Encoding of Probabilistic Rewarding and Aversive Events by Pallidal and Nigral Neurons

Mati Joshua,1,2 Avital Adler,1,2 Boris Rosin,1 Eilon Vaadia,1,2 and Hagai Bergman1,2,3

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

Submitted 14 July 2008; accepted in final form 1 December 2008

The neural network of the basal ganglia (BG) is commonly viewed as two functionally related subsystems (e.g., Bar-Gad and Bergman 2001; Gurney et al. 2004): the neuromodulator subsystem and the main-axis subsystem. The neuromodulators (e.g., midbrain dopaminergic neurons and cholinergic tonically active interneurons of the striatum, DANs and TANs, respectively) control plasticity of the corticostriatal synapse (Calabresi et al. 2000; Reynolds et al. 2001). The main-axis subsystem includes connections between all neocortical areas, the amygdala and the hippocampus and the BG input structures, i.e., the striatum (caudate, putamen, and ventral striatum) and the subthalamic nucleus. These project both directly and indirectly to the frontal cortex (via the thalamus) and brain stem premotor nuclei (Haber and Gdowski 2004).

Previous studies on primates have shown that BG neuromodulator activity is modulated by expectation, delivery, and omission of rewards (Morris et al. 2004; Nakahara et al. 2004; Ravel et al. 2001; Schultz 1998). These data have been modeled in a reinforcement framework in which the dopamine neurons could signal prediction error (Schultz et al. 1997). Reward modulation of the main axis has mainly been studied at the level of the striatum (Apicella et al. 1992; Lau and Glimcher 2007; Lauwereyns et al. 2002; Samejima et al. 2005). Several studies have revealed discharge modulation of pallidal and SNr neurons by reward (Gdowski et al. 2001; Handel and Glimcher 2000; Pasquereau et al. 2007; Turner and Anderson 2005) and even by the probability of future reward (Arkadir et al. 2004). Nevertheless, understanding the full domain of value encoding by a neural network calls for study of neuronal responses to expectation, delivery, and omission of predicted aversive events as well. We recently reported that the responses of DANs and TANs of monkeys engaged in a probabilistic conditioning task involving both aversive and appetitive events (Joshua et al. 2008). The BG main axis may be affected by other neuromodulator systems, e.g., serotonin (Daw et al. 2002; Parent et al. 1995), and thus may have a broader encoding domain than that of the TANs and the DANs. However, there are no studies on the responses of the primate BG main-axis high-frequency discharge (HFD) neurons to expectation of deterministic or probabilistic aversive events.

We therefore used the same classical conditioning paradigm with aversive and rewarding probabilistic outcomes used in a previous study (Joshua et al. 2008) and recorded the activity of GPi, GPe, and SNr neurons in the same two monkeys that served as subjects for the recording of DANs and TANs activity. This enabled us to compare the different structures of the main axis and these structures and the main BG neuromodulators. We limited this study to the major neuronal population of these BG structures: the HFD neurons (DeLong 1971; Elias et al. 2007; Schultz 1986).

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.
tional animal care and use committee. Methods are explained in detail in a previous study (Joshua et al. 2008). Here we present a brief summary of these methods but describe in detail methods not used in the previous study.

Behavioral task

Two monkeys (L and S, *Macaca fascicularis*, female 4 kg and male 5 kg) were introduced to seven different fractal visual cues, each predicting the outcome in a probabilistic manner. Three cues (reward cues) predicted a food outcome (L: 0.4 ml, 100-ms duration; S: 0.6 ml, 150-ms duration) with delivery probabilities of 1/3, 2/3, and 1; three cues (aversive cues) predicted an airpuff outcome (100- and 150-ms duration for L and S, respectively; 50–70 psi; split and directed 2 cm from each eye; Airstim System, San Diego Instruments) with delivery probabilities of 1/3, 2/3, and 1. The seventh cue (the neutral cue) was never followed by a food or an airpuff outcome. The full-screen cues were presented on a 17-in. monitor (located 50 cm from the monkeys’ eyes) for 2 s and were immediately followed by an outcome (food, airpuff) or no outcome, according to the probabilities associated with the cue. Outcomes and outcome omissions were signaled by one of three sounds that discriminated the three possible events: a drop of food, an airpuff, or no outcome. Trials were followed by a variable intertrial interval (ITI, monkey S: 3–7 s; monkey L: 4–8 s; Fig. 1A).

Recording and data acquisition

During the acquisition of the neuronal data, two experimenters (MJ and AA) controlled the position of eight coated tungsten microelectrodes (impedance 0.2–0.8 MΩ at 1,000 Hz), and the real-time spike sorting (AlphaMap, Alpha Spike Detector, Alpha-Omega Engineering) of the eight electrodes. 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 segments of stable data were selected for further analysis. Second, cells in which >0.02 of the total interspike intervals were <2 ms were excluded from the database. Third, only units with an isolation score (Joshua et al. 2007) >0.8 were included in the database. Finally, only cells that met the above-cited inclusion criteria for >20 min during the performance of the behavioral task were included in the neural database (average 56 min and 307 trials). Table 1 provides the statistics for the cells that were included in the analysis database.

GPe neurons were identified according to their stereotaxic coordinates (based on magnetic resonance imaging [MRI] and primate atlas data) and their real-time physiological identification. These physiological parameters included the characteristic symmetric, narrow, and high-amplitude spike shape; the typical firing rate and pattern (DeLong 1971); and the neuronal activity of the striatum obtained earlier.
in the same electrode trajectory to the GPe. The GPe cells can be categorized into two subgroups (DeLong 1971): one with a high-frequency discharge rate (in this study categorized into two subgroups (DeLong 1971): one with a high-frequency discharge rate (HFD) and the other with a low-frequency discharge rate (LFD). Typically the discharge of the HFD neurons was found to be interrupted by long intervals of total silence (Elias et al. 2007) and the LFD firing pattern usually included short bursts with the amplitude of the spike declining along the burst.

Palilidar border cells (Bezard et al. 2001; DeLong 1971; Mitchell et al. 1987b) were identified by their typical regular firing pattern and broad action potentials and were excluded from the study database. Cells were also recorded from the output structures of the basal ganglia: the GPi and the SNr. Neurons of both structures were identified according to their stereotactic coordinates (based on MRI and primate atlas data) and real-time physiological recordings. For GPi neurons, the identification criteria constituted the depth of the electrode, the physiological identification of border cells between the GPe and the GPi (DeLong 1971), and the real-time assessment of the firing pattern of the cell. SNr neurons were identified according to the electrophysiological characteristics (narrow spike shape and high firing rate) of the cells (DeLong et al. 1983; Schultz 1986) and the firing characteristics of neighboring neurons and fibers (e.g., fibers of the internal capsule, SN pars compacta (SNc) dopaminergic neurons, and fibers of the oculomotor nerve).

We estimated the stereotactic coordinates of the physiological recordings within the basal ganglia nuclei by alignment of MRI scans and the primate atlas (Martin and Bowden 2000) sections. By using these anatomical and physiological criteria we attempted to sample all territories of the three studied BG nuclei.

Three computerized digital video cameras recorded the monkey’s face and upper limbs at 50 Hz. Video analysis was carried out on custom software to identify periods when the monkeys closed their eyes. Briefly, the monkey’s 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); a classification of eye states (open or closed) was made based on the number of dark pixels in the eye area. The eye state detection (ESD) 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. Mouth movements were monitored by an infrared reflection detector (Dr. Bouis Devices, Karlsruhe, Germany). The infrared signal was filtered between 1 and 100 Hz by a band-pass four-pole Butterworth filter and sampled at 1.56 kHz. Based on these recordings we detected times in which the monkeys moved their mouths by implementing a threshold-based method. We compared mouth-movement detection with the video of the monkeys’ faces over several recording days and found that they were consistent.

At the end of the experiment the chamber and head holder of both monkeys were removed, the skin was sutured, and following a recovery period the monkeys were sent to a primate sanctuary (http://monkeypark.co.il).

### Statistical analysis of population responses

Responses of the HFD neurons in the GP (Arkadir et al. 2004; Georgopoulos et al. 1983; Mink and Thach 1991b; Mitchell et al. 1987a; Turner and Anderson 2005) and SNr (Nevet et al. 2007; Sato and Hikosaka 2002) to behavioral events are composed of either increases or decreases in discharge rate. For this reason, responses of BG main-axis neurons were calculated as the absolute deviation from the baseline of the firing rate ($baseline_{FR}$) and then averaged across the population. However, this statistic does not have a natural zero baseline. To obtain such a baseline we calculated the average of the same statistic (i.e., absolute deviation from baseline) in the last 3 s of the ITI when using the same number of trials as those used for the calculation of the cell response and denoted it as $baseline_{abs}$.

First, we define $baseline_{FR}$ as

$$baseline_{FR} = \frac{1}{\text{Recorded Trials}} \sum_{t} [psth_{ITI-END}(t)]$$

Then $baseline_{abs}$ is defined as

$$baseline_{abs} = \frac{1}{\text{Recorded Trials}} \sum_{t} \text{abs}[psth_{ITI-END}(t) - baseline_{FR}]$$

Note that $baseline_{abs}$ calculates the mean fluctuations of the baseline firing rate around $baseline_{FR}$.

We then subtract this value from the response, i.e.

$$\text{response}(t) = \text{abs}[psth(t) - baseline_{FR}] - baseline_{abs}$$

The average population response was defined as the average of the responses of all units (Figs. 3A, 5A, and 7A). To validate results obtained using this statistic we divided each cell’s response into 1-ms bins with either increases or decreases in firing rate. We then averaged these responses separately across the populations. This analysis yielded the same qualitative result as the former (data not shown). In addition, we calculated the average peristimulus time histogram (PSTH) without the absolute operation (Supplemental Fig. S1). Finally, some of the neurons had sustained ITI activity after reward delivery; we analyzed the population responses to cues following trials with no reward; however, analysis yielded the same results as those of the whole population analysis (data not shown).
To determine significant responses in the single-unit 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 target PSTH and identified time segments in which the response exceeded threefold the ITI SD (3-sigma rule). A response was considered significant only if the duration of the deviant segment was >60 ms (threefold the SD of the smoothing filter).

To obtain the number of time bins in which a cell had a significant response to an event, we calculated the fraction of cells that had a significant response in each 1-ms time bin after an event. We divided these responses into increases and decreases in the firing rate and calculated the fraction of cells that increased their firing rate and the fraction of cells that decreased their firing rate during the response epoch (Figs. 3B, 5B, and 7B and Table 2).

The latency of a response was defined as the first bin in which a significant (3-sigma rule) response was detected. This conservative estimate of response latency enables comparison of the relative latencies of different neuronal populations; however, other methods (e.g., Berenyi et al. 2007; Ritov et al. 2002) might yield different estimates of the response latencies. For each population we calculated the median of the response latency and the confidence interval (CI) of this median. The CI was calculated by resampling (bootstrapping with repetitions) the latencies and recalculating the median of these surrogates. We repeated this process 1,000 times and the 95% CI was determined as the boundary values that included 95% of the median surrogates (excluding 2.5% above and 2.5% below the boundaries). Calculating the CI with bias correction gave similar results.

Statistical analysis of single-unit responses

We defined the difference index between the responses of a single cell to two events as the mean absolute difference between the corresponding PSTHs and used resampling (bootstrap) methods to test the significance of this index (Joshua et al. 2008). We calculated two difference indices: the first, the response index, measures the difference between the reward or aversive event and the neutral event. The second, the probability coding index, measured the difference between responses to the events with a high probability (P = 2/3 and 1) of receiving an outcome and responses to the events with a low probability (P = 1/3) of receiving the same outcome. We also calculated the temporal evolution of the fraction of cells with significant probability discrimination. Responses were binned in nonoverlapping 100-ms bins and tested for significance (ANOVA test, P < 0.01) at each time bin (Supplemental Fig. S2).

Note that the statistical significance of the response and probability coding index analyses depends on the number of trials. In the response index analysis we compare the reward or aversive responses to the neutral trial response; however, there are relatively fewer neutral than aversive or rewarding trials. In the probability coding index analysis we compare the high and low probabilities that are usually introduced more often than the neutral cue (threefold more for the low-probability cue and fourfold more for the high-probability cue). Due to these limitations we did not compare between the response and probability coding indices but only between the same indices when the number of trials was similar (e.g., we compared the response index for the reward and aversive trials).

The responses of most HFD pallidal and SNr neurons to the cues were sustained and thus the deviation from rate baseline in the outcome and no-outcome epochs could be the result of a slow decay from the sustained cue-related activity. We tested whether activity after the ending of the cue (average rate in 1 s) differed significantly (t-test, P < 0.05) from both the activity before the cue (0.5 s pre-cue) and from the activity at the end of the cue epoch (0.5 s before cue ending). Cells in which both of these tests were significant and activity did not fall between the pre-cue and end of the cue activity were considered to have a response that was not suspected to be due to decay of their discharge back to baseline level.

RESULTS

Monkey behavior reflected expectation of rewarding and aversive events

We recorded the monkeys’ behavior during performance of a probabilistic classical conditioning task (Fig. 1A) with food or airpuff as the rewarding and aversive outcomes, respectively. We tested how extensive conditioning (several months, 5 days/week, ~1,000 trials/day) affected the monkeys’ behavior by monitoring licking and blinking responses during neural recordings (Fig. 1, B and C).

Figure 1 shows the average frequency of blinking and licking in all trial epochs. The frequency of licking increased in response to cues predicting food but only slightly to the aversive and neutral cues (Fig. 1C). Similarly the monkeys’ frequency of blinking increased to cues predicting airpuff but only slightly to reward and neutral cues (Fig. 1B). The increase in blinking and licking during the cue epoch was maximal in trials where the probability of outcome was 2/3 or 1 and smaller in trials where the probability was 1/3. The frequency of the behavioral responses to reward and aversive events was only slightly larger for the licking versus the blinking responses. For example, at the end of the R1 (Reward, P = 1) cue presentation the monkeys increased their licking frequency from baseline by 40%, whereas in the A1 (Aversive, P = 1) cue the monkeys increased their blinking frequency by 35% (t-test, P = 0.057).

<table>
<thead>
<tr>
<th>CUE</th>
<th>POPULATION</th>
<th>REWARD INCREASE</th>
<th>REWARD DECREASE</th>
<th>AVERSIVE INCREASE</th>
<th>AVERSIVE DECREASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPc</td>
<td>17.3</td>
<td>4.0</td>
<td>5.9</td>
<td>5.7</td>
<td>12.0</td>
</tr>
<tr>
<td>GPl</td>
<td>12.0</td>
<td>3.2</td>
<td>6.5</td>
<td>4.8</td>
<td>14.5</td>
</tr>
<tr>
<td>SNr</td>
<td>25.3</td>
<td>11.1</td>
<td>13.9</td>
<td>14.5</td>
<td>23.8</td>
</tr>
</tbody>
</table>

For each time bin (1 ms) the percentage of cells that responded to a given event by a significant (3-sigma rule) increase (Inc) or decrease (Dec) in firing rate was calculated; this percentage was then averaged across all time bins of each of the three epochs. Epoch duration of 2,000 ms, starting at the event, was used for the three epochs. For example, in the GPc the average percentage of cells that responded with an increase in firing rate at the cue epoch was 17.3% (Cue–Reward–Inc–GPc entry). Note that this measure gives a smaller percentage from the overall number of responding cells since it is dependent on the fraction of cells with a significant modulation. Each neuron might make a different contribution to this average according to the number of significant response bins in the relevant epoch.
The behavioral responses to food or airpuff delivery were not dependent on their previous predictions (Fig. 1, B and C, outcome). Food and airpuff omission, as well as the final (no-outcome) event of the neutral trials, were indicated to the monkeys by a “no-outcome” sound. When expected food or airpuff was not delivered (no outcome of the \( P = 1/3 \) or \( P = 2/3 \) trials) the licking and blinking frequency increased, respectively; this increase was in line with the previously instructed probability. Licking and blinking increased slightly to the neutral trials (Fig. 1, C), no outcome, green line).

Analysis of the behavioral responses indicates that the monkeys could distinguish between aversive, reward, and neutral cues and between the high \( (P = 2/3 \) and 1) and low \( (P = 1/3 \) outcome probabilities. Accordingly, we grouped the events with high probability \( (P = 2/3 \) and \( P = 1 \) for the neural activity analysis.

**Neuronal database**

We recorded 592 GPe, 267 GPi, and 226 SNr units during the performance of the probabilistic conditioning task (Fig. 1A); out of these, 310 GPe, 149 GPi, and 145 SNr units passed the quality inclusion criteria (see METHODS) and their responses were further analyzed (Table 1).

Figure 2 shows examples of the responses of neurons form GPe, GPi, and SNr to the 18 events of our behavioral task. The GPe neuron in Fig. 2A had a large response in the reward-

**FIG. 2.** Neural activity of neurons of the globus pallidus external and internal segments (GPe and GPi, respectively) and substantia nigra pars reiculata (SNr). A: peristimulus time histograms (PSTHs) of a single GPe cell 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 airpuff. Columns are aligned according to the trial epoch. First column: cue presentation epoch (0.5 to 2 s after cue onset). Second column: outcome epoch (−0.5 to 2 s after delivery of food or airpuff). Third column: trials in which no outcome was delivered; outcome omission was signaled to the monkey by the no-outcome sound (−0.5 to 2 s after sound onset). The first 0.5 s of the 2nd and 3rd columns overlaps the last 0.5 s of the left column. Gray-level codes are marked on the middle plot (A, Aversive; N, Neutral; R, Reward; the number is the outcome probability). PSTHs were constructed by summing activity across trials in 1-ms resolution and then smoothing with a Gaussian window (SD of 20 ms). Total number of trials in this example = 511: isolation score = 0.98; fraction of spikes in first 2 ms of the interspike interval (ISI) histogram = 0.0007. B: same conventions as in A for a GPi neuron. Total number of trials = 530; isolation score = 0.99, fraction of spikes in first 2 ms of the ISI histogram = 0.0002. C: same conventions as in A for a SNr neuron. Total number of trials = 234; isolation score = 0.98, fraction of spikes in first 2 ms of the ISI histogram 0.0007. D: example of 2 raster plots from the SNr neuron in C. Top: raster of the R1 cue. Bottom: raster of the A1 cue. Vertical black arrows mark the cue onset. E: an example of the analog data (after digital 250- to 6,000-Hz band-pass filter) for a single trial (marked by a gray horizontal arrow in D). The last row contains a magnified 0.75-s segment from the 2.5-s analog segment above.
predicting cue conditions (top left) that differentiates between reward probabilities. This neuron had a small response in the aversive and neutral conditions (middle and bottom left). Discharge rate returned rapidly to baseline in the outcome and no-outcome (omission) phases (middle and right columns). Figure 2B shows a GPe neuron; with respect to the GPe example, responses of the GPI neuron were largest in the reward-cue conditions (top left). Unlike the neuron in Fig. 2A, this neuron also responded to the aversive cue (bottom left); however, this response was similar to the response to the neutral cue (middle left). Finally, the SNr neuron in Fig. 2C also responded to the reward cue and only slightly to the aversive cue (left column). However, unlike the other two neurons, this neuron responded mainly with a decrease to the reward cue. Notably this neuron had a very large response when reward was omitted (top right). To summarize, all the neurons had larger responses to the reward cues than to the aversive cue. Furthermore, reward probability, but not aversive probability, was encoded by these neurons. The neurons that did respond to the aversive cue responded similarly to the aversive and neutral cues. In the following text we provide further analyses of both the population PSTH and the single-cell responses of all recorded neurons at the three BG structures.

Activity was asymmetrically modulated by expectation of aversive events and reward in the cue epoch

Figure 3 shows the population analysis of the absolute response (deviation from the background discharge rate) of GPe, GPI, and SNr neurons to the cues. The absolute population response (see METHODS) was sustained and spanned the complete (2-s) duration of the cue epoch. The GPe and SNr population responses to reward cues were significantly larger than the responses to aversive and neutral cues (Fig. 3A). Furthermore, in the beginning of the cue epoch, responses were larger for the cues indicating a high probability of future reward than for the low-probability cue; however, this probability-dependent difference was not observed for aversive cues. Compared with the large differential modulation of the GPe and SNr, the difference in GPe population response between reward and aversive events was small and the population response did not robustly differentiate reward probabilities (Fig. 3A).

We used the absolute operator to examine the deviation of the discharge rate of the BG main axis from their baseline (ITI) discharge rate since the high-frequency tonic discharge (Table 1) of these neurons enables them to respond with both increases and decreases in their discharge rate. Absolute population analysis assumes that opposite modulations can be detected by the nervous system (for example, due to specificity in connectivity); however, this may not be the case. Thus we also performed the population analysis without using absolute operator. This analysis assumes that target structures are homogeneously innervated by neurons of the studied structure and do not keep labeled lines for individual neurons with increases or decreases in discharge rate. This standard population analysis revealed the same trends of larger responses for reward cues (Supplemental Fig. S1).

![Figure 3](http://jn.physiology.org/). Population response in the cue epoch. A: population responses (average ± SE) to the task cues. The PSTHs were calculated with the absolute operator and show the mean deviation from the background activity. Top: GPe (n = 310 neurons). Middle: GPI (n = 149). Bottom: SNr (n = 145). Color coding: dark blue, responses to high-probability (P = 1) and P = 2/3) reward cues; light blue, low-probability (P = 1/3) reward cue; green, neutral cue; orange, aversive low-probability cue; red, aversive high-probability cue. B: fraction of cells with significant (3-sigma rule) modulations of firing rate in the cue epochs. Blue, responses to all reward predicting cues; red, responses to all aversive predicting cues. Neutral events are not included because of their relatively lower number and to enable inclusion of all rewarding/aversive events in the statistical tests. The ordinate is the fraction of cells that had a significant response at each time bin (1 ms). The values above zero are the fraction of cells that significantly increased their firing rate; the values below zero are the fraction of cells that significantly decreased their firing rate.
This is probably due to the larger fraction of these neurons that responded with increases rather than with decreases in their firing rate (Fig. 3B and Table 2).

Figure 3B shows the fraction of cells that increased or decreased their rate at each time after the cue presentation. Unlike the population PSTH analysis, this analysis uses a cutoff (3-sigma rule; see METHODS) for the identification of bins with a significant deviation from the background discharge rate. In line with the population PSTH analysis, the fraction of cells that significantly modulated their firing rate in each of the 1-ms bins of the cue epoch was larger for the reward cue than for the aversive cue. This difference in the number of cells with significant responses to the reward versus the aversive cue was larger for increases in firing rate than for decreases (Fig. 3B and Table 2). Comparing the patterns of response bins with the increase versus decrease in discharge rate showed that, unlike the BG neuromodulators (Joshua et al. 2008), these opposing responses were coincident (Fig. 3B); i.e., some of the cells increased their firing rate whereas others decreased it at the same time. Finally, both the population (Fig. 3A and Supplemental Fig. S1) and the fractional analyses (Fig. 3B) showed that activity in the main axis was sustained, which contrasts with the phasic responses of the neuromodulators (Joshua et al. 2008). Note that in the population PSTH analysis we found a substantial response to the aversive cue. However, in the next sections we show that these responses were similar to the response to the neutral cue and thus do not reflect the expectation of an aversive event.

The population PSTH is an average measure and therefore may be biased by a few neurons with an extreme response and, likewise, opposite effects may be averaged out. On the other hand, the fractional analysis classifies the responding bins in a binary rather than a graded way. We therefore formulated the difference index as a measure of the modulations of a single neuron to different events. For the response index, we grouped responses across probabilities and tested whether single-cell responses to reward and aversive cues were different from their response to the neutral cue. Figure 4A shows the scatterplots comparing the response index for the reward and aversive trials. Many of the BG main-axis neurons had a significant reward and/or aversive response index (GPe: 34.8%; GPi: 33.6%; and SNr: 51.7% of the total number of recorded neurons), indicating a significant difference between these responses and the responses to the neutral cue. In all populations, the response index for the reward trials of most neurons was larger than the response index for aversive trials (Fig. 4A). A substantial fraction of the BG units showed a significant response index for reward cues, whereas

FIG. 4. Single-cell responses in the cue epoch. A: log-log scatterplots comparing the response index of individual neurons to reward and aversive cues. The response index was calculated for each cell (310 GPe, 149 GPi, and 145 SNr neurons) 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 the line represent cells with a response index that was larger for the reward cues than for aversive cues. Top: GPe; Middle: GPi; Bottom: SNr. 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 0–2,000 ms from cue presentation. Inset: pie chart of the fraction of cells with a significant index for reward (blue), aversive (red), and both (green) cues out of all cells with a significant response index (number of responding cells is given in the text at inset, top). B: log-log scatterplots comparing the probability coding of rewarding and aversive events by individual GPe, GPi, and SNr 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. Color code: blue, probability coding index significant only for reward cues; red, probability coding index significant only for aversive cues; green, both response indices were significant; gray, neither response index was significant. Points below the identity line represent cells with a probability coding index that was larger for the reward cues than for aversive cues. Inset: pie chart of the fraction of cells with a significant probability coding index for reward only (blue), for aversive only (red), and for both (green) cues out of all cells with a significant probability index.
only a small number of cells had a significant response index for aversive cues (Fig. 4A, insets).

The probability coding index compares the difference between reward and aversive probability coding. For this purpose we classified the cues into high-probability \( (P = 2/3 \text{ and } 1) \) and low-probability \( (P = 1/3) \) cues (in accordance with the monkeys’ behavior; Fig. 1). In Fig. 4B we show scatterplots of the probability coding indices of three neuronal populations. In addition to the larger reward response index (Fig. 4A), coding of the reward probability was larger (Fig. 4B) and more frequent (Fig. 4B, insets) than coding of the aversive probability in the three neuronal populations. Supplemental Fig. S2 shows the time course of the probability encoding (100-ms bins, ANOVA). In most cases, a sustained encoding is seen that is greater for the rewarding than that for the aversive trials.

In both of these difference index analyses (response index and probability coding index) the fraction of cells with a significant index was larger for SNr (51–53% of the cells) than that for GPe and GPi (32–34%; Fig. 4, inset text; \( \chi^2 \) test, \( P < 0.01 \)). The difference in the fraction of cells between the GPe and GPi was not significant (\( \chi^2 \) test, \( P = 0.78 \)).

In the outcome epoch neurons responded both to food and air puff delivery but, unlike in the cue epoch, they did not consistently encode the probability of these events

Figure 5 shows the population PSTH and fraction of responding cells for the outcome epoch. PSTH population analysis of the outcome epoch showed that all BG main-axis populations responded to both reward and aversive outcomes (Fig. 5A). Responses in this epoch to the neutral trials (i.e., when no reward or airpuff was expected) were small (Fig. 5A, green traces, and next paragraph). In the GPe and GPi the peak response was larger for the food outcome than that for the airpuff, whereas in the SNr the magnitude of the peak response to aversive and reward outcomes was similar. Unlike the population cue responses, the population responses to the outcomes that followed cues indicating different outcome probabilities were similar and the SNr population alone showed a slight difference at food delivery time (Fig. 5A). As in the cue epoch, the BG responses to the different outcomes contained both increases and decreases, with more cells increasing than decreasing their firing rate (Fig. 5B and Table 2) and the differences between the average responses to reward versus aversive outcomes were due to differences in the fraction of cells responding with increases in discharge rate (Table 2).

Figure 6 shows the response index and probability coding index analysis in the outcome epoch. The GPe and GPi responses to the reward outcome were larger and more frequent than the responses to aversive outcome (Fig. 6A, top subplots). However, many SNr cells responded to both food and airpuff outcomes (Fig. 6A, bottom). Contrary to the population analysis (Fig. 5A), many GPe, GPi, and SNr cells did in fact encode the difference between high- and low-reward probabilities (Fig. 6B). These differences between the average population and the single-unit analysis suggest that the absence of significant probability coding in the population analysis can be attributed to opposite modulation effects; i.e., some cells had a

![Figure 5](http://jn.physiology.org/)

**FIG. 5.** Population response in outcome epoch. A: population responses at the time of outcome delivery (blue, food; red, airpuff) and the response to the neutral noise in the trials when no outcome was expected (green, neutral trials). The PSTHs are calculated with the absolute operator and show mean deviation from baseline. B: fraction of cells with significant modulations of firing rate. Same conventions as in Fig. 3.
A larger response to the high probability, whereas others had a larger response to the lower-probability trials. Finally, the fraction of SNr neurons with a significant response index (84%) was greater than the corresponding fraction of GPe and GPi cells (58%; Fig. 6; $\chi^2$ test, $P < 0.01$). The fraction of SNr cells with probability coding indices (34%) was greater than the corresponding fraction of GPe and GPi cells (25 and 18%, respectively, Fig. 6B). However, this difference in fraction of cells was significant only for the GPi ($\chi^2$ test, $P < 0.01$).

Encoding of reward prediction error would predict the opposite trend in the coding of reward probability in the cue and outcome (Fiorillo et al. 2003; Morris et al. 2004). To probe this possibility we tested for correlations between the difference in response to the high and low probabilities at the cue epoch versus the difference at the outcome epoch. For the GPe and GPi we found a small positive correlation coefficient (CC 0.16 and 0.34, respectively; $t$-test, $P < 0.01$); for the SNr we found a small negative correlation that did not reach significance (CC 0.08; $P = 0.32$). Thus we conclude that HFD neurons of the main axis of the BG do not encode the prediction error.

Neural response in the no-outcome epoch

Figure 7 shows the PSTH population and fraction of responding cells for the no-outcome epoch. As in the cue epoch, population analysis in the no-outcome epoch showed that responses to reward omission trials were larger (Fig. 7A) and more frequent (Fig. 7B) than responses to aversive omission trials. As in the outcome epoch the difference between the population responses to omission of high- and low-probability outcomes was small (Fig. 7A). The population response (Fig. 7A) and the fraction of units with significant changes in their discharge rate (Fig. 7B) to outcome omission declined rapidly and reached the baseline within <1.5 s. This contrasts with the outcome responses where the response did not decline to the background (ITI) level even after 2 s (Fig. 5, A and B).

Figure 8 shows the response index and probability coding index analysis in the no-outcome epoch. This single-cell analysis shows that, comparable to the population analysis, cell responses to the reward omission were larger (Fig. 8A) and more frequent (Fig. 8B) than responses to aversive omission trials. As in the outcome epoch the difference between the population responses to omission of high- and low-probability outcomes was small (Fig. 7A). The population response (Fig. 7A) and the fraction of units with significant changes in their discharge rate (Fig. 7B) to outcome omission declined rapidly and reached the baseline within <1.5 s. This contrasts with the outcome responses where the response did not decline to the background (ITI) level even after 2 s (Fig. 5, A and B).
with a significant index (22 and 19%, respectively; $\chi^2$ test, $P < 0.05$).

Activity in the outcome and no-outcome epochs did not only reflect decay from sustained cue activity

Activity in the cue epoch is sustained and continues until the end of the cue epoch (Fig. 3 and Supplemental Fig. S1). Thus activity after the cue epoch (i.e., at outcome and no-outcome epochs) could reflect a slow decay of cue-related activity to the tonic discharge level of these neurons. For example, the response of the GPe neuron in Fig. 2A at the outcome epoch (Fig. 2A, top middle plot) could be attributed to a slow decay from cue activity. A contrasting example is the response of the SNr neuron in Fig. 2C at the no-outcome epoch (Fig. 2C, top right plot). This response cannot be attributed to a slow decay since it shows a clear increase after reward omission (no outcome).

In Fig. 9 we show the percentage of cells whose activity in the outcome/no-outcome epochs was significantly different from the precue activity and the percentage of cells from these groups in which activity did not reflect decay (see METHODS). We found that many of the responses to the reward outcome could not be attributed to a slow decay of the sustained cue activity (Fig. 9A, black bars; GPe: 28%; GPi: 22%; and SNr: 40% out of the whole population). The number of responses (that cannot be attributed to decay of cue activity) to aversive outcome was smaller than the number of responses to reward outcome (Fig. 9A, gray bars; GPe: 4%; GPi: 6%; and SNr: 20%). Very few GPe and GPi cells responded to reward omission itself (Fig. 9B, black bars; GPe: 9%, GPi: 6%); however, in the SNr a larger fraction of cells responded (decay excluded) to reward omission (Fig. 9B, black bar; SNr: 21%). In all the structures the number of cells that responded (decay excluded) to aversive omission was smaller than the fraction of cells that responded to reward omission (Fig. 9B, white bar; GPe: 1%; GPi: 1%; and SNr: 8%).

In summary, we found that activity in the outcome/no-outcome epoch did not only reflect the decay from sustained cue-related activity and that BG HFD cells clearly encode outcome and no-outcome events. Note that the fraction of cells of which we could rule out the possibility of decay from sustained activity is a lower limit of the actual number of responding cells. This is because our method for testing the null hypothesis—that activity is not due to decay—is very conservative (i.e., the discharge at the outcome or the no-outcome epoch may fall between the ITI and the end of cue discharge level and still reflect a valid response to the outcome or no-outcome events). Other methods that include interpolation of the whole temporal pattern of the response may report a larger number of responding cells to the outcome and no-outcome events.

SNr neurons responded with shorter latencies than those of GPe and GPi neurons

Figure 10 shows the analysis of the response latency to the reward and aversive cues. The latency of SNr responses was significantly shorter than the responses of the GPe and GPi
No difference between the GPe and GPi populations was found ($P > 0.93$). We grouped the responses to reward and aversive cues and the increase and decrease responses since we did not find any significant difference between these parameters (Fig. 10, B and C). Although not significant, the GPi decrease response tended to be earlier than the increase response (Turner and Anderson 1997).

We did see similar trends in the responses in the outcome and the no-outcome epochs. However, the persistent but nevertheless nonsteady activity of the BG neurons during the cue response (Fig. 3A) prevented us from establishing a reliable baseline for testing the outcome epoch responses and thus we carried out only the latency analysis for the cue epoch.

**DISCUSSION**

In this report we extended our previous study (Joshua et al. 2008) to the study of the responses of BG main-axis HFD neurons to expectation, delivery, and omission of appetitive (food), aversive (airpuff), and neutral (sound only) events. We found that the responses of GPe, GPi, and SNr neurons were longer in duration and less stereotypic than the responses of the main BG neuromodulators (TANs and DANs). As with the TANs and DANs, the responses of the BG...
main-axis neurons were larger and usually encoded reward better than aversive-related events. We found substantial differences between the three populations of BG main-axis neurons. Most notably, SNr responses were more frequent, had shorter latencies, and encoded the airpuff delivery better than the corresponding responses of GPe and GPi neurons.

Neural responses were larger for the reward than for the aversive trials

We found preferential activation to reward versus aversive events. One possible explanation for this asymmetric neural activity is that the asymmetry arises from differences in the relative value of the rewarding/aversive stimuli that we used. An alternative possibility is that the encoding of reward/aversive expectation is asymmetric in the BG. We find the second possibility more likely since the population responses to the aversive predicting cue and to the neutral cue were remarkably similar (Fig. 3) and very few cells encoded the cue predicting airpuff (Fig. 4). It could be argued that the monkeys ignored the air puff; we have shown this is not the case since there were large behavioral responses to cues predicting the airpuff (Fig. 1). In a previous experiment (Mirenowicz and Schultz 1996), in which the subjective values were calibrated, the authors compared a reward of 0.15 ml of juice and an aversive 28- to 58-psi airpuff directed to the hand. Similar airpuff intensities have been used in other studies comparing the responses of amygdala neurons (40- to 60-psi airpuff vs. 0.1–0.9 ml of liquid food) (Belova et al. 2007; Paton et al. 2006) and lateral prefrontal cortex neurons (29-psi airpuff, 10 cm from the monkey’s face) (Kobayashi et al. 2006) to both rewarding and aversive events. The airpuff in the current experiment was larger (50–70 psi) and delivered 2 cm from the monkey’s eyes. Thus this larger and closer airpuff must have had a negative subjective value. We further discuss the possibility of asymmetric encoding in the following text.

Preferential control over reward-related behavior

We have shown that just before the end of the cue, the fraction of trials in which the monkey licked in expectation of future reward and the fraction of trials in which the monkey blinked in expectation of future airpuff were similar in magnitude. In addition we found a large blinking response even when the airpuff was omitted (Fig. 1C). Finally, with the exception of the outcome epoch, the licking and the blinking behaviors reflected the expected (low vs. high) probability of the reward and the aversive events. Nevertheless, the BG single-cell activity was found to be biased toward the encoding of reward-related events and encoding of aversive events was very weak. This difference in activity may be compensated by the difference in synaptic connections between the BG and their targets; however, such differences have yet to be described. Several studies have used similar paradigms to compare neural responses to reward food and aversive airpuff (Kobayashi et al. 2006; Mirenowicz and Schultz 1996; Paton et al. 2006). Paton et al. (2006) showed that in the amygdala, expectations of food and airpuff are represented symmetrically. Our research shows that, in contrast to the amygdala, food and airpuff expectations are represented asymmetrically in the basal ganglia. Thus we found comparable aversive- and reward-related behaviors; however, whereas the activity in the basal ganglia strongly reflects reward behavior and encodes probability, aversive-related events and their probability are only weakly encoded in basal ganglia activity.

Although we found similarity in the behavioral responses (Fig. 1, B and C), in this study we did not calibrate the subjective value (utility) of food versus airpuff; however, we did manipulate the expectation of aversive outcome. In previous instrumental conditioning experiments, including both reward and aversive events, the monkey could avoid the aversive airpuff by a correct response (Mirenowicz and Schultz 1996; Yamada et al. 2004, 2007). In the current experiment the airpuff was unavoidable and thus the aversive cue led to direct expectation of aversion.

In a previous study (Joshua et al. 2008), we reported that the responses of midbrain DANs and striatal TANs (of the same monkeys engaged in the same behavioral task) are biased toward the encoding of rewarding events. The BG main axis is affected by additional neuromodulator systems, e.g., serotonin (Lavoie and Parent 1990). Theoretical studies have suggested that the phasic serotonin signal might report the prediction error for future punishment (Daw et al. 2002; Dayan and Huys 2008) and therefore could compensate for the biased encoding of the value domain by the TANs and the DANs. The current study of the BG output structures indicates that the BG main-axis neurons have a bias toward control of reward-related behavior similar to that of TANs and DANs. Thus even if there are BG modulators other than the cholinergic and dopaminer-
gic striatal inputs, the activity of BG output neurons follows the same trend as that of the TANs and DANs and is biased toward rewarding events. We therefore suggest that the other modulators do not extend the basal ganglia encoding to aversive events and that there are neuronal systems other than the BG that have control over aversive-related behavior.

**BG main-axis responses were long-lasting and diverse**

In contrast to the short (<0.7 s) responses of the BG modulators (Apicella 2007; Joshua et al. 2008; Morris et al. 2004; Schultz 1998), the responses of the BG main-axis HFD neurons lasted throughout the 2-s-cue epoch. This is in line with previous descriptions of pallidal (Arkarid et al. 2004) and SNr (Wichmann and Kliem 2004) responses. Long-duration, set-related responses have frequently been described in the cortex (Fuster 1999; Miyashita 1988; Wise and Kurata 1989), where they have been attributed to short-term memory or action-preparation processes. We cannot rule out similar processes in the basal ganglia and the experimental design does not allow us to dissociate set-related versus cue-evoked responses. However, the encoding of probability by the BG main-axis neurons (Figs. 3 and 4) and the dissociation between actions and neural response (for example, no neural encoding of the probability of aversive trials, the early decay of the neural activity compared with licking behavior after reward delivery) suggests that the activity of these neurons may encode the value of the current state or state–action pairs (Lau and Glimcher 2007; Samejima et al. 2005).

The tonic discharge rate of the HFD neurons (population average: 45.1–88.1 spikes/s in this study) endows them with a better dynamic range for responses with a decrease in discharge rate. Nevertheless, consistent with many previous studies (Georgopoulos et al. 1983; Mink and Thach 1991a; Mitchell et al. 1987a; Turner and Anderson 1997) we found that the BG HFD neurons respond to behavioral events more frequently with increases than with decreases in discharge rate. The latencies and the temporal distribution of the responses with increases and decreases in discharge rate were similar (Figs. 3B, 5B, 7B, and 10C), thus leading to highly diverse BG encoding, with different polarities and different amplitudes of the responses. The differences between the population responses with no encoding of the a priori probability of outcome (Fig. 5) versus the single-unit encoding of this probability (Fig. 6) are in line with such a balanced diversity of the responses of BG single units. These diverse responses augment the information capacity of the BG output structure (Bar-Gad et al. 2003).

**Different response characteristics of the main-axis nuclei**

In this study we found several major differences between the GPe, GPi, and the SNr. We found more intense changes in the responses of the SNr compared with the responses of the GPe and the GPi. SNr neurons responded with shorter latencies to the cue (Fig. 10A) and encode the airpuff outcome better than the pallidal neurons (Figs. 5 and 6). A simple explanation for the enhanced encoding is the orofacial (licking and blinking) motor behavior of the monkeys in this experiment. Initial studies emphasized the role of the SNr in the control of orofacial movements (DeLong et al. 1983; Hikosaka and Wurtz 1983). Although this separation is not clear-cut (DeLong et al. 1985; Wichmann and Kliem 2004) our results may reflect this organization. Thus the small and less-frequent responses in the GPi could reflect the smaller representation of orofacial movements in the GPi. This could be also the reason for the activation of the SNr to aversive events, but as noted earlier this does not explain the asymmetric value representation in the SNr.

At the circuitry level, one possibility is that the origins of the difference in pallidal versus SNr responses could be a result of different projections from the striatum or the subthalamic nucleus (Huber and Gdowski 2004). Another possibility is that the GPe has different pathways to the GPi and SNr and those GPe neurons that do project to the SNr are the neurons with the short latency and larger response. Nevertheless, we did not find any topographic organization in the responses of the GPe that supports this hypothesis (data not shown). Finally, another putative explanation for the differences between the GPi and the SNr is the direct effects of somatodendritic release of dopamine on SNr, but not on pallidal, neurons. The similar latencies of SNc and SNr responses support the hypothesis that SNc neurons may drive SNr responses by somatodendritic release of dopamine (Cragg et al. 2001; Windels and Kiyatkin 2006).

Finally, the neural recordings were made after the monkey was highly familiar with the task and thus activity might not be the same as activity that occurs during learning. Previous studies of dopaminergic neurons have shown that activity in a familiar probabilistic task does resemble the activity in a learning task (Fiorillo et al. 2003; Hollerman and Schultz 1998; Morris et al. 2004). A functional MRI study has shown that striatal activity underlies novelty-based choice in humans (Wittmann et al. 2008). Whether this is the case for other basal ganglia populations and the single-cell activity that underlies novelty representation should be investigated by future studies.

**Concluding remarks**

In this study we extend our previous work on BG neuro-modulators (Joshua et al. 2008). We found a similar bias of GPe, GPi, SNr, TANs, and DANs for the encoding of expectation of rewarding versus aversive events. Thus the BG main axis may mainly reflect the teaching message (and corticostriatal plasticity control) of the TANs and DANs and may not be significantly affected by additional modulators with broader or different messages. Our results show a complex and different encoding by GPe, GPi, and SNr neurons. Moreover, they indicate a different encoding by GPi and SNr neurons and therefore suggest that there are many functional differences between these two BG output nuclei, despite their similar biochemical and physiological characteristics. Future models and studies of the computational physiology of the basal ganglia and their disorders should therefore attempt to disentangle the different functions of GPi and SNr.

**Acknowledgments**

We thank Dr. Bryon Gomberg for MRI; M. Levi and M. Rivlin for help in preparing the experimental setup; Y. Renernt and I. Finkes for monkey training and general assistance; and G. Schoenbaum, Y. Shaham, and G. Morris for critical reading of earlier versions of this manuscript.

Downloaded by 10.220.33.1 on September 13, 2017 from http://jn.physiology.org/
GRANTS

This study was partly supported by a “Fighting against Parkinson” Grant from the Hebrew University Netherlands Association and a Max Vorst Family Foundation grant.

REFERENCES


