

Midbrain Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning Trials

Mati Joshua,^{1,2} Avital Adler,^{1,2} Rea Mitelman,^{1,2} Eilon Vaadia,^{1,2} and Hagai Bergman^{1,2,3}

¹Department of Physiology, The Hebrew University–Hadassah Medical School, Jerusalem 91120, Israel, and ²The Interdisciplinary Center for Neural Computation and ³Eric Roland Center for Neurodegenerative Diseases, The Hebrew University, Jerusalem 91904, Israel

Midbrain dopaminergic neurons (DANs) typically increase their discharge rate in response to appetitive predictive cues and outcomes, whereas striatal cholinergic tonically active interneurons (TANs) decrease their rate. This may indicate that the activity of TANs and DANs is negatively correlated and that TANs can broaden the basal ganglia reinforcement teaching signal, for instance by encoding worse than predicted events. We studied the activity of 106 DANs and 180 TANs of two monkeys recorded during the performance of a classical conditioning task with cues predicting the probability of food, neutral, and air puff outcomes. DANs responded to all cues with elevations of discharge rate, whereas TANs depressed their discharge rate. Nevertheless, although dopaminergic responses to appetitive cues were larger than their responses to neutral or aversive cues, the TAN responses were more similar. Both TANs and DANs responded faster to an air puff than to a food outcome; however, DANs responded with a discharge elevation, whereas the TAN responses included major negative and positive deflections. Finally, food versus air puff omission was better encoded by TANs. In terms of the activity of single neurons with distinct responses to the different behavioral events, both DANs and TANs were more strongly modulated by reward than by aversive related events and better reflected the probability of reward than aversive outcome. Thus, TANs and DANs encode the task episodes differentially. The DANs encode mainly the cue and outcome delivery, whereas the TANs mainly encode outcome delivery and omission at termination of the behavioral trial episode.

Key words: primate; basal ganglia; spike train; reinforcement; substantia nigra; striatum

Introduction

The neural network of the basal ganglia (Bar-Gad and Bergman, 2001; Gurney et al., 2004) is commonly viewed as two functionally related subsystems. The main axis includes fast neurotransmissions (glutamate and GABA) between the cortex, striatum, and the basal ganglia output structures. The second subsystem is composed of neuromodulators that adjust the activity along the main axis by regulation of plasticity at the corticostriatal synapse (Calabresi et al., 2000; Reynolds et al., 2001). The primary basal ganglia neuromodulators are dopamine [from midbrain dopaminergic neurons (DANs) (Arbutnot and Wickens, 2007)] and acetylcholine [from striatal cholinergic tonically active interneurons (TANs) (Calabresi et al., 2000)].

Previous studies have shown that DANs encode the prediction

error in the positive domain; (i.e., they respond when conditions are better than expected) (Schultz et al., 1997). Consistent with the classical concept of dopamine–acetylcholine balance (Barbeau, 1962), the DANs and the TANs have opposite responses. DANs typically increase their discharge rate in response to appetitive predictive cues and outcomes, whereas TANs suppress their tonic discharge (Graybiel et al., 1994). Thus, it has been suggested that some of the dopamine influence on striatal projections neurons is mediated through inhibition of the TANs (Wang et al., 2006).

In contrast to the extensive research on reward-related activity, only a few studies have explored whether basal ganglia neurons encode the negative domain (e.g., aversive outcome or omission of rewards, which might not be identically encoded by the nervous system). Dopamine neurons decrease their firing rate in response to reward omission (Schultz et al., 1997). However, this suppression is limited because firing rate is truncated at zero. Other groups (Morris et al., 2004; Bayer and Glimcher, 2005) have reported that the discharge rate of dopaminergic neurons does not demonstrate instantaneous incremental encoding of reward omission, and an alternative encoding scheme, based on response duration, has been proposed (Bayer et al., 2007). There are even fewer studies and less agreement on basal ganglia re-

Received Aug. 13, 2008; accepted Sept. 16, 2008.

This work was partly supported by the “Fighting against Parkinson” grant from the Hebrew University Netherlands Association. We thank Dr. Bryon Gombert for MRI; Michael Levi and Michal Rivlin for help in preparing the experimental setup; Yael Renert and Inna Finkes for monkey training and general assistance; and Geoffrey Schoenbaum, Yavin Shaham, and Genela Morris for critical reading of early versions of this manuscript.

Correspondence should be addressed to Mati Joshua, Department of Physiology, The Hebrew University–Hadassah Medical School, P.O. Box 12272, Jerusalem 91120, Israel. E-mail: mati@alice.nc.huji.ac.il.

DOI:10.1523/JNEUROSCI.3839-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/2811673-12\$15.00/0

sponses to aversive events. Classical and instrumental conditioning studies suggest that some of the dopamine neurons increase their firing rate after a cue that predicts aversive outcomes (Mirenowicz and Schultz, 1996; Guarraci and Kapp, 1999). However, that increase in firing rate may be a result of reward generalization. Studies on anesthetized rats have shown that DANs mainly decrease their discharge rate after aversive stimulus (Ungless et al., 2004; Coizet et al., 2006). There are reports that TAN activity differentiates appetitive and aversive stimuli (Ravel et al., 2003; Yamada et al., 2004), but it remains unclear whether and how TANs respond to expectation of aversion.

Here, we designed a classical conditioning paradigm with aversive and rewarding probabilistic outcomes. Symmetric manipulations of expectation of food (rewarding event) or an air puff (aversive event) enable the comparison of neural responses to expectation of positive and negative outcomes. To provide additional controls for sensory, arousal-related, and generalization responses, our behavioral task included neutral trials, which had the same structure as the rewarding and the aversive trials but never yielded positive or negative outcomes.

Materials and Methods

All experimental protocols were performed in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and with Hebrew University guidelines for the use and care of laboratory animals in research, supervised by the institutional animal care and use committee.

Behavioral task. Two monkeys (L and S; *macaque fascicularis*; female, 4 kg; male, 5 kg) were engaged in a probabilistic delay classical-conditioning task (see Fig. 1*b*). The monkeys were seated in a primate chair facing a 17 inch computer screen placed at a distance of 50 cm. Seven different fractal cues (Chaos Pro 3.2 program; www.chaospro.de), stretched on the entire screen, were introduced to the monkey, each predicting the outcome in a probabilistic manner. Three cues (reward cues) predicted a liquid food outcome (L, 0.4 ml, 100 ms duration; S, 0.6 ml, 150 ms) with a delivery probability of 1/3, 2/3, and 1. Three other cues (aversive cues) predicted an air puff outcome (L, 100 ms duration; S, 150 ms; 50–70 psi; split and directed 2 cm from each eye; Airstim; San Diego Instruments) with a delivery probability of 1/3, 2/3, and 1. The seventh cue (the neutral cue) was never followed by a food or air puff outcome. Cues were presented for 2 s and were immediately followed by a result epoch, which could include an outcome (food, air puff) or no outcome according to the probabilities associated with the cue. The beginning of the result epoch was signaled by one of three sounds that discriminated the three possible events: a drop of food, an air puff, or no outcome (see Fig. 1*b*). Sounds were normalized to the same intensity and duration. These sounds were additional to the background device sounds (air puff solenoid and food pump). All trials were followed by a variable intertrial interval (ITI) (monkey S, 3–7; monkey L, 4–8 s). Because of the probabilistic structure of the behavioral task and to equalize the average occurrence of each outcome the nondeterministic cues ($p \neq 1$ for reward or aversive outcome) were introduced three time more than the deterministic ones. With this occurrence ratio, all trials were randomly interleaved.

During a behavioral session (usually five sessions per week), the monkeys performed 900–1300 trials/d before losing their motivation for food. Over the weekend, the monkeys were given *ad libitum* access to food. Water was available *ad libitum* during all training and recording periods. After the training period (L, 6; S, 2 months), we recorded the behavior and the basal ganglia neural activity while the monkeys were engaged in the behavioral task. The same images and sounds were used both for training and for the recording periods (L, 6; S, 5 months); however, the visual and the auditory stimuli were shuffled between monkeys.

Surgery, magnetic resonance imaging, and rehabilitation. After the training period, a magnetic resonance imaging (MRI)-compatible Cilux head holder and a square Cilux recording chamber with a 27 mm (inner

side) were attached to the monkey's head. The head holder enabled the immobilization of the head during recording. The recording chamber was attached to the skull tilted 40° laterally in the coronal plane with its center targeted at the stereotaxic coordinates of the GPe (A15, L7, H1) (Szabo and Cowan, 1984; Martin and Bowden, 2000). Analgesia and antibiotics were administered during surgery and continued for 2 d postoperatively. Recording began after a postoperative recovery period of 5 d.

We estimated the stereotaxic coordinates of the physiological recordings within the basal ganglia nuclei with MRI scans (see Fig. 1*a*). The MRI scan (General Electric 1.5 Tesla system; fast spin echo inversion recovery sequence; dual surface coil; repetition time, 3 s; echo time, 0.044 s; inversion time, 0.3 s; echo train length, 8; coronal slices, 2 mm wide) (Matsui et al., 2007) was performed with five tungsten electrodes at accurate coordinates of the chamber [$Y, X = (6, 0), (0, -6), (0, 0), (0, 6),$ and $(-6, 0)$ in mm from the chamber center]. We then aligned the two-dimensional MRI images with the sections of the atlas of *Macaca fascicularis* (Martin and Bowden, 2000). We performed an additional MRI scan at the final stage of the recording period of monkey L to verify our coordinate system. At the end of the experiment, the chamber and head holder of both monkeys were removed, the skin was sutured, and after a recovery period the monkeys were sent to a primate sanctuary (<http://monkeypark.co.il>).

All surgical procedures were performed under aseptic conditions and general isoflurane and N₂O deep anesthesia. MRI procedure was performed under Dormitor and ketamine light anesthesia.

Recording and data acquisition. During recording sessions, the monkey's head was immobilized and eight glass-coated tungsten microelectrodes (impedance, 0.2–0.8 MΩ at 1000 Hz), confined within a cylindrical guide (1.65 mm inner diameter), were advanced separately (EPS; Alpha Omega Engineering) into the targets in the basal ganglia. The electrical activity was amplified with a gain of 5K and bandpass filtered with a 1–6000 Hz four-pole Butterworth filter and continuously sampled at 25 kHz by 12 bits \pm 5 V analog-to-digital (A/D). Spike activity was sorted and classified on-line using a template-matching algorithm (ASD; Alpha Omega Engineering). Spike detection pulses and behavioral events were sampled at 25 kHz (AlphaLab; Alpha Omega Engineering).

Mouth movements were monitored by an infrared reflection detector (see Fig. 2*a*) (Dr. Bouis Devices). The infrared signal was filtered between 1 and 100 Hz by a bandpass four-pole Butterworth filter, and sampled at 1.56 kHz. In addition, three computerized digital video cameras recorded the monkey's face and upper limbs at 50 Hz. Video analysis was performed on home-made custom software to identify periods when the monkeys closed their eyes (see Fig. 2*b*). Briefly, monkey eye location was identified by a human observer (once for a daily recording session in which the monkey's head was immobilized by connecting the head holder to an external metal frame), and a classification of eye states (open or closed) was made based on the number of dark pixels in the eye area. The algorithm was tested by random samples from several recording days and found to be consistent with the judgments of a human observer for >99% of the images. In representing recording sessions, we recorded the monkeys' spontaneous vocalizations, their arm movements with an accelerometer, eye position using infrared reflection, and heartbeat by electrocardiogram (ECG) (veterinary ECG 5 leads system; Palco Laboratories).

During the acquisition of the neuronal data, two experimenters (M.J., A.A.) controlled the position (2–50 μ m steps) of the eight electrodes and the on-line spike sorting (ASD; Alpha Omega). Quality of detection and spike sorting was estimated and graded on-line every 3 min. The on-line quality estimation was based on the superimposed analog traces of the recently (20–100) sorted spikes, the waveforms of events that crossed an amplitude threshold set by the experimenter above the noise level of each electrode, the cumulative distribution of the distances from the detected events to the detection template, and the stability of the discharge rate.

The first step in the neuronal data analysis targeted verification of the real-time isolation quality (Joshua et al., 2007) and stability of the discharge rate (Gourévitch and Eggermont, 2007). Recorded units were subjected to off-line quality analysis that included tests for rate stability, refractory period, waveform isolation, and recording time. First, firing rate as a function of time during the recording session was graphically displayed, and the largest continuous segment of stable data were selected

Table 1. The neural database

Population	No. of cells	Isolation score	Fraction ISI <2 ms	Discharge rate (spike/s)	Recorded time (s)	No. of recorded trials	No. of spikes/recorded cell
TAN	L: 71	0.92 ± 0.06 [0.8–0.99]	0.0002 ± 0.0003 [0–0.0016]	6.8 ± 1.4 [3.1–9.4]	3647 ± 1891 [1210–9901]	309 ± 168 [111–828]	25,112 ± 14,270 [5026–66,410]
	S: 109	0.91 ± 0.06 [0.8–0.99]	0.0002 ± 0.0003 [0–0.0016]	5.0 ± 1.4 [1.93–9.2]	3414 ± 1920 [1260–10,801]	378 ± 204 [123–1221]	17,147 ± 10,948 [2442–63,364]
DAN	L: 41	0.76 ± 0.16 [0.5–0.99]	0.0008 ± 0.0009 [0–0.0039]	4.5 ± 1.9 [0.69–9.4]	2992 ± 1372 [1260–7195]	254 ± 118 [110–621]	13,211 ± 10,817 [3952–67,895]
	S: 65	0.8 ± 0.13 [0.5–0.99]	0.0007 ± 0.001 [0–0.005]	3.7 ± 1.6 [0.77–9.4]	3984 ± 2037 [1260–11,161]	393 ± 200 [126–1098]	15,070 ± 10,327 [2324–46,617]

The recording statistics were calculated separately for each neural population. Each cell in the table contains the mean and SD and in brackets the range of the scores. The range of the isolation score is 0 to 1. "Fraction ISI <2 ms" is the fraction of interspike intervals (ISIs) <2 ms of all ISIs of a cell. Recording time and number of recorded trials represent only the part of the recording satisfying the inclusion criteria and included in the analysis database.

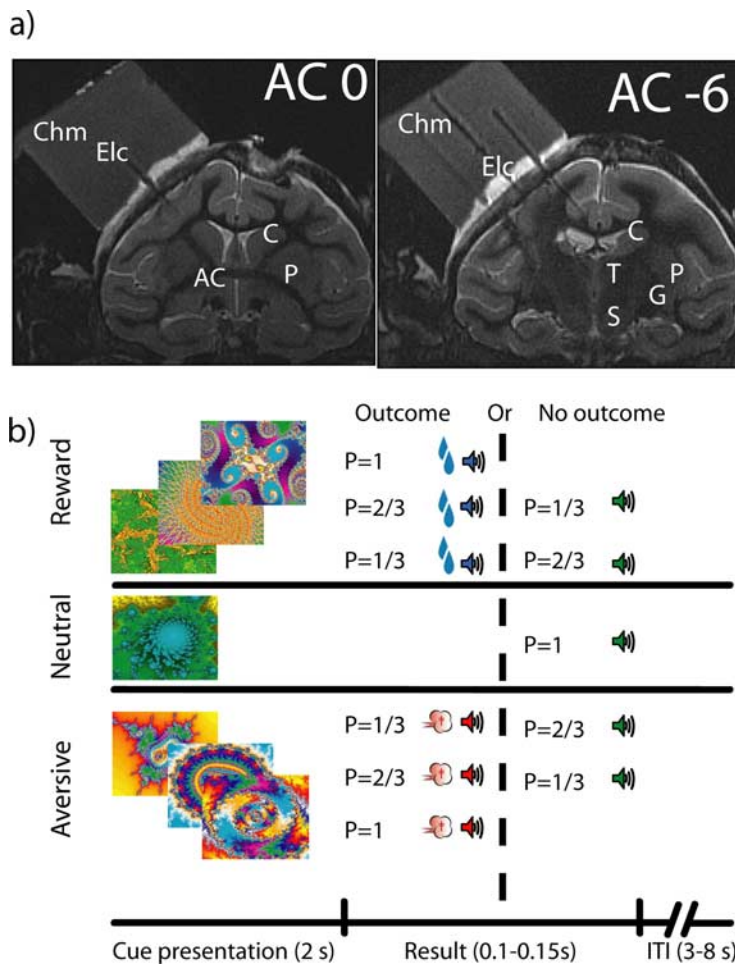


Figure 1. MRI and task. **a**, MRI identification of recording coordinates. Coronal MRI images numbered with respect to distance (in millimeters) from anterior commissure. Tungsten microelectrodes are inserted at known chamber coordinates. Identification of the brain structures is based on alignment of the MRI images with the monkey atlas. Abbreviations: AC, Anterior commissure; C, caudate; Chm, recording chamber (filled with 3% agar); Elc, electrode; G, globus pallidus; P, putamen; S, substantia nigra; T, thalamus. **b**, Behavioral task. Top, Reward trials; middle, neutral trials; bottom, aversive trials. Cues are shown for monkey L. Different speaker colors represent different sounds.

for additional analysis. Second, cells in which >0.02 of the total interspike intervals were <2 ms were excluded from the database. Third, only TANs with an isolation score (Joshua et al., 2007) >0.8 and DANs with an isolation score >0.5 were included in the database. The lower threshold used for the DANs is attributable to the highly dense cellular structure of the substantia nigra pars compacta (SNc) which makes single-cell isolation difficult. We tested the subgroup of DANs with an isolation score >0.8 ($N = 52$) and found the same qualitative population results as reported for the larger DAN population. Finally, only cells that met the

above inclusion criteria for >20 min during performance of the behavioral task were included in the neural database (average, 59 min and 346 trials). Table 1 provides the statistical details of the cells that were included in the analysis database.

During recording, units were classified according to anatomical location, extracellular waveform, firing rate and pattern, background activity, and in some cases response to free reward and to injection of dopamine agonists. To validate classification, we performed off-line analyses of the extracellular waveform shape, firing rate, and firing pattern of the neurons (see Figs. 3*b*, 4*b*). Waveform shape was quantified as the duration from the first negative peak to the next positive peak; rate was defined as the average of the overall firing rate; firing pattern was quantified by the coefficient of variation (SD/mean) of the interspike intervals. To further validate the DAN population response, we repeated the population analysis on a subset of DANs with a firing rate <8 Hz and peak-to-peak duration of >0.5 ms. The results of this analysis were similar to those of the whole recorded population (data not shown). Finally, apomorphine (0.1 mg/kg) was injected in a few cases (see Fig. 4*c*) to test for suppression of DAN activity (Aebischer and Schultz, 1984). We quantified this suppression as the root mean square (RMS) of the high-pass-filtered signal (300–6000 Hz). We used the RMS and not the spike rate to avoid possible errors and biases induced by spike detection and sorting (Moran et al., 2006), which are enhanced after apomorphine intramuscular injection because of monkey movements.

Statistical analysis. Neuronal responses to behavioral events were first characterized by their poststimulus time histogram (PSTH). The histograms were calculated in 1 ms bins and smoothed with a Gaussian window with a SD of 20 ms. The baseline firing rate was calculated by averaging the firing rate in the last 3 s of the variable (4–8, 3–7 s; monkey L and S, respectively) ITI and was denoted as $baseline_{FR}$ as follows:

$$baseline_{FR} \equiv \text{mean}_{0 \leq t \leq 3}(\text{psth}_{ITI_END}(t)).$$

To determine significant responses in the single PSTH analysis, we calculated the SD of the PSTH of the last 3 s of the ITI using the same number of trials as in the studied PSTH and identified time segments in which the deviation from $baseline_{FR}$ exceeded three times the ITI-SD (3 σ rule). A response was considered significant only if the

duration of the deviant segment was >60 ms (three times the SD of the smoothing filter). A cell was considered to have a significant response on a trial epoch if at least one of its PSTHs (e.g., one of the three PSTHs after reward aversive or neutral cue) had a significant response. To check that this analysis was not biased by multiple comparison confounds, we performed the same analysis on ITI epoch and found that none of the cells was significantly modulated at this epoch.

We defined the difference index between two events as the mean absolute difference between their PSTHs, i.e., the following:

$$\text{difference index}(\text{event}_1, \text{event}_2) \equiv \text{mean}_t(\text{abs}(\text{psth}_1(t) - \text{psth}_2(t))).$$

This index is a mean difference between rate functions and hence has units of spike per second. To test the significance of this index, we used resampling (bootstrap) methods. Single-trial responses were shuffled and resampled repeatedly into two groups, and the difference index was then calculated between them. This process was repeated 500 times. A difference index was considered significant if it was larger than the difference indices of a given fraction of these surrogates ($1 - p$ where p is the test confidence level). To cross-check the difference index results, we performed MANOVAs ($p < 0.05$) using 50 and 100 ms time bins. We also bootstrapped the MANOVA statistics and found that all these analyses yielded similar results. In this manuscript, we elected to show the difference index because it gives an intuitive range of difference [i.e., the average difference (in spike rate) between the responses to two events].

We derived two indices from the difference index; the first was the response index that was defined as the difference index when one of the events was the neutral event, i.e. the following:

$$\text{response index}(\text{event}) \equiv \text{difference index}(\text{event}, \text{neutral event}).$$

The second was the probability coding index. This index was defined as the difference index between the events with a high probability ($p = 2/3$ and 1) of receiving an outcome and the event with a low probability ($p = 1/3$) of receiving the same outcome. The clustering of the events into high and low probability followed the behavioral responses of the monkey (see Results) and allowed us to generate a simple graphic representation of our results. A MANOVA of the responses to all the three different probabilities yielded similar results.

In addition to the single-cell analysis, we performed population analyses. The responses of striatal TANs and DANs are very stereotypic (Graybiel et al., 1994; Schultz, 1998). Hence, the average population response was estimated by averaging the PSTH deviation from baseline_{FR} across the whole population. To determine whether the population response was significant, we first constructed the single-cell PSTH at bins of 20 ms, and then averaged across the population to obtain the population PSTH. Finally, we performed a t test to check bin by bin whether the population response was significantly different from zero ($p < 0.01$). If the population PSTH was significant for more than three consecutive bins, it was considered a significant population response.

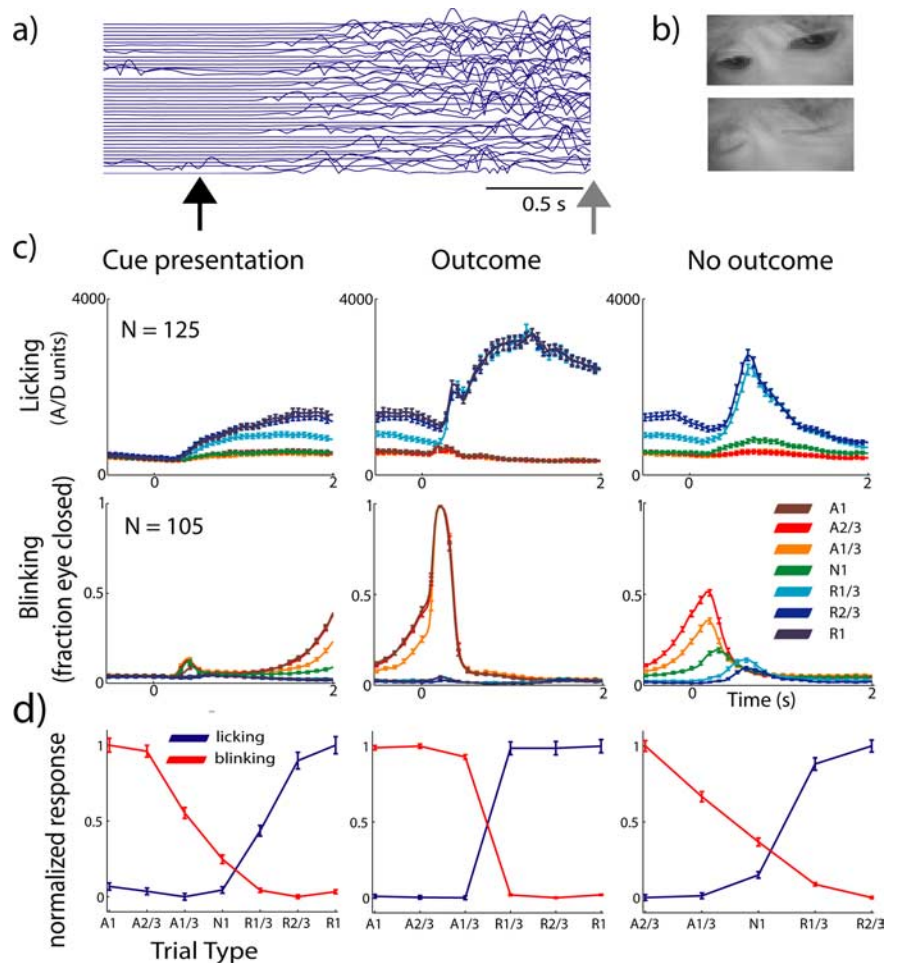


Figure 2. Behavioral monitoring and results. *a*, Mouth signal: example from the reward cue epoch of the licking signal, monitored by an infrared reflection detector. The black arrow indicates time of cue presentation, and the gray arrow indicates cue offset and reward tone onset. *b*, Image of monkey's eyes. Video signal was processed and each frame was classified according to the state of the eyes [i.e., open (top) or closed (bottom)]. *c*, Behavioral results. Top, Licking (average \pm SEM) as recorded by an infrared reflection detector directed at the monkey's mouth. The voltage output of the detector was sampled by A/D converter and the y-scale is given in arbitrary A/D units. Bottom, Fraction of trials with eyes closed (average \pm SEM) as recorded by computerized video processing. Columns correspond to trial epoch (cue; outcome, food or air puff; no outcome, sound only) aligned to event onset (time = 0). Note the overlap of 0.5 s between the start of the outcome and the no-outcome epochs and the last 0.5 s of the cue epoch. Data were averaged for each session and then across sessions (N , number of recording sessions). Color coding of trial types is given at bottom right (A, aversive; N, neutral; R, reward; the number is the outcome probability). *d*, Normalized behavioral response. Licking (blue) and blinking (red) response (average \pm SEM, number of sessions as in *c*) in a time window around the behavioral event (cue, 500–0 ms before cue end; outcome and no outcome, 0–500 ms after cue end for blinking response and 500–1000 ms for licking response). The responses are normalized between 0 and 1 [i.e., in each epoch a response (X) is transformed by $(X - \min)/(\max - \min)$, where \min and \max are the minimal and maximal values of the response in this epoch]. Abscissa, Different behavioral conditions (A, aversive; N, neutral; R, reward; the number is the outcome probability).

The data of the two monkeys were grouped unless a significant difference between the individual monkeys was detected. Data analysis was performed on custom software using MATLAB V7 (Mathworks).

Results

We recorded the neuronal activity of TANs and DANs (Fig. 1*a*, Table 1) in parallel with the monitoring of the monkeys' behavior (Fig. 2). During recordings, the monkeys performed a probabilistic classical conditioning task (Fig. 1*b*) with food or air puff as the rewarding and aversive outcomes, respectively. This task design provides a symmetric expectation of a rewarding or aversive event after cue presentation and therefore served to test the following three hypotheses.

First, DAN and TAN activity reflects expectation, delivery, and omission of reward and of aversive events. The alternative is

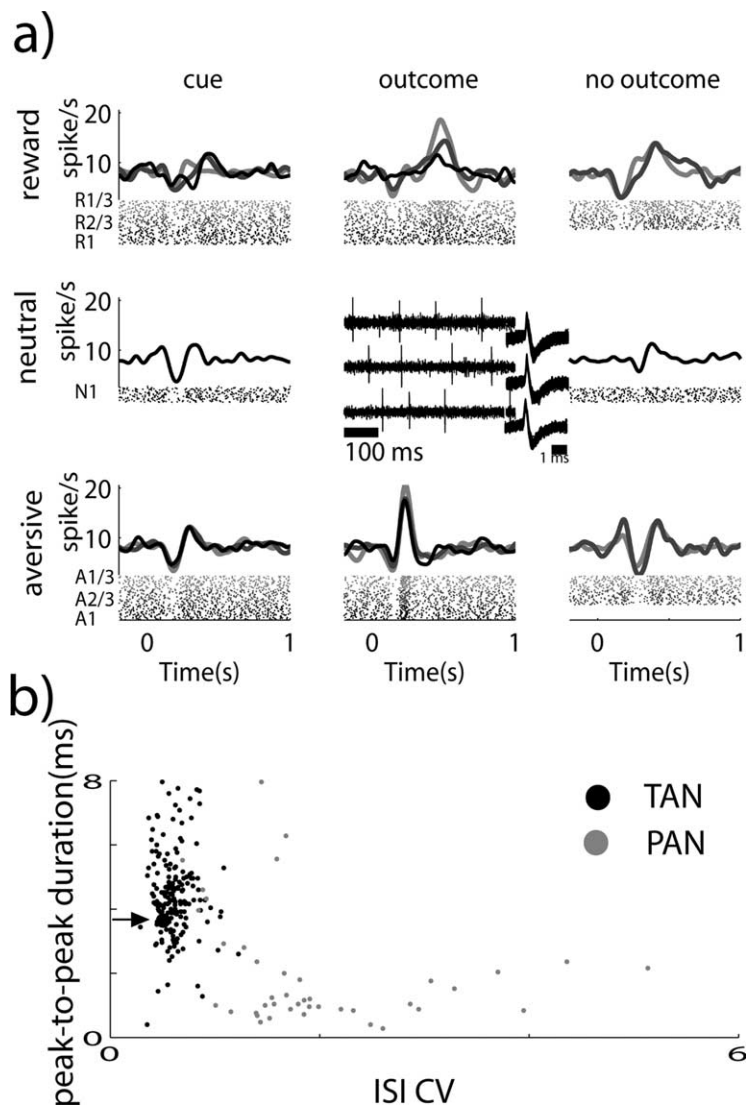


Figure 3. An example of neural activity of a single striatal TAN and identification of striatal cell types. *a*, Rasters and PSTHs of a single TAN of monkey L aligned to the trial behavioral events. The rows are separated according to the expected outcome. First row, Trials with cues that predict the delivery of food. Second row, Trials with the neutral cue (a cue always followed by no outcome). Third row, Trials with cues that predict an air puff. Columns are aligned according to the trial epoch. First column, Cue presentation epoch (−0.2 to 1 s after cue onset). Second column, Outcome epoch (−0.2 to 1 s after delivery of food or air puff). Third column, Trials in which no outcome was delivered; outcome omission was signaled to the monkey by the no-outcome sound (−0.2 to 1 s after sound onset). Color codes are marked at the left side of the cue rasters (A, aversive; N, neutral; R, reward; the number is the outcome probability). For the graphic presentation, rasters were randomly pruned and adjusted to contain the same number of trials. The total number of trials (before pruning) was 708. PSTHs were constructed by summing activity across trials in 1 ms resolution and then smoothing with a Gaussian window (SD, 20 ms). Examples from three 500 ms segments of the analog signal (from first, second, and last third of the recording session) are shown in the middle plot. Examples of spike waveforms are shown next to the 500 ms analog segment. The spike waveform plot includes 100 superimposed waveforms selected randomly around the time of the corresponding analog trace. Isolation score was 0.98; the fraction of spikes in first 2 ms of the interspike interval (ISI) histogram was 0.00002. *b*, Off-line analysis of striatal cell identification based on firing pattern (abscissa) and spike peak-to-peak duration (ordinate). Color code: Black, TANs; gray, phasic active neurons (PANs). Off-line analysis of neuron shape and coefficient of variance (CV) of the time of interspike interval shows that striatal neurons are separated into two clusters in which the PANs have a large CV with comparably narrow waveforms and TANs have a small CV with very wide waveforms. The cell in *a* is plotted in a large black circle and marked with an arrow.

that only reward-related events are represented by the activity of one or both basal ganglia neuromodulators.

Second, DAN and TAN activity encode an error in the temporal prediction (TD) of reward and aversive events (Sutton and Barto, 1998). The TD hypothesis suggests opposite modulations for positive (i.e., delivery and expectation of reward and omission of aversive events) and negative errors (i.e., expectation and deliv-

ery of aversive events and omission of predicted rewarding event).

Third, TAN activity mirrors DAN activity. In a previous study, it was shown that the pause response of TANs was coincident with the increase in DAN activity; however, DANs but not TANs incrementally encoded reward probability (Morris et al., 2004). Here, we test whether this simultaneous opposite response also appears in a task that includes expectation and delivery of aversive events. We also examine whether TANs and DANs discriminate between reward and aversive events during the same parts of the task episode.

Monkey behavior reflects expectation of rewarding and aversive events

We recorded the monkeys’ behavior during performance of a probabilistic classical conditioning task (Fig. 1*b*) with food or air puff as the rewarding and aversive outcomes, respectively. We tested how extensive (several months; 5 d/week; ~1000 trials/d) conditioning affected the monkeys’ behavior by monitoring licking and blinking responses during neural recordings (Fig. 2*a,b*).

The monkeys increased their licking in response to cues predicting food but only slightly to the aversive and neutral cues (Fig. 2*c*, top row). Similarly, the monkeys’ frequency of blinking increased to cues predicting air puff but only slightly to reward and neutral cues (Fig. 2*c*, bottom row). The increase in blinking and licking during the cue epoch was maximal in trials in which the probability of outcome was 2/3 or 1 and smaller in trials in which the probability was 1/3 (0–500 ms before cue ending; $p < 0.01$, Tukey’s HSD *post hoc*).

The behavioral responses to food or air puff delivery (and their corresponding sounds) were not dependent on their previous predictions (Fig. 2*c*, outcome column). Food and air puff omission, as well as the final (no outcome) event of the neutral trials were indicated to the monkeys by an additional “no outcome” sound. When expected food or air puff were not delivered (no outcome on the $p = 1/3$ or $p = 2/3$ trials), licking and blinking increased, respectively; this increase was in accordance with the previously instructed probability. The increase in the licking and blinking behavior was smaller and shorter than the increase after food or air puff outcomes (Fig. 2*c*, no outcome). Licking and blinking increased slightly to the neutral trials (Fig. 2*c*, no outcome, green line).

Normalization of the behavioral responses (Fig. 2*d*) reflects the opposite trends of the response to aversive versus rewarding events. It suggests that the monkeys mainly categorized the

high-probability ($p = 2/3$ and 1) versus the low-probability ($p = 1/3$) cues. Heart rate analysis can discriminate between high- and low-arousal states (Berntson et al., 1997). However, analysis of the heart rate and its variability did not reveal significant differences between the epochs after aversive versus reward predicting cues, suggesting a symmetric effect on monkey arousal.

In sum, the analysis of the behavioral responses indicates the monkeys could distinguish between aversive, reward, and neutral cues and between the cues with high- and low-outcome probabilities. According to these behavioral findings, we grouped the events with high probability ($p = 2/3$ and $p = 1$) for the neural activity analysis.

The neuronal database

We recorded 191 DANs from the SNc and 313 TANs from the putamen; of these, 106 DANs and 180 TANs passed the quality criteria (see Materials and Methods) and their response was further analyzed (Table 1). Figures 3*a* and 4*a* show examples of the activity of a TAN and a DAN, respectively, recorded during the performance of the behavioral task. The TANs and DANs were identified on-line (see Materials and Methods) and identity was verified by off-line clustering of the spike waveforms, spike train pattern (Figs. 3*b*, 4*b*), and occasionally by analysis of the responses to apomorphine injection (Fig. 4*c*).

We found that, in each trial epoch (cue, outcome, and no outcome), most of the cells had a significant response to at least one event (Fig. 5). Below, we provide additional analysis both of the population and the single-cell responses at each epoch and compare the responses to the aversive, neutral, and reward-related events and between the DANs and TANs.

TAN and DAN activity is asymmetrically modulated by expectation of aversive events and reward in the cue epoch

Population analysis of the neuronal activity in the cue epoch shows that, whereas TAN average responses to the aversive, neutral, and reward predicting cues tended to overlap (Fig. 6*a*, top), the population average activity of the DANs was highly discriminative both between reward and aversive events and between cues with high ($p = 2/3$ and $p = 1$) and low ($p = 1/3$) prediction probability of reward delivery. The DANs positive response to aversive cues was smaller than the response to the reward cues (Fig. 6*a*, bottom). The suppression of TAN activity after high-probability reward cues tended to be longer than after low-probability cues (Fig. 6*a*, compare blue with light blue lines) (for similar trends, see Shimo and Hikosaka, 2001; Ravel et al., 2003). As previously reported (Morris et al., 2004), comparison of the time of reward cue modulation showed that the DAN increase and TAN decrease of discharge rate were coincident (Fig. 6*b*, top)

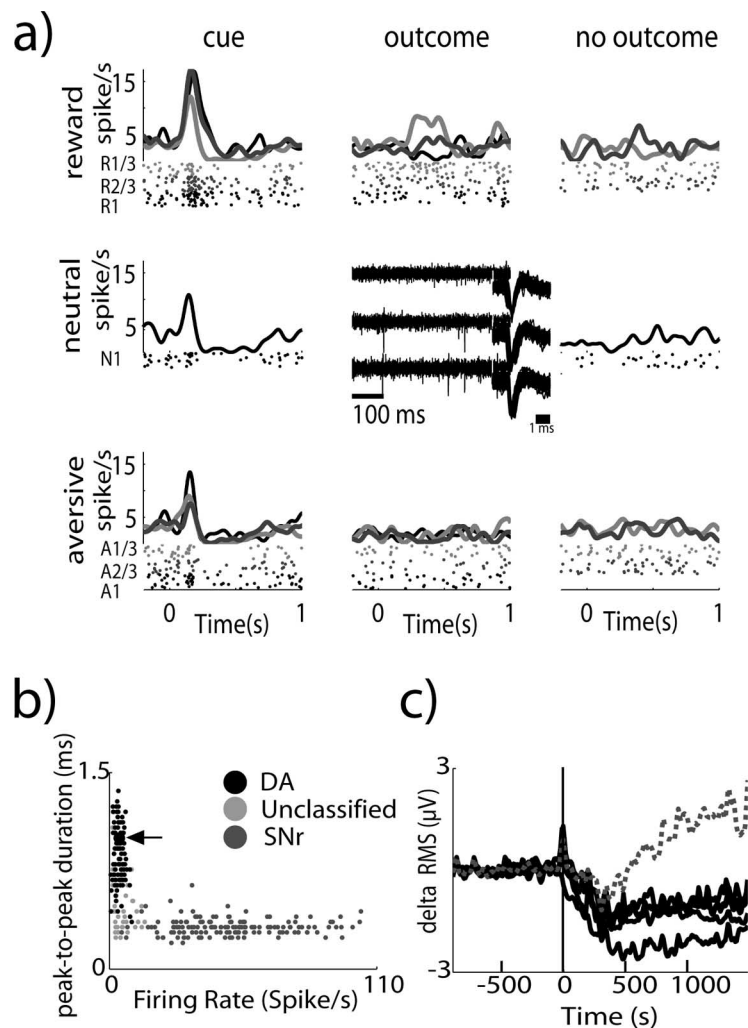


Figure 4. An example of neural activity from a single DAN and identification of substantia nigra cell types. *a*, Same conventions as in Figure 3*a*. Total number of trials was 271; isolation score was 0.67; fraction of spike in first 2 ms of the ISI histogram was 0.0001. *b*, Off-line analysis of substantia nigra (SN) cell identification based on firing rate (abscissa) and spike peak-to-peak duration (ordinate). Color code: Black, DANs; gray, substantia nigra pars reticulata (SNr) neurons; light gray, unclassified SN neurons. Off-line analysis of the spike shape and firing rate shows that nigral neurons are separated into two clusters in which the SNr cells have a high firing rate with narrow waveforms and DANs have a low firing rate with wide waveforms. Cells that were not classified as DAN or SNr tended to be between clusters. The cell in *a* is plotted in a large black circle and marked with an arrow. *c*, Example of neuronal responses to apomorphine injection in a single recording day. The continuous line is the RMS of the bandpass-filtered analog signal (300–6000 Hz) in bins of 10 s. Color code: Black, Electrodes in which a DAN was identified; dotted gray, electrode with a SNr neuron.

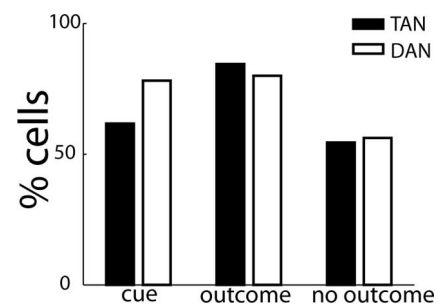


Figure 5. Percentage of TANs and DANs with significant responses to the different behavioral events. The percentage of neurons with significant responses to the cue, outcome, and no-outcome tone events of the total number of studied neurons ($n = 180$ TANs and 106 DANs). Color code: Black, TAN; white, DAN. For each epoch, we grouped trials according to trial type (aversive, reward, and neutral). A cell was considered to be significantly modulated in an epoch if at least one of the responses in that epoch was significant.

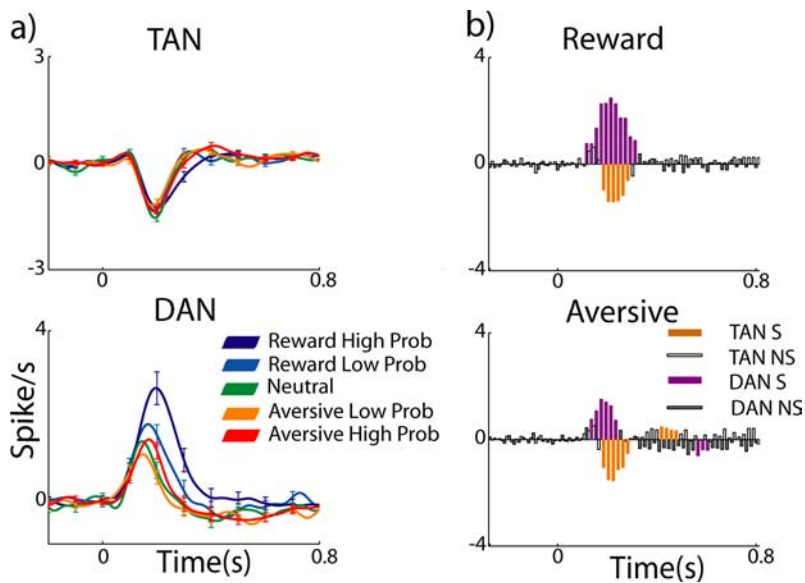


Figure 6. TAN and DAN population response at cue epoch. *a*, Population average response to behavioral cues. Only the first 0.8 s after the cue is shown to highlight the short duration of the responses. Top, TANs ($n = 180$ neurons). Bottom, DANs ($n = 106$). Color coding: Dark blue, Responses to high-probability ($p = 1$ and $p = 2/3$) reward cues; light blue, reward low ($p = 1/3$)-probability cues; green, neutral cue; orange, aversive low-probability cues; red, aversive high-probability cues. *b*, TAN versus DAN population response. The populations and timescale are the same as in *a*. The population response was considered significant if it passed the significance criteria (t test, $p < 0.01$) for at least three consecutive 20 ms bins. For this analysis, all trials of the same type (aversive or reward) were grouped. Top, TAN versus DAN in the reward trials. Bottom, TAN versus DAN in the aversive trials. Color coding: Orange, TAN significant bins; white, TAN nonsignificant bins; purple, DAN significant bins; gray, DAN nonsignificant bins.

with slightly shorter lags for the DAN responses. Finally, unlike the responses to the reward cue, the TAN and DAN responses to aversive cues had a significant second phase in which the TANs increased their activity and DANs decreased their activity (Fig. 6*b*, bottom).

The population average PSTH can be biased by a few neurons with an extreme response or opposite effects may be averaged out. We therefore formulated the difference index (see Materials and Methods) as a measure of the modulations of a single neuron to different events. We grouped responses across probabilities and tested whether the single-cell responses to reward and aversive cues were different from the response to the neutral cue. We found that, in both TAN and DAN populations, the response index (absolute deviation from the neutral response) for the reward trials was larger than the response index for aversive trials (Fig. 7*a*). A substantial fraction of TANs and DANs showed a significant response index to reward cues, whereas only a small number of cells had a significant response index to aversive cues (Fig. 7*a*, inset).

When separating the DAN and the TAN responses into high-probability ($p = 2/3$ and 1) and low-probability ($p = 1/3$) cues, coding of the reward probability was larger and more frequent than coding of the aversive probability (Fig. 7*b*). A multivariate ANOVA in which we did not group the high probability ($p = 2/3$ and 1) cues yielded similar results (data not shown). The difference between TAN single-cell responses and the TAN population results suggests that single-cell responses had opposite trends and were averaged out in the population analysis.

To summarize, we found larger and more frequent single-cell modulation of TAN and DAN discharge after reward than after aversive predicting cues (hypothesis 1). As expected from the TD hypothesis, DANs code reward probability both at level of the population and the single-unit responses; however, the TAN

population did not robustly encode the outcome probability. Disconfirming the TD hypothesis, DANs also increased their activity in response to aversive cues and there was a small (but significant) difference between the DAN responses to aversive cues with different predictions of a future aversive event (hypothesis 2). Finally, the first phase of the responses of the TANs to the visual cues generally mirrored the response of the DANs. DANs but not the TAN population discriminated robustly between reward and aversive cues (hypothesis 3).

Both TANs and DANs respond to aversive and reward outcome

Population analysis of the neural activity at the time of outcome delivery and coincident sounds showed that both TANs and DANs respond to food and air puff delivery, but with a faster response to the air puff (Fig. 8*a*). Whereas the cumulative responses of the TANs to aversive and reward outcome were similar in magnitude, the response of the DANs was larger for the reward events. However, although smaller in magnitude, the DANs respond with an excitation to the aversive outcome (Fig. 8*a*). TANs and DANs activity at reward

delivery was larger for the low-probability trials than for the high-probability trials (Fig. 8*a*). Comparison of the modulation time showed that the TAN and the DAN responses at the outcome epoch did not mirror each other. The large significant increase in the second phase of the TAN response to the reward outcome was coincident with only a small nonsignificant decrease in DAN activity (Fig. 8*b*, top). Furthermore, in the aversive outcome, the second phase of the TAN response overlapped the end of the first phase of the DAN response (i.e., the increase in the discharge rate of the TAN overlapped the increase in DAN activity) (Fig. 8*b*, bottom).

Single-cell analysis shows that a large fraction of the cells responded to reward or aversive events (Fig. 9*a*). However, as in the cue epoch, reward probability was better encoded than air puff probability in the outcome epoch as well (Fig. 9*b*), further showing that TAN and DAN activity was more strongly modulated by expectation of reward.

To summarize, we found a large modulation of TAN and DAN discharge rate after delivery of food or air puff, but with probability coding only for reward (hypothesis 1). We found larger TAN and DAN activity at reward delivery for the low-probability trials than for the high-probability trials. This trend was opposite to the trend found in the cue epoch (large response to the large probability) as expected according to the TD hypothesis (hypothesis 2). However, contrary to the naive TD hypothesis (which predicts that activity moves from outcome to cue during conditioning), we found that many TANs and DANs encoded the aversive outcome, whereas only very few encoded the aversive cue (compare Figs. 8, 9 with Figs. 6, 7). Furthermore, as opposed to the TD hypothesis, which predicts a decrease, DANs increased activity in response to the aversive outcome. In addition, the multiphasic response of the TANs, and the major changes observed in the second excitatory phase of the TAN responses, does

not enable a straightforward comparison with the TD predictions. Finally, the TAN and DAN activity were not completely coincident, and both populations discriminated between reward and aversive trials (hypothesis 3).

TAN but not DAN populations robustly differentiates between omission of rewards and omission of aversive events

In contrast to our previous study (Morris et al., 2004), outcome omission was explicitly notified to the monkeys by a typical sound (Fig. 1*b*). In the 2004 study, the responses of both DANs and TANs to the reward omission were small. Population analysis of the no-outcome events in the current study showed that TANs, but not DANs, had large modulations of their discharge rates. During this epoch, the TAN population response (Fig. 10*a*) differentiated between the reward trials (food omission), the aversive trials (air puff omission), and the neutral trials (expected no outcome). Furthermore, the suppression of TAN activity was slightly longer after omission of the high-probability reward (Fig. 10*a*). Analysis of the population PSTHs shows that only the average response of the TANs, but not the DANs to the outcome omission events was significant (Fig. 10*b*). The multiphase TAN response was not coincident with the phases of the insignificant DAN modulation (Fig. 10*b*). Finally, we did not find a significant difference between the duration of the DAN responses (Bayer et al., 2007) to the outcome omission after high- and low-probability cues (data not shown).

Single-cell analysis shows that, as in the cue epoch, the response index of both TANs and DANs for the reward trials was larger than the response index for aversive trials (Fig. 11*a*). A substantial fraction of TANs and DANs showed a significant response index to reward omission, whereas only a smaller number of cells had a significant response index to aversive omission (Fig. 11*a*, inset). Coding of reward probability was larger and more frequent than coding of the probability of the aversive events (Fig. 11*b*). The difference between DAN single-cell responses and the population results suggests that some single-cell responses had opposite trends and were averaged out in the population analysis.

To summarize, TAN and DAN single-cell modulations were larger for reward than for aversive omission, with probability coding only for reward omission (hypothesis 1). As shown previously, after outcome omission (no outcome) the

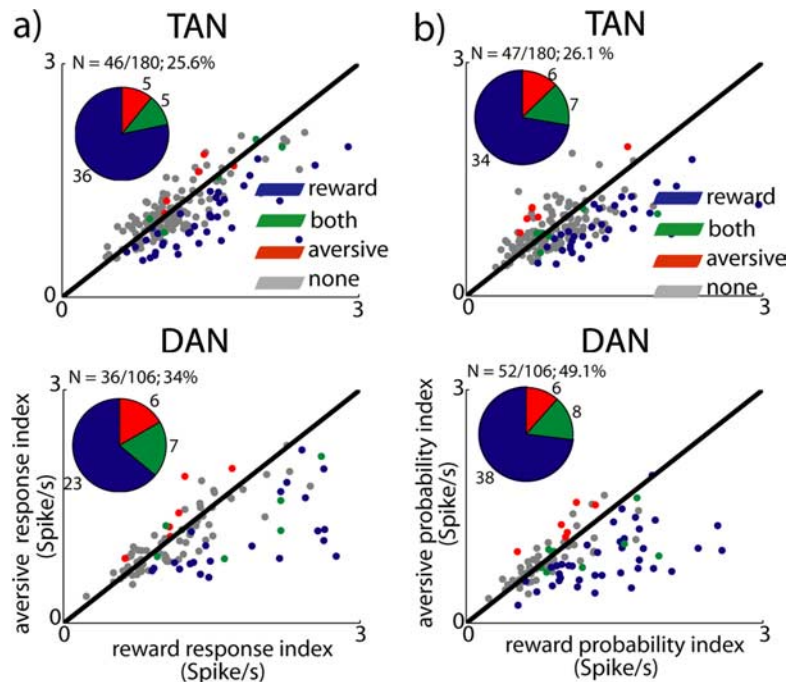


Figure 7. TAN and DAN single-cell response at cue epoch. *a*, Scatter plots comparing the response index of individual neurons to reward and aversive cues. Response index was calculated for each cell ($n = 180$ TANs and 106 DANs) as the absolute difference between the aversive or reward cue-aligned PSTH and the PSTH of the neutral cue. The black line is the identity ($Y = X$) line. Points below this line represent cells with a response index that is larger for the reward cues than for aversive cues. Top, TAN. Bottom, DAN. Color code: Blue, Response index significant only for reward cues; red, response index significant only for aversive cues; green, both response indices were significant; gray, neither response index was significant. Significance level was $p < 0.05$. The time window used for this analysis was 0–1000 ms from cue presentation. Inset, Pie chart of the fraction of cells with a significant response index for reward (blue), aversive (red), and both (green) cues of all cells with significant response index (number of responding of total number of cells is given in the text at inset top). *b*, Scatter plot comparing the probability coding of individual TAN and DAN neurons. The index was calculated as the difference between the grouped response to the high-probability ($p = 2/3$ and $p = 1$) and the low-probability ($p = 1/3$) events. The format and color code are the same as in *a*. Points below the identity line represent cells with a probability-coding index that is larger for the reward cues than for aversive cues.

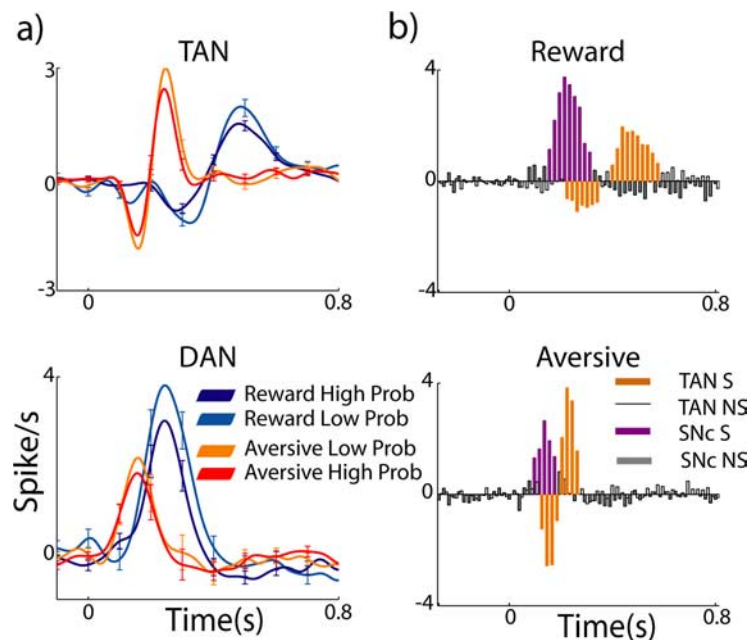


Figure 8. TAN and DAN population response at outcome delivery. *a*, Population responses at the time of outcome (food or air puff) and the corresponding sounds delivery. *b*, Comparison between the responses of TANs and DANs. The conventions are the same as in Figure 6.

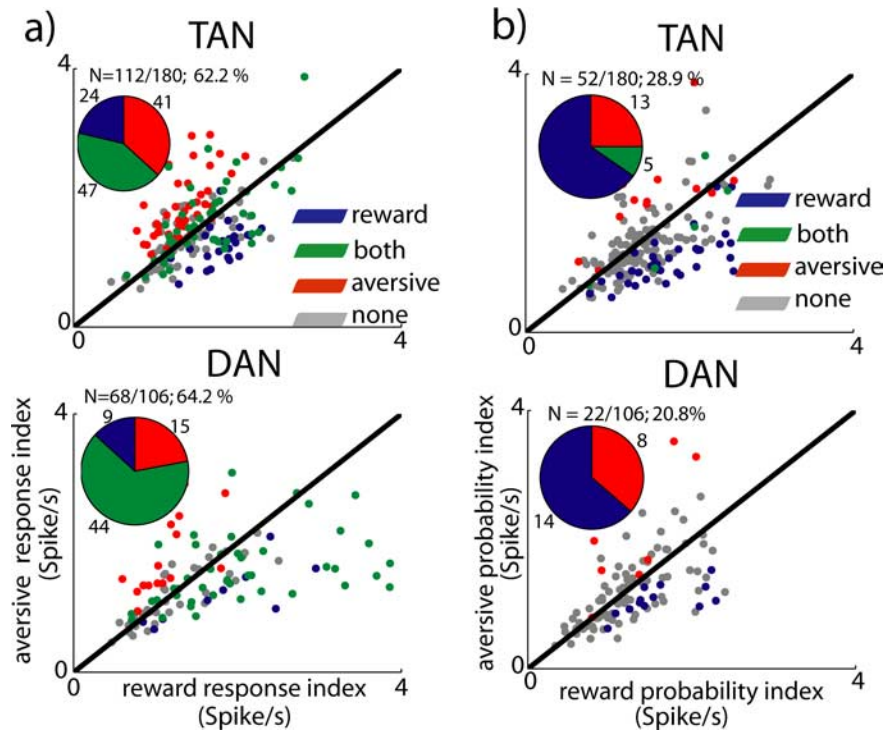


Figure 9. TAN and DAN single-cell response at outcome epoch. *a*, Scatter plots comparing the response index of individual neurons to reward and aversive outcomes. *b*, Scatter plot comparing the probability-coding index of the single neuron response to reward and aversive outcome. The conventions are the same as in Figure 7.

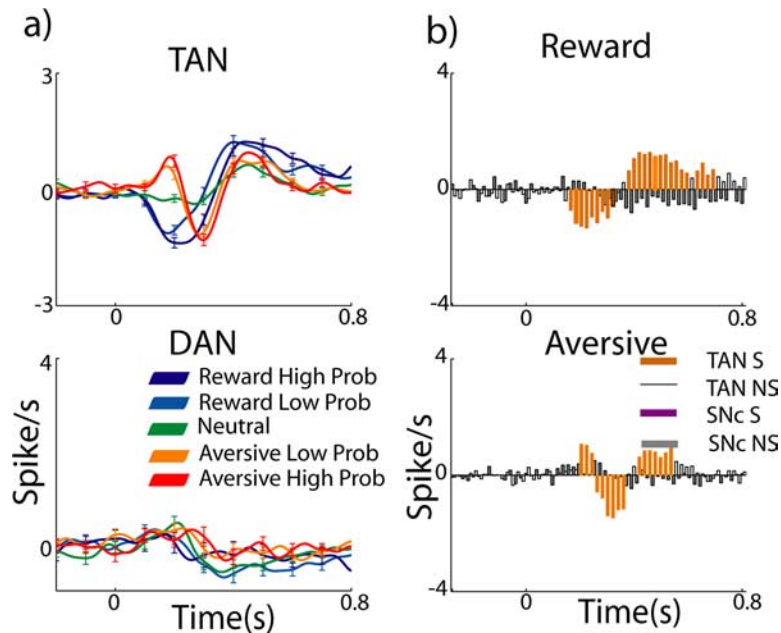


Figure 10. TAN and DAN population response at no outcome. *a*, Population responses in trials with no food or air puff delivery. The same no-outcome tone is given at time = 0. *b*, Comparison between the responses of TANs and DANs. The conventions are the same as in Figure 6.

DANs encode the TD error weakly (hypothesis 2). The TAN and DAN activity were not coincident, and the TAN, but not the DAN, population coded the difference between reward, aversive, and neutral trials robustly (hypothesis 3).

Discussion

In this manuscript, we have shown that DAN and TAN encoding is not limited to encoding of reward prediction error but also

reflects other psychological and behavioral processes. We found that rate modulations of TANs and DANs to expectation of reward were larger than the modulation, which followed predictions of aversive events. Furthermore, these neurons encode the expectation level (or the previous probability) of reward better than the expectation of aversive events. Finally, TAN responses were not coincident with DAN responses in all trial epochs. DANs encode the difference between reward and aversive trials in the cue and outcome epoch, whereas the TAN population encodes this difference in the outcome and no-outcome epochs. Therefore, complementary coding of TANs and DANs expands the encoding scope of the basal ganglia neuromodulators.

TANs and DANs strongly encode aversive outcome but not aversive expectations

There is no consensus regarding the responses of DANs to aversive events. Some studies suggest that at least some of the DANs increase their firing rate after aversive outcome (for review, see Horvitz, 2000), whereas others have evidence of a decrease (Ungless et al., 2004; Coizet et al., 2006). The reported increase in the firing rate of the DANs and striatal dopamine levels to negative events might be attributable to reward generalization (Mirenovic and Schultz, 1996; Kakade and Dayan, 2002; Day et al., 2007). However, the blinking and licking behavior observed here indicate that the monkeys were able to reliably discriminate between reward, neutral, and aversive cues. Second, the calculation of the response index as the difference between the responses to appetitive/aversive and the neutral events overcomes the confounding effects of generalization. Finally, we found that the neuronal response to the aversive outcome was faster than responses to reward trials.

Thus, DAN responses to aversive events may reflect multiple sources of modulation (see below for error prediction encoding). This may explain some of the inconsistencies between previous experiments. Whereas in the behaving animal there can be positive modulations of DAN discharge in response to aversive events because of attention/arousal processes, when the animal is anesthetized, only discomfort-related activity can be demonstrated (Ungless et al., 2004; Coizet et al., 2006).

As for the TANs, our study confirms the pioneering works showing fast and robust TAN responses to an aversive outcome (Ravel et al., 1999, 2003). However, our study extends our previous work (Morris et al., 2004) showing minimal differences between the TAN population responses to cues predicting future

rewards, to cues predicting future aversive events, and neutral cues. However, we found that unlike the population results that probably represent the average of opposing effects, many single TANs differentiate between reward and neutral trials and encode reward probability (Fig. 7). We further found that the population TANs encode the difference in the reward probability at outcome delivery (Fig. 8*a*) and that these cells have a large response to outcome omission (Fig. 10*a*). In the previous work (Morris et al., 2004), we found no discriminative response at outcome delivery and only a low response to reward omission. Future studies should explore whether this lack of concordance is attributable to differences in behavioral paradigms (for example, explicit vs implicit notification of trial termination, operant vs classical conditioning, and only rewarding outcomes vs rewarding and aversive outcomes) or to the animals' behavioral strategy and confidence in the prediction of future outcome.

DANs encode more than reward prediction errors

Recent studies have shown that DAN activity encodes the mismatch between prediction and reality. Most of these studies have focused on the mismatch in the positive domain (i.e., when conditions are better than expected) (Schultz et al., 1997). DANs typically increase their discharge rate in response to appetitive predictive cues and outcomes. In line with the predictions of reinforcement learning theories, the DAN discharge decreases with omission of predicted rewards (Schultz et al., 1993; Fiorillo et al., 2003; Matsumoto and Hikosaka, 2007). However, this discharge suppression is limited because the neuronal firing rate is truncated at zero. Indeed, several groups (Morris et al., 2004; Bayer and Glimcher, 2005) have reported that the instantaneous firing of DANs does not demonstrate incremental encoding of reward omission, and it was suggested that omission is encoded by duration of the discharge decrease (Bayer et al., 2007). In this experiment, however, we found to find any significant coding of reward omission by response amplitude or duration.

Naive reinforcement learning models categorize events as having positive or negative errors and would suggest opposite sign modulation to reward and aversive trials (Schultz et al., 1997). However, we found similar trends for DAN responses to predictions, outcomes, and omission of reward and aversive-related events (Figs. 6*a*, 8*a*, 10*a*). In particular, we found a substantial increase to both reward and aversive outcome. Furthermore, responses of the DANs to reward omission and aversive outcome (Figs. 8*a*, 10*a*, respectively) were very different (decrease vs increase), although in both cases there was a negative reinforcement error.

To summarize, our results reveal an increase in the complexity of the encoding by the DANs of value. This does not rule out their role in the temporal difference hypothesis. On the contrary, our working hypothesis holds that the discharge rate of DANs and TANs reflects changes in reward prediction as well as changes in

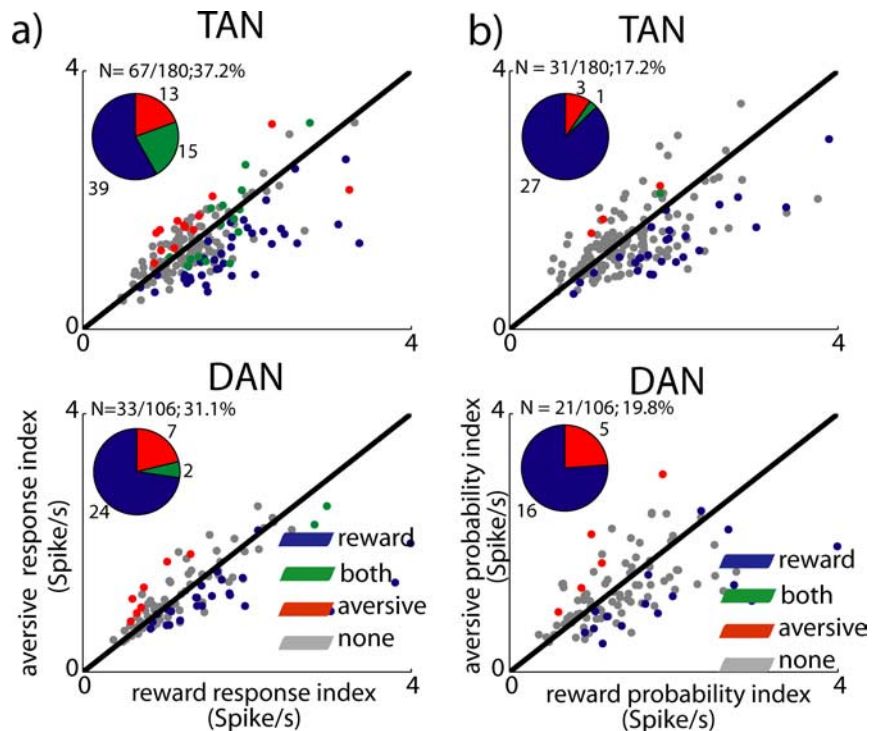


Figure 11. TAN and DAN single-cell response at no outcome. *a*, Scatter plots comparing the response index of individual neurons in trials in which food and air puff were not delivered. *b*, Scatter plot comparing the probability-coding index. The conventions are the same as in Figure 7.

attention/arousal levels (Horvitz, 2000; Ravel and Richmond, 2006; Redgrave and Gurney, 2006).

Asymmetric encoding of positive and negative expectations by the basal ganglia

Previous primate instrumental conditioning experiments in which DANs and TANs were recorded did not include expectation of aversive outcome because the air puff could be avoided by a correct response (Mirenovic and Schultz, 1996; Yamada et al., 2004). The symmetric classical conditioning paradigm of this study, which included reward predicting, aversive predicting, and a neutral cue, enabled us to explore whether there was symmetry in the encoding of expectation of rewarding versus aversive events by the TANs and the DANs.

Single-cell analysis revealed that TAN and DAN encoding of reward expectation and omission was larger and more frequent than encoding of expectation and omission of aversive events (Figs. 7*a*, 11*a*). Furthermore, we found that TAN and DAN encoding of the reward probability was larger and more frequent than their encoding of the probability of the air puff-related events (Figs. 7*b*, 9*b*, 11*b*). The preferential activation to reward was also apparent in the population response of the DANs at the cue and outcome epoch in which the activity of these cells was larger and coded the probabilities better (Figs. 6*a*, 8*a*). Thus, in line with previous studies (Mirenovic and Schultz, 1996; Yamada et al., 2007), we show that even in a classical conditioning task in which the air puff is unavoidable, expectation of aversive events is weakly represented in the basal ganglia activity.

TANs do not mirror the DAN responses

The anatomical demonstration of dopaminergic innervations of striatal cholinergic interneurons (Lehmann and Langer, 1983)

and the suppression of acetylcholine efflux from striatal slice by dopamine (Stoof et al., 1992) suggest that DANs directly inhibit the TANs (Wang et al., 2006). TANs might mediate the dopaminergic message to the D₁ and D₂ dopamine receptor containing striatal projection neurons.

The opposite and coincident responses of the TANs and DANs to predictive cues (Fig. 6) support direct inhibition. However, TAN responses at the terminal stage of the trial (Figs. 8*b*, 10*b*) include major positive deflections that do not mirror any phase of the dopaminergic response. Notably, after outcome omission, DANs respond similarly to the neutral outcome, reward, and air puff omissions, whereas the TANs robustly discriminate between the three events (Fig. 10). Thus, DANs may better encode the cue predicting events and the TANs may provide more information at the completion of the trial. This is consistent with the findings of subpopulations of striatal projection neurons with selective evaluative encoding of trial results (Lau and Glimcher, 2007, 2008). In any case, these differential responses indicate that the TAN discharge is not totally governed by its dopaminergic inputs; neither are the TANs and DANs driven by a common source (Matsumoto et al., 2001) with opposite effects on the two systems.

Concluding remarks

In this study, we showed that the dopaminergic and the cholinergic neuromodulators of the basal ganglia encode the positive domain of behavior in a nonredundant manner. This asymmetric encoding of behavior suggests that the basal ganglia collaborate with other neuronal systems to shape the animal's response to diverse environmental events. The characteristics and interactions of these different neuronal systems may provide the basis for asymmetric, irrational human attitudes toward rewarding and aversive events (Tversky and Kahneman, 1981). Finally, the stronger involvement of the basal ganglia in positive reinforcement learning is congruent with the findings that parkinsonian patients are better at learning to avoid choices that lead to negative outcomes than they are at learning from positive outcomes (Frank et al., 2004).

References

- Aebischer P, Schultz W (1984) The activity of pars compacta neurons of the monkey substantia nigra is depressed by apomorphine. *Neurosci Lett* 50:25–29.
- Arbuthnott GW, Wickens J (2007) Space, time and dopamine. *Trends Neurosci* 30:62–69.
- Barbeau A (1962) The pathogenesis of Parkinson's disease: a new hypothesis. *Can Med Assoc J* 87:802–807.
- Bar-Gad I, Bergman H (2001) Stepping out of the box: information processing in the neural networks of the basal ganglia. *Curr Opin Neurobiol* 11:689–695.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- Bayer HM, Lau B, Glimcher PW (2007) Statistics of midbrain dopamine neuron spike trains in the awake primate. *J Neurophysiol* 98:1428–1439.
- Berntson GG, Bigger JT Jr, Eckberg DL, Grossman P, Kaufmann PG, Malik M, Nagaraja HN, Porges SW, Saul JP, Stone PH, van der Molen MW (1997) Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34:623–648.
- Calabresi P, Centonze D, Gubellini P, Pisani A, Bernardi G (2000) Acetylcholine-mediated modulation of striatal function. *Trends Neurosci* 23:120–126.
- Coizet V, Dommert EJ, Redgrave P, Overton PG (2006) Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat. *Neuroscience* 139:1479–1493.
- Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* 10:1020–1028.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Gourévitch B, Eggermont JJ (2007) A simple indicator of nonstationarity of firing rate in spike trains. *J Neurosci Methods* 163:181–187.
- Graybiel AM, Aosaki T, Flaherty AW, Kimura M (1994) The basal ganglia and adaptive motor control. *Science* 265:1826–1831.
- Guarraci FA, Kapp BS (1999) An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav Brain Res* 99:169–179.
- Gurney K, Prescott TJ, Wickens JR, Redgrave P (2004) Computational models of the basal ganglia: from robots to membranes. *Trends Neurosci* 27:453–459.
- Horvitz JC (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96:651–656.
- Joshua M, Elias S, Levine O, Bergman H (2007) Quantifying the isolation quality of extracellularly recorded action potentials. *J Neurosci Methods* 163:267–282.
- Kakade S, Dayan P (2002) Dopamine: generalization and bonuses. *Neural Netw* 15:549–559.
- Lau B, Glimcher PW (2007) Action and outcome encoding in the primate caudate nucleus. *J Neurosci* 27:14502–14514.
- Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior. *Neuron* 58:451–463.
- Lehmann J, Langer SZ (1983) The striatal cholinergic interneuron: synaptic target of dopaminergic terminals? *Neuroscience* 10:1105–1120.
- Martin RF, Bowden DM (2000) Primate brain maps: structure of the macaque brain. Amsterdam: Elsevier Science.
- Matsui T, Koyano KW, Koyama M, Nakahara K, Takeda M, Ohashi Y, Naya Y, Miyashita Y (2007) MRI-based localization of electrophysiological recording sites within the cerebral cortex at single-voxel accuracy. *Nat Methods* 4:161–168.
- Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447:1111–1115.
- Matsumoto N, Minamimoto T, Graybiel AM, Kimura M (2001) Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J Neurophysiol* 85:960–976.
- Mirenovic J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–451.
- Moran A, Bar-Gad I, Bergman H, Israel Z (2006) Real-time refinement of subthalamic nucleus targeting using Bayesian decision-making on the root mean square measure. *Mov Disord* 21:1425–1431.
- 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.
- Ravel S, Richmond BJ (2006) Dopamine neuronal responses in monkeys performing visually cued reward schedules. *Eur J Neurosci* 24:277–290.
- Ravel S, Legallet E, Apicella P (1999) Tonicly active neurons in the monkey striatum do not preferentially respond to appetitive stimuli. *Exp Brain Res* 128:531–534.
- Ravel S, Legallet E, Apicella P (2003) Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. *J Neurosci* 23:8489–8497.
- Redgrave P, Gurney K (2006) The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci* 7:967–975.
- Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. *Nature* 413:67–70.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
- Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13:900–913.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.

- Shimo Y, Hikosaka O (2001) Role of tonically active neurons in primate caudate in reward-oriented saccadic eye movement. *J Neurosci* 21:7804–7814.
- Stoof JC, Drukarch B, de Boer P, Westerink BH, Groenewegen HJ (1992) Regulation of the activity of striatal cholinergic neurons by dopamine. *Neuroscience* 47:755–770.
- Sutton RS, Barto AG (1998) Reinforcement learning—an introduction. Cambridge, MA: MIT.
- Szabo J, Cowan WM (1984) A stereotaxic atlas of the brain of the cynomolgus monkey (*Macaca fascicularis*). *J Comp Neurol* 222:265–300.
- Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211:453–458.
- Ungless MA, Magill PJ, Bolam JP (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303:2040–2042.
- Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J, Tkatch T, Lovinger DM, Surmeier DJ (2006) Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 50:443–452.
- Yamada H, Matsumoto N, Kimura M (2004) Tonicly active neurons in the primate caudate nucleus and putamen differentially encode instructed motivational outcomes of action. *J Neurosci* 24:3500–3510.
- Yamada H, Matsumoto N, Kimura M (2007) History- and current instruction-based coding of forthcoming behavioral outcomes in the striatum. *J Neurophysiol* 98:3557–3567.