Neuron, Vol. 43, 133–143, July 8, 2004, Copyright ©2004 by Cell Press

Coincident but Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons

Genela Morris,[1,2,*] David Arkadir,[1] Alon Nevet,[1] Eilon Vaadia,[1,2] and Hagai Bergman[1,2]
[1]Department of Physiology
Hadassah Medical School
[2]The Interdisciplinary Center for Neural Computation
The Hebrew University
Jerusalem 91120
Israel

Summary

Midbrain dopamine and striatal tonically active neurons (TANs, presumed acetylcholine interneurons) signal behavioral significance of environmental events. Since striatal dopamine and acetylcholine affect plasticity of cortico-striatal transmission and are both crucial to learning, they may serve as teachers in the basal ganglia circuits. We recorded from both neuronal populations in monkeys performing a probabilistic instrumental conditioning task. Both neuronal types respond robustly to reward-related events. Although different events yielded responses with different latencies, the responses of the two populations coincided, indicating integration at the target level. Yet, while the dopamine neurons' response reflects mismatch between expectation and outcome in the positive domain, the TANs are invariant to reward predictability. Finally, TAN pairs are synchronized, compared to a minority of dopamine neuron pairs. We conclude that the striatal cholinergic and dopaminergic systems carry distinct messages by different means, which can be integrated differently to shape the basal ganglia responses to reward-related events.

Introduction

Striatal dopamine and acetylcholine are intertwined in anatomy, physiology, and pathology. The striatum, the primary input stage of the basal ganglia, displays the densest staining in the central nervous system for both dopaminergic (DA) (Lavoie et al., 1989; Jones et al., 2001) and cholinergic (ACh) (Holt et al., 1997) markers. Effective treatments of Parkinson's disease involve manipulation of DA and ACh in an opposing fashion, either by elevation of the extracellular level of striatal dopamine or alternatively by reduction of striatal acetylcholine (Lang and Lees, 2002; Pisani et al., 2003). These observations have led investigators of the basal ganglia to postulate the DA/ACh balance hypothesis, which states that the two transmitters act antagonistically in the striatum (Barbeau, 1962; Nisenbaum and Kitai, 1995). This is presumably achieved by reciprocal inhibition of striatal acetylcholine release by the nigro-striatal DA pathway (Pisani et al., 2000, 2003; Drukarch et al., 1989; DeBoer and Abercrombie, 1996) and of DA release by ACh (Kudernatsch and Sutor, 1994), in addition to their opposing effects on the excitability of striatal projection neurons (Zhou et al., 2003).

Striatal acetylcholine, commonly believed to be secreted by local tonically active neurons (TANs) (Wilson et al., 1990; Bennett and Wilson, 1999; Aosaki et al., 1995), and striatal dopamine, released by midbrain neurons, both play a crucial role in the control of motivation and learning. Deficit in either substance has been shown to disrupt reward-related procedural learning processes (Kitabatake et al., 2003; Matsumoto et al., 1999; Knowlton et al., 1996). On the cellular level, DA and ACh play a crucial role in cortico-striatal plasticity (Reynolds et al., 2001; Calabresi et al., 1998, 1999, 2000; Centonze et al., 1999a, 1999b, 2003). When presented with an unpredicted reward or with stimuli that predict reward, midbrain dopaminergic neurons (Hollerman and Schultz, 1998; Schultz et al., 1997; Waelti et al., 2001) and TANs (Apicella et al., 1998; Shimo and Hikosaka, 2001; Graybiel et al., 1994; Blazquez et al., 2002) display stereotypical responses consisting of a phasic deviation from their tonic firing rate. In accordance with the DA/ACh balance hypothesis, these typical responses seem to be opposites, in that the dopaminergic neurons elevate their firing, whereas the TANs' firing is mainly decreased. The dopaminergic response has been recently interpreted as an error signal that informs the cortico-striatal system of the discrepancy between the prediction of a reward and its actual occurrence (Schultz et al., 1997; but see Redgrave et al., 1999; Horvitz, 2000). This hypothesis is congruent with the computational temporal difference (TD) model for reinforcement learning (Sutton and Barto, 1998; Suri and Schultz, 2001). Striatal TANs display a similar activity pattern, shifting their response to the earliest stimuli predicting future rewards (Aosaki et al., 1994; Apicella et al., 1997; Ravel et al., 2001; Shimo and Hikosaka, 2001), suggesting that they may act in the same manner. A monotonic relationship between the degree of predictability and neuronal responses has recently been described for midbrain DA neurons in monkeys performing a classical conditioning task (Fiorillo et al., 2003). However, the responses of the striatal TANs were never tested in a formal probabilistic task.

Despite these similarities, it is unlikely that the information conveyed by responses of two populations affecting a single target is redundant. To examine the role of striatal TANs and DA neurons in a task involving association of environmental input to motor output, we devised a probabilistic instrumental conditioning task (Figure 1A) that enabled us to manipulate the degree of reward predictability. In each trial, one of a set of 4–5 visual cues was briefly presented to monkeys in one of two possible locations on a computer screen. After a constant delay, a go signal instructed the monkeys to indicate the cued location by pressing one of two keys. Correct performance was rewarded in a probabilistic manner, depending on the preceding visual cue (Figure 1B). We recorded the single unit activity of TANs in the putamen nucleus of the striatum and DA neurons in the substantia nigra pars compacta (SNc) (Figure 1C) from

*Correspondence: genela@md.huji.ac.il

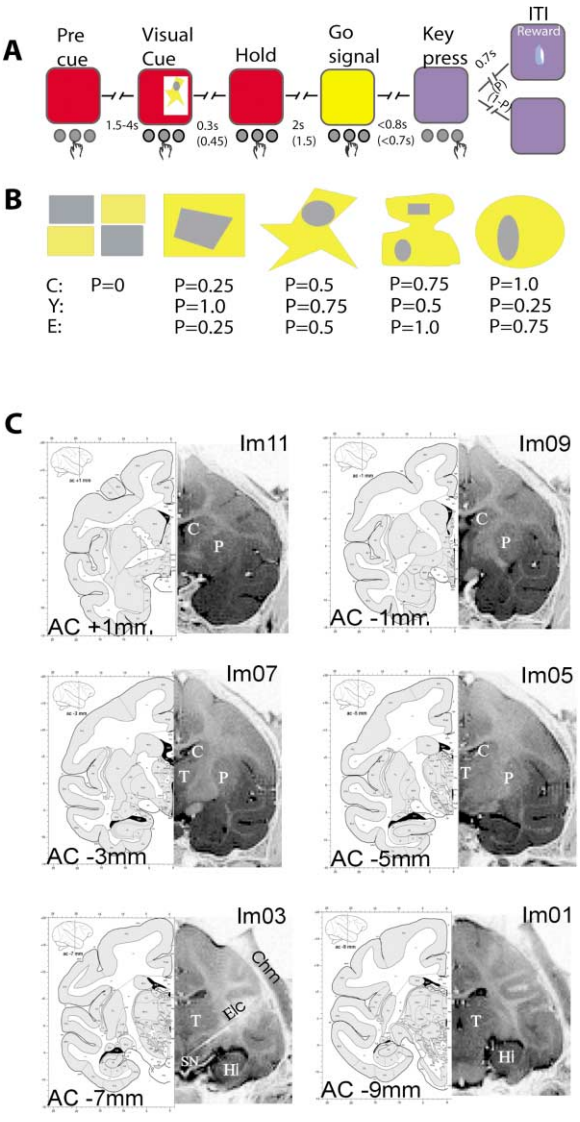

Figure 1. Behavioral Paradigm and Recording Site

(A) Trial flow. Beneath the stage definitions, screen displays are depicted with the corresponding time periods. Numbers in parentheses indicate cases in which the conditions differed between monkeys. The circles at the bottom of the panel represent the three keys, and the hand represents the desired action of the monkey.

(B) Visual cues. The cues are presented with the associated probability of reward (conditioned on correct task performance) for the three monkeys.

(C) SNc recording coordinates. Coronal MRI images taken from monkey C, numbered with respect to anterior commissure, with a tungsten microelectrode inserted at the SNc level. Abbreviations: Chm, recording chamber; C, caudate; P, putamen; AC, anterior commissure; T, thalamus; Elc, electrode; S, substantia nigra. The MRI sections shown are of the left hemisphere for Im11 to Im05 and of the right hemisphere for Im03 and Im01 to enhance the dynamic range of the MRI gray levels for those sections without the SN electrode. The black/white coloring has been inverted for easier comparison with the atlas. The identification of the brain structures are based on this alignment.

three monkeys (Y, E, and C) performing this task. Only correct trials were used for analysis (see Experimental Procedures for a detailed description of recording sites and technique). Thus, the experimental design allowed us to address the question of encoding of reward predictability in conditions involving the mapping of sensory information to action and at the same time to differentiate the two neuronal populations under identical conditions.

Results

Behavior

The monkeys were trained to the point that their behavioral responses were independent of trial condition, despite the fact that the different visual cues were associated with reward at different predictability levels. This enabled us to rule out differences in neuronal activity due to kinematical, behavioral, or motivational differences. This control was particularly vital in this study, since TANs have been shown to respond differentially according to the probability of behavioral response (Blazquez et al., 2002) and DA neurons responses have recently been shown to correlate with reaction time (Satoh et al., 2003) as well as due to the motor nature of the putamen (Crutcher and DeLong, 1984; Alexander and DeLong, 1985; Lee and Assad, 2003). To reduce motor variability, we restricted the monkeys' allowed response times to 700–800 ms, which was almost at the limit of their ability. Reaction times and movement times in all recorded trials (for all monkeys) are plotted in Figures 2A and 2B, respectively. The distribution of these parameters was independent of reward probability in each of the three monkeys ($p > 0.4$, one-way ANOVA and Kruskal-Wallis nonparametric ANOVA). Note, however, that the strict constraint on the response time imposed a regime of time pressure for performance of the correct response. In such cases it has been shown both theoretically and experimentally (Reddi and Carpenter, 2000; Carpenter and Williams, 1995; Roitman and Shadlen, 2002; Goodie and Crooks, 2004) that performance is suboptimal, similar to the speed/accuracy tradeoff effect in motor performance. For our purposes, it was important that this reduction in performance should remain independent of trial condition. Indeed, the percentage of correct choices (Figure 2C), as well as self-aborted trials (i.e., trial break error and response omission error [pooled together in Figure 2D]), was invariant across all conditions (one-way ANOVA, $p > 0.7$; Kruskal-Wallis nonparametric ANOVA, $p > 0.4$).

Since no behavioral differences were permitted, a different measure was required to ensure that the monkeys learned the probabilities associated with the different cues and were able to utilize these to predict an upcoming reward correctly. To this end, we introduced within each recording session several probe trials (5%–10% of total trials, randomly interleaved between the single cue trials) in which two visual cues were presented simultaneously at both positions. In these trials pressing either key could yield a reward whose probability depended on the key that was pressed. Figure 2E depicts the monkey's choices. The color matrix covers all combinations presented to monkey Y. Note the clear gradient of left key preference from the upper right corner ($P_{left} = 1.0$, $P_{right} = 0.25$) to the lower left corner ($P_{left} = 0.25$, $P_{right} = 1.0$) and the lack of preference at the diago-

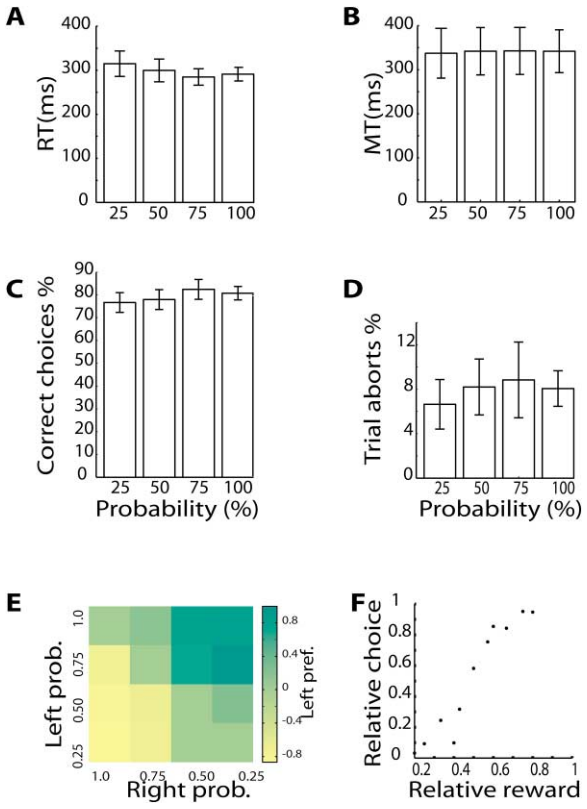

Figure 2. Behavioral Results

(A) Reaction times (RT) according to reward probability (mean ± SEM).

(B) Movement times (MT) according to reward probability (mean ± SEM).

(C) Percent correct choices, according to reward probability (mean ± SEM).

(D) Percent self-aborted trials, consisting of errors due to early release of the central key or to failure to release the central key.

(E) Left key preference in double-cued trials. For each combination of probabilities, the left key preference was quantified as the relative excess of the left responses

$$\frac{R_{left} - R_{right}}{R_{left} + R_{right}}$$

and is color coded (monkey Y).

(F) Probability matching in monkeys' preferences. Relative choice of right key

$$\frac{R_{right}}{R_{right} + R_{left}},$$

as a function of the associated relative reward probability

$$\frac{P_{right}}{P_{right} + P_{left}}.$$

nal $P_{left} = P_{right}$. Similar results were obtained for monkeys E and C. The results for all the monkeys pooled together are summarized by the relative rate of the monkeys' choices, as a function of the ratio of the relative reinforcement probability. The figure shows approximate probability matching (with a slight overshoot), indicating that the monkeys had indeed learned the probabilities associated with each cue. Note that this behavior is suboptimal. Interestingly, this behavior is generally observed in adult human subjects (Vulkan, 2000; Wolford et al., 2000) and only rarely in rats and pigeons, which are known to maximize utility (Herrnstein, 1970; Baum, 1979).

Neuronal Responses to Behavioral Events

We recorded 132 (79, 19, and 34) DA cells and 97 (17, 35, and 45) TANs from monkeys E, C, and Y, respectively, during task performance. We report only results obtained during correctly performed trials. Trials in which two visual cues were presented simultaneously were not included in the current analysis. A total of 114 DA neurons and 93 TANs showed significant responses (Mann-Whitney, $p < 0.05$) to at least one of the three reward-related events (visual cue, reward, and reward omission) and were selected for further analysis. Figure 3 depicts the representative responses of one TAN (Figure 3A) and one DA neuron (Figure 3B). Each row represents one type of trial, classified according to the probability that a reward would follow a correct response. All three reward-related events are distinctly represented in the activity of both neurons. However, whereas the DA neuron responds oppositely to reward versus its omission (Hollerman and Schultz, 1998; Satoh et al., 2003), the TAN responded to these opposing events with the same polarity, although the magnitude of response to omission was smaller. As in this example, in all cases of TAN responses to reward omission, the gross features of the reward omission response were similar to those following the visual cue and the reward.

The TAN's response to the three behavioral events differs in terms of latency, as well as magnitude of the surrounding excitation. However, within each event, the similarity between the responses for all reward probability conditions is high. In striking contrast, the DA neuron's responses (Figure 3B) clearly differentiate between the various reward probability conditions, reaching reversal of the response to reward omission (Hollerman and Schultz, 1998). This response is in line with the previous finding obtained in classical conditioning (Fiorillo et al., 2003) and the TD signal hypothesis for DA neurons, predicting that the cue response will increase with its rewarding value (i.e., with the increase in cue-associated reward probability) and the reward response will decrease with its predictability (increase in probability).

For more elaborate quantitative analysis, cell responses to behavioral events were parameterized as the difference in average firing rate at the 400 ms following the event compared to the preceding 400 ms. For the TANs, this procedure was preceded by half-wave rectification of the response (insets in Figure 3). A large proportion of DA neurons (102, 110, and 74) and TANs (76, 80, and 48) displayed significant changes in discharge rate and pattern following the visual cue, reward, and reward omission, respectively (Mann-Whitney, $p < 0.05$). Most DA neurons and approximately half of the TANs responded to all three events. The pattern shown in Figure 3, whereby DA neurons hold specific information regarding reward expectation, along with the present reality, whereby TANs provide general information regarding a potentially significant event, was consistent

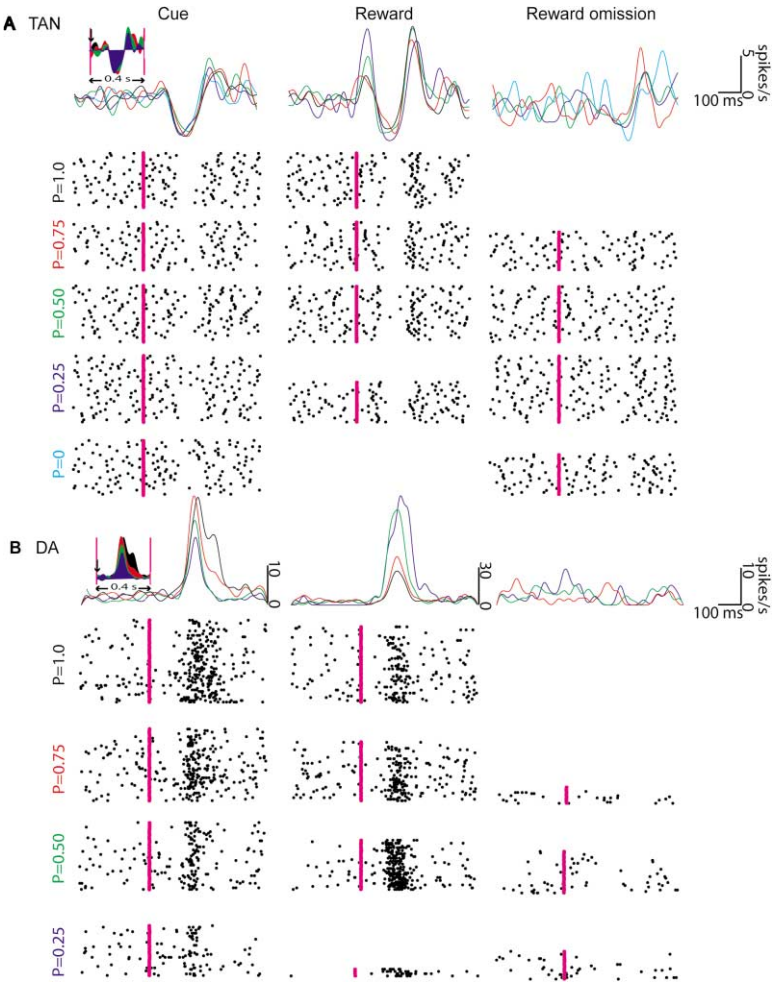

Figure 3. Neuronal Responses to Behavioral Events

(A) Example of TAN. Responses to visual cue (left), reward (middle), and reward omission (right) in correctly performed trials. Bottom: raster displays for each event (columns) divided into blocks of different probabilities (rows). Top: mean firing rates aligned on the behavioral events (peristimulus time histogram, PSTH). Bin size = 1 ms. PSTHs are smoothed with a Gaussian window, $\sigma = 10$ ms. The responses for the different reward probabilities are superposed. Line colors correspond to the probabilities indicated at the sides of the raster displays. For illustration purposes, in conditions with over 35 trials, only a subset of the first 35 trials is shown in the raster plots. Inset: TAN responses were quantified as the half wave rectified mean firing rate in 400 ms period following behavioral event. The colored area represents the number of excess emitted/omitted spikes in that time period.

(B) Example of DA neuron. Same conventions as in (A). Inset: DA neuron responses were quantified as the mean change firing rate in 400 ms period following the behavioral event. The colored area represents the number of emitted spikes in that time period compared to the background firing.

for neurons of both populations. Statistical examination of all neurons revealed that while the TAN responses were not significantly different when comparing trials of different probabilities (Kruskal-Wallis nonparametric ANOVA, $p > 0.3$ in all cases), the observed DA response pattern differed significantly across the different probability conditions ($p < 0.001$ for visual cue and reward). The differences between the responses to reward omission in the different probability conditions were found to be nonsignificant ($p > 0.2$). Neither type of neuron exhibited sustained change from baseline activity in response to (or preceding) any behavioral event. Neuronal activity preceding the reward did not depend on trial condition ($p > 0.8$), indicating that uncertainty level (Fiorillo et al., 2003) did not affect firing at this period.

Population Responses to Behavioral Events

Each of the three behavioral events (cue, reward, and reward omission) elicited responses that appeared to be characteristic of the vast majority of neuronal responses within every population, with highly similar latencies, patterns, and polarity of response, as well as a relatively constant tonic (background) firing rate. Figure 4 shows mean population responses to the visual cue (left), reward (middle), and reward omission (right) of all recorded DA neurons (Figure 4A) and TANs (Figure 4B), classified according to the reward probability. On the population level, not only did the DA neurons respond differentially to the various reward conditions, but these responses were graded in agreement with the TD hypothesis (Figure 5A): as the probability of reward increased, the cue responses became larger and the reward responses became smaller. No significant differences were observed in the reward omission responses across the different trial types. As in the case of single neurons, the population averages did not display sustained responses at any time.

In contrast to the DA case, the TANs did not respond significantly differently with respect to different probabilities of reward, nor did they follow any consistent trend, either at the single neuron level or as a population (Figure 4B). To quantify this apparent difference between the DA and TAN response, we conducted a linear regression analysis of the mean changes in firing rate during responses to the visual cue, reward, and reward omission in relation to the different reward probabilities. The results are plotted in Figure 5A, showing that the DA response to the visual cue and reward is highly correlated with the probability of reward, in sharp contrast to the TAN response. Since the TAN response consists of several distinct phases, which may or may not occur (initial rise, pause, and second rise), we conducted a

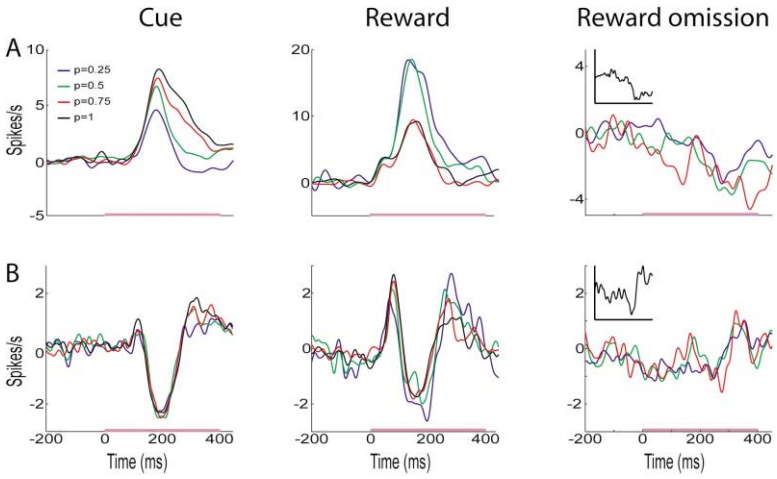

Figure 4. DA and TAN Population Responses

(A) DA population. Averages of mean firing rates in response to behavioral events. The responses are illustrated as superposed PSTHs aligned to the time of the event, for reward probabilities 0.25, 0.5, 0.75, and 1.0. The left panel shows responses to visual cue, the middle panel shows responses to reward, and the right panel shows responses to reward omission. Population PSTHs were calculated with 1 ms bins and smoothed with a Gaussian window, $\sigma = 6$ ms. Magenta bar: time period for response quantification. Inset: pooled population response (over probabilities) to reward omission of responding cells.

(B) TAN population. Same conventions as in (A).

phase by phase analysis of the response (Figure 6). The phases were identified as the time of the first peak, the trough, and the second peak in the average responses of the TANs to the cue and reward. As seen in the figure, in all three phases the response did not differ significantly across the reward probabilities ($p > 0.1$, one-way ANOVA).

Note that the DA neurons' response to reward following a $p = 1.0$ cue was not null. This finding may be reconciled with the TD models through the difference between probabilistic instrumental conditioning (as in this task) and classical conditioning. In our task, the probability of reward was conditional on correct performance (which is imperfect in the strict time constraints imposed in this study). The effective probabilities of reward are therefore reduced by the performance factor. Since performance was similar in all conditions (Figures 2C and 2D), the general trend should be unaffected. To show the dependence of the neuronal responses on the effective probabilities of reward, we computed (Figure 5B) these probabilities using the average percent correct choices by each monkey as the a priori probability. Since these varied somewhat between the monkeys, the regression analysis now has 12 rather than 4 points per event and cell type. Indeed, this figure shows that while maintaining the linearity with regard to reward probability shown in Figure 5A, the DA response is closer to null when the prediction error, normalized by the effective probability, is extrapolated to zero.

Response Latency

The response latencies of DA neurons as well as that of the TANs varied depending on events: response to reward was usually fastest, while the visual cue yielded slower responses, and those for reward omission were even slower. However, when comparing the two types of neurons, we found that the responses of both populations to each event coincide. Figure 7A shows the covariation matrix of a pair of simultaneously recorded TAN and DA neuron, triggered on the visual cue (arrowhead). Here, only bins in which the two neurons respond (with a consistent lag, at the single trial level) are expected to display significant covariation. This covariation should be positive if the two responses are of identical polarity and negative for opposing polarities. The matrix diagonal (denoting zero lag or coincidence) and the ap-

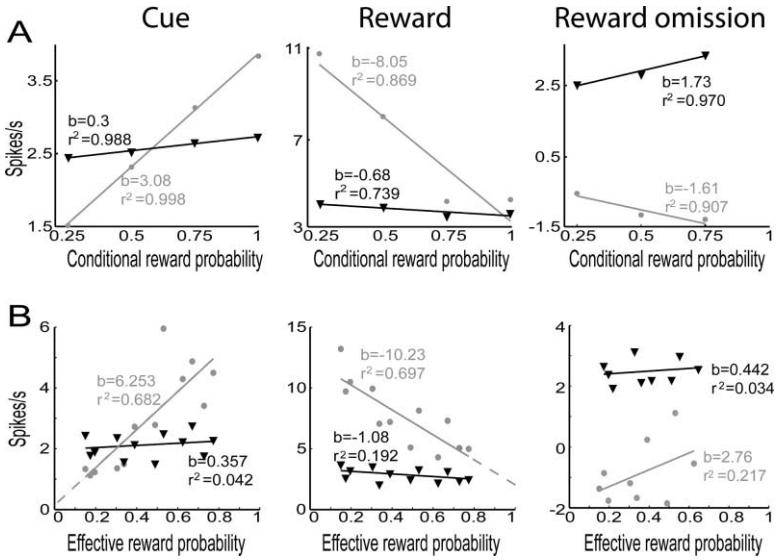

Figure 5. Linear Regression Analysis of DA and TAN Responses

(A) Response as a function of reward probability, reflecting the conditional probability *P*(*reward*|*correct performance*). Linear regression plots of mean spike rates during responses of all TANs (black) and DA neurons (gray) to the corresponding events. Values on ordinate refer to change in mean spike rate during the 400 ms (indicated by the bar in Figured 4A and 4B following the event). b indicates slope of regression line, r^2 indicates correlation coefficient.

(B) Responses as a function of effective reward probability computed as the probability of receiving reward following either correct or erroneous performance. Effective probability was computed as the joint probability *P*(*correct choice*)∩*P*(*reward*). Same conventions as in (A).

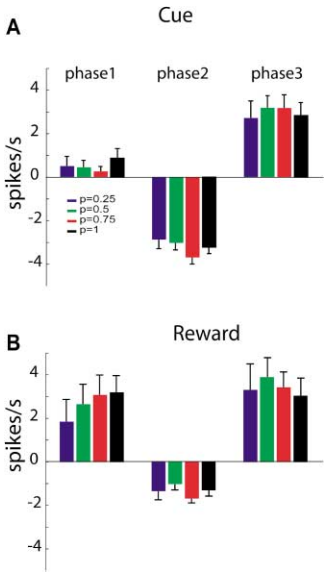

Figure 6. TAN Response Phases

(A) Visual cue response. Change in firing rate (mean + SEM) calculated at the times of first peak of population average (phase 1), negative peak of population average (phase 2), and second peak of population average (phase 3).

(B) Reward response. Same conventions as in (A).

proximate epoch of responses of the two neurons are marked.

This example is indeed typical of the entire responding population. Figures 7B and 7C show superimposed plots of mean responses of both populations to the visual cue and reward, respectively. As the figure shows, the increase in firing of the DA neurons and the pause in firing of the TAN population largely overlap. Figures 7D and 7E, showing the corresponding distributions of single neuron latencies in response to each event, indicates that the overlap in population response is due to an overlap in the latencies of responses of the single neurons comprising each population.

Temporal Correlations

The similarity of responses within each population of neurons, as well as the electrical coupling between dopaminergic neurons in the SNc (Grace and Bunney, 1983), raise the issue of functional connectivity within the DA neurons. These features suggest that there could be a high degree of temporal correlation within this population as is the case with the TANs (Raz et al., 1996; Kimura et al., 2003). We performed cross-correlation analysis for spike trains of 110 simultaneously recorded TAN pairs and 65 simultaneously recorded pairs of DA neurons. An example of two such days of recording is plotted in Figure 8. The results from the TAN data (Figure 8A) are consistent with the previous findings, with 60% of the pairs significantly correlated, irrespective of the behavioral epoch chosen for analysis and following shift predictor normalization (Perkel et al., 1967). By contrast, the same analysis of pairs of dopaminergic neurons yielded the much smaller value of 27% significantly correlated pairs (Figure 8B). Thus, a second major difference between the DA and the TAN systems is their

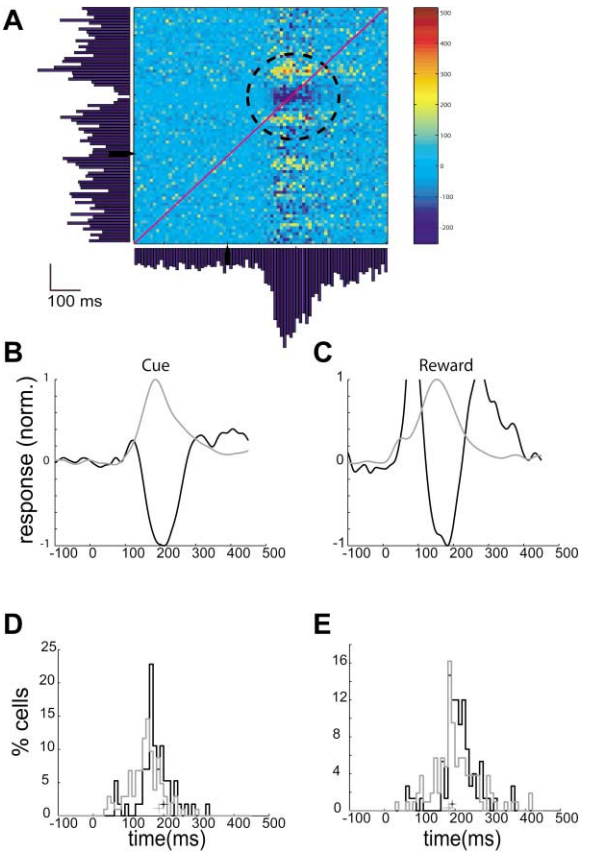

Figure 7. Response Latencies

(A) Covariation matrix of simultaneously recorded TAN (rows) and DA neuron (columns) triggered on the appearance of a visual cue (arrowhead). Each pixel depicts the mean number of spikes that both neurons emitted in deviation from baseline consistently, on a trial-by-trial basis in a 10 × 10 ms lag bin. The average response (PSTH) of the TAN and the DA neurons are given adjacent to the *y* and *x* axes, respectively.

(B) Visual cue response. Gray: DA population response (baseline subtracted) was averaged for all probabilities pooled together, normalized to the peak of each response, so that all neurons are weighted equally in the average. Black: TAN population responses (baseline subtracted) were normalized to maximum trough and averaged for all probabilities.

(C) Reward response. Same conventions as in (B).

(D) Distribution of latencies to middle of response to the visual cue. The middle of DA response (gray) was taken as the midpoint between the point when the firing increased beyond the 0.01 significance level (for at least 3 consecutive 1 ms time bins) and the point of decrease back to the 0.01 significance level. The middle of TAN pause (black) was calculated in the same fashion, using decreases and subsequent increases in firing rate to the 0.01 significance level. Plus sign denotes the values for the neurons depicted in (A).

(E) Distribution of latencies to middle of response to the visual cue. Same conventions as in (D).

synchronization level. The striatal cholinergic system displays a high level of correlation in the spiking activity of the TANs, whereas the DA neurons are largely independent in their firing.

Discussion

The present results provide clues for the co-involvement of the striatal dopaminergic and cholinergic systems in

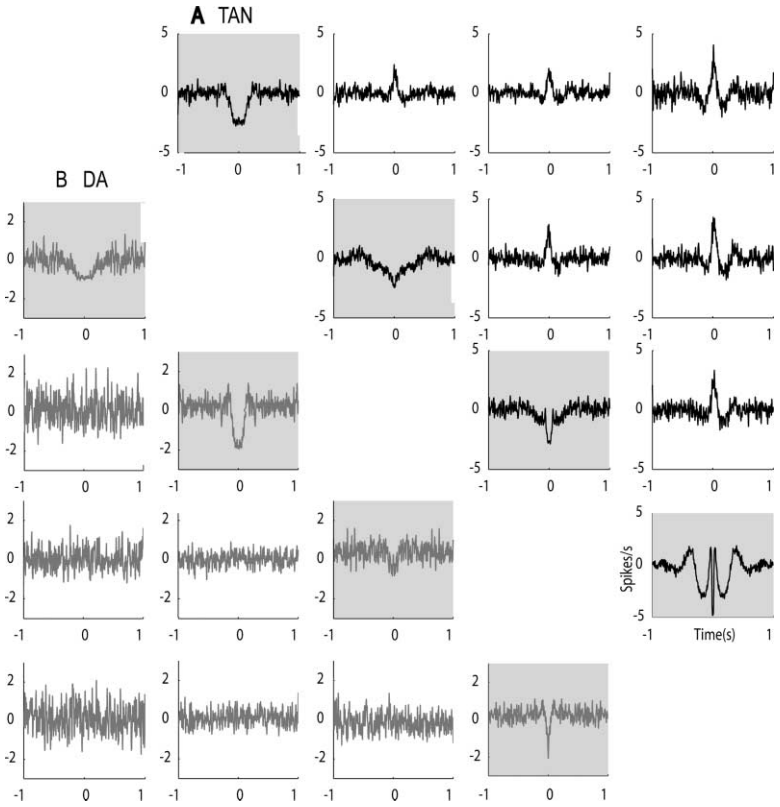

Figure 8. Cross-Correlograms of Simultaneously Recorded TANs and DA Neurons

(A) Correlation functions of four simultaneously recorded TANs. Correlation functions were computed for the 4 s time period preceding and following the "trial begin" signal and corrected by shift predictor. The corrected correlograms were calculated only for pairs recorded with different electrodes with 1 ms bins and smoothed with a Gaussian window (σ = 2 ms). Each half matrix displays the set of all possible correlation pairs, with autocorrelograms on the main diagonal, with shaded background.

(B) Correlation functions of four simultaneously recorded dopaminergic neurons. Same method of calculation and conventions as in (A).

instrumental learning. We recorded from TANs in the putamen and DA neurons in the SNc in the same monkeys, under identical conditions, occasionally simultaneously. Therefore, the demonstrated dissociation between the two neuro-modulatory systems must reflect an inherent distinction between the information they handle and, consequently, between their effects on the physiology of the cortico-striatal complex. On the one hand, both TANs and DA neurons emit coincident robust signals following reward-related events, simultaneously exerting their respective changes in the extracellular levels of striatal ACh and DA. Yet, the messages conveyed by each, as well as the manner in which they are transmitted, are fundamentally different. This type of interplay suggests complementary roles for these two teaching systems.

The responses of the TANs and DA neurons coincide, both at the level of the population average and at the level of single neurons. This is particularly evident in view of the differences in response latencies to the different events. Since the DA neurons and cholinergic interneurons innervate the same population of striatal neurons (Zhou et al., 2003), exerting both short-term and long-term effects on the efficacy of cortico-striatal synapses (Reynolds et al., 2001; Calabresi et al., 2000; Flores-Hernandez et al., 2000; Kitai and Surmeier, 1993), it is particularly appealing to consider possible modes of interplay between the effects of the striatal TANs and those of the DA afferents within the context of learning.

The present results demonstrate that the probability of reward is encoded in the activity of DA neurons. It is important to note, however, that probability, in this experimental context, is merely a reflection of the predictive value of various events in relation to reward, rather than a dimension represented in the basal ganglia as such. Our results indicate that, in line with the TD reinforcement learning rule (Schultz et al., 1997; Sutton and Barto, 1998; Suri and Schultz, 2001), the DA neurons' response to reward and reward-predicting events is a monotonic function of the mismatch between expectation of rewarding events and actual outcome (the TD error). One may note that the DA response does not reach zero when the apparent prediction error should be null (reward following the p = 1 cue). However, it is closer to null when the prediction error is estimated by the effective probability, taking into account the actual performance parameters of the monkeys, than by the more conservative conditional probability. Nevertheless, the intercept still does not reach null values. This could point to the overall context imposed by the probabilistic regime of reinforcement, i.e., the possibility of correct behavior that is unrewarded. Thus, the monkeys may generalize over the stimuli. Alternatively, misidentification of the stimuli may occur. However, positive responses to reward following the p = 1 cue could also be attributed to the method of computing the prediction. The issue of the confining probability used for estimating the effective probability is not a simple one, as it inevitably introduces arbitrary decisions regarding the algorithm (which kind of errors to use or ignore, the effective observation period, etc.) used by the monkey, or the neuronal apparatus, to estimate effective probability. Our choice of factoring the conditional probability by the overall probability of correct performance maintains the concept of probability of reward following a certain visual cue. Still, this may not have been optimal, since at least some of the factors comprising the degree of performance are presumably fully predicted by the mon-

key (and DA neurons) and therefore should not influence reward predictability.

Differential responses of DA neurons associated with reward predictability have recently been demonstrated by Schultz and coworkers in a classical conditioning paradigm (Fiorillo et al., 2003) and are presently being extended in our instrumental conditioning task. Interestingly, by contrast to the findings reported by Fiorillo et al. (2003), DA neurons did not exhibit sustained change from baseline activity in response to (or preceding) any behavioral event, including the visual cue signaling reward with $p = 0.5$, a condition in which uncertainty concerning future reward is maximal. A key difference between the experimental paradigms is our use of trace conditioning (in which a cue appears briefly and has to be remembered at the GO signal), while Fiorillo et al. used delay conditioning (in which the cue persists until the GO signal). Accordingly, Fiorillo et al. report that when using trace-conditioning, statistically significant uncertainty related activity was not reproduced. Other differences may exist between the two studies, including different sampling bias in recording DA neurons from VTA and SNc, use of instrumental rather than classical conditioning, different interstimulus intervals, etc.

As predicted by the TD models and shown in previous studies (Hollerman and Schultz, 1998; Fiorillo et al., 2003; Satoh et al., 2003), the DA response reverses in polarity, displaying decreased firing for negative prediction errors. However, the present results suggest that in this domain the nature of the signal is qualitatively different than that described in the positive domain. Whereas for positive errors (actual outcome better than predicted) the DA neurons response is linear to the prediction error, the response to negative errors is unaffected by the extent of the mismatch between prediction and reality, merely signaling its existence. This finding is inconsistent with the straightforward prediction associated with the TD model. However, it is not surprising given the low-frequency spontaneous discharge of the DA neurons that is further decreased following negative errors. Such conditions make efficient coding of the negative domain difficult. We therefore suggest that negative errors may be more accurately signaled by a third teacher. Notably, in experimental conditions where the monkeys were explicitly informed that a negative prediction error has occurred, DA neurons seemed to report negative errors far less efficiently than positive errors, e.g., smaller fraction of neurons with pure depression of discharge and smaller magnitude or gain of the negative responses (Tobler et al., 2003; Satoh et al., 2003).

In close resemblance to the DA neurons, the TANs responded to the same environmental events, signaling those holding rewarding value and at the same time displaying an intrinsic sense of timing. In particular, note their responses to omission of (partially) expected reward. The importance of this finding is 2-fold: first, since the timing of expected reward was not externally indicated in any way in our experimental setup, the response indicates that TANs, like DA neurons, encode information regarding the expected timing of events. This point deserves special attention, as TANs have not been known to respond to a virtual event: i.e., one that is not explicitly cued, yet the timing of which has internal cognitive meaning (Ravel et al., 2001). However, in contrast to the DA neurons, this response is of the same polarity as to reward delivery. This finding is indicative of a more general phenomenon, which is the main difference found between the two types of neuron. As opposed to the DA neurons, the TANs' response to the reward-related events is indifferent to reward predictability. A prominent pause in firing follows all reward-related events, good or bad, surprising or fully predicted. Moreover, neither phase of the typical triphasic template of TAN is statistically sensitive to reward predictability. This finding is consistent with previous studies that show that even aversive events yield similar TAN responses (Ravel et al., 1999; Blazquez et al., 2002; Shimo and Hikosaka, 2001; however, see Ravel et al., 2003).

The above distinction points to a difference in the possible teaching roles of the dopaminergic and cholinergic systems. We suggest that while the dopaminergic response appears to code the predictive value of various events in relation to reward, providing a rectified TD error signal, the TAN message may serve as a temporal frame, defining the time in which the dopamine signal will be processed. The proposed account for the dissociation in the nature of the two signals is further supported by the difference in pair-wise temporal correlations between the two populations. If indeed the cholinergic signal provides timing information, it is imperative that all cholinergic neurons serving this function emit this signal at the same time. The functional anatomy of the basal ganglia ensures this by constant synchronization of the TAN activity (at least within the spatial sampling limit imposed by our multiple-electrode setup). By contrast, in order for the system to utilize correctly the dopaminergic signal as an error signal, it must be able to assess its magnitude accurately. As theoretical analysis has shown, this is optimally accomplished by averaging across a population of maximally independent and asynchronous neurons (Zohary et al., 1994).

A separate signal defining a time frame, on top of the already timed DA signal, seems superfluous at first sight. However, while the change in firing of DA neurons is well timed to the significant events, the time of elevated extracellular dopamine may be much less precise due to the inefficiency of DA removal from the synapses. Extracellular dopamine levels have been shown to remain elevated in the order of hundreds of milliseconds to seconds following short bursts of firing (Cragg et al., 2000; Roitman et al., 2004; Venton et al., 2003). Thus, the local DA concentration provides an inaccurate time frame for reinforcing the cortico-striatal network states. This is in sharp contrast to striatal ACh, which is rapidly degraded by the extremely dense AChE (Zhou et al., 2003). It is crucial, therefore, to limit the time during which the DA signal can be utilized to modify cortico-striatal synaptic efficacy. It was shown that induction of LTP in the cortico-striatal pathway is mediated by activation of dopamine D1/D5 receptors (Reynolds et al., 2001; Kerr and Wickens, 2001). Activation of M2 (colocalized with D1/D5 receptors [Zhou et al., 2003]) reduces LTP at cortico-striatal synapses (Calabresi et al., 1998). Release from this block by the pause in the TAN response can serve as the window for DA-dependent plasticity. Finally, ACh reduces the sensitivity of striatal projection neurons to their cortical inputs by

fixing their up/down state (Akins et al., 1990). The TAN pause enhances this sensitivity in the critical periods in which this entire circuit is susceptible to long-term change. This may enable striatal neurons to adjust their state according to the cortical inputs, thereby ensuring that the DA teacher reinforces the correct state of the network.

To conclude, the different responses and correlation patterns of DA and TANs suggest that the two systems do not mirror one another. The data presented here, along with previous data, may suggest an intriguing form of cooperation between the two systems: the ACh signal informs the basal ganglia actors when to learn, the DA signal tells them how to learn, and the nature of the immediately preceding cortico-striatal activity defines what will be learned.

Experimental Procedures

Animals and Behavioral Task

Three macaque (*Macaca fasicularis*) monkeys (2 female, monkeys C and E, and 1 male, monkey Y) weighing 2.5–4 kg were used in this study. The monkeys' care and surgical procedures were in accordance with the NIH Guide for the Care and Use of Laboratory Animals (1996) 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. The monkeys were trained to perform an instrumental conditioning task (Figures 1A and 1B) with a probabilistic reward schedule. The probability of receiving reinforcement for correct performance depended on the presented visual cue. The monkeys were seated facing a screen with a panel consisting of three keys. Trials were initiated when the monkey touched the central key. After a variable delay (1.5–2.5 s in monkeys C and E; 2–4 s in monkey Y), a visual cue appeared for a short period (0.3 s in monkeys C and E; 0.45 s in monkey Y) on a randomly chosen side of the screen. The monkeys were well acquainted with a set of 4 (Y and E) or 5 (C) possible cues. Each cue was associated with a different probability of reward (0 [C], 0.25, 0.5, 0.75, and 1.0). The cue presentation was followed by a fixed hold period of 2 s (monkeys Y and C) or 1.5 s (monkey E), after which a GO signal appeared. The monkeys were required to press either the left or right key, according to the location of the memorized cue within an allowed response time of 800 ms for monkeys C and E and 700 ms for monkey Y. Correct response was followed (with an interval of 700 ms) by a liquid reward at the probability associated with the visual cue. No external cue indicated the expected time of the reward. All trials (incorrect, correct, rewarded, and unrewarded) were followed by a variable intertribal interval (ITI) (3–6 s in monkeys E and C; 5–7 s in monkey Y).

MRI Localization of Recording Targets

We estimated the stereotaxic coordinates of the recorded structures according to MRI scans aligned with an anatomical atlas of a *Macaca fascicularis* (Szabo and Cowan, 1984; Martin and Bowden, 2000). After training, a square recording chamber with a 27 mm (inner) side was attached to the skull to allow access to the basal ganglia targets. The recording chamber was tilted 40°–50° laterally in the coronal plane, with its center targeted at stereotaxic coordinates of the SNc. The chamber's coordinates were adjusted according to MRI imaging. An MRI scan (Biospec Bruker 4.7 Tesla animal system, fast-spin echo sequence; effective TE = 80 ms and TR = 2.5 s, 13 coronal slices 1 or 2 mm wide) was performed with a 150 μm diameter tungsten electrode at the estimated location of the target. We aligned the two-dimensional MRI images with the sections of the atlas. Matched MRI (of monkey E) and atlas sections are shown in Figure 1C. All surgical and MRI procedures were performed under general and deep anesthesia.

Recording and Data Analysis

During recording sessions, the monkeys' heads were immobilized, and eight glass-coated tungsten microelectrodes (impedance 0.3–1.2 MΩ at 1000 Hz), confined within a cylindrical guide (1.65 mm inner diameter), were advanced separately (EPS, Alpha-Omega Engineering, Nazareth, Israel) to the recording targets (Figure 1C). The signal from the electrodes was amplified with a gain of 10K and band-pass filtered with a 300–6000 Hz 4-pole Butterworth filter (MCP+, Alpha-Omega Engineering). This electrical activity was sorted and classified online using a template-matching algorithm (MSD, Alpha-Omega Engineering). The sampling rate of spike detection pulses and behavioral events was 12 KHz (AlphaMap, Alpha-Omega Engineering).

Upon reaching the target areas, as judged by the stereotaxic and MRI coordinates, the recorded cells were identified according to their physiological properties. TANs were identified by their characteristic spike shape and firing pattern (Apicella et al., 1998; Aosaki et al., 1995; Shimo and Hikosaka, 2001; Raz et al., 1996). DA neurons were judged by their long-duration, polyphasic spikes and were additionally examined for firing elevation in response to free reward (Hollerman and Schultz, 1998; Waelti et al., 2001). After recording, the electrode tracks were generally continued to the neighboring structures, further aiding verification of the recorded structures.

Only spike trains that were considered to be emitted by a single cell during real-time sorting were subjected to rate stability analysis in non-task-related time segments. In the rate stability analysis the instantaneous rate of a neuron as a function of time during the ITI period was displayed for the entire period of recording, and the largest continuous segment of stable data was selected for further analysis. Cells were chosen for the database after examination for response to at least one of the behavioral events (visual cue, reward, and reward omission) using a Mann-Whitney U test, $p < 0.05$ after a Bonferroni adjustment to compensate for multiple comparisons. Throughout this report, significance is accepted at the $p < 0.01$ level, nonsignificance when $p > 0.1$, unless otherwise explicitly mentioned. We used both parametric (t test and ANOVA) and nonparametric (Mann-Whitney and Kruskal-Wallis) statistical tests.

Cell responses to behavioral events for purposes of analysis of variance and linear regression analysis were parameterized as the difference in average firing rate at the 400 ms following the event compared to the 400 ms preceding it. Due to the complex nature of the TAN response, this procedure was preceded by half-wave rectification of the response. This consisted of smoothing the spike trains with a Gaussian window, $\sigma = 10$ ms, subtracting the baseline rate, and computing the absolute values of this baseline subtracted spike train. This response vector was then subjected to the same "difference in firing rate" analysis.

Timing analysis of the simultaneously recorded TAN and DA neuron was performed by computing the covariation matrix of both neurons triggered on the behavioral event using 10 ms $\times$ 10 ms bins. For this, we computed the baseline subtracted spike count of each neuron in 10 ms bins in every trial (from 300 ms before the behavioral event till 500 ms after the event) and multiplied the resulting matrices of both neurons. The baseline was calculated as the mean of the 300 ms prior to the trigger event. The result is an 80 $\times$ 80 bin matrix describing all possible time lags between the firing of both neurons. Since the purpose of the analysis is to compare latencies, this matrix was not normalized for the behavioral response. In bins in which neither neuron responds, the values should be noise level, varying around zero. Moreover, if only one neuron responds and the other is firing around its baseline, the obtained values should still be zero. Only in bins describing the responses of both would one expect to see significant covariation (positive if both respond with the same polarity and negative if they respond with opposite polarities). Coincidence, or zero time lag, is represented on the main diagonal of the matrix.

Cross-correlation analysis was performed only on spike trains emitted by cells recorded by different electrodes so as to avoid effects of spike sorting. The analysis was performed by calculating the correlation between all cell pairs in 1 ms bins over $\pm$1 s time periods. The result was then smoothed with a Gaussian window ($\sigma = 2$ ms). Cross-correlation functions were considered significant (with peaks, troughs, or oscillations) if they differed from the shift predictor (Perkel et al., 1967) by $\pm$0.995 confidence intervals.

Acknowledgments

We thank S. Haber, Y. Ben-Shaul, and Y. Prut for fruitful discussions and comments on earlier versions of this manuscript; W. Schultz

and C. Fiorillo for sharing their method of recording DA neurons; G. Goelman for MRI; and V. Sharkansky for technical assistance. This study was partly supported by a center of excellence grant administered by the ISF, by the United States-Israel Binational Science Foundation, by the German-Israel Binational Foundation (GIF), and by the BMBF Israel-Germany collaboration in medical research. G.M. was supported by the Horowitz fellowship.

Received: September 8, 2003
Revised: March 22, 2004
Accepted: May 24, 2004
Published: July 7, 2004

References

Akins, P.T., Surmeier, D.J., and Kitai, S.T. (1990). Muscarinic modulation of a transient K+ conductance in rat neostriatal neurons. Nature *344*, 240–242.

Alexander, G.E., and DeLong, M.R. (1985). Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. J. Neurophysiol. *53*, 1417–1430.

Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A.M., and Kimura, M. (1994). Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J. Neurosci. *14*, 3969–3984.

Aosaki, T., Kimura, M., and Graybiel, A.M. (1995). Temporal and spatial characteristics of tonically active neurons of the primate's striatum. J. Neurophysiol. *73*, 1234–1252.

Apicella, P., Legallet, E., and Trouche, E. (1997). Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. Exp. Brain Res. *116*, 456–466.

Apicella, P., Ravel, S., Sardo, P., and Legallet, E. (1998). Influence of predictive information on responses of tonically active neurons in the monkey striatum. J. Neurophysiol. *80*, 3341–3344.

Barbeau, A. (1962). The pathogensis of Parkinson's disease: a new hypothesis. Can. Med. Assoc. J. *87*, 802–807.

Baum, W.M. (1979). Matching, undermatching, and overmatching in studies of choice. J. Exp. Anal. Behav. *32*, 269–281.

Bennett, B.D., and Wilson, C.J. (1999). Spontaneous activity of neostriatal cholinergic interneurons in vitro. J. Neurosci. *19*, 5586–5596.

Blazquez, P.M., Fujii, N., Kojima, J., and Graybiel, A.M. (2002). A network representation of response probability in the striatum. Neuron *33*, 973–982.

Calabresi, P., Centonze, D., Gubellini, P., Pisani, A., and Bernardi, G. (1998). Blockade of M2-like muscarinic receptors enhances long-term potentiation at corticostriatal synapses. Eur. J. Neurosci. *10*, 3020–3023.

Calabresi, P., Centonze, D., Gubellini, P., and Bernardi, G. (1999). Activation of M1-like muscarinic receptors is required for the induction of corticostriatal LTP. Neuropharmacology *38*, 323–326.

Calabresi, P., Centonze, D., Gubellini, P., Pisani, A., and Bernardi, G. (2000). Acetylcholine-mediated modulation of striatal function. Trends Neurosci. *23*, 120–126.

Carpenter, R.H., and Williams, M.L. (1995). Neural computation of log likelihood in control of saccadic eye movements. Nature *377*, 59–62.

Centonze, D., Gubellini, P., Bernardi, G., and Calabresi, P. (1999a). Permissive role of interneurons in corticostriatal synaptic plasticity. Brain Res. Brain Res. Rev. *31*, 1–5.

Centonze, D., Gubellini, P., Picconi, B., Calabresi, P., Giacomini, P., and Bernardi, G. (1999b). Unilateral dopamine denervation blocks corticostriatal LTP. J. Neurophysiol. *82*, 3575–3579.

Centonze, D., Gubellini, P., Pisani, A., Bernardi, G., and Calabresi, P. (2003). Dopamine, acetylcholine and nitric oxide systems interact to induce corticostriatal synaptic plasticity. Rev. Neurosci. *14*, 207–216.

Cragg, S.J., Hille, C.J., and Greenfield, S.A. (2000). Dopamine release and uptake dynamics within nonhuman primate striatum in vitro. J. Neurosci. *20*, 8209–8217.

Crutcher, M.D., and DeLong, M.R. (1984). Single cell studies of the primate putamen. II. Relations to direction of movement and pattern of muscular activity. Exp. Brain Res. *53*, 244–258.

DeBoer, P., and Abercrombie, E.D. (1996). Physiological release of striatal acetylcholine in vivo: modulation by D1 and D2 dopamine receptor subtypes. J. Pharmacol. Exp. Ther. *277*, 775–783.

Drukarch, B., Schepens, E., Schoffelmeer, A.N., and Stoof, J.C. (1989). Stimulation of D-2 dopamine receptors decreases the evoked in vitro release of [3H]acetylcholine from rat neostriatum: role of K^+ and Ca^{2+}. J. Neurochem. *52*, 1680–1685.

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

Flores-Hernandez, J., Hernandez, S., Snyder, G.L., Yan, Z., Fienberg, A.A., Moss, S.J., Greengard, P., and Surmeier, D.J. (2000). D(1) dopamine receptor activation reduces GABA(A) receptor currents in neostriatal neurons through a PKA/DARPP-32/PP1 signaling cascade. J. Neurophysiol. *83*, 2996–3004.

Goodie, A.S., and Crooks, C.L. (2004). Time-pressure effects on performance in a base-rate task. J. Gen. Psychol. *131*, 18–28.

Grace, A.A., and Bunney, B.S. (1983). Intracellular and extracellular electrophysiology of nigral dopaminergic neurons. 3. Evidence for electrotonic coupling. Neuroscience *10*, 333–348.

Graybiel, A.M., Aosaki, T., Flaherty, A.W., and Kimura, M. (1994). The basal ganglia and adaptive motor control. Science *265*, 1826–1831.

Herrnstein, R.J. (1970). On the law of effect. J. Exp. Anal. Behav. *13*, 243–266.

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.

Holt, D.J., Graybiel, A.M., and Saper, C.B. (1997). Neurochemical architecture of the human striatum. J. Comp. Neurol. *384*, 1–25.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience *96*, 651–656.

Jones, I.W., Bolam, J.P., and Wonnacott, S. (2001). Presynaptic localisation of the nicotinic acetylcholine receptor beta2 subunit immunoreactivity in rat nigrostriatal dopaminergic neurones. J. Comp. Neurol. *439*, 235–247.

Kerr, J.N., and Wickens, J.R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. J. Neurophysiol. *85*, 117–124.

Kimura, M., Matsumoto, N., Okahashi, K., Ueda, Y., Satoh, T., Minamimoto, T., Sakamoto, M., and Yamada, H. (2003). Goal-directed, serial and synchronous activation of neurons in the primate striatum. Neuroreport *14*, 799–802.

Kitabatake, Y., Hikida, T., Watanabe, D., Pastan, I., and Nakanishi, S. (2003). Impairment of reward-related learning by cholinergic cell ablation in the striatum. Proc. Natl. Acad. Sci. USA *100*, 7965–7970.

Kitai, S.T., and Surmeier, D.J. (1993). Cholinergic and dopaminergic modulation of potassium conductances in neostriatal neurons. Adv. Neurol. *60*, 40–52.

Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. Science *273*, 1399–1402.

Kudernatsch, M., and Sutor, B. (1994). Cholinergic modulation of dopamine overflow in the rat neostriatum: a fast cyclic voltammetric study in vitro. Neurosci. Lett. *181*, 107–112.

Lang, A.E., and Lees, A.J. (2002). Anticholinergic therapies in the treatment of Parkinson's disease. Mov. Disord. *17*, S7–S12.

Lavoie, B., Smith, Y., and Parent, A. (1989). Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase immunohistochemistry. J. Comp. Neurol. *289*, 36–52.

Lee, I.H., and Assad, J.A. (2003). Putaminal activity for simple reactions or self-timed movements. J. Neurophysiol. *89*, 2528–2537.

Martin, R.F., and Bowden, D.M. (2000). Primate Brain Maps: Structure of the Macaque Brain (Amesterdam: Elsevier Science).

Matsumoto, N., Hanakawa, T., Maki, S., Graybiel, A.M., and Kimura, M. (1999). Nigrostriatal dopamine system in learning to perform sequential motor tasks in a predictive manner. J. Neurophysiol. *82*, 978–998.

Nisenbaum, L.K., and Kitai, S.T. (1995). A new concept in the action of anticholinergic drugs in an animal model of Parkinson's disease. In Molecular and Cellular Mechanisms of Neostriatal Function, M.A. Ariano and D.J. Surmeier, eds. (Austin, TX: R.G. Landes Company), pp. 273–282.

Perkel, D.H., Gerstein, G.L., and Moore, G.P. (1967). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. Biophys. J. *7*, 419–440.

Pisani, A., Bonsi, P., Centonze, D., Calabresi, P., and Bernardi, G. (2000). Activation of D2-like dopamine receptors reduces synaptic inputs to striatal cholinergic interneurons. J. Neurosci. (Online) *20*, RC69.

Pisani, A., Bonsi, P., Centonze, D., Gubellini, P., Bernardi, G., and Calabresi, P. (2003). Targeting striatal cholinergic interneurons in Parkinson's disease: focus on metabotropic glutamate receptors. Neuropharmacology *45*, 45–56.

Ravel, S., Legallet, E., and Apicella, P. (1999). Tonically active neurons in the monkey striatum do not preferentially respond to appetitive stimuli. Exp. Brain Res. *128*, 531–534.

Ravel, S., Sardo, P., Legallet, E., and Apicella, P. (2001). Reward unpredictability inside and outside of a task context as a determinant of the responses of tonically active neurons in the monkey striatum. J. Neurosci. *21*, 5730–5739.

Ravel, S., Legallet, E., and Apicella, P. (2003). Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. J. Neurosci. *23*, 8489–8497.

Raz, A., Feingold, A., Zelanskaya, V., Vaadia, E., and Bergman, H. (1996). Neuronal synchronization of tonically active neurons in the striatum of normal and parkinsonian primates. J. Neurophysiol. *76*, 2083–2088.

Reddi, B.A., and Carpenter, R.H. (2000). The influence of urgency on decision time. Nat. Neurosci. *3*, 827–830.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? Trends Neurosci. *22*, 146–151.

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

Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci. *22*, 9475–9489.

Roitman, M.F., Stuber, G.D., Phillips, P.E.M., Wightman, R.M., and Carelli, R.M. (2004). Dopamine operates as a subsecond modulator of food seeking. J. Neurosci. *24*, 1265–1271.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. J. Neurosci. *23*, 9913–9923.

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

Shimo, Y., and Hikosaka, O. (2001). Role of tonically active neurons in primate caudate in reward-oriented saccadic eye movement. J. Neurosci. *21*, 7804–7814.

Suri, R.E., and Schultz, W. (2001). Temporal difference model reproduces anticipatory neural activity. Neural Comput. *13*, 841–862.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning—an Introduction (Cambridge, MA: MIT Press), pp. 3–322.

Szabo, J., and Cowan, W.M. (1984). A stereotaxic atlas of the brain of the cynomolgus monkey (*Macaca fascicularis*). J. Comp. Neurol. *222*, 265–300.

Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. J. Neurosci. *23*, 10402–10410.

Venton, B.J., Zhang, H., Garris, P.A., Phillips, P.E., Sulzer, D., and Wightman, R.M. (2003). Real-time decoding of dopamine concentration changes in the caudate-putamen during tonic and phasic firing. J. Neurochem. *87*, 1284–1295.

Vulkan, N. (2000). An economist's perspective on probability matching. J. Econ. Surv. *14*, 101–118.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. Nature *412*, 43–48.

Wilson, C.J., Chang, H.T., and Kitai, S.T. (1990). Firing patterns and synaptic potentials of identified giant aspiny interneurons in the rat neostriatum. J. Neurosci. *10*, 508–519.

Wolford, G., Miller, M.B., and Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. J. Neurosci. *20*, RC64.

Zhou, F.M., Wilson, C., and Dani, J.A. (2003). Muscarinic and nicotinic cholinergic mechanisms in the mesostriatal dopamine systems. Neuroscientist *9*, 23–36.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.