Novelty encoding by the output neurons of the basal ganglia

Mati Joshua[1,2]*, Avital Adler[1,2] and Hagai Bergman[1,2]

[1] Department of Medical Neurobiology, The Hebrew University-Hadassah Medical School, Jerusalem, Israel
[2] The Interdisciplinary Center for Neural Computation, The Hebrew University, Jerusalem, Israel

*Correspondence:
Mati Joshua, Department of Medical
Neurobiology, The Hebrew University –
Hadassah Medical School, POB 12272,
Jerusalem 91120, Israel.
e-mail: mati@alice.nc.huji.ac.il

Reinforcement learning models of the basal ganglia have focused on the resemblance of the dopamine signal to the temporal difference error. However the role of the network as a whole is still elusive, in particular whether the output of the basal ganglia encodes only the behavior (actions) or it is part of the valuation process. We trained a monkey extensively on a probabilistic conditional task with seven fractal cues predicting rewarding or aversive outcomes (familiar cues). Then in each recording session we added a cue that the monkey had never seen before (new cue) and recorded from single units in the Substantia Nigra pars reticulata (SNpr) while the monkey was engaged in a task with new cues intermingled within the familiar ones. The monkey learned the association between the new cue and outcome and modified its licking and blinking behavior which became similar to responses to the familiar cues with the same outcome. However, the responses of many SNpr neurons to the new cue exceeded their response to familiar cues even after behavioral learning was completed. This dissociation between behavior and neural activity suggests that the BG output code goes beyond instruction or gating of behavior to encoding of novel cues. Thus, BG output can enable learning at the levels of its target neural networks.

Keywords: reinforcement learning, substantia nigra pars reticulata, primate, spikes

INTRODUCTION

Experimental and modeling studies have emphasized the involvement of the basal ganglia in reinforcement learning (Schultz et al., 1997; Sutton and Barto, 1998). Dopaminergic neurons respond when there is a reward prediction error (Hollerman and Schultz, 1998; Morris et al., 2004; Bayer and Glimcher, 2005; Joshua et al., 2008); the dopaminergic signal enables reshaping of cortico-striatal mapping (Reynolds et al., 2001; Shen et al., 2008) and modification of behavior. In fact studies of the activity of striatal neurons during learning have shown that both behavior and neural activity rapidly adapt to new events (Lauwereyns et al., 2002; Pasupathy and Miller, 2005; Williams and Eskandar, 2006; Kimchi and Laubach, 2009).

Early studies provided evidence that the Substantia Nigra pars reticulata (SNpr) neurons have both sensory and motor related responses (DeLong et al., 1983; Hikosaka and Wurtz, 1983a,b; Nishino et al., 1985; Schultz, 1986). More recent studies have shown that the SNpr motor and sensory signals depend on the context of the movement (Handel and Glimcher, 2000). One contextual modulation is the reinforcement associated with a movement (Gulley et al., 2002; Sato and Hikosaka, 2002; Wichmann and Kliem, 2004). These studies of the SNpr and most studies of the basal ganglia during animal conditioning have focused on activity after animals had been extensively trained and reward associations established. In this manuscript we analyzed the activity of SNpr neurons during the learning of new associations.

A number of computational models posit that the basal ganglia output encodes future actions and projects behavioral instructions (e.g., by gating or enabling selected actions) to the cortex and brainstem motor centers (Chevalier and Deniau, 1990; Mink, 1996). This suggests that SNpr activity would be modified along the same time course as behavior. Alternatively, the output of the basal ganglia may provide a learning signal which differs from the behavioral instruction. This could lead to a dissociation of the activity of SNpr neurons from the behavior. To probe this issue, we recorded extracellular spiking activity of neurons in the SNpr, one of the two output structures of the basal ganglia, while assessing modifications in a monkey's oro-facial learning behavior. New cues were intermingled with highly familiar stimuli, enabling comparison of behavioral and neural responses to novel and familiar events.

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 the Hebrew University guidelines for the use and care of laboratory animals in research, supervised by the institutional animal care and use committee. All procedures have been described in more detail in our previous reports (Joshua et al., 2008, 2009b). Here we present a summary of these methods and describe methods not used in the previous manuscripts in detail.

BEHAVIORAL TASK

A monkey (Macaca Fascicularis, female, 4 kg) was trained on a classical conditioning task with seven different fractal visual cues. Three cues (reward cues) predicted a liquid food outcome (0.4 ml, 100 ms duration) with a delivery probability of 1/3, 2/3 and 1; three cues (aversive cues) predicted an airpuff outcome (100 ms duration; 50–70 psi; split and directed 2 cm from each eye) with a delivery probability of 1/3, 2/3 and 1. The 7th cue (the neutral cue) was never followed by a food or an airpuff outcome.

After 6 months of training on the seven cue task, we implanted a recording chamber and recorded neural activity from the basal ganglia (Joshua et al., 2008, 2009b). In every recording session (usually two sessions per day) we added a fractal visual cue that had never been introduced to the monkey before (new cue). The cue could predict an aversive or reward outcome with a probability of 1/3 or 2/3. **Figure 1A** shows an example of the flow of the task. All cues (new and familiar) occupied the full screen of 17" LCD monitor located 50 cm from the monkey's eyes and were presented for 2 s. The cues were followed by an outcome/no outcomes (indicated also by different sounds) and then by a variable inter- trial interval (4–8 s).

Due to the probabilistic structure of the behavioral task and in order to equalize the average occurrence of each outcome, the familiar and new cues (that had an uncertain outcome; i.e. p not equal 1) were introduced three time more than the familiar cues with $p = 1$. All trials (familiar and new cues) were randomly interleaved with the occurrence ratio noted above. The analysis of the responses of the same group of cells to the familiar events is reported elsewhere (Joshua et al., 2009b).

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 1,000 Hz), confined within a cylindrical guide (1.65-mm inner diameter), were advanced separately (EPS, Alpha-Omega Engineering, Nazareth, Israel) into the target. The electrical activity was amplified with a gain of 5 K and band-pass filtered with a 1–6,000 Hz four-pole Butterworth filter and continuously sampled at 25 KHz by 12 bits ± 5 V A/D (AlphaLab, Alpha-Omega Engineering). Spike activity was sorted and classified online using a template-matching algorithm (ASD, Alpha-Omega Engineering). Spike-detection pulses and behavioral events were sampled at 25 kHz (AlphaLab, Alpha-Omega Engineering).

Recorded units were subjected to offline quality analysis that included tests for rate stability, refractory period (less than 2% of the inter spike intervals were less than 2 ms), waveform isolation (Isolation score > 0.8 Joshua et al., 2007) and recording time (more than 20 min continuously).

SNpr neurons were identified during recording according to the electrophysiological characteristics (narrow spike shape and high firing rate) of the cells (DeLong et al., 1983) and the firing characteristics of neighboring neurons and fibers (e.g., fibers of the internal capsule, SN pars compacta dopaminergic neurons, and fibers of the oculomotor nerve). To validate classification we carried out offline analyses of the neurons' extracellular waveform shape and firing rate. Waveform shape was quantified as the duration from the first negative peak to the next positive peak; rate was defined as the average rate during the whole recording session. The results of this analysis were reported in previous manuscript (see **Figure 4** in Joshua et al., 2008).

While the monkey was engaged in this same task with the familiar and new cues we also recorded the activity of cells from the external and internal segment of the globus pallidus (GPe and GPi respectively) and from the tonic active neurons of the striatum (TANs) and the midbrain dopaminergic neurons. The analysis of these populations did not reveal the difference between behavior and neural activity that we found for SNpr neurons and therefore was not included in this report.

STATISTICAL ANALYSIS

Analysis of behavior

A computerized digital video camera recorded the monkey's face at 50 Hz. Video analysis was carried out on custom software to identify periods when the monkey closed its eyes (Mitelman et al., 2009). The mouth was monitored by an infrared reflection detector. Based on these recordings we detected times in which the monkey moved its mouth by implementing a threshold based method. We compared mouth movement detection with the video movies over several recording days and found that they were consistent.

For each trial we defined two variables:

$$\text{Lick} = \begin{cases} 1 & \text{if monkey licked in the last 500 ms of cue epoch} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Blink} = \begin{cases} 1 & \text{if monkey closed its eyes in the last 500 ms of} \\ & \text{cue epoch} \\ 0 & \text{otherwise} \end{cases}$$

Note that the outcome (airpuff, food or their omission) immediately followed the cue epoch, and therefore the behavior of the monkey at the end of the cue epoch reflected outcome anticipation.

Behavior on trial t was defined as the difference between the licking and blinking response:

Behaviour(t) = Lick(t)−Blink(t)

This definition may appear ambiguous since behavior was scored zero in trials when the monkey did not lick or blink, and in trials where it both licked and blinked. However we found coincident licking and blinking (in the last 500 ms of a trial) to be rare in our data set (less than 5% of the trials). Analysis of behavior that excluded these trials yielded similar results (data not shown).

The reduction of the continuous behavior to a single value for each trial enabled an ordered comparison of the different responses. It should be noted, separating the behavior into lick and blink responses would have required two different tests to determine behavioral changes; leading to multi comparison difficulties which would also limit direct comparison with the neural data.

To reduce the trial by trial fluctuations in the behavioral responses we smoothed the behavior vector by a moving average of 20 new trials. To compare the responses to the new events to the responses in the familiar trials we grouped the familiar trials that were introduced between the first and last trials of the group of 20 new trials (the same grouping was applied to the neural data – see below). Finally, to enable comparison between different sessions, which occasionally differed in the baseline of the behavior, we normalized all smoothed behavioral responses between 0 and 1; i.e., in each session the behavior response [$X(t)$] was transformed to a behavior-index by ($X(t)$-min)/(max-min), where min and max are the minimal and maximal values of the smoothed behavioral responses to all events in that session. We repeated the behavioral analysis with different time windows and threshold detections for licking, and found similar results to those reported here (data not shown).

Comparing responses to the new cue with responses to familiar cues in the same category

Our main goal was to compare behavior and neural activity. To minimize the confounding effects that could arise from differences in the analysis methods we ran the same statistical analysis on both cell activity and behavior. We used the familiar trials that were recorded between the first and last of the group of 20 new trials to test whether responses (firing rate for cells or the behavioral response) to the new cue resembled the response to any of the familiar cues with the same outcome (t-test, $p < 0.05$). A response to a new cue was considered to be significantly different from the responses to the familiar cues of the same category if it was significantly different from all responses to the familiar events and if the response to the new event did not fall between the responses to the familiar events. We repeated this test without the latter condition and obtained similar results. In addition, we only ran the test on the new and familiar cues with the same outcome probability; this analysis yielded similar results (i.e. dissociation of cells from behavior) to those reported below. We therefore decided to report the results of our most conservative test; i.e., the response to the new cue was considered to be significantly different from the responses to the familiar cue only if it was significantly different from all responses to the familiar events of the same category and if the response to the new event did not fall between the responses to the familiar events.

In the analysis of the neural data we did not try to characterize the learning curve of each cell separately since only a small fraction of our cells reached the quality criteria for the whole session. Some cells reached the quality criteria from the first trial but not in later parts of the session while others were isolated only in the middle of the session.

To test for the temporal structure of the modulations we repeated the test for differences between familiar and new events in bins of 200 ms after cue presentation during the cue presentation epoch (2 s). In this analysis we were interested in quantifying the time course of modulation after behavior reached saturation and hence we performed this analysis only on trials that were recorded after more than 50 presentations of the new cue.

For all the above analyses we did not use non-probabilistic ($p = 1.0$) cue events for several reasons. First, previous studies have shown that basal ganglia cells may encode the uncertainty of probabilistic events (Fiorillo et al., 2003; Tan and Bullock, 2008). We therefore used only familiar cues with the same certainty (note that for $p = 1/3$ and $2/3$ the outcome uncertainty, which may be defined as the outcome variability or entropy, is equal). Second, since the new cues were always probabilistic ($p = 1/3$ or $2/3$) the monkey could have learned this meta-rule. Finally, the non-probabilistic cues were presented much less often than the probabilistic events (to enable an equal number of outcome delivery and omission trials) and hence tests involving these cues had lower statistical power.
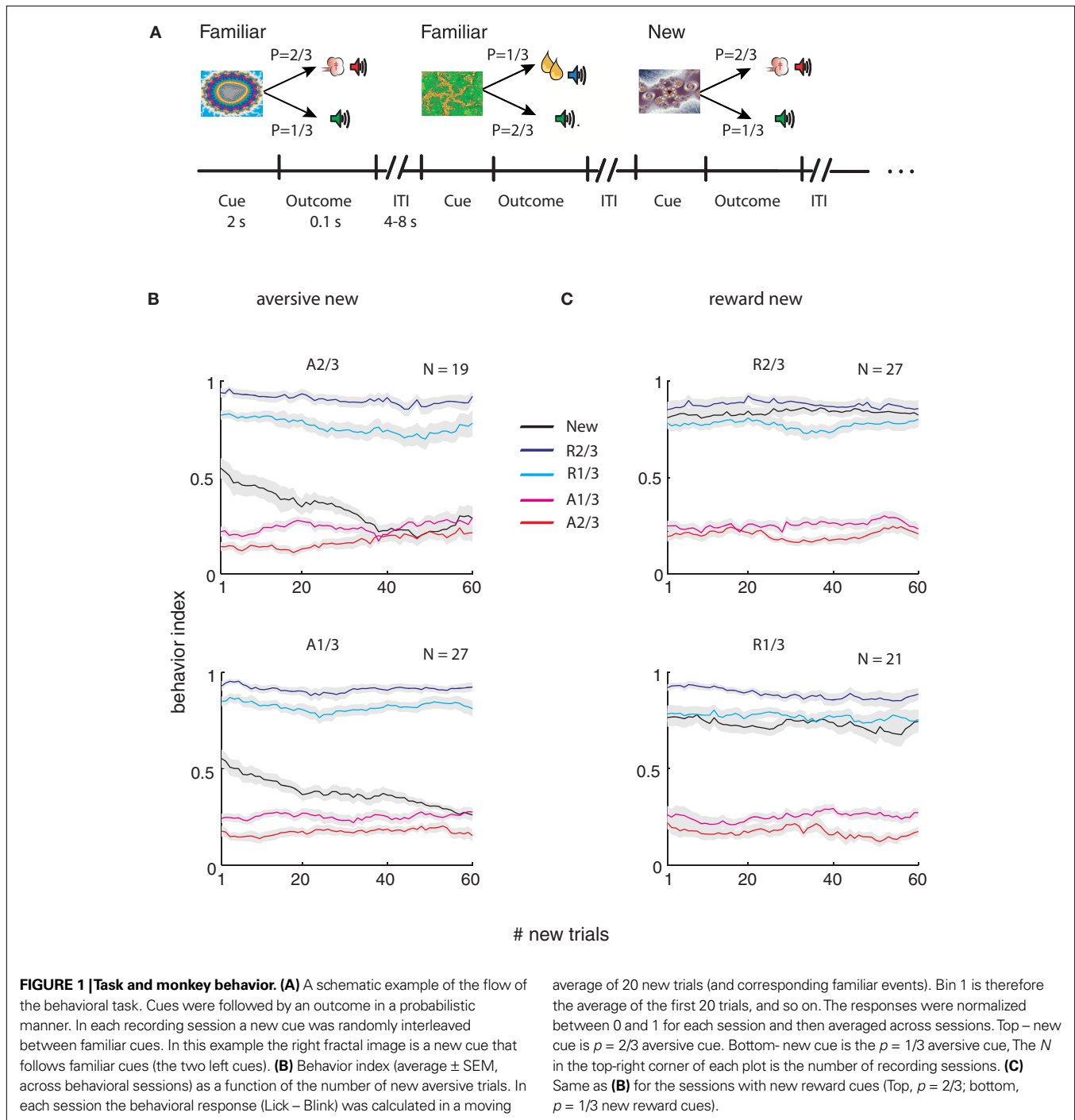
RESULTS

We trained a monkey extensively on a classical conditioning task (5 days/week for 6 months, Joshua et al., 2008, 2009b). During the training period the monkey learned to reliably associate the seven visual cues with the probability of food (reward cues) or airpuff (aversive cue) outcomes. Then, after implantation of the recording chamber and a recovery period we resumed the behavioral session

in parallel with the neural recordings. In every recording session we introduced one new cue that had never been introduced before. The monkey learned to associate the new cue with a probabilistic rewarding or aversive outcome (**Figure 1A**). Data were collected from 101 SNpr cells, of which 66 passed the study inclusion criteria, while the monkey was engaged in the behavioral task. Of the 66 cells, 14 were recorded for two recording sessions (but on the same day) and were used twice in the analysis, yielding $n = 80$ in the analysis database. We repeated the analysis including only one recording session per cell and obtained similar results (data not shown). The average analysis time was 54 min per neuron, which included on average 57 new trials intermingled within 289 familiar trials.

Figure 1A shows the flow of the behavioral task; new cues were randomly interleaved with familiar cues. **Figure 1B** shows average behavioral response to new and familiar cues. We found that behavior adapted very rapidly to the rewarding events and more slowly to the aversive events. In both cases, after 50 new trials the behavioral responses to the new event resembled the oro-facial responses to the corresponding familiar event. Note that in **Figure 1** each data point is an average of 20 trials; thus the difference between reward and aversive responses in the first bin reflects the different time scale of learning and not the ability of the monkey to predict the first trial outcome.
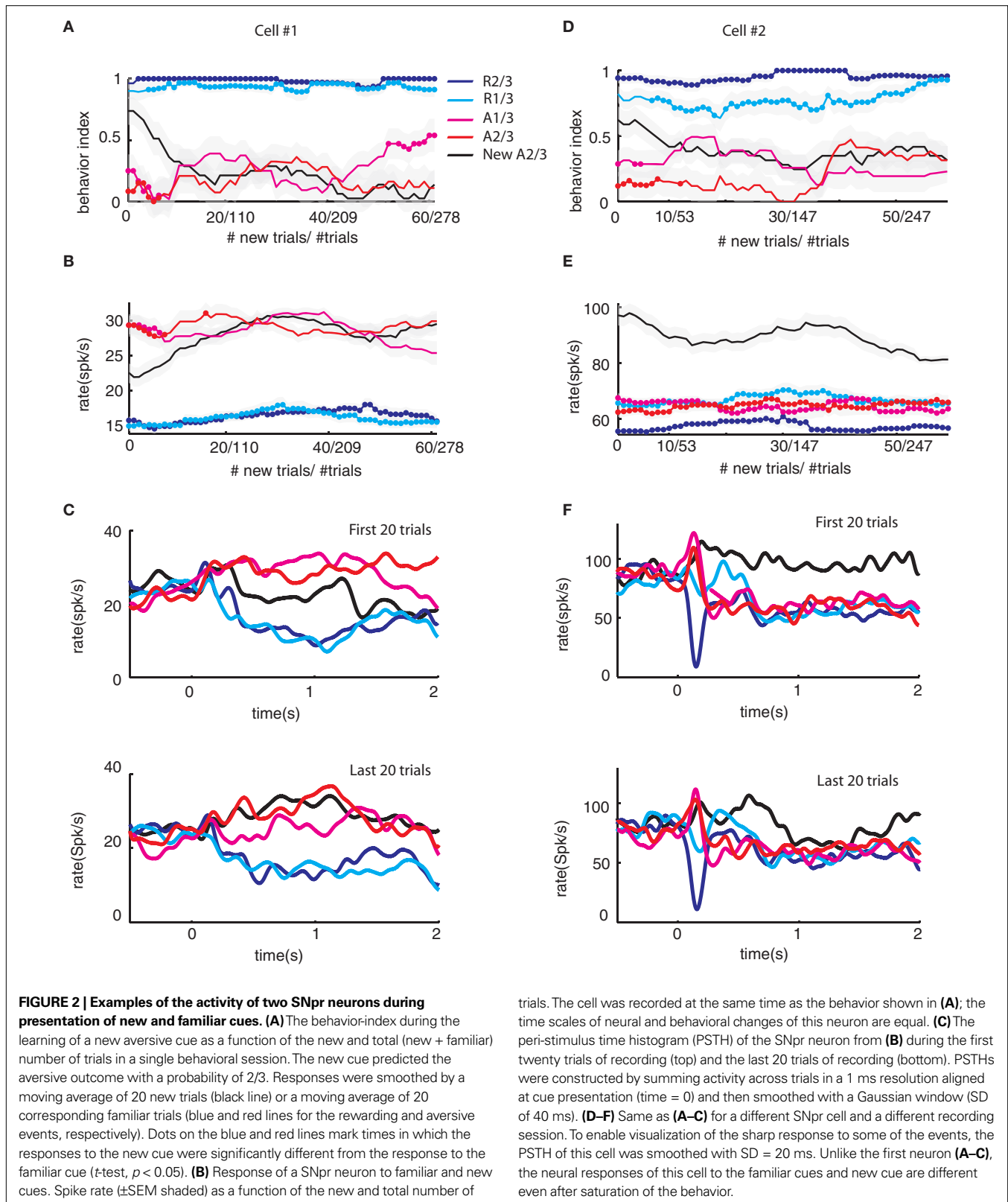
Figure 2 shows examples of the neural responses of two SNpr cells and the associated licking and blinking behavior. The changes in activity in the first cell (**Figures 2B,C**) matched the modification in the monkey's behavior (**Figure 2A**). At first, both neural activity and the behavioral response to the new aversive stimulus were between the responses to the familiar rewarding and aversive cues. After ~15 new trials both resembled the response to the familiar aversive predicting cues. However, a significant fraction of SNpr cells did not show the same time course for behavior and neural learning. In **Figures 2D–F** we show an example of a SNpr cell that did not follow the behavioral learning. Even though the monkey's behavioral response after several new trials was similar to its response to the familiar aversive cue (**Figure 2D**), the response of this SNpr cell (**Figures 2E–F**) to a new aversive cue differed from the responses to all familiar cues during the entire recording session.

We found that these two different profiles of responses were observed in other SNpr neurons. **Figure 3** shows an analysis of the fraction of cells and the fraction of behavior sessions in which the responses to the new cue were different from the responses to the familiar cue of the same category. One possible problem in comparing behavioral and neural data is the difference in the statistical methods. In **Figure 3** we tried to minimize this confounding factor by performing the same tests on both the behavioral and neural data. **Figure 3** shows that after 50 new trials the behavioral responses to the new aversive event resembled the oro-facial responses to the corresponding familiar event (black line). This conclusion was valid not only for the average response (**Figure 1B**) but also for individual sessions (**Figure 3A**). By contrast to the behavioral response, many SNpr neurons (**Figure 3A**, red curve) responded differentially to the new and familiar aversive cues even after behavioral learning had reached its ceiling. Similarly, a large fraction of SNpr neurons maintained different responses to the new rewarding cue even after 50 new trials, despite an almost immediate behavioral adaption to the new rewarding trials (**Figure 3B**).

FIGURE 1 | Task and monkey behavior. (A) A schematic example of the flow of the behavioral task. Cues were followed by an outcome in a probabilistic manner. In each recording session a new cue was randomly interleaved between familiar cues. In this example the right fractal image is a new cue that follows familiar cues (the two left cues). **(B)** Behavior index (average ± SEM, across behavioral sessions) as a function of the number of new aversive trials. In each session the behavioral response (Lick – Blink) was calculated in a moving average of 20 new trials (and corresponding familiar events). Bin 1 is therefore the average of the first 20 trials, and so on. The responses were normalized between 0 and 1 for each session and then averaged across sessions. Top – new cue is $p = 2/3$ aversive cue. Bottom- new cue is the $p = 1/3$ aversive cue, The N in the top-right corner of each plot is the number of recording sessions. **(C)** Same as **(B)** for the sessions with new reward cues (Top, $p = 2/3$; bottom, $p = 1/3$ new reward cues).

To facilitate comparison of behavior and neuron activity, in **Figure 3** we disregarded the temporal profile of the responses of SNpr neurons and used a single number (total spike count during cue presentation) to quantify the neural response. To further quantify the response of SNpr neurons to new events we investigated the temporal pattern of the neural responses to the new cues after behavioral learning had saturated (i.e. after more than 50 presentation of the new cue, see examples at **Figures 2C,F**). **Figure 4** shows the analysis of the temporal pattern of the average responses to the new events. As found previously for average responses to familiar events (Joshua et al., 2009a,b) the responses to the new cues were diverse. They include both transient and sustained responses (**Figures 4A,D**) and both increases and decreases in discharge rate (**Figures 4B,E**). To test for the specific contribution of the new events to the cell response we compared the response to the new cue and the response to familiar cues of the same category at different times after cue presentation (in a 200-ms bin). This analysis used the response to familiar events as a dynamic baseline; thus any modulation of the activity beyond these responses can be attributed to the novelty of the new cue. We found that most of the

FIGURE 2 | Examples of the activity of two SNpr neurons during presentation of new and familiar cues. (A) The behavior-index during the learning of a new aversive cue as a function of the new and total (new + familiar) number of trials in a single behavioral session. The new cue predicted the aversive outcome with a probability of 2/3. Responses were smoothed by a moving average of 20 new trials (black line) or a moving average of 20 corresponding familiar trials (blue and red lines for the rewarding and aversive events, respectively). Dots on the blue and red lines mark times in which the responses to the new cue were significantly different from the response to the familiar cue (*t*-test, *p* < 0.05). **(B)** Response of a SNpr neuron to familiar and new cues. Spike rate (±SEM shaded) as a function of the new and total number of

trials. The cell was recorded at the same time as the behavior shown in **(A)**; the time scales of neural and behavioral changes of this neuron are equal. **(C)** The peri-stimulus time histogram (PSTH) of the SNpr neuron from **(B)** during the first twenty trials of recording (top) and the last 20 trials of recording (bottom). PSTHs were constructed by summing activity across trials in a 1 ms resolution aligned at cue presentation (time = 0) and then smoothed with a Gaussian window (SD of 40 ms). **(D–F)** Same as **(A–C)** for a different SNpr cell and a different recording session. To enable visualization of the sharp response to some of the events, the PSTH of this cell was smoothed with SD = 20 ms. Unlike the first neuron **(A–C)**, the neural responses of this cell to the familiar cues and new cue are different even after saturation of the behavior.

cells that responded differently to new and familiar events encoded the difference at the first second of the cue presentation epoch (**Figures 4C,F**). Furthermore the majority of the cells increased

their firing rate beyond the responses to the familiar cues and only very few showed a discriminative (between the new and the familiar cue) decrease in discharge rate (**Figures 4C,F**).
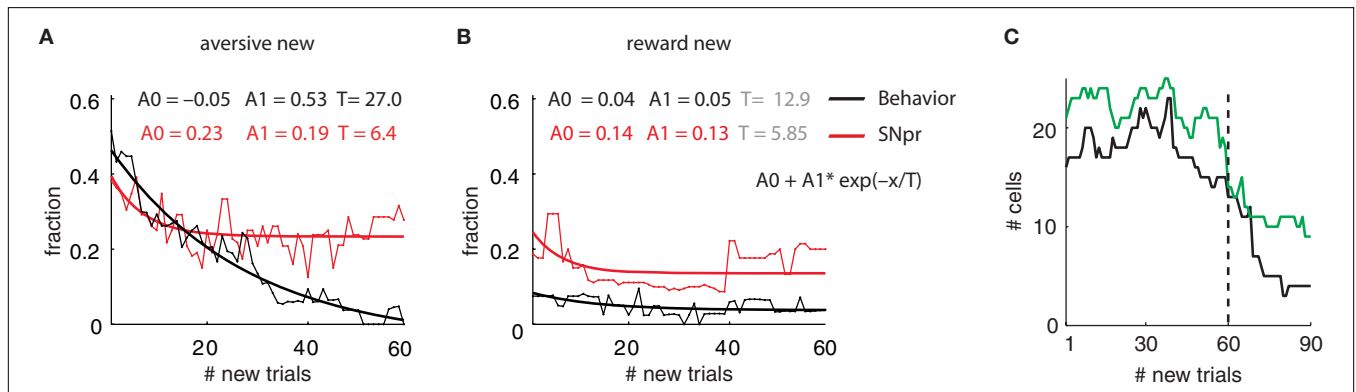
FIGURE 3 | Monkey's behavior but not SNpr neural activity achieves similarity between responses to new and familiar cues – Population analysis. (A) Fraction of SNpr cells (red) and behavioral sessions (black) which differentiate between responses to a new aversive cue and familiar aversive cues. Data were fitted to an exponential model: $Y = A_0 + A_1 \cdot \exp(-X/T)$; X = number of new trials, weighted according to the number of cells or sessions in each trial. T is the time constant of the exponential fit in units of number of new trials. The fit values appear at the top of the figure; fit values not significantly different from zero (t-test, $p > 0.05$) are in gray. **(B)** Same as **(A)** for the new reward cues. **(C)** The number of cells used in each bin for the analysis of fraction of cells. After the 60th bin (# new trials 60–79) there was a considerable reduction in the number of recorded cells and hence we limited the analysis in **(A)** and **(B)** to the first 60 bins. Note that the increases in the number of cells are due to isolation of new cells during task performance.
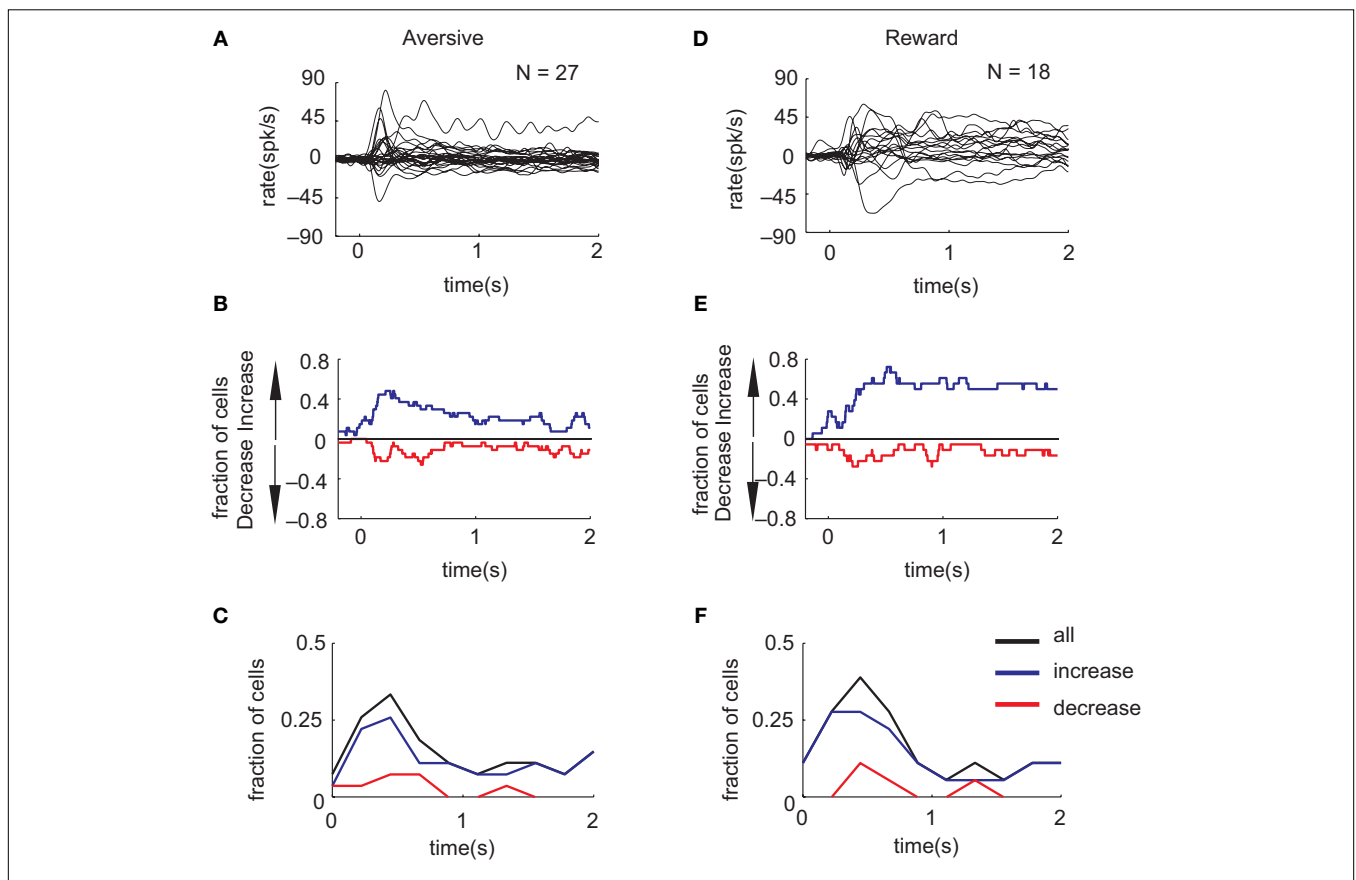
FIGURE 4 | The temporal pattern of SNpr neural responses to new cues after saturation of behavioral learning. (A) The PSTHs of the responses to the new cues predicting aversive outcome superimposed for all cells. For this analysis we used only responses after the 50th presentation of the new cue and excluded cells with fewer than 20 new trials. PSTH were smoothed with a Gaussian window with SD = 20 ms and the rate baseline was subtracted to enable comparison between responses. **(B)** The fraction of SNpr neural responses that showed a significant (2σ rule) increase (blue) or decrease (red) in their discharge rate in response to the new cue after saturation of behavioral learning. **(C)** The temporal pattern of SNpr encoding of new cues. The fraction of cells in which the response to the new aversive cue was significantly different than the response to familiar aversive events as a function of the time after cue presentation in bins of 200 ms. Black – all responses, blue/red – responses with increase/decrease of discharge rate. **(D–F)** Same as **(A–C)** for the familiar and new cues predicting reward outcome.

DISCUSSION

Reinforcement learning depends on reliable valuation of behavioral states and actions. Previous studies have shown that input and internal basal ganglia activity are related to valuation or action selection (e.g., Arkadir et al., 2004; O'Doherty et al., 2004; Lau and Glimcher, 2008). However it is not known whether activity in the output structures of the basal ganglia represents the behavioral instruction or is part of the valuation process. The connection between SNpr neurons to brainstem motor areas (Hikosaka and Wurtz, 1983c; Redgrave et al., 1992) suggests that modulations in the SNpr activity should be tightly related to motor performance. Previous work has shown that the relation between actual movement and the activity of SNpr neurons depends on the context of the movement, which includes the association of a movement with rewards (Handel and Glimcher, 2000; Sato and Hikosaka, 2002). In this manuscript, we have shown that the activity of a large fraction of SNpr neurons is even further dissociated from behavior. We found that although the behavior does not distinguish between new and familiar events, the SNpr neural activity does dissociate these events. Extensive presentation of the new cue would probably lead to similarity in the responses to the new and familiar cues; hence, we have shown that SNpr neural activity continues to change even after behavior saturates. This dissociation indicates that the basal ganglia can play a role in modulating the activity of their targets beyond instruction or gating of behavior.

Accurate valuation may not be obligatory to achieve the optimal policy (Sutton and Barto, 1998). In this study, as in many cases of behavioral learning, the learning process is not complete even after the establishment of the behavior for the new stimuli. For example learning might still be needed to better estimate the outcome probability (Bach et al., 2009). The SNpr may signal the noisier (ambiguous) estimate of the outcome probability of the novel stimuli, which would lead to downstream updating of the outcome probability. Thus, the basal ganglia output may adjust the ongoing learning process of other neural structures (thalamo-cortical or brainstem motor centers) by signaling new events. In the current experiment we did not identify the SNpr neurons according to their target neurons. Thus the differences between cells response properties found here could be reflected in their targets; e.g. cells that are dissociated from behavior project to the cortex via the thalamus and those cells that resemble behavior project to the brain stem motor areas. Finally, our results concur with other studies showing novelty encoding in the basal ganglia (Ljungberg et al., 1992; Redgrave and Gurney, 2006; Wittmann et al., 2008). These studies focused on the initial learning period. Since we only had one or two new cues per day, our task design is more suitable for exposing the long term effects of the new cue.

How does the basal ganglia network generate the patterns of activity we have observed? One hint might come from comparing the SNpr to other populations in the basal ganglia. During the same task we also recorded the activity from the GPe, GPi, TANs and the midbrain dopaminergic neurons. The analysis of these populations did not reveal the difference between behavior and activity that we found for SNpr neurons. The lack of encoding in GPe neurons raises the possibility that novelty encoding in the SNpr is due to the direct projection from the striatum to the SNpr. However axonal tracing studies have shown that neurons that project from the striatum to the SNpr have collaterals that terminate in the GPe (Levesque and Parent, 2005; Kita, 2007). This suggests that striatal novelty encoding should also be found in the GPe. Nevertheless novelty encoding in the GPe could be much weaker (and hence not detected by our methods) and the convergence of many GPe cells on SNpr might lead to amplification of the novelty signal. Similarly, we do not reject the possibility that the dopaminergic neurons are involved in generating the SNpr unique signal. Previous studies have shown that novelty is encoded in the dopaminergic neurons in the first few trials of a new task (Ljungberg et al., 1992). In the current manuscript we focused on the effects of a new cue after more than 50 trials. Although we did not find effects for the dopaminergic neurons, these cells have a very low firing rate and slight changes in their discharge rate may not be detected by our methods. Finally, the difference between GPi and SNpr is consistent with the larger modulations of the SNpr for the familiar events (Joshua et al., 2009b). Further studies are therefore needed to shed light on the sources of the responses of the SNpr neurons. In any case, the current study shows that activity in the output structures of the basal ganglia represents part of the valuation process and does not only encode the behavioral instruction.

REFERENCES

Arkadir, D., Morris, G., Vaadia, E., and Bergman, H. (2004). Independent coding of movement direction and reward prediction by single pallidal neurons. *J. Neurosci.* 24, 10047–10056.

Bach, D. R., Seymour, B., and Dolan, R. J. (2009). Neural activity associated with the passive prediction of ambiguity and risk for aversive events. *J. Neurosci.* 29, 1648–1656.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Chevalier, G., and Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci.* 13, 277–280.

DeLong, M. R., Crutcher, M. D., and Georgopoulos, A. P. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J. Neurosci.* 3, 1599–1606.

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Gulley, J. M., Kosobud, A.-E. K., and Rebec, G. V. (2002). Behavior-related modulation of substantia nigra pars reticulata neurons in rats performing a conditioned reinforcement task. *Neuroscience* 111, 337–349.

Handel, A., and Glimcher, P. W. (2000). Contextual modulation of substantia nigra pars reticulata neurons. *J. Neurophysiol.* 83, 3042–3048.

Hikosaka, O., and Wurtz, R. H. (1983a). Visual and oculomotor functions of monkey substantia nigra pars reticulata. *I. Relation of visual and auditory responses to saccades. J. Neurophysiol.* 49, 1230–1253.

Hikosaka, O., and Wurtz, R. H. (1983b). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *J. Neurophysiol.* 49, 1268–1284.

Hikosaka, O., and Wurtz, R. H. (1983c). Visual and oculomotor functions of monkey substantia nigra pars reticulata. *IV. Relation of substantia nigra to superior colliculus. J. Neurophysiol.* 49, 1285–1301.

Hollerman, J. R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J. Neurosci.* 28, 11673–11684.

Joshua, M., Adler, A., Prut, Y., Vaadia, E., and Bergman, H. (2009a). Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events. *Neuron* 62, 695–704.

Joshua, M., Adler, A., Rosin, B., Vaadia, E., and Bergman, H. (2009b). Encoding of probabilistic rewarding and aversive events by pallidal and nigral neurons. *J. Neurophysiol.* 101, 758–772.

Joshua, M., Elias, S., Levine, O., and Bergman, H. (2007). Quantifying the isolation quality of extracellularly recorded action potentials. *J. Neurosci. Methods* 163, 267–282.

Kimchi, E. Y., and Laubach, M. (2009). Dynamic encoding of action selection by the medial striatum. *J. Neurosci.* 29, 3148–3159.

Kita, H. (2007). Globus pallidus external segment. *Prog. Brain Res.* 160, 111–133.

Lau, B., and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.

Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature* 418, 413–417.

Levesque, M., and Parent, A. (2005). The striatofugal fiber system in primates: a reevaluation of its organization based on single-axon tracing studies. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11888–11893.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50, 381–425.

Mitelman, R., Joshua, M., Adler, A., and Bergman, H. (2009). A noninvasive, fast and inexpensive tool for the detection of eye open/closed state in primates. *J. Neurosci. Methods* 178, 350–356.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.

Nishino, H., Ono, T., Fukuda, M., and Sasaki, K. (1985). Monkey substantia nigra (pars reticulata) neuron discharges during operant feeding. *Brain Res.* 334, 190–193.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876.

Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.

Redgrave, P., Marrow, L., and Dean, P. (1992). Topographical organization of the nigrotectal projection in rat: evidence for segregated channels. *Neuroscience* 50, 571–595.

Reynolds, J. N., Hyland, B. I., and Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70.

Sato, M., and Hikosaka, O. (2002). Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *J. Neurosci.* 22, 2363–2373.

Schultz, W. (1986). Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory, and complex events. *J. Neurophysiol.* 55, 660–677.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851.

Sutton, R. S., and Barto, A. G. (1998). Reinforcement learning – an introduction. Cambridge, MA, The MIT Press.

Tan, C. O., and Bullock, D. (2008). A local circuit model of learned striatal and dopamine cell responses under probabilistic schedules of reward. *J. Neurosci.* 28, 10062–10074.

Wichmann, T., and Kliem, M. A. (2004). Neuronal activity in the primate substantia nigra pars reticulata during the performance of simple and memory-guided elbow movements. *J. Neurophysiol.* 91, 815–827.

Williams, Z. M., and Eskandar, E. N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat. Neurosci.* 9, 562–568.

Wittmann, B. C., Daw, N. D., Seymour, B., and Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973.