

# Shaping of Motor Responses by Incentive Values through the Basal Ganglia

Benjamin Pasquereau,<sup>1,4</sup> Agnes Nadjar,<sup>1,4</sup> David Arkadir,<sup>1,3,4</sup> Erwan Bezdard,<sup>1</sup> Michel Goillandeau,<sup>1</sup> Bernard Bioulac,<sup>1,2,4</sup> Christian Eric Gross,<sup>1,2,4</sup> and Thomas Boraud<sup>1,4</sup>

<sup>1</sup>Basal Gang, Laboratoire de Neurophysiologie, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5227, Université Victor Segalen, 33076 Bordeaux, France, <sup>2</sup>Centre Hospitalier Universitaire de Bordeaux, 33800 Bordeaux, France, <sup>3</sup>Interdisciplinary Center for Neural Computation, The Hebrew University, Jerusalem 91905, Israel, and <sup>4</sup>Laboratoire Franco-Israelien de Neurophysiologie et Neurophysique des Systemes, 33076 Bordeaux, France

The striatum is a key neural interface for cognitive and motor information processing in which associations between reward value and visual stimulus can be used to modify motor commands. It can guide action–selection processes that occur farther downstream in the basal ganglia (BG) circuit, by encoding the reward value of an action. Here, we report on the study of simultaneously recorded neurons in the dorsal striatum (input stage of the BG) and the internal pallidum (output stage of the BG) in two monkeys performing a center-out motor task in which the visual targets were associated with different reward probabilities. We show that the tuning curves of motor-related neurons in both structures are modulated by the value of the action before movement initiation and during its execution. The representations of values associated with different actions change dynamically during the task in the internal globus pallidus, with a significant increase in the number of encoding neurons for the chosen target at the onset of movement. This report sheds additional light on the functional differences between the input and output structures of the BG and supports the assertion that the dorsal basal ganglia are involved in movement-related decision-making processes based on incentive values.

**Key words:** electrophysiology; striatum; globus pallidus; monkey; behavioral task; reward

## Introduction

In a visually guided motor task, decision making is a distributed neural process that involves the basal ganglia (BG), which presumably act as a “helper system” (Opris and Bruce, 2005; Samejima et al., 2005; Morris et al., 2006). Studies that focused on the primary input stage of the BG, the striatum, suggest that it is modulated by reward-related activity in midbrain dopaminergic neurons (Hollerman et al., 1998; Fiorillo et al., 2003; Satoh et al., 2003; Morris et al., 2004) and is thus able to determine the potential reward probability and quantity associated with a specific stimulus (Cromwell et al., 2005; Schultz, 2005). In turn, this information can be used to help select an appropriate behavioral response. The encoding of the incentive value of a specific stimulus by dopaminergic (Satoh et al., 2003; Tobler et al., 2005; Morris et al., 2006) or striatal (Arkadir et al., 2004; Morris et al., 2004) neurons is well documented, but interactions between such incoming cognitive information and a behavioral–motor outcome are poorly understood. In a recent paper, Samejima et al. (2005) showed that the specific reward value of an action may be

represented by striatal neurons and predicted that it would guide action selection in the downstream BG circuit. To address this hypothesis, we investigated whether the representation of reward value is indeed transmitted farther downstream in the basal ganglia and whether this cognitive information interacts with representations of motor parameters. This combined representation of action (motor) and cognitive (reward value) parameters could contribute to the decision-making process. Neural responses were thus recorded simultaneously in the major input structure of the BG, the motor striatum (dorsolateral putamen), and the major output nucleus [the internal globus pallidus (GPI)] of monkeys involved in a center-out visually guided motor task, in which each visual cue was associated with different reward values (see Fig. 1A,B).

## Materials and Methods

**Animals and surgery.** The study was conducted on two female rhesus monkeys (*Macaca mulatta*, weighing 5.6 and 4 kg). All experiments were performed during daytime. Although food was available *ad libitum*, the primates were kept under hydric restriction to increase their motivation during the task learning and recording sessions. A veterinarian skilled in the healthcare and maintenance of nonhuman primates supervised all aspects of animal care. Surgical and experimental procedures were performed in accordance with the Council Directive of 24 November 1986 (86/609/EEC) of the European Community and the National Institute of Health *Guide for the Care and Use of Laboratory Animals*. The monkeys were given daily training for them to learn a behavioral task. Once they had reached a success rate >0.95 for a series of >400 trials (9 and 6 months training, respectively), a recording chamber was implanted on

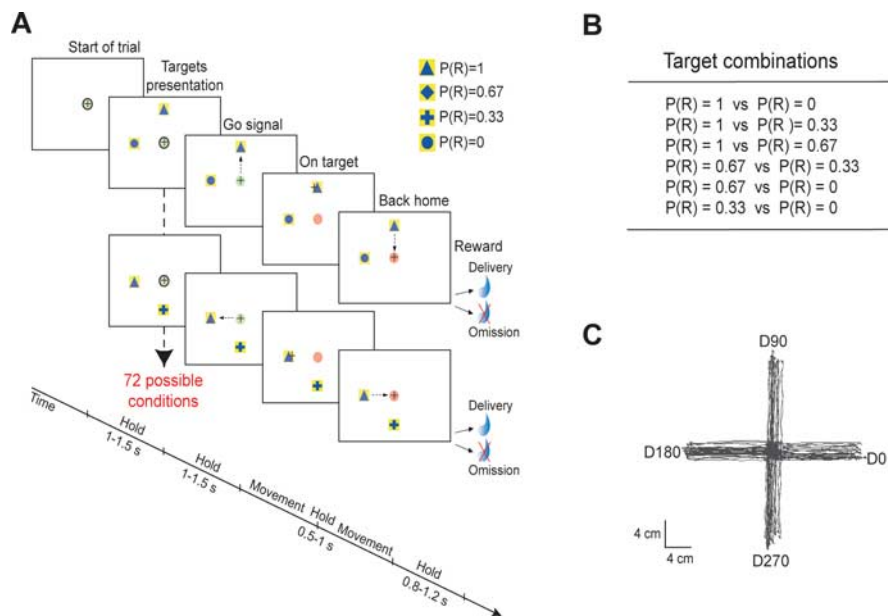
Received Aug. 28, 2006; revised Dec. 26, 2006; accepted Dec. 27, 2006.

B.P. was supported by Ministry of the Research Grant 14333-2004. We thank Rony Paz for his helpful scientific comments. The English of the manuscript was proofread by Techtrans Consulting, Meudon, France.

Correspondence should be addressed to Thomas Boraud, Laboratoire de Neurophysiologie, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5227, 146, rue Leo Saignat, 33076 Bordeaux Cedex, France. E-mail: tboraud@u-bordeaux2.fr.

DOI:10.1523/JNEUROSCI.3745-06.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/271176-08\$15.00/0



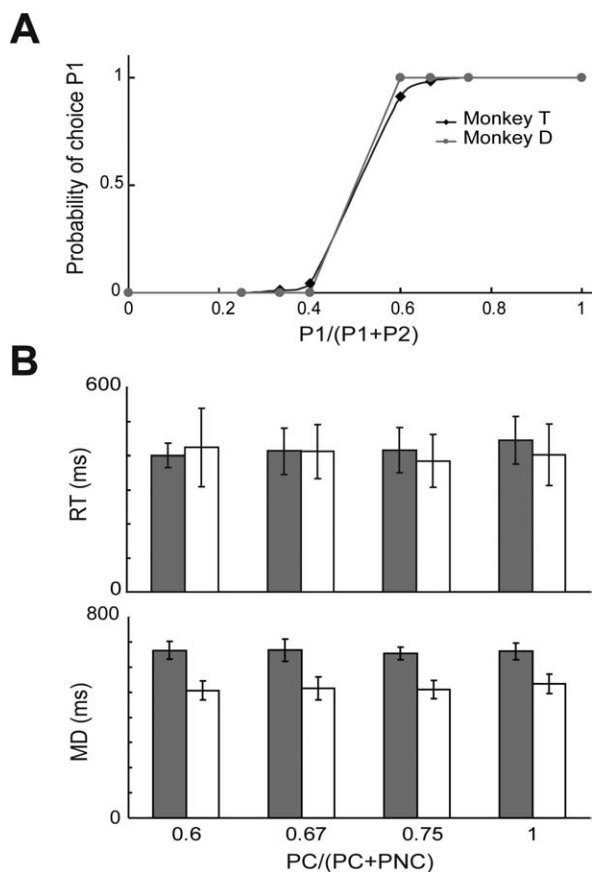
**Figure 1.** Behavioral paradigm. **A**, The reward probability-based free-choice task. Two targets (P1 and P2) are displayed simultaneously during each trial, in four possible positions in random order (6 target combinations) and in random locations (4 × 3 possibilities). **B**, Combinations of displayed targets during the task. **C**, Example of movement trajectories executed during 50 successful trials by monkey T.

the skull of each animal. The surgical procedure for attaching the recording chamber has been extensively described in previous publications (Bezard et al., 2001; Boraud et al., 2001).

**Behavioral task.** The task was monitored using Labview (National Instruments, Austin, TX). Monkeys were trained to move a custom-made manipulandum in a horizontal plan (26 × 26 cm) with their right hand. This manipulandum moved a cursor on a computer screen placed 50 cm in front of the animal’s face. The monkeys initiated a trial by keeping the cursor inside a central green circle (precue interval) (Fig. 1A). After a random period (1–1.5 s), two different targets (chosen from four possible cues) were displayed simultaneously on the screen (Fig. 1B). Each cue appeared in one of four possible directions (0, 90, 180, and 270°) and was associated with a specific probability of a reward being delivered at the end of a successful trial. The associated probabilities were as follows:  $P(R) = 0, 0.33, 0.67,$  and  $1$ . The monkeys were overtrained to identify the geometric figures depicted inside the different targets. Cue type and direction were determined randomly for each trial. To induce a situation in which there was always a “better” choice, a single trial could not include two identical cues or two targets in the same location. After a random period (1–1.5 s), the “Go” signal is given by the disappearance of the black central circle, and the monkeys had to initiate a movement toward one of the two targets (Fig. 1C). The cursor had to be maintained on the target for a random period in the range of 0.5–1 s, before being moved back to the central circle. To complete the trial, the monkeys had to maintain the cursor inside the central circle for a minimum random duration (0.8–1.2 s). Disappearance of the central circle indicated to the animal that it had succeeded. The reward was delivered (0.3 ml of fruit juice) according to the probability associated with the selected target. For each successful trial, if the monkey chose the target associated with the highest probability of receiving a reward, its choice was defined as optimal. If not, it could still receive its reward with a probability equal to that determined for the chosen target. The trials were separated by 2–2.5 s intertrial intervals (ITIs), during which the screen was black. In the case of an error, the trial was aborted and followed by an ITI.

**Training.** During the first learning steps, only one target [ $P(R) = 1$ ] was displayed to both monkeys. The cue appeared randomly in any of the four directions. After each monkey had learned to execute the different movements, a different target associated with another reward probability [ $P(R) = 0.67, P(R) = 0.33,$  and then  $P(R) = 0$ ] was presented on the screen. Through repetition of the same trial, the monkeys learned the reward value associated with each target. Finally, the six target combinations were used randomly, and the monkeys were free to select any direction of movement.

**Recordings and data acquisition.** During the recording session, the monkeys’ heads were immobilized. The single-unit activity was recorded using eight individually driven glass-coated tungsten microelectrodes (0.5 MΩ at 1 kHz). Recordings were performed simultaneously in the left striatum (four electrodes) and in the left GPi (four electrodes). Microelectrode sets were initially placed in two guides with an inner diameter of 1.3 mm. At the beginning of each session, both guides were lowered manually to a 10 mm depth below the dura. Subsequently, electrodes were individually lowered (Electrode Positioning System; Alpha Omega Engineering, Nazareth, Israel) until the typical signal of striatal neurons or pallidum neurons (GPi) was detected. Each electrode signal was amplified with a gain of  $10^3$  and filtered with a bandpass of 300 Hz to 6 kHz (Multi-Channel Processor+; Alpha Omega Engineering). When the spike-to-noise ratio was approximately  $>3$ , the electrical activity was sorted and classified on-line using a template-matching algorithm (Multi-Spike Detector; Alpha Omega Engineering) and stored by means

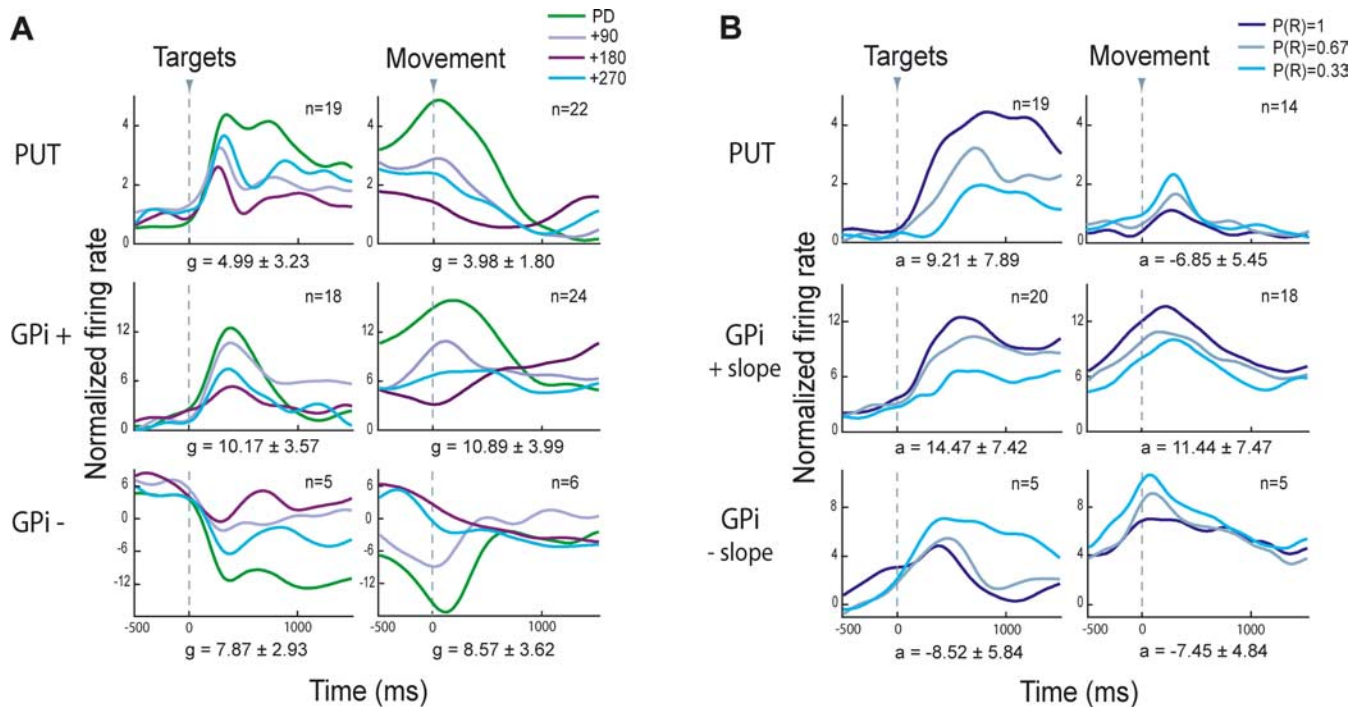


**Figure 2.** Behavioral results. **A**, Probability of choosing target P1, according to its relative reward value. P1 and P2 were target electronic labels with no visible differentiation between the two as perceived by the monkeys. The monkeys optimized their strategy by choosing cues associated with the highest reward probabilities. **B**, Mean and SD of the RT and MD of successful trials. The gray and white columns represent monkeys T and D, respectively. The ratio  $PC/(PC+PNC)$  represents the degree of choice complexity between the possibly chosen and the nonchosen cue probabilities.

**Table 1. The distribution of the responding neurons**

	Direction-sensitive only	Probability-sensitive only	Both parameters sensitive	Interaction
Target presentation				
PUT	15 of 148 (10%)	13 of 148 (9%)	20 of 148 (14%)	14 of 148 (9%)
GPI	13 of 111 (12%)	11 of 111 (10%)	14 of 111 (13%)	9 of 111 (8%)
Movement				
PUT	18 of 148 (12%)	6 of 148 (4%)	23 of 148 (16%)	14 of 148 (9%)
GPI	39 of 111 (35%)	5 of 111 (4%)	18 of 111 (16%)	10 of 111 (9%)

The activity of neurons, in the two task phases, is modulated by the direction of movement and by reward prediction (two-way ANOVA,  $p < 0.05$ ). Only a part of the neurons sensitive to both parameters shows an interaction between both encodings.



**Figure 3.** Population perievent histograms. *n* refers to the number of neurons for each population selected according to the corresponding regression. **A**, Neural discrimination of movement direction. The encoding of movement direction was modeled by a tuning curve (nonlinear regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ). The superimposed responses were related to PD, +90, +180, and +270°. GPI neurons fired according to target location during excitatory (GPI+) or inhibitory (GPI-) response peaks to task epochs. Below each plate, the distribution (mean  $\pm$  SD) for the gain (*g*) of the individual tuning curves is provided. **B**, Representative reward-value encoding neurons. Probability encoding is achieved by means of a monotonic variation in firing rate, with increasing reward probability (linear regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ). This encoding can be characterized by an increasing (positive slope) or decreasing (negative slope) firing rate. Below each plate, the distribution (mean  $\pm$  SD) for the slope (*a*) of the individual linear regressions is provided.

of an analog-to-digital converter at 12 kHz (AlphaMap; Alpha Omega Engineering). During the lowering of the electrode, the first neurons observed had a low tonic firing rate typical of cells in the putamen (PUT) (0.5–5 spikes/s). We recorded extracellular spike activity of presumed projection neurons, which showed very little spontaneous activity (Hikosaka et al., 1989), although no such activity was detected with presumed interneurons, which showed irregular tonic discharge (Aosaki et al., 1995). The passage of the electrode into the pallidum was then marked by a sharp increase in background neural activity, and isolated action potentials characteristic of pallidal neurons were of short duration ( $\sim 0.3$  ms from onset of initial negativity to peak positivity) and were characterized by high tonic firing (DeLong, 1971). Neurons of GPI fired with similar proportions of regular, irregular, and bursting patterns (Boraud et al., 2001), and no spontaneous pauses were found, contrary to neurons of the external globus pallidus (DeLong, 1971). Behavioral events and the *x* and *y* position of the manipulandum were merged on-line with the recording files.

**Data analysis.** The analyses were performed with custom-made Matlab (MathWorks, Natick, MA) programs and NeuroExplorer (Nex Technologies, Littleton, MA). The data were expressed in the form of mean  $\pm$  SD. The task design encompassed six possible combinations of cue pairs (we

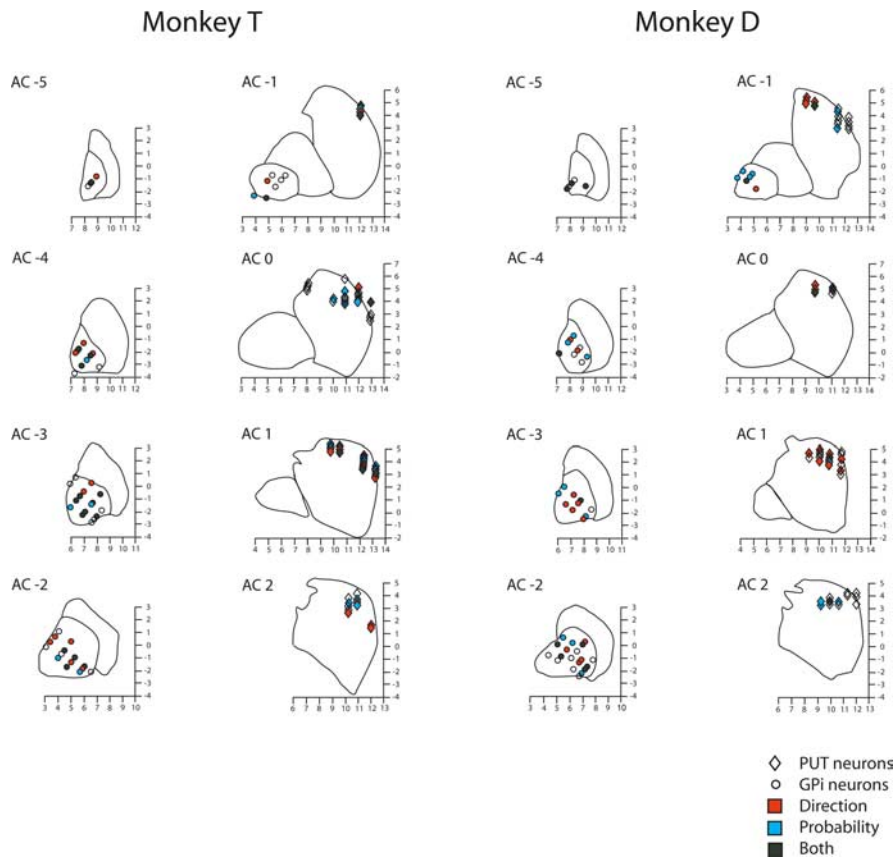
have electronic labels for target P1 and P2, but the monkeys had no cue to differentiate them), with four directions for one target and three for the other (Fig. 1A, B), thus making a total of 72 combinations (supplemental Fig. 1, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). However, because a single trial could not include two identical cues or two targets in the same location, cues and target locations were not independent. For this reason, it was not possible to perform straightforward multiway ANOVA. To overcome this obstacle, we collapsed the conditions and performed sequential analysis.

Because our first goal was to assess the hypothesis of motor response shaping by incentive value, we focused our efforts on a chosen direction for the statistical analysis.

The first step was to compute whether task parameters (reward probability and movement direction) modulated the behavior of both monkeys with the same characteristics. Reaction time (RT) and movement duration (MD) for correct responses were analyzed using a three-way ANOVA (reward probability  $\times$  movement direction  $\times$  monkey,  $p < 0.05$ ). The effect of the complexity of choice during the task on behavioral parameters (RT and MD) was assessed using a one-way ANOVA ( $p < 0.05$ ).

Then, for each neuron, average neural responses to task epochs were





**Figure 4.** Theoretical reconstruction of recording sites. The dots show the estimated locations of the electrode tips during recording sessions (scale in millimeters). These are identified according to the encoded parameter (movement direction, reward probability, or both) during either the preparatory epoch or movement. If overlapping occurs and neurons respond to different parameters, the location is considered to be the multiparameter type (black dots).

obtained using peristimulus time histograms (PSTHs) with 20 ms bins and smoothed with a Gaussian window,  $\sigma = 60$  ms. The average firing rate was normalized by subtracting base activity recorded during ITIs. To study which information was conveyed during each neural response (0: 500 ms after target presentation and  $-200:300$  ms centered on movement onset), a two-way ANOVA was used (reward probability  $\times$  movement direction,  $p < 0.05$ ). This test also assessed interactions between both behavioral parameters. To evaluate the neuronal parameter encoding, we used the peak value of a Gaussian-smoothed PSTH for each condition described below. The relationship between neural firing and movement direction (direction encoding) was subsequently modeled by nonlinear regression analysis with the use of the first-degree periodic function:  $y = a + g \cdot \cos(\theta - \theta_0)$ , where  $\theta$  is the target direction,  $a$  is the baseline discharge rate of the cell,  $g$  is the half-wave amplitude of the sinusoidal function (i.e., gain), and  $\theta_0$  is the direction of the sinusoidal function associated with the greatest change in firing rate (compared with baseline), i.e., the regression preferred direction (PD). The parameter  $\gamma$  thus represents the predicted sinusoidal discharge rate model. We then computed the associated nonlinear regression during both epochs, for  $P(R) = 1$ , i.e., a target with no uncertainty. The coefficient of determination ( $R^2 > 0.8$ ) produced by the regression analysis of the tuning curve was taken as an estimator for direction encoding. The gain of the tuning curve was significant when its value exceeded at least 2.5 SDs from the baseline. This threshold corresponds to the 95% confidence limit of neural activity, assuming the variations to be attributable to Gaussian noise. The tendency of striatal and pallidal neurons to monotonically increase or decrease their firing rate with increasing reward probability (reward probability encoding) was tested using linear regression ( $R^2 > 0.8$ ;  $p < 0.05$ ) for each neuron selected by a previous ANOVA. For the population of neurons that encoded direction and probability, we then computed the modulation, by reward value, of the population tuning curve using a

one-way ANOVA (probability,  $p < 0.05$ ), followed by a Holm–Sidak *post hoc* test.

Our second approach was to assess neural activity in accordance with the probabilistic context. For this purpose, the probability associated with each of the two targets presented to the animal at each trial was determined as described in the following: we computed a three-dimensional analysis of response peaks with respect to target presentation (preparatory activity) and movement onset (movement-related activity), as a function of chosen (PC) and nonchosen (PNC) cue probability. To detect any relationship between the firing rate of neurons and reward probabilities associated with both targets, we used linear regressions with  $R^2 > 0.8$  and  $p < 0.05$ . This allowed the detection of significant encoding of only one target (PC or PNC) or of both targets at the same time.

**Histology.** After recording was completed, the two monkeys were anesthetized with an overdose of pentobarbital sodium and perfused through the left ventricle with 4% paraformaldehyde. Their brains were equilibrated with 20% sucrose. Frozen sections were cut and stained with acetylcholine iodide. The anatomical location of penetrations was reconstructed by creating lesions after completion of the last recording session, in the form of electrophysiological landmarks, as well as dark lines of gliosis, which revealed the recording tracks. The approximate location of each recorded neuron could be estimated by comparing the location of the relevant penetration in the chamber of the lesion to electrophysiologically identified borders, and its position relative to the marking lesions.

## Results

### Behavioral results

During the recording sessions, the two monkeys (T and D) consistently chose the target leading to the highest reward probability (97% success rate for monkey T and 100% for monkey D) (Fig. 2A). Reward probability and movement direction had no effect on RT (the period between presentation of the Go signal and movement initiation) and MD (three-way ANOVA, probability  $\times$  direction  $\times$  monkey,  $p > 0.05$ ). The complexity of choice was measured as the ratio of the probability of being rewarded for the chosen target to the total probability of being rewarded and was found to have no influence on movement execution (one-way ANOVA,  $p > 0.05$ ) (Fig. 2B). Both monkeys initiated their arm movements with the same RT ( $412 \pm 74$  ms). Although monkey D (MD of  $542 \pm 55$  ms) completed its movements more promptly than monkey T (MD of  $665 \pm 58$  ms), no other difference was demonstrated.

### Database

We recorded the responses of 268 projection neurons in the left frontolateral striatum and 238 neurons in the GPi, during 44 recording sessions. The mean number of successful trials per session was  $494 \pm 55$ . The present study is based on the 148 PUT cells (88 in monkey T, 60 in monkey D) and the 111 GPi neurons (52 in monkey T, 59 in monkey D), which displayed a stable discharge rate and an acceptable signal-to-noise ratio. Here, we

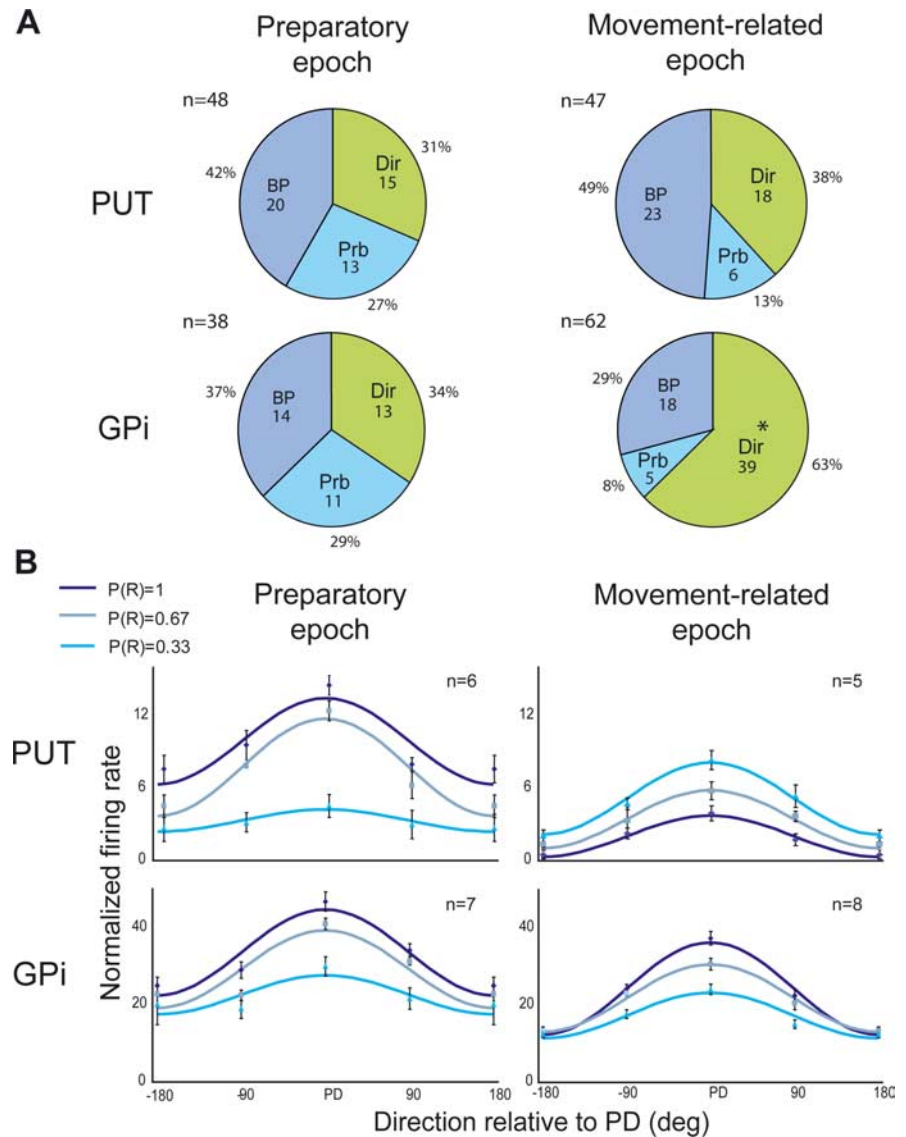
report on the responses obtained from 500 ms before to 1500 ms after two task events; (1) target onset and (2) movement onset. There was no difference in the proportion of neurons responding to the various task parameters between the two animals ( $\chi^2$  test,  $p < 0.05$ ), such that data from both could be pooled for the purposes of analysis. For each neuron, the changes in firing rate associated with movement direction and/or reward probability were assessed with two-way ANOVA ( $p < 0.05$ ).

### Putamen and pallidal neurons encode multiple parameters

We defined two epochs to analyze the response of the neurons: (1) the preparatory activity, which is recorded after the appearance of the target, and (2) the movement-related activity, which is recorded during the onset of movement. The number of encoding neurons in each phase is summarized in Table 1. The preparatory activity of 24% of PUT neurons (15 of 148 were sensitive to direction only and 20 of 148 were sensitive to both parameters) and 24% of GPi neurons (13 of 111 and 14 of 111) was modulated by the location of the chosen cue (i.e., the direction of movement after the Go signal). The movement-related activity of 28% of the PUT neurons (18 of 148 and 23 of 148) and 51% of the GPi cells (39 of 111 and 18 of 111) was significantly affected by the direction of movement. Furthermore, the firing rates of 22% of PUT neurons (13 of 148 and 20 of 148) and 23% of GPi neurons (11 of 111 and 14 of 111) were modulated by the reward probability value during the preparatory epoch. During the movement-related phases, 20% of PUT neurons (6 of 148 and 23 of 148) and 21% of GPi neurons (5 of 111 and 18 of 111) fired according to reward probability. In addition, 18% of PUT neurons (26 of 148) and 26% of GPi neurons (29 of 111) responded to either direction or reward probability during preparatory and movement-related epochs.

### Shaping of the motor information in putamen and GPi

Neuronal discharge was modulated according to movement direction in both epochs (Fig. 3A). To establish directional sensitivity, sinusoidal tuning curves (regression,  $R^2 > 0.8$ ;  $p < 0.05$ ) were constructed for each responding neuron (Georgopoulos et al., 1982) selected through a previous ANOVA ( $p < 0.05$ ). We first computed the PD for each neuron, i.e., the direction associated with the largest change in firing rate. Directional tuning was observed in PUT neurons and GPi neurons during the movement-related epoch but also during the preparatory time, when the animal was deciding which movement to make. As reported previously for an ocular saccade task, PUT neurons in-



**Figure 5.** Interaction between reward–value and encoding of direction. **A**, Pie charts of neurons that are direction-sensitive only (Dir), reward-predicting only (Prb), or sensitive to both parameters (BP), during the preparatory and movement-related epochs (two-way ANOVA, probability  $\times$  direction,  $p < 0.05$ ).  $n$  refers to the number of responding neurons, and the percentages represent populations that encode each parameter according to responding neurons. \* indicates that the ratio of GPi neurons that encode direction during movement execution is significantly larger than during preparatory activity or PUT neurons ( $\chi^2$  test,  $p < 0.001$ ). **B**, The PUT and GPi population directional tuning curves were modulated by different reward probabilities [ $P(R) = 0.33, 0.67, \text{ and } 1$ ]. Each neuron selected in these populations encoded the direction with a cosinusoidal function ( $R^2 > 0.8$ ;  $p < 0.05$ ) and the value with a linear function ( $R^2 > 0.8$ ;  $p < 0.05$ ). The tuning curves were obtained with nonlinear regressions, with mean  $\pm$  SEM parameters that generate a sinusoidal model for discharge rate. The curves for single-neuron analysis were aligned with their PD to plot populations.  $n$  refers to the number of neurons in each population that simultaneously encode both parameters.

crease their discharge activity in response to task events (Hikosaka et al., 1989; Kawagoe et al., 1998; Hikosaka et al., 2000; Watanabe et al., 2001; Itoh et al., 2003), whereas GPi neurons encode information about target location with either increases (GPi+) or, less frequently, decreases (GPi-) in activity.

### Linear encoding of the reward value

To analyze the encoding of reward probability by previously selected neurons (two-way ANOVA,  $p < 0.05$ ), linear regressions of the peak responses to task events ( $R^2 > 0.8$ ;  $p < 0.05$ ) were computed (Fiorillo et al., 2003). Figure 3B shows representative responses of neural populations in putamen and pallidum that

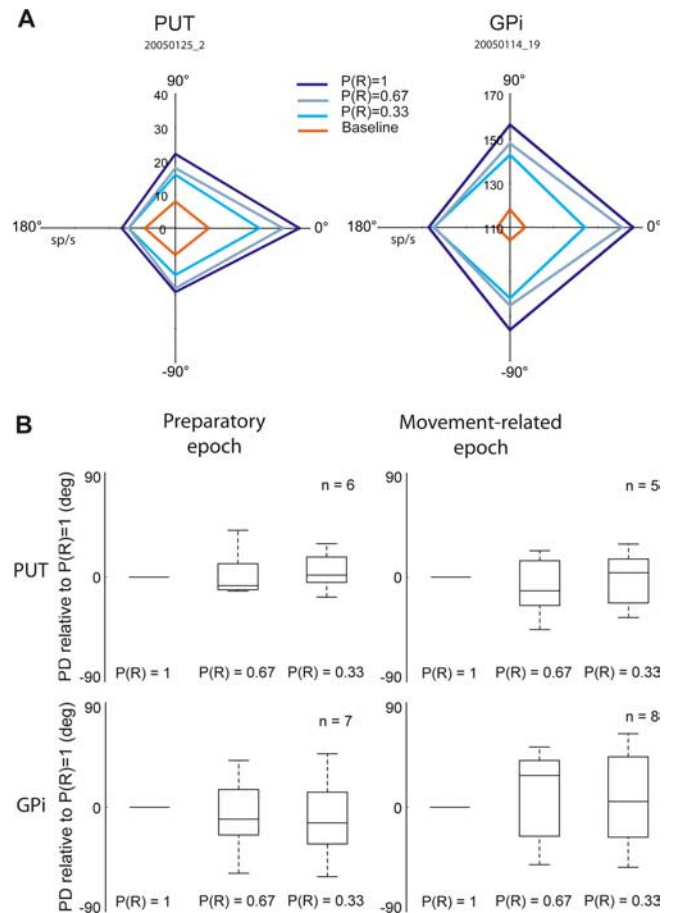
encode the reward value. During the preparatory epoch, 86% of PUT neurons that encode the reward value linearly (19 of 22, positive slope) showed increases in firing rate with increasing reward probability, whereas during the motor act, this encoding was reversed for 78% of neurons (14 of 18, negative slope). In the GPi, 78–80% of the neurons (20 of 25 for the preparatory phase and 18 of 23 for the movement onset) encoded the reward value linearly with an increase (positive slope) and 20–22% with a decrease (negative slope). In the pallidum, although experimental bias might not be excluded because of the small number of GPi– neurons recorded (16 of 111), only GPi+ neurons encoded the reward value. Finally, neurons that processed movement and cognitive parameters were not topologically segregated in the dorsal striatum and in the GPi (Fig. 4).

### Shaping of motor information through reward values

Interactions between reward value of the chosen target and directional tuning were also analyzed (two-way ANOVA,  $p < 0.05$ ) (Fig. 5A). The preparatory activity of 14% of PUT neurons (20 of 148) and 13% of GPi neurons (14 of 111) was simultaneously modulated by the location and the reward probability of the chosen cue (Table 1). During the movement-related phases, 16% of PUT neurons (23 of 148) and 16% of GPi neurons (18 of 111) were also sensitive to both parameters. So, it was found that 25 and 29% of neurons, respectively, in the putamen (37 of 148) and in the GPi (32 of 111) encoded both parameters during at least one investigated epoch. Among these neurons, we extracted single neurons that encoded, respectively, the direction and the reward value with cosinusoidal and linear functions (regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ) to observe interactions between both encodings. The tuning curves for selected units were aligned with their PD, and each point was averaged (Fig. 5B). Importantly, the directional tuning curves for these populations of neurons were significantly modulated by the reward probability. The reward-related modulation is significantly higher for the PD (one-way ANOVA;  $F_{(3,6)} = 23.89$ ,  $p < 0.001$ ;  $F_{(3,5)} = 7.15$ ,  $p = 0.009$ ;  $F_{(3,7)} = 17.17$ ,  $p < 0.001$ ;  $F_{(3,8)} = 36.59$ ,  $p < 0.001$ ). Interestingly enough, these neurons belong to the population of neurons that showed statistically significant interactions (Table 1). It was also observed that the action value did not affect the PD of single neurons (Fig. 6A). At population level, PD of neurons remained constant for three reward values (one-way ANOVA;  $F_{(3,6)} = 0.07$ ,  $p = 0.92$ ;  $F_{(3,5)} = 0.59$ ,  $p = 0.56$ ;  $F_{(3,7)} = 0.06$ ,  $p = 0.94$ ;  $F_{(3,8)} = 0.18$ ,  $p = 0.84$ ) (Fig. 6B). As a whole, these results demonstrate that motor-related neural activity is shaped by reward probability. This, in turn, suggests a role for the BG in a decision-making process, in which future actions are chosen according to cognitive values.

### Encoding of the probabilistic context

To further investigate this role, neuronal responses were correlated with decisional complexity, i.e., the difficulty of choosing between the two presented targets. Activity was analyzed for the entire recorded neural population as a function of all possible combinations of probability pairs and of the choice actually made. This yielded six possible combinations of chosen and ignored targets [ $P(R) = 1$  vs  $P(R) = 0$ ,  $P(R) = 1$  vs  $P(R) = 0.33$ , etc.]. For each task epoch (preparatory time and movement, respectively), 19 and 16% of PUT neurons (29 and 25 of 148) and 35 and 40% of GPi neurons (39 and 45 of 111) were related to just one target (linear regression,  $R^2 > 0.8$ ;  $p < 0.05$ ) (Fig. 7A,B,D). The preparatory activity of 5% of PUT neurons (8 of 148) and 14% of GPi neurons (16 of 111) was modulated with respect to



**Figure 6.** *A*, Polar diagrams illustrating the dual influence of movement direction and reward probability on the firing rate of two cells during target display (PD of 0°). *B*, Box plots showing PD related to three action values for neural populations that encode simultaneously both parameters (regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ). For each reward probability, the PD was computed with nonlinear regression, and it was normalized according to the PD when  $P(R) = 1$ . The central line of the box plots represents the median, the edges of the box show the interquartile range, and the edges of the whiskers show the full extent of the overall distributions.  $n$  refers to the number of neurons in each population.

both cues (Fig. 7C,D). In response to the onset of movement, when a decision had already been made, the ratio dropped significantly ( $\chi^2$  test,  $p < 0.05$ ) to  $< 5\%$  in the GPi (6 of 111) but not in the striatum. The influence of the context on the response of neurons is variable: the modulation by the nonchosen cue depends on the chosen one (Fig. 8).

### Discussion

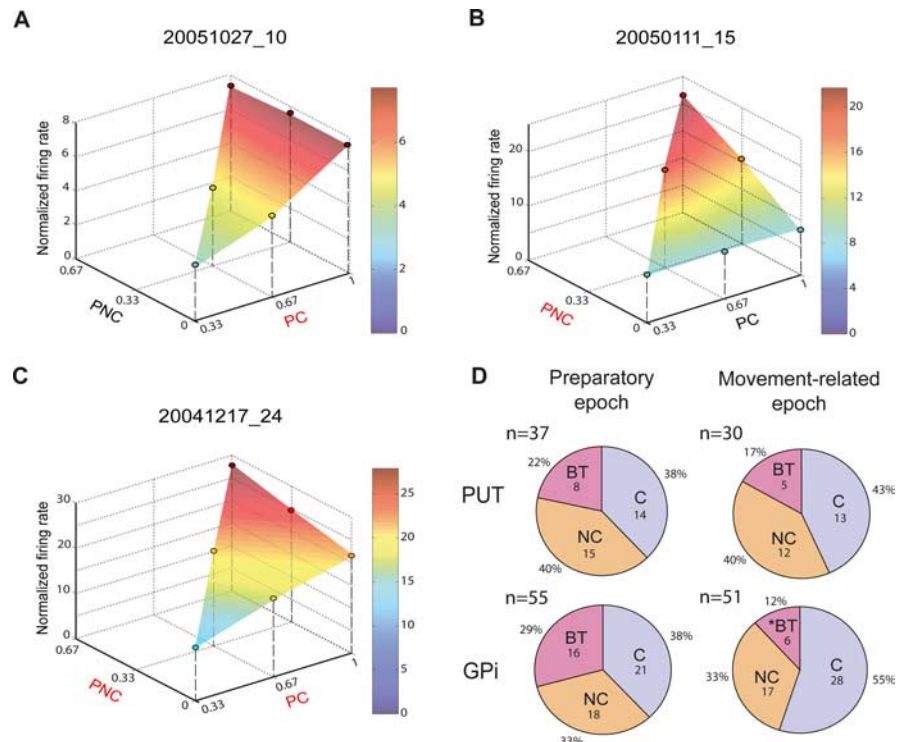
Here we showed that the activity of  $\sim 15\%$  of the input and output neurons of the basal ganglia can be modulated simultaneously by movement parameters (direction) and cognitive parameters (reward probability) before and after an instruction to move. We also confirmed that BG neurons display dynamic encoding of the parameter of a motor task (Arkadir et al., 2004). As an example, a subpopulation of PUT and GPi neurons encode the probabilistic context (a target combination). This encoding remained stable in the input nucleus when the targets appeared and when the monkeys moved, but a dramatic change occurred at the output nucleus, in which the chosen action is preferentially encoded at movement onset. This implies that the latter does not just transmit the information of the putamen but that it also computes data required to assist the chosen action.



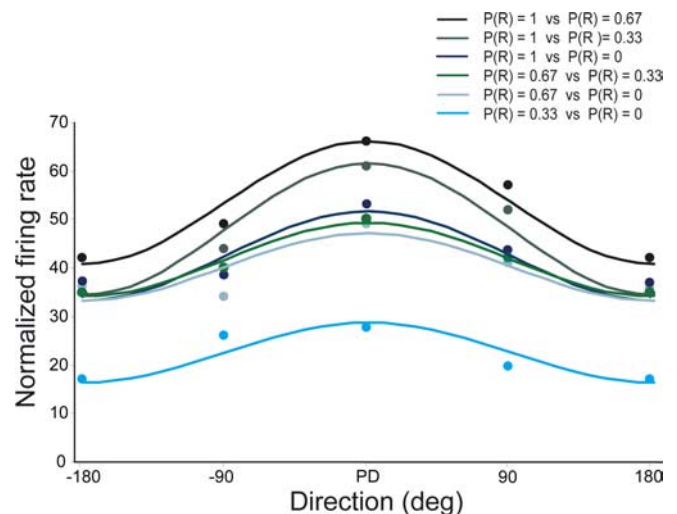
During tasks, the monkeys optimized their reward-obtaining strategy by consistently selecting the target that was associated with the highest reward probability, independently of the difficulty of choice. Note that this behavior contrasts with several studies of striatal or dopaminergic neurons, which describe an approximate probability matching strategy (Morris et al., 2004, 2006; Bayer and Glimcher, 2005). In our case, this difference may be attributable to overtraining or to the fact that five of the six possible probabilistic combinations in the involved task included events with a certainty (Fig. 1*B*). The only uncertain condition was indeed 0.67 versus 0.33.

In accordance with previous reports, GPi neurons encode movement direction according to a cosine tuning function (Georgopoulos et al., 1982; Turner and Anderson, 1997). We show that PUT neurons follow the same rule, which is consistent with previous results from the processing of saccadic eye movements (Hikosaka et al., 1989, 2000). Furthermore, this directional tuning was observed not only in response to movement onset but also just after target appearance, in the preparatory epoch, when the animal was actually deciding which movement to perform. In this study, only two-dimensional movement vectors were investigated, and no other kinematic parameters such as velocity, viscosity, posture adjustments, etc. were analyzed. The latter may account for discrepancies observed in the case of neurons for which activity is significantly related to movement (Table 1) but which do not encode tuning curves (Fig. 3). At variance with other studies performed in oculomotor system (Kawagoe et al., 1998), it was observed that the PD of neurons remained constant. This discrepancy between the two studies can be explained by the differences between the two behavioral paradigms. In the memory-guided saccade task, they used the movement direction as a direct indication of the action value, whereas in our experimental design, the encoding of the reward is only carried by the shape of the visual cue. The direction of the movement is independent of the action value. The PD does not carry any information about reward value and is only related to the executive part of the action to perform.

Action-value encoding in the striatum has been proposed as a core feature of information processing in the BG (Cromwell et al., 2005; Samejima et al., 2005). Specifically, this structure is heavily innervated by dopaminergic axons, which show increasing phasic responses to stimuli that predict reward with increasing probability (Fiorillo et al., 2003; Morris et al., 2004, 2006). Previous studies have already demonstrated reward encoding in the striatum during target presentation (Morris et al., 2004; Cromwell et al., 2005) and in the external part of the globus pallidus during movement-related periods (Arkadir et al., 2004). The data disclosed here confirm and extend these findings, showing that both striatal and pallidal neurons encode reward probability during decision making and during the movement initiation phase. Importantly, it is also shown that, apart from pure co-modulation,



**Figure 7.** *A–C*, Examples of three-dimensional representations of preparatory activity as a function of the PC and PNC. *A*, PUT neurons whose firing rate was modulated only according to PC. *B*, GPi neuron that was modulated only by PNC. *C*, GPi neuron that discharged according to probabilities related to both targets. Cues (PC and/or PNC) related to neural activity are shown in red. *D*, Distribution pie charts of neurons responding to cue values. The firing rate of some neurons was linearly modulated (regression,  $R^2 > 0.8$ ;  $p < 0.05$ ) by the chosen cue probability (C), the nonchosen cue probability (NC), or by both targets at the same time (BT). *n* refers to the number of responding neurons, and the percentages represent populations that encode each parameter according to responding neurons. \* indicates that GPi neurons encode the probabilistic context significantly more frequently in the preparatory epoch than during movement ( $\chi^2$  test,  $p < 0.05$ ).



**Figure 8.** Example of GPi neuron that encodes simultaneously the reward value related to both targets (linear regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ) and the location of the chosen cue (nonlinear regression analysis,  $R^2 > 0.8$ ;  $p < 0.05$ ) during the preparatory epoch. Each sinusoidal curve was built for one target combination. In this example, the reward value of the nonchosen cue modulates the response for a chosen target with  $P(R) = 1$  but not for  $P(R) = 0.67$ .

statistically significant interactions exist between reward value and directional encoding at the single-cell and network levels in both structures (Table 1). The reshaping of the tuning curve by the reward probability of the chosen target is involved in these

interactions. Other interactions may have an influence on other motor parameters such as velocity or viscosity, which have not been taken into account here.

The computation of reward probability is not restricted to the chosen target but also takes into account the nonchosen target, highlighting the crucial role of the probabilistic context, as reported previously for striatal neurons (Samejima et al., 2005). The striatum and GPi can encode reward values as a two-dimensional algebraic process. During the preparatory epoch, numerically equivalent populations of BG neurons encode a chosen target, a nonchosen target, or both. The reward values of each target are estimated and processed according to the whole context of the task for a decision to be made. During the motor phase, GPi neurons are significantly more involved in the encoding of movement direction (Fig. 5A) and chosen target (Fig. 7D). This finding might explain why the primary function of GPi neurons has often been linked to motor activity (Mink, 1996; Turner and Anderson, 1997; Arkadir et al., 2004). This dynamic processing of the probabilistic context does not occur in the putamen, which maintains a constant representation of the task parameters during movement preparation and execution. Thus, the GPi is a more versatile structure and can be involved in the dynamic evaluation of a motivational context. These results not only reinforce the hypothesis that the BG network plays a role in reinforcement-driven decision making (Samejima et al., 2005) but also show that, whereas the putamen maintains constant encoding during a given task, the GPi is more profoundly involved in the executive part of the task.

The data reported on here provide the following view of the computations that take place in the BG network during the decision-making process. The interaction between reward value and direction of movement start in the striatum. Although the mechanisms still remain to be clarified, dopaminergic signals should be a major contributing factor (Schultz et al., 1997; Reynolds et al., 2001; Kawagoe et al., 2004). This shaping of motor information by the reward value is then transmitted to the GPi, the main output structure of the basal ganglia. It allows population encoding for different preferred directions to be differentiated. The population whose PD corresponds to the location of the target with the highest reward probability has the greatest weight. The GPi integrates this information and can choose the optimal response through a competitive mechanism (Mink, 1996) or by symmetry breaking (Leblois et al., 2006) and thus facilitate movements toward the best target but inhibit others. Basal ganglia thus compute the target value and location, then informs the premotor cortical area during the decision phase of the best choice (best target and where it is), and subsequently assist in the execution of a movement toward the selected target.

## References

- Aosaki T, Kimura M, Graybiel AM (1995) Temporal and spatial characteristics of tonically active neurons of the primate's striatum. *J Neurophysiol* 73:1234–1252.
- Arkadir D, Morris G, Vaadia E, Bergman H (2004) Independent coding of movement direction and reward prediction by single pallidal neurons. *J Neurosci* 24:10047–10056.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- Bezard E, Boraud T, Chalon S, Brotchie JM, Guilloteau D, Gross CE (2001) Pallidal border cells: an anatomical and electrophysiological study in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-treated monkey. *Neuroscience* 103:117–123.
- Boraud T, Bezard E, Bioulac B, Gross CE (2001) Dopamine agonist-induced dyskinesias are correlated to both firing pattern and frequency alterations of pallidal neurons in the MPTP-treated monkey. *Brain* 124:546–557.
- Cromwell HC, Hassani OK, Schultz W (2005) Relative reward processing in primate striatum. *Exp Brain Res* 162:520–525.
- DeLong MR (1971) Activity of pallidal neurons during movement. *J Neurophysiol* 34:414–427.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT (1982) On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J Neurosci* 2:1527–1537.
- Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J Neurophysiol* 61:780–798.
- Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev* 80:953–978.
- Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol* 80:947–963.
- Itoh H, Nakahara H, Hikosaka O, Kawagoe R, Takikawa Y, Aihara K (2003) Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior. *J Neurophysiol* 89:1774–1783.
- Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1:411–416.
- Kawagoe R, Takikawa Y, Hikosaka O (2004) Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. *J Neurophysiol* 91:1013–1024.
- Leblois A, Boraud T, Meissner W, Bergman H, Hansel D (2006) Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *J Neurosci* 26:3567–3583.
- Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 50:381–425.
- Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43:133–143.
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H (2006) Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci* 9:1057–1063.
- Opris I, Bruce CJ (2005) Neural circuitry of judgment and decision mechanisms. *Brain Res Brain Res Rev* 48:509–526.
- Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. *Nature* 413:67–70.
- Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310:1337–1340.
- Satoh T, Nakai S, Sato T, Kimura M (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci* 23:9913–9923.
- Schultz W (2005) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645.
- Turner RS, Anderson ME (1997) Pallidal discharge related to the kinematics of reaching movements in two dimensions. *J Neurophysiol* 77:1051–1074.
- Watanabe M, Cromwell HC, Tremblay L, Hollerman JR, Hikosaka K, Schultz W (2001) Behavioral reactions reflecting differential reward expectations in monkeys. *Exp Brain Res* 140:511–518.