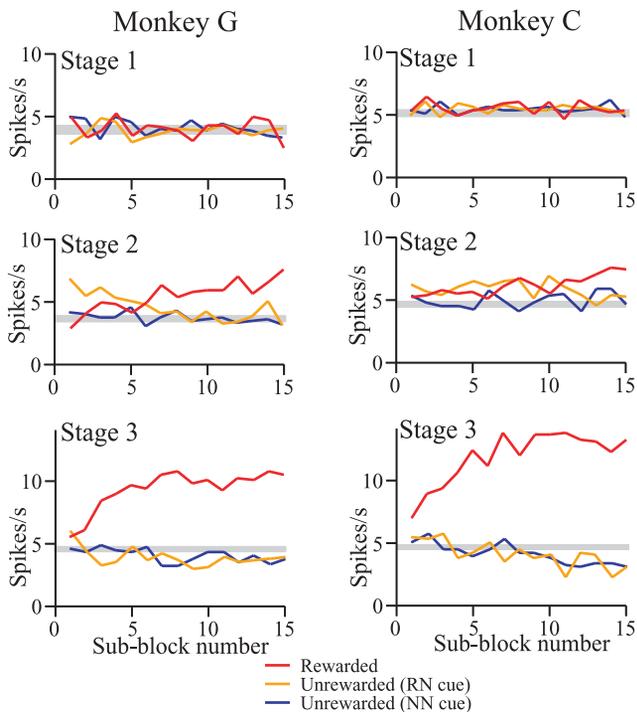


FIG. 6. Changes in population cue responses of DA neurons depend on the current and preceding reward conditions. Data are shown separately for 3 training stages and separately for 3 different conditions: R, rewarded; RN, rewarded in the preceding block and unrewarded in the current block; NN, unrewarded in both the preceding and current blocks. Abscissa indicates the sub-block number in 1 block of 1DR. Gray line indicates the mean \pm SE of spontaneous activity.

learning stages. The population response to the fixation point plotted against trial numbers (Fig. 8) also showed that the fixation response was weakly present throughout trials in every stages in both monkeys (i.e., increase from the spontaneous activity). However, the response changed considerably within a block of 1DR. In stage 3, the fixation response was very strong in the first trial and was diminished in the second or later trials. This prominent fixation response was common in both monkeys. Interestingly, the fixation response in the first trial was stronger than the cue or reward response (cf. Figs. 8 and 4) and was consistent across different DA neurons. This first trial effect was less clear in stages 1 and 2.

DISCUSSION

Long- and short-term shifts of DA neuronal activity from reward to reward predictor

In a previous study (Kawagoe et al. 2004), we recorded from DA neurons in well-trained monkeys. They responded strongly to reward predictor but not to reward itself except for the first or second trial. In rewarded trials, DA neurons responded to the delivery of a reward at the beginning of a 1DR block, but this response was soon replaced by a positive response to the cue stimulus that indicated an upcoming reward. In unrewarded trials, activity of DA neurons was suppressed at the time of reward at the beginning of a 1DR block, but the response was soon replaced by a negative response to the cue stimulus that indicated no reward. Every time the rewarded direction was changed, DA neurons changed their responses from reward to

cue in a few trials and adapted to the altered position-reward mapping.

However, it was unknown how DA neurons adapted to the altered position-reward mapping so quickly. In this study, we found that the speed of the short-term shift of DA neuronal responses became quicker with long-term training of 1DR. In the early stage of training, the dominant activity of DA neurons was the response to reward. The short-term shift first appeared in the intermediate stage, but the shift was slow, in that activity of DA neurons transferred from reward to reward-indicating cue gradually within a block. The short-term shift became quicker in the advanced stage, and therefore the dominant activity of DA neurons was the response to the reward-indicating cue. Consequently, the overall activity of DA neurons shifted from reward to the reward-indicating cue during the long-term training (long-term shift).

Our results are in accordance with a pioneering work by Ljungberg et al. (1992), Schultz et al. (1993), and Mirenovic and Schultz (1994). In their study, DA neurons in naive monkeys tended to respond to a reward, whereas those in well-trained monkeys responded to the earliest stimulus that indicated the reward. After further training, however, the response of DA neurons to the reward-predicting stimulus became weaker. In contrast, DA neurons in our experiments continued to respond to the cue stimulus even after long-term experience of 1DR. The difference may originate from a difference in reward schedule. In the study of Schultz et al., a reward was given in every trial and was fully expected, whereas in our 1DR task, a reward was given selectively for one out of four positions. The stimulus-reward mapping was

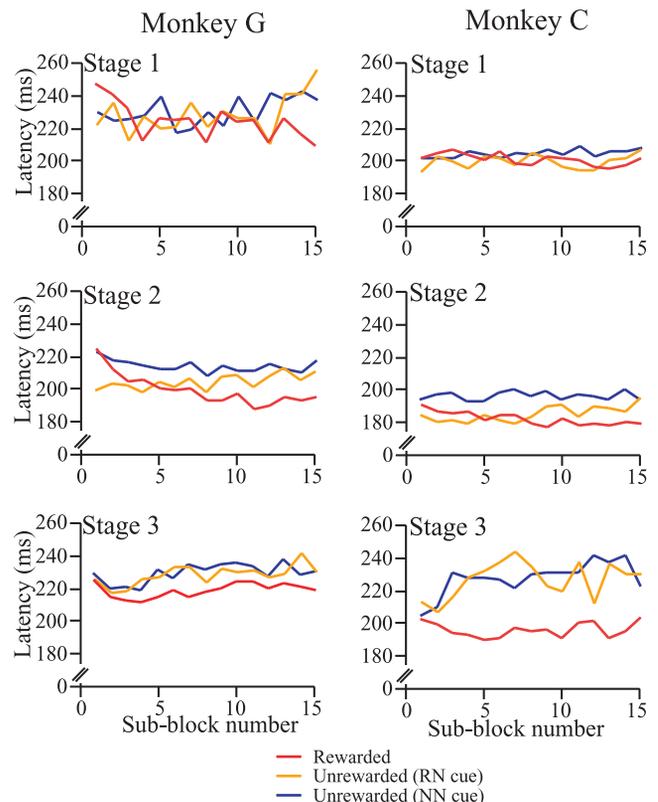


FIG. 7. Changes in saccade latency depend on the current and preceding reward conditions. Each data point indicates the mean saccade latency for each reward condition and for each training stage. Same format as Fig. 6.

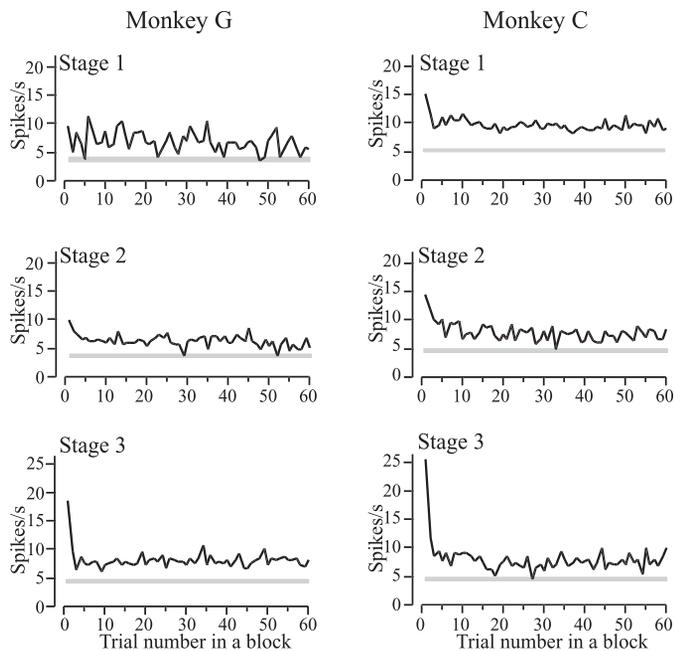


FIG. 8. Responses of DA neurons in 1DR to the onset of the fixation point across 3 training stages in 2 monkeys. Data include both rewarded and unrewarded trials, since no information on reward was given when the fixation point came on. For the same reason, abscissa shows the trial numbers for all 4 directions together. Gray line indicates the mean \pm SE of spontaneous activity.

fixed in the study of Schultz et al., whereas it was variable in our 1DR task. In the study of Schultz et al., DA neurons responded either only to reward itself or only to a reward-predicting stimulus, whereas in 1DR, the responses of DA neurons shifted from reward to a reward predictor within a block. In this sense, our data provided a new aspect in the function of DA neurons, namely, their relationship to behavioral switching.

Emergence of behavioral bias with short- and long-term shifts of DA neuronal activity

We previously found that the 1DR task gave rise to a strong behavioral bias: saccade latency was shorter and saccade velocity was higher when the saccade was followed by a reward (rewarded condition) than when it was followed by no reward (unrewarded condition) (Takikawa et al. 2002). The comparison of DA neurons and caudate (CD) projection neurons in well-trained monkeys suggested that DA neurons, with their connection to CD neurons, modulate the spatially selective signals in CD neurons in the reward-predicting manner (Kawagoe et al. 2004), and CD neurons in turn modulate saccade parameters (Itoh et al. 2003; Watanabe et al. 2003) with their connections to the substantia nigra pars reticulata and the superior colliculus (Hikosaka et al. 2000). The present data are consistent with this hypothesis. Both saccade latency and DA neuronal activity changed within a block of 1DR trials, usually in the opposite directions (Figs. 6 and 7).

However, the related changes in DA neuronal activity and saccade latency may not necessarily indicate a causal relationship. An alternative possibility is that DA neuronal activity and saccade latency are modulated by a common input. For example, CD neurons could be a source of the common input. While CD neurons may influence saccade parameters, they may also

influence activity of DA neurons with their direct connections or indirect connections through substantia nigra pars reticulata neurons (Parent and Hazrati 1994; Tepper et al. 1995). In this case, the reward-predicting activity of DA neurons would originate from CD neurons.

The two hypotheses stated above are contrasting and will remain an important topic in future studies on basal ganglia mechanisms.

Relation to behavioral switching

The performance of well-experienced monkeys in the 1DR task may be compared with behavioral switching rather than learning. For every block of 1DR, the rewarded direction was changed, and the monkey changed its behavior so that saccades were quicker and faster to the currently rewarded direction. The monkey acquired the ability of behavioral switch while DA neurons changed their cue responses quickly to adapt to the current position-reward mapping. Both behavioral and neuronal changes became faster as the monkey experienced more 1DR blocks.

It has been suggested that the basal ganglia play an important role in switching behavior (see Redgrave et al. 1999b). A large body of evidence is derived from studies on patients with Parkinson's disease (Brown and Marsden 1990). The basal ganglia are equipped with parallel mechanisms that may act antagonistically on target structures such as the superior colliculus and the thalamus (Albin et al. 1989; Chevalier and Deniau 1990). The concurrent and sequential actions of these mechanisms, respectively, would be suitable for selection (Mink 1996; Nambu et al. 2002) and switching (Hikosaka et al. 1993) of motor programs or behavioral sets.

However, there have been few studies indicating that changes in neuronal activity in the basal ganglia were correlated with (or caused) behavioral switching. A series of studies in our laboratory using 1DR task have shown that, when position-reward mapping was altered, neurons in the caudate (Kawagoe et al. 1998), substantia nigra (Sato and Hikosaka 2002), and superior colliculus (Ikeda and Hikosaka 2003) changed their activity, together with saccade parameters (Takikawa et al. 2002), to match the position-reward mapping. In this study, we showed that the cue responses of DA neurons and the differentiation of saccade latency emerged after repeated experiences of different sets of position-reward mapping and became quicker jointly until they may be called switching. These results suggest that the reward-predictive activity of DA neurons may play an important role in adapting to frequent switches of position-reward mapping.

Response of DA neurons to the initiation signal

A robust response of DA neurons occurred to the onset of the fixation point which signals the start of a task trial. The fixation response also underwent long-term changes: in the advanced stage it became prominent in the first couple of trials in a block of experiments (Fig. 8). Remarkably, this response was one of the strongest responses emitted by DA neurons. It has been reported that DA neurons respond to novel sensory stimuli (Ljungberg et al. 1992; Horvitz 2000), suggesting that a major signal carried by DA neurons is novel (Redgrave et al. 1999a). The fixation response may signal novelty because there

was a clear break before the fixation point appeared in the first trial in a block. We also noticed that DA neurons typically responded to a visual or auditory stimulus when it was presented unexpectedly, but stopped responding if the stimulus was repeated; a subtle sound outside the monkey's view was particularly effective. The onset of the fixation point in the first trial may be regarded as a novel, unexpected stimulus, while the same stimulus in the second trial would no longer be novel. On the other hand, the fixation point in the first trial informed the monkey of a change or switch in the learned set, provided that the monkey was well trained. This is consistent with the hypothesis that DA neurons contribute to the processes of terminating current selections and opening new ones, as Redgrave et al. (1999a) argued previously.

REFERENCES

- Albin RL, Young AB, and Penney JB.** The functional anatomy of basal ganglia disorders. *Trends Neurosci* 12: 366–375, 1989.
- Brown RG and Marsden CD.** Cognitive function in Parkinson's disease: from description to theory. *Trends Neurosci* 13: 21–29, 1990.
- Chevalier G and Deniau JM.** Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci* 13: 277–280, 1990.
- Dickinson A and Balleine B.** Motivational control of goal-directed action. *Anim Learn Behav* 22: 1–18, 1994.
- Hikosaka O and Wurtz RH.** Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *J Neurophysiol* 49: 1268–1284, 1983.
- Hikosaka O, Matsumura M, Kojima J, and Gardiner TW.** Role of basal ganglia in initiation and suppression of saccadic eye movements. In: *Role of the Cerebellum and Basal Ganglia in Voluntary Movement*, edited by Mano N, Hamada I, and DeLong MR. Amsterdam: Elsevier, 1993, p. 213–219.
- Hikosaka O, Takikawa Y, and Kawagoe R.** Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev* 80: 953–978, 2000.
- Hollerman JR and Schultz W.** Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1: 304–309, 1998.
- Horvitz JC.** Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96: 651–656, 2000.
- Ikeda T and Hikosaka O.** Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39: 693–700, 2003.
- Itoh H, Nakahara H, Hikosaka O, Kawagoe R, Takikawa Y, and Aihara K.** Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior. *J Neurophysiol* 89: 1774–1783, 2003.
- Judge SJ, Richmond BJ, and Chu FC.** Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980.
- Kawagoe R, Takikawa Y, and Hikosaka O.** Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1: 411–416, 1998.
- Kawagoe R, Takikawa Y, and Hikosaka O.** Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. *J Neurophysiol* 91: 1013–1024, 2004.
- Ljungberg T, Apicella P, and Schultz W.** Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67: 145–163, 1992.
- Matsumura M, Kojima J, Gardiner TW, and Hikosaka O.** Visual and oculomotor functions of monkey subthalamic nucleus. *J Neurophysiol* 67: 1615–1632, 1992.
- Mink JW.** The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 50: 381–425, 1996.
- Mirenowicz J and Schultz W.** Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72: 1024–1027, 1994.
- Nambu A, Tokuno H, and Takada M.** Functional significance of the cortico-subthalamo-pallidal “hyperdirect” pathway. *Neurosci Res* 43: 111–117, 2002.
- Parent A and Hazrati L-N.** Multiple striatal representation in primate substantia nigra. *J Comp Neurol* 344: 305–320, 1994.
- Redgrave P, Prescott TJ, and Gurney K.** Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22: 146–151, 1999a.
- Redgrave P, Prescott TJ, and Gurney K.** The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience* 89: 1009–1023, 1999b.
- Robinson DA.** A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10: 137–145, 1963.
- Sato M and Hikosaka O.** Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *J Neurosci* 22: 2363–2373, 2002.
- Schultz W.** Predictive reward signal of dopamine neurons. *J Neurophysiol* 80: 1–27, 1998.
- Schultz W.** Getting formal with dopamine and reward. *Neuron* 36: 241–263, 2002.
- Schultz W and Romo R.** Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *J Neurophysiol* 57: 201–217, 1987.
- Schultz W and Dickinson A.** Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500, 2000.
- Schultz W, Apicella P, and Ljungberg T.** Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13: 900–913, 1993.
- Schultz W, Dayan P, and Montague PR.** A neural substrate of prediction and reward. *Science* 275: 1593–1599, 1997.
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, and Hikosaka O.** Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* 142: 284–291, 2002.
- Tepper JM, Martin LP, and Anderson DR.** GABA_A receptor-mediated inhibition of rat substantia nigra dopaminergic neurons by pars reticulata projection neurons. *J Neurosci* 15: 3092–3103, 1995.
- Watanabe K, Lauwereyns J, and Hikosaka O.** Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *J Neurosci* 23: 10052–10057, 2003.