Behavioral/Systems/Cognitive

Mechanisms of Action Selection and Timing in Substantia Nigra Neurons

David Fan,[1,2] **Mark A. Rossi,**[1,2] **and Henry H. Yin**[1,2]

Departments of [1]Psychology and Neuroscience, and [2]Neurobiology, Center for Cognitive Neuroscience, Duke University, Durham, North Carolina 27708

The timing of actions is critical for adaptive behavior. In this study we measured neural activity in the substantia nigra as mice learned to change their action duration to earn food rewards. We observed dramatic changes in single unit activity during learning: both dopaminergic and GABAergic neurons changed their activity in relation to behavior to reflect the learned instrumental contingency and the action duration. We found the emergence of "action-on" neurons that increased firing for the duration of the lever press and mirror-image "action-off" neurons that paused at the same time. This pattern is especially common among GABAergic neurons. The activity of many neurons also reflected confidence about the just completed action and the prospect of reward. Being correlated with the relative duration of the completed action, their activity could predict the likelihood of reward collection. Compared with the GABAergic neurons, the activity of dopaminergic neurons was more commonly modulated by the discriminative stimulus signaling the start of each trial, suggesting that their phasic activity reflected sensory salience rather than any reward prediction error found in previous work. In short, these results suggest that (1) nigral activity is highly plastic and modified by the learning of the instrumental contingency; (2) GABAergic output from the substantia nigra can simultaneously inhibit and disinhibit downstream structures, while the dopaminergic output also provide bidirectional modulation of the corticostriatal circuits; (3) dopaminergic and GABAergic neurons show similar task-related activity, although DA neurons are more responsive to the trial start signal.

Introduction

Whether pressing an elevator button or hitting a tennis ball, the timing of actions is critical. Animals can modify action timing through experience with feedback (Platt et al., 1973; Yin, 2009; Yu et al., 2010), but the neural mechanisms underlying such learning remain poorly understood (Mita et al., 2009).

In this study we examined the activity of midbrain substantia nigra neurons as mice learn to time their actions for food rewards. The substantia nigra pars reticulata (SNr) is a major output nucleus of the basal ganglia, a set of subcortical nuclei critical for action selection and instrumental learning (Redgrave et al., 1999a; Yin and Knowlton, 2006; Hikosaka, 2007; Yin, 2010). The neurons from the SNr project to downstream motor control structures, releasing GABA and providing tonic inhibition of target structures (Hikosaka, 2007; Tepper and Lee, 2007). Previous work on eye movements in monkeys has shown that these neurons pause at the time of voluntary saccades, thus disinhibiting tonically active motor control circuits downstream (Hikosaka, 2007). Next to the SNr, in the substantia nigra pars compacta (SNc), are dopamine (DA) neurons that project to the striatum and frontal cortex. DA is a critical modulator of the basal ganglia

circuits. Degeneration of SNc DA cells in Parkinson's disease results in severe motor impairments, such as the inability to initiate voluntary actions (akinesia) and the slowing of actions (bradykinesia), symptoms also found in animal models using experimental DA depletion (Zeiler, 1985; Hudzik et al., 2000).

Although SNr GABA neurons are known to project to SNc DA neurons (Tepper and Lee, 2007), the functional significance of the SNr output and dopaminergic modulation in action selection remains unclear. How do they contribute to the learning and performance of a new action? How does environmental feedback (e.g., success or failure) affect neural activity in DA and GABA neurons?

To address these questions, we simultaneously recorded from GABA neurons in the SNr and DA neurons in the SNc using a discrete-trial temporal differentiation task. In this task, mice are required to press a lever and hold it down for a minimum duration to earn a food reward. In operant conditioning, an arbitrary aspect of behavior can be defined as the "operant" in the feedback function between the organism and the environment. The use of action duration as the operant, or temporal differentiation (Skinner, 1938; Platt et al., 1973), can help us identify the neural circuits responsible for the generation of the action by "tagging" the neural activity with the required action duration. Since the action duration, shared by the neural activity and behavior, can also be modified by instrumental learning, this method also allows us to monitor the changes in neural activity as animals learn to time their actions. We hypothesized that both dopaminergic and GABAergic neurons will alter their activity during the learning to reflect the action duration required by the instrumental contingency.
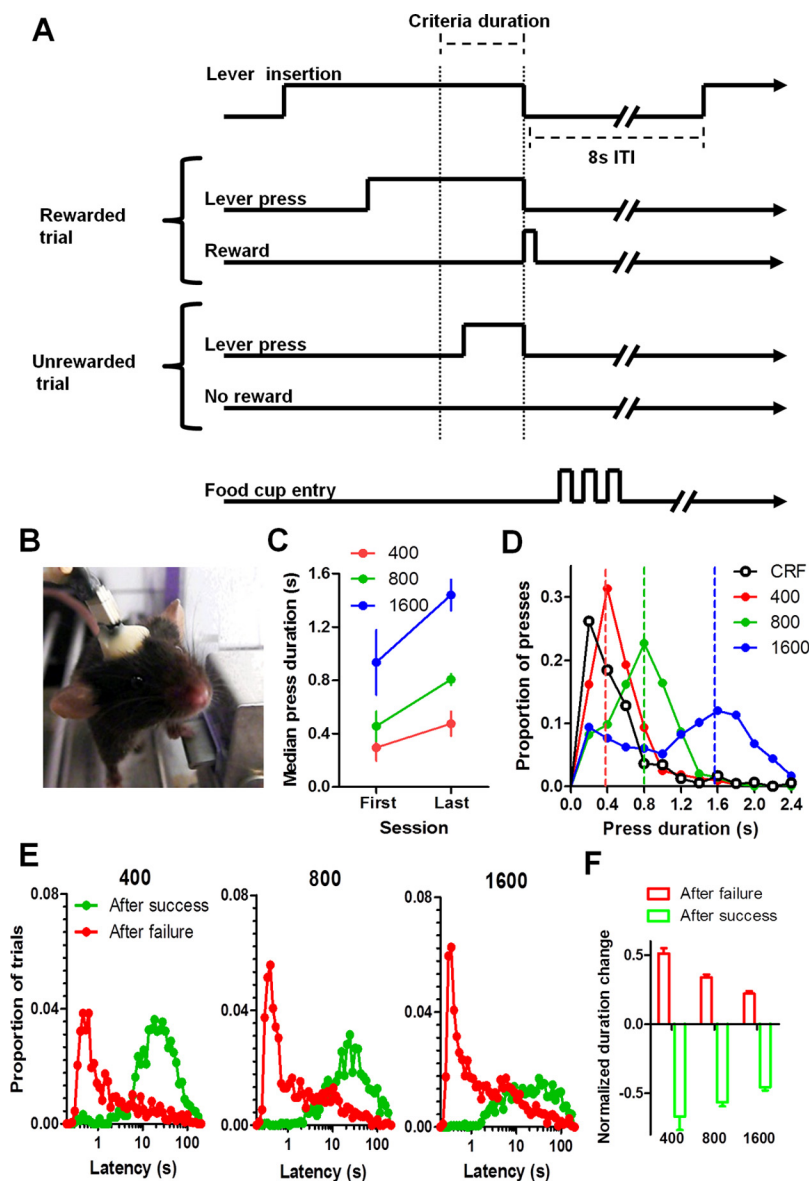
Figure 1. Experimental design and behavioral results. **A**, Illustration of action timing task. Each trial starts with the insertion of a lever. The mouse must press and hold down the lever for a minimum duration to earn a food pellet, which is delivered into the food cup as soon as the lever is released. If the press is too brief, then the trial ends with no reward delivery. The lever is retracted immediately after any press, followed by an ITI of 8 s. **B**, Photograph of a mouse pressing a lever, showing the chronically implanted 16-channel multielectrode array connected to a miniaturized headstage. **C**, Median press duration on the first and last sessions of training for each criterion duration (normally 6 daily sessions at each criterion duration). With learning, the median press duration increased and approached the criterion durations. **D**, Relative frequency distribution of press durations on the last day of training for each criterion duration. CRF denotes continuous reinforcement with no duration requirement (any press exceeding 20 ms earns a reward); "400" denotes a duration requirement of >400 ms; "800", >800 ms; and "1600", >1600 ms. Dotted lines indicate criterion durations. **E**, Latency to initiate action after trial start. Latency is on average significantly longer after a successful (rewarded) trial than after a failed (unrewarded) trial. **F**, The average change in press durations for all sessions (expressed as proportion of the criterion duration) as a function of previous trial outcome. After success, the press duration is reduced relative to the previous trial, whereas after failure the duration is increased.

Materials and Methods

Subjects and surgery. All procedures were approved by the Institutional Animal Care and Use Committee at Duke University. Five male C57BL/6J mice (2–6 months old at the time of experiments) were used. The surgery procedure was similar to that described in previous studies (Costa et al., 2004; Yin et al., 2009). Mice were anesthetized with isoflurane, and placed in a stereotaxic frame. After creating a craniotomy (~1 mm by 2 mm), electrode arrays were lowered at the following coordinates (in relation to bregma): 3.0 mm posterior, 1.0 mm lateral, and 4.6 mm below brain surface. In all experiments, 16-channel microwire arrays (Innovative Neurophysiology) were used. The arrays consisted of micro-polished tungsten wires, 50 μm in diameter and 5–7 mm in length, arranged in a 2 by 8 configuration, and attached to an Omnetics connector (Omnetics Connector Corporation). Row spacing was 200 μm and pitch was 150 μm.

Following the completion of experiments, all mice were anesthetized with isoflurane and transcardially perfused with 0.9% saline followed by 10% buffered formalin solution. Brains were fixed in formalin solution for ~24 h and sliced into 80–100 μm coronal sections with a microtome (Vibratome 1000 Plus). Brain sections were then stained with thionin, and examined under a microscope to verify placement of the electrode tips within the substantia nigra.

Behavioral task. Three days before the start of behavioral training, all mice were food deprived. They were fed 1.5–3 g of their home chow ~1 h after each behavioral session. Their weight was maintained at 85–90% of *ad libitum* feeding weight throughout the experiments, and they were weighed each day immediately after the experimental session to ensure adequate body weight was maintained. Training and testing took place in Med Associates operant chambers. Each chamber contained two retractable levers, a food magazine that delivers food rewards (Bio-Serv 14 mg Dustless Precision Pellets), and a house light (3 W, 24 V) mounted on the wall opposite the levers and magazine. An infrared beam was used to record head entries into the food magazine.

We first trained mice to press a lever to earn food pellets on a continuous reinforcement schedule (CRF) for three 90 min sessions. There was no explicit time requirement for each lever press, although the shortest detectable press was 20 ms. Each press was immediately followed by a food pellet.

The mice then received discrete trial temporal differentiation training. Only presses exceeding the criterion duration were rewarded. Each trial began with the insertion of the lever, and ended only after a press meeting the duration criterion was produced, at which point the lever was retracted and house light was turned off for 8 s (Fig. 1A). The mice were trained with progressively longer criterion durations (>400 ms, >800 ms, and >1600 ms), with six daily sessions on each criterion duration (with the exception of one mouse due to the loss of neural activity early in training). Each session ended after either 100 rewards were earned or 180 min passed.

Multielectrode recording and spike sorting. Single-unit activity was recorded using the Cerebrus data acquisition system (Blackrock Microsystems). The data were sampled at 30 kHz, after filtering with both analog and digital bandpass filters (analog high-pass first order Butterworth filter at 0.3 Hz, analog low-pass third order Butterworth filter at 7.5 kHz). Single unit data was separated with a high-pass digital filter (fourth order Butterworth filter at 250 Hz), while local field potential (LFP) signals were filtered with a third order high-pass filter and seventh order low-

pass filter (0.1 Hz–5 Hz cutoffs). Spikes were sorted using Offline Sorter (Plexon) and single-unit activity was isolated on the basis of principal component analysis. Only single-unit activity with a clear separation from noise was used for the analysis (Fan et al., 2011). The isolated single units were then classified into GABAergic or dopaminergic cells. Because the shape of the waveforms can vary depending on the location of the electrode wire in relation to the cell (Gold et al., 2006), we relied mainly on spike duration in cell type classification (Fig. 2).

Perievent histogram spike rate estimation. A Gaussian kernel method was used to estimate perievent firing rates for the interval from −3 s before the start of a lever press to 10 s after the end of a lever press. Because the firing rate estimate depends on the width of the kernel used, a method to calculate the optimal kernel width was used for each neuron (Shimazaki and Shinomoto, 2010). This method is superior to traditional binning techniques for perievent time histograms because it avoids manual selection of bin sizes, which can lead to poor estimation of firing rates, especially if one bin size is applied to many histograms with very different firing rates and dynamics. We limited this optimal kernal width calculation to 3000 firing events within the described interval, which provides a good estimate when compared with a width estimate obtained with no restrictions. Once the optimal kernal width for each histogram was calculated, the firing rate was estimated with a sampling resolution of 10 ms. Confidence intervals for the firing rate estimation were calculated using a bias-corrected, 200-sample bootstrap method. The confidence interval was found by assuming a normal distribution of the estimator (Abeles, 1982). Bootstrapping was done using the set of individual trials as the sample pool. For sessions with <30 trials, all spikes from all trials were placed into one sampling pool instead.

To account for the different press durations of each trial when estimating the perievent firing rate, each press duration was normalized to the criterion duration before firing rate estimation calculations (Mukamel et al., 2011). This made it possible to calculate a reliable perievent firing rate estimate for press durations, even when a session contains many trials with different press durations. The time interval within a press was "stretched" or "shrunk" until its length matched the criterion duration for that session. Weights were assigned for every spike event. For spikes occurring within a modulated time window, the weighting factor was set to the criterion duration divided by the original press duration. These weights were used when estimating firing rates and calculating bootstrap confidence intervals with Gaussian kernels.

Classifying significantly modulated perievent firing rates. We identified several common features of firing dynamics and designated intervals to classify significant changes in firing rate. By using confidence limits to compare the firing rates during different intervals, we identified neurons that significantly increased or decreased their firing rate in an interval, relative to the surrounding intervals. We compared the lower confidence bounds within an interval to the upper confidence bounds of one or more surrounding intervals to identify significant changes in firing rate. At least one-fourth of the interval must satisfy this
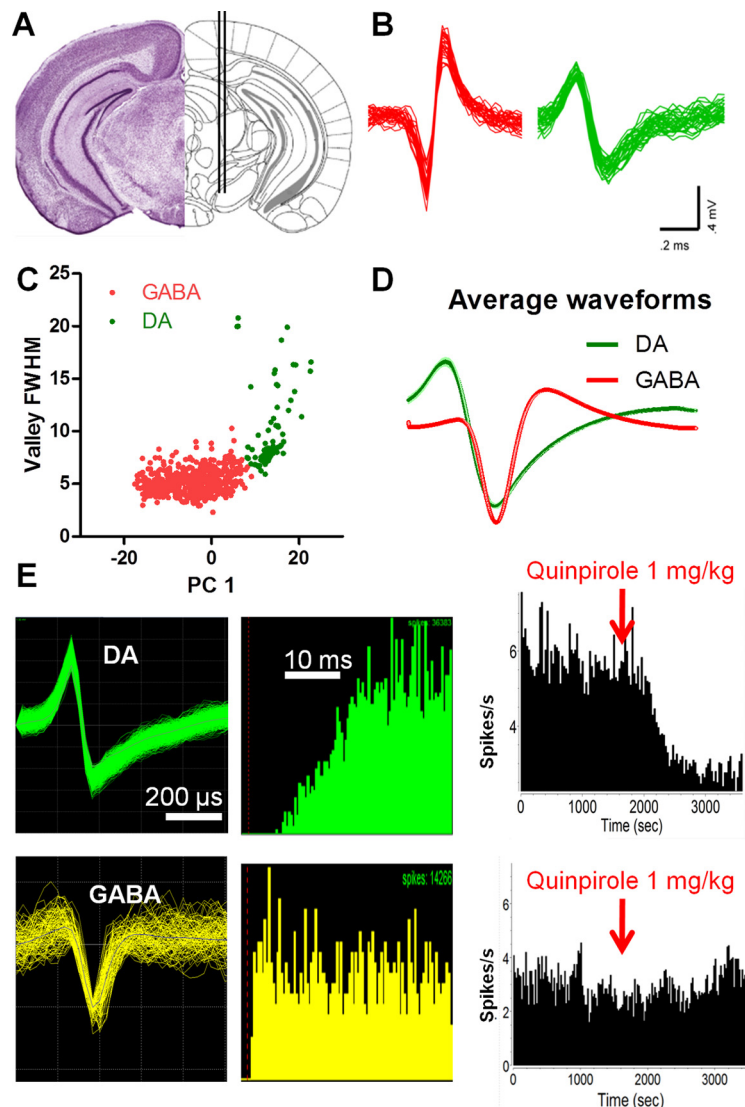

Figure 2. Classification of neurons. ***A***, Placement of 16-channel micro-electrode arrays (2 by 8) into the substantia nigra as illustrated by a coronal brain section. ***B***, Two examples of recorded waveforms: putative GABAergic neuron (left) and putative dopaminergic neuron (right). ***C***, Principal component analysis of the waveforms of DA and GABA cells. The *x*-axis is the first principal component and the *y*-axis is the valley full-width half-max, i.e., the width of the valley at half the depth. ***D***, Average waveforms of all recorded cells, normalized by waveform height. Shaded region represents SEM. DA neurons have significantly wider waveforms (longer spike duration) than GABA neurons. ***E***, Waveform and interspike interval distribution of a putative DA neuron and a putative GABA neuron classified according to the waveform characteristics. The firing rate of the DA neuron is greatly reduced following the injection of quinpirole (1 mg/kg, i.p.), a selective D_2-receptor agonist that activates the D_2 autoreceptors on the axon terminals of the DA neurons. In contrast, no significant reduction in firing rate was observed in the putative GABA neuron following quinpirole injection.

requirement for the entire interval to be considered significantly modulated.

To classify "trial start" neurons, the 500 ms window after lever insertion was compared with a baseline 1000 ms window before the event. To classify "action boundary" neurons, the window from 1000–100 ms before the start of the lever press was used for measuring neural activity related to movement initiation, and the 100–1000 ms window after the end of the lever press was used for measuring neural activity after movement termination. The time windows used are based on visual inspection of the data. For action initiation, firing rate was compared with a 1000 ms baseline window immediately before it, while for action termination, firing rate was compared with both the firing rate during the press and the 1000 ms window after the action termination window. Action-on and action-off neurons were classified by comparing the firing rate within the press duration with both the action initiation and the action termination windows.

Table 1. Summary of overlap in cell categories

	TS	Before press	Duration	After press	Reward	Confidence	TS outcome
DA							
TS	42	15	31	29	20	11	7
Before press	15	36	20	18	12	15	5
Duration	31	20	45	39	24	14	4
After press	29	18	39	47	23	13	2
Reward	20	12	24	23	29	13	5
Confidence	11	15	14	13	13	20	2
TS outcome	7	5	4	2	5	2	9
GABA							
TS	187	142	116	72	91	84	51
Before press	142	342	222	135	122	126	84
Duration	116	222	298	140	117	123	71
After press	72	135	140	203	85	102	55
Reward	91	122	117	85	171	82	65
Confidence	84	126	123	102	82	170	51
TS outcome	51	84	71	55	65	51	112

Table showing the number of pairwise overlaps of different categories of neurons. The diagonal cells are highlighted, showing the total number of cells for that category. For example, of 42 DA cells modulated by the trial start event (TS), 15 also responded to before press, 31 to the duration, 29 to after press, etc. TS, Trial start neurons; TS outcome, trial start outcome-modulated neurons.

Receiver operating characteristic analysis. The area under the receiver operating characteristic (ROC) curve (AUC) indicates the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (Macmillan and Creelman, 2004). An AUC value of 0.5 is no better than chance, while an AUC value closer to 1 or 0 indicates a classifier with better discriminating power. The Mann–Whitney U statistic was used to calculate both the AUC as well as the *p*-value for the AUC statistic. Thus, the following:

$$AUC_1 = \frac{U_1}{n_1 n_2}, \quad (1)$$

where U_1 is the Mann–Whitney U for sample 1 and n_1 and n_2 are the sample sizes of samples 1 and 2, respectively. For large samples, U is approximately normally distributed, and thus a *p*-value can be estimated from the standardized value z, as follows:

$$z = \frac{U - m_U}{\sigma_U}, \quad (2)$$

where m_U and σ_U are the mean and SD of U respectively. m_U and σ_U are given by the following:

$$m_U = \frac{n_1 n_2}{2}, \quad (3)$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}. \quad (4)$$

To classify neurons that "encode" decision confidence (see Fig. 11), the firing rate of the neurons during a short window (see below) after the lever press was used as the parameter, and whether or not a food cup entry occurred within 4 s after the press was used to determine the actual prediction value. AUC values with a corresponding *p*-value of <0.01 were considered to be statistically significant. Neurons with an AUC > 0.5 were classified as upmodulated neurons, i.e., neurons that increased firing right after lever press if the animal entered the food cup within the next 4 s; while those with an AUC < 0.5 were grouped as downmodulated. To determine the firing rate window to use after the lever press, a discrimination index (DI) was calculated for varying windows from 200–600 ms after the lever press, in increments of 50 ms (Kepecs et al., 2008), as follows: $DI = |2(AUC - 0.5)|$.
For each neuron, the window which yielded the highest discrimination index was used for the analysis.

A similar ROC analysis was used to classify trial start/outcome-modulated neurons as well as reward-modulated neurons. For trial start/
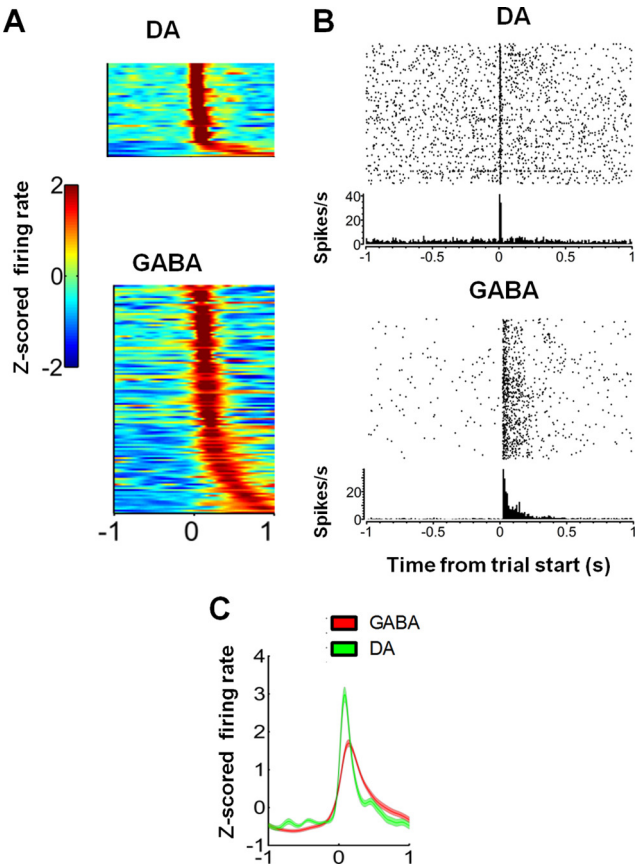

Figure 3. Neural activity in response to the trial start signal (lever insertion). *A*, Spike density functions sorted by latency to peak response time for DA ($n = 41$) and GABA ($n = 122$) neurons. *x*-axis indicates time from trial start (lever insertion), and each row represents activity from a single neuron from a temporal differentiation session. Firing rates are *z*-score normalized. *B*, Raster plots and perievent histograms of representative neurons showing short-latency responses to the lever insertion that signals the start of a trial (discriminative stimulus). Each row represents a single trial from a temporal differentiation session. *C*, Population averages of all trial start neurons with shaded region indicating SEM. Firing rates are *z*-score normalized.

outcome-modulated neurons, the firing rate in a 200 ms window after the starting of a trial was used to compare the activity on rewarded and unrewarded trials. For reward encoding, if the first magazine entry occurred within 4 s of press termination, the 1 s window after the first magazine entry was used to compare the activity on rewarded and unrewarded trials. For trials without a magazine entry within 4 s, the 1 s window after the mean latency to the first magazine entry for that session was used instead.

Results
Training
Temporal differentiation training started after mice had learned to press the lever for food pellet rewards on a continuous reinforcement schedule with no minimum action duration requirement (CRF). Each trial began with the insertion of the lever into the chamber, and ended with its retraction once it is pressed, with an intertrial-interval (ITI) of 8 s (Fig. 1*A*). The delivery of a 14 mg food pellet was contingent upon the production of a minimum action duration: if the lever was held down long enough, then immediately after the lever was released, a food pellet was delivered into a cup adjacent to the lever (Fig. 1*B*). There is a tradeoff between prematurely releasing the lever (thus having to wait for the next trial to make another attempt) and holding the lever down longer than necessary to ensure reward collection. Ideally, it is best to press the lever just long enough to earn a reward. As
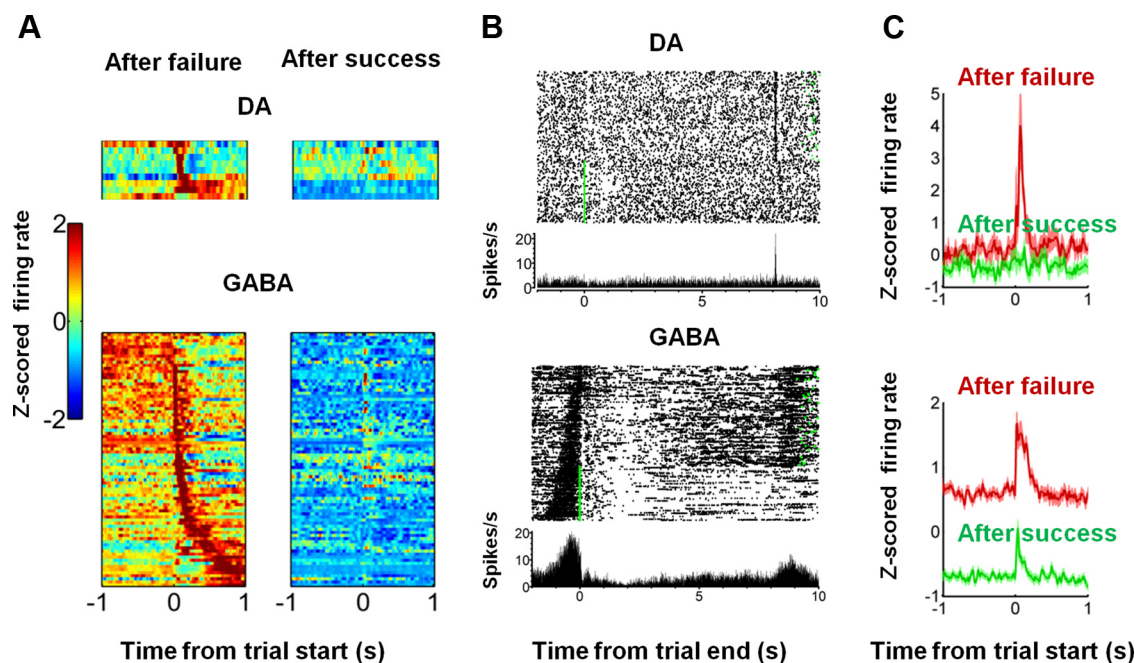
Figure 4. Outcome modulation of trial start activity. **A**, Spike density functions sorted by peak response time for DA ($n = 9$) and GABA ($n = 82$) neurons (bin size = 50 ms, smoothed with 6-tap Gaussian filter) after unrewarded (failure) and rewarded (success) trials. *x*-axis indicates time from trial start (lever insertion) and each row represents the spike density function for a single neuron during a temporal differentiation session. Firing rates are *z*-score normalized. **B**, Raster plots showing representative neurons. Each row is a single trial from a temporal differentiation session. Green markers indicate the end of presses that were rewarded; the rest of the trials were unrewarded. These cells show a burst of firing at the beginning of the next trial, but only if the previous press was unrewarded. They decreased firing after a rewarded lever press. **C**, Population averages of all trial start related cells with shaded region indicating SEM. Firing rates are *z*-score normalized.

shown in Figure 1*C*, mice readily learned to time their lever pressing so that the duration just exceeded the criterion duration, as found in a recent study (Yin, 2009). A two-way ANOVA revealed no interaction between criterion duration and session ($F_{(2, 8)} = 1.97$, $p > 0.05$), a main effect of duration ($F_{(1, 8)}$, $p < 0.001$), and a main effect of session ($F_{(1, 8)}$, $p < 0.001$). As shown in Figure 1*D*, in accord with previous work (Platt et al., 1973; Yin, 2009; Yu et al., 2010), the median duration approached the criterion duration with training. As the duration requirement increased, the mice also failed more often, producing more prematurely released lever presses (percentage of correct trials: >400 ms, 52 ± 4%; >800 ms, 39 ± 4%; >1600 ms, 28 ± 3%).

Figure 1*E* shows the latency to initiate a lever press after the start of each trial during the final session. Latency is defined as the time between the insertion of the lever at the start of the trial and the start of the lever press. After a successful trial, the latency to initiate the next action is longer. By contrast, latency after a failed trial was shorter (planned comparisons, $p < 0.001$ for all criterion durations).

How does the outcome of any trial affect performance on the next trial? As shown in Figure 1*F*, after an unrewarded trial, the next press is on average much longer in duration, but after a rewarded trial, the next press is shorter (Fig. 1*F*, planned comparisons, $p < 0.001$ for all criterion durations). Thus, the reward not only affected the latency to initiate the next action, but also the duration of the next action. These results suggest that, contrary to traditional accounts of reinforcement learning (Thorndike, 1911; Sutton and Barto, 1998), the food reward does not simply reinforce a particular action duration, i.e., cause it to be repeated. Rather, the animal dynamically adjusts its behavioral output based on previous feedback. After earning a reward with a sufficiently long lever press, often the animal reduced the duration of the next press, as if to test whether a shorter duration would be sufficient. Such data illustrate the impact of reward

feedback on motor exploration in mice (Brainard and Doupe, 2000; Sober et al., 2008; Andalman and Fee, 2009).

Single-unit activity in the substantia nigra

Using multielectrode arrays, we recorded single-unit neural activity from 666 neurons in five mice. The histological analysis shows clear marks of electrode tracks; all the electrode wires were implanted in the substantia nigra, in both the pars reticulata and pars compacta (Fig. 2*A*). We did not find any indication that any of the electrode wires were outside of these two regions.

Because the duration of the action potential from DA neurons is known to be significantly longer than that of GABA neurons (Grace and Bunney, 1983), we were able to classify the recorded neurons based on spike duration (Fig. 2). The neurons were classified as either GABAergic (GABA, $n = 580$) or dopaminergic (DA, $n = 86$) based on their waveforms (Wilson, 2004; Tepper and Lee, 2007). To verify that the long-duration spike waveforms are indeed from DA neurons, we compared firing rates before and after the injection of quinpirole, a D_2-like dopamine receptor agonist (1 mg/kg, i.p.; 3 experiments in 2 mice, 14 neurons). Because D_2 autoreceptors are expressed on the terminals of DA neurons, quinpirole is expected to reduce the firing rate of DA neurons, but not of GABA neurons. Quinpirole dramatically reduced firing rate in putatively classified DA neurons, but did not alter the firing rate of putative GABA neurons (Fig. 2*E*). In addition, the average firing rate of GABA neurons were also higher than that of DA neurons (DA: 6.3 ± 1 Hz, $n = 86$; GABA: 10.2 ± 0.6 Hz, $n = 580$; unpaired *t* test, $p < 0.001$).

We classified the single-unit recording data into the following task-relevant categories: trial start, press duration, action boundaries (start and end of lever press), decision confidence, outcome-modulated, and reward (see Fig. 16 for summary). However, there is considerable overlap among these classes. The activity of most neurons is modulated by more than one behaviorally rele-
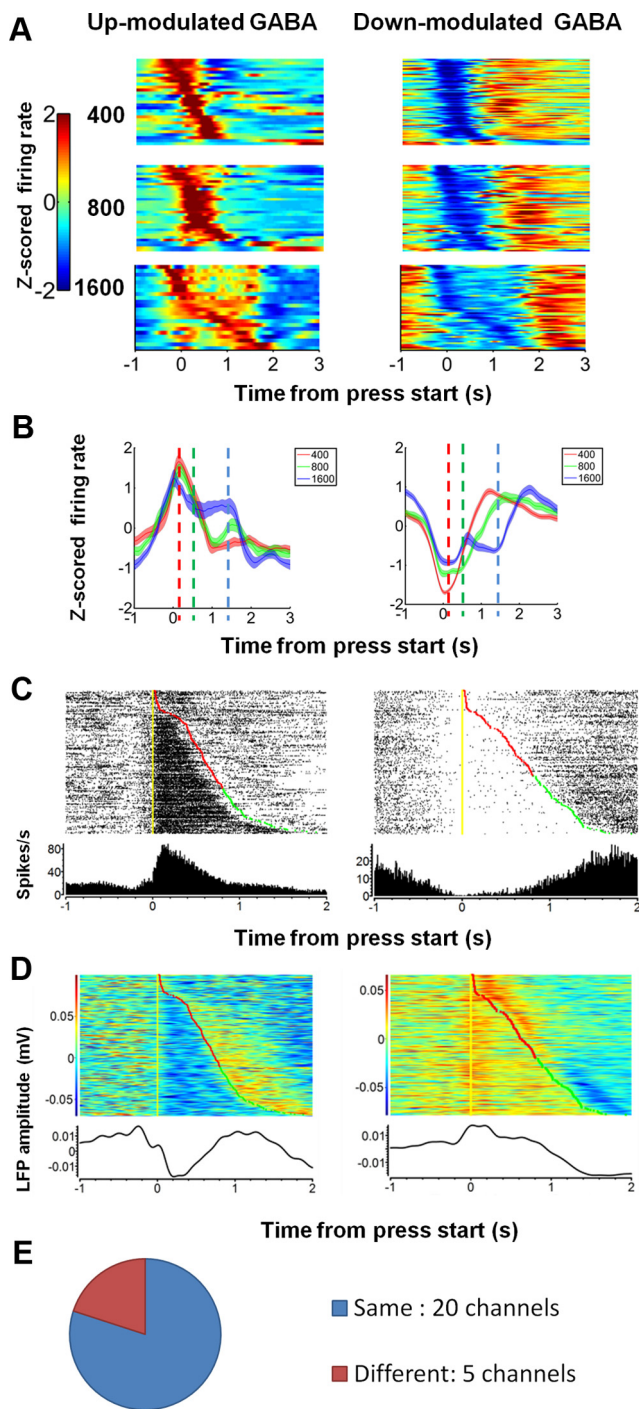
Figure 5. Neural activity reflecting action duration in GABA neurons. **A**, Spike density functions of upmodulated GABA neurons (n = 97) and downmodulated GABA neurons (n = 201) sorted by peak response time and duration criterion value (400, 800, and 1600 ms; bin size = 100 ms). Firing rates are z-score normalized. **B**, Normalized average firing rate of upmodulated and downmodulated neurons for each criterion duration (red, 400; green, 800; blue, 1600). Dotted line of corresponding color marks the duration criteria. **C**, Example raster plots. Each row is a single trial from a temporal differentiation session. Yellow markers indicate "Lever Start"; red markers indicate "Lever End Unrewarded"; green markers indicate "Lever End Rewarded." The trials are sorted according to the duration of the lever press, starting with trials with the shortest action durations on top. **D**, Perievent raster plots of LFP recorded from the same electrodes as the neurons shown in **C**. Left, In the same channel showing upmodulated activity, we also observed depolarization. Right, In the same channel showing downmodulated activity, we observed hyperpolarization. **E**, Pie chart showing the proportion of channels containing neurons with the same or different direction of modulation (each channel has more than one duration-modulated neuron). Same: a given channel contains duration coding neurons
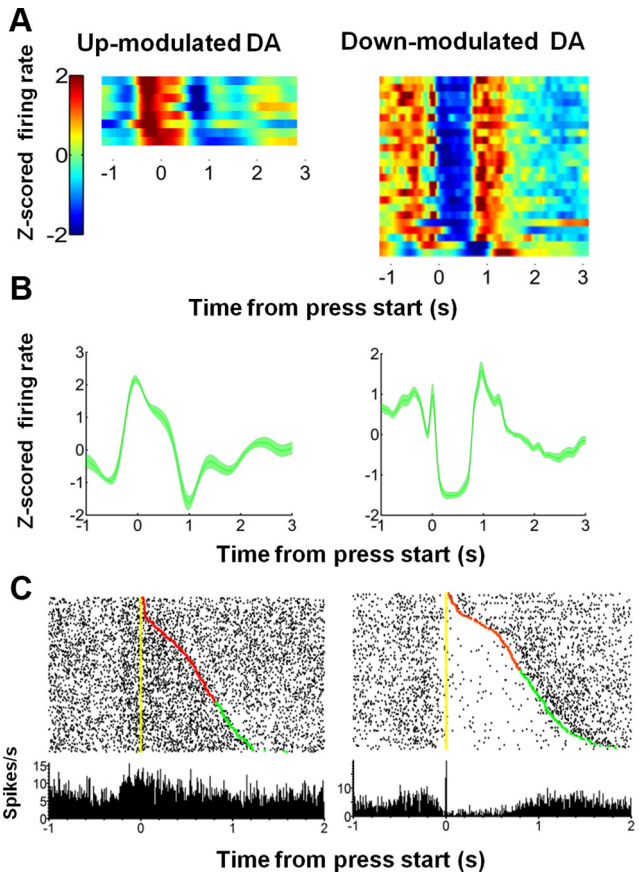
Figure 6. Neural activity reflecting action duration in DA neurons. **A**, Spike density functions of upmodulated (n = 8) and downmodulated DA neurons (n = 24) sorted by peak response time. Only neurons recorded during the >800 ms are shown here, as there were not enough DA neurons from the other sessions. Firing rates are z-score normalized. **B**, Population averages of all duration modulated cells with shaded region representing SEM. Firing rates are z-score normalized. **C**, Example raster plots. Each row is a single trial from a temporal differentiation session. Yellow markers indicate "Lever Start"; red markers indicate "Lever End Unrewarded"; green markers indicate "Lever End Rewarded." The trials are sorted according to the duration of the lever press, starting with trials with the shortest action durations on top.

vant event. The extent of overlap is illustrated in Table 1. It is therefore important to stress that, in most cases, neurons in the substantia nigra are modulated by more than one behavioral event as the animal performs the action.

Trial start

The activity of many neurons changed at the time of lever insertion, which signals the start of each trial. We compared the average firing rate before and after the insertion of the lever, and classified cells with significant changes in firing rate (p < 0.01) as "trial start" neurons. As shown in Figure 3, this type of neuron shows a short latency (< 50 ms) response after the lever insertion.

⟵

showing the same direction of modulation (either all upmodulated or all downmodulated). Different: a given channel contains neurons showing both up and down modulation. Of the 25 channels with more than one duration-modulated neurons, 20 contained neurons with the same direction of modulation during the lever press (i.e., neurons from that electrode are all upmodulated or all downmodulated). The remaining five channels had single units with both upmodulation and downmodulation. This pattern suggests that neurons that are spatially adjacent (recorded from the same electrode) tend to share the direction of modulation—i.e., there are clusters of action-on and action-off neurons.

A much higher proportion of DA neurons increased firing rate in response to the trial start signal (see Fig. 16 A, 48%, 41/86 of DA neurons vs 21%, 122/580 of GABA neurons, $\chi^2 = 34.1$, $p < 0.0001$). Neurons that decreased firing in response to trial start are more common among GABA neurons than DA neurons (GABA: 11.2%, 65/580; DA: 1.2%, 1/86, $\chi^2 = 8.46$, $p < 0.01$).

In many cells, however, the trial start response was not unconditionally elicited by lever insertion; rather, it depended on the outcome of the previous trial. In other words, neural activity in response to the start of the trial reflected previous success and failure (Fig. 4). Figure 4 B shows cells that fire at the trial start only when the previous trial was a failure (unrewarded). We used a ROC analysis (see Materials and Methods) to quantify the outcome discrimination shown by these neurons (see Fig. 15). If outcome selectivity is low, the AUC is close to 0.5, i.e., the firing rate is the same regardless of the previous outcome. Only cells with significant outcome selectivity (AUC significantly greater or <0.5) were classified as "outcome-modulated" neurons ($p < 0.01$).

Neurons that increased their firing rates at the trial start signal, after a previous failure (trial start outcome modulated), were equally common among DA and GABA neurons (14%, 82/580 of GABA neurons and 11%, 9/86 of DA neurons, $\chi^2 = 0.86$, $p > 0.05$). Thus, the outcome of a trial can have a lasting impact on neural activity on the next trial. Previous work on the prefrontal cortex in monkeys has found similar activity (Histed and Miller, 2006), but to our knowledge this is the first report of such activity in the substantia nigra in any species.

Press duration

Most nigral neurons changed their firing pattern during the acquisition of the temporal differentiation task. Some neurons increased firing (action-on), whereas others decreased firing (action-off) during the lever press (Figs. 5, 6). Action-on neurons were equally common among DA and GABA neurons (17%, 97/580 of GABA neurons; 12%, 10/86 of DA neurons, $p > 0.05$), as were action-off neurons (35%, 201/580 of GABA neurons; 41%, 35/86 of DA neurons, $p > 0.05$).

LFP recordings from the same channels showed a similar pattern. As shown by Figure 5D, LFP recorded from the same electrode as the action-on unit in Figure 5C showed a significant reduction in voltage, which indicates a net depolarization of the neuronal pop-
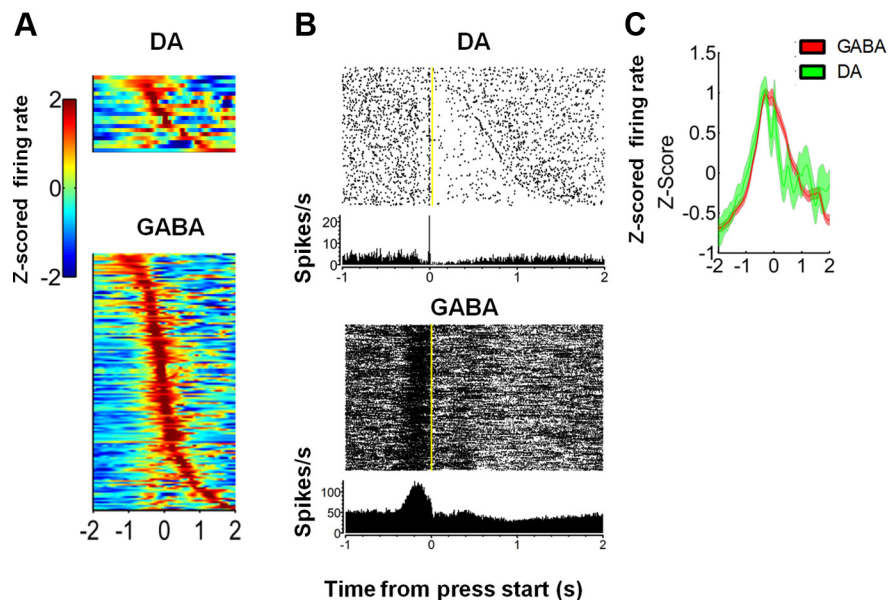

Figure 7. Neural activity in relation to press start (upmodulated). **A**, Spike density functions sorted by latency to peak response time for DA ($n = 19$) and GABA ($n = 145$) neurons. x-axis indicates time from the start of the lever press and each row represents activity from a single neuron. Firing rates are z-score normalized. **B**, Raster plots showing representative upmodulated neurons with respect to the start of the press. Each row is a single trial from a temporal differentiation session. Note the increase in firing rate just before the start of the press. Yellow markers indicate "Lever Start." **C**, Population averages of all upmodulated neurons with respect to the start of the press with shaded region representing SEM. Firing rates are z-score normalized.
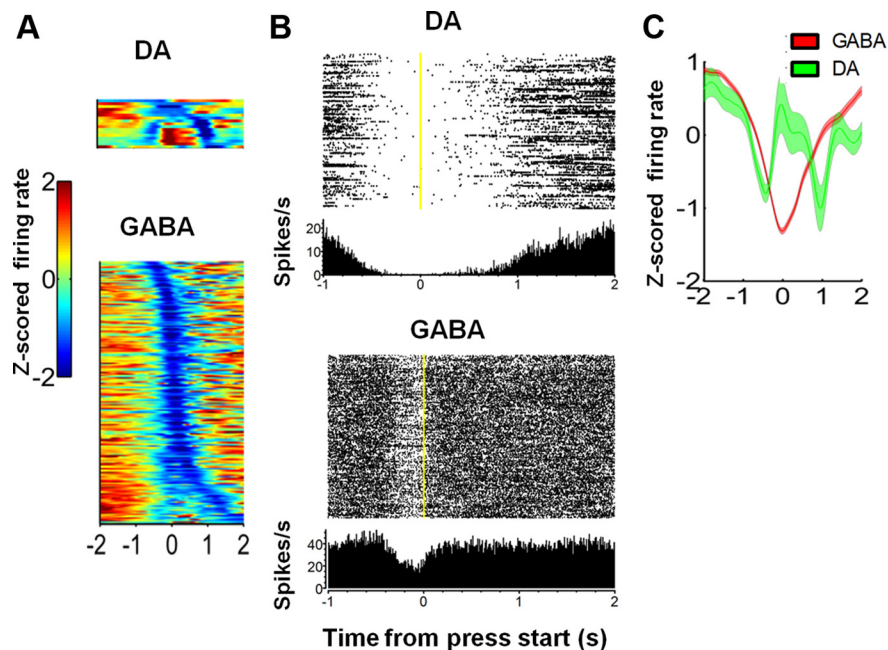

Figure 8. Neural activity in relation to press start (downmodulated). **A**, Spike density functions of DA ($n = 17$) and GABA ($n = 197$) neurons with respect to the start of the lever press, sorted by the latency to the minimum value. Firing rates are z-score normalized. x-axis indicates time from trial start (lever insertion) and each row represents activity from a single neuron from a temporal differentiation session. Firing rates are z-score normalized. **B**, Raster plots showing representative downmodulated neuron with respect to the start of the press. Each row is a single trial from a temporal differentiation session. Note the decrease in firing rate just before the start of the press. Yellow markers indicate "Lever Start." **C**, Population averages of all downmodulated press initiation cells with shaded region representing SEM. Firing rates are z-score normalized.

ulation contributing to the LFP. On the other hand, from the electrode recording the action-off single unit in Figure 5C, the LFP showed a general increase in voltage during the lever press, suggesting a net hyperpolarization of the local neuronal popula-
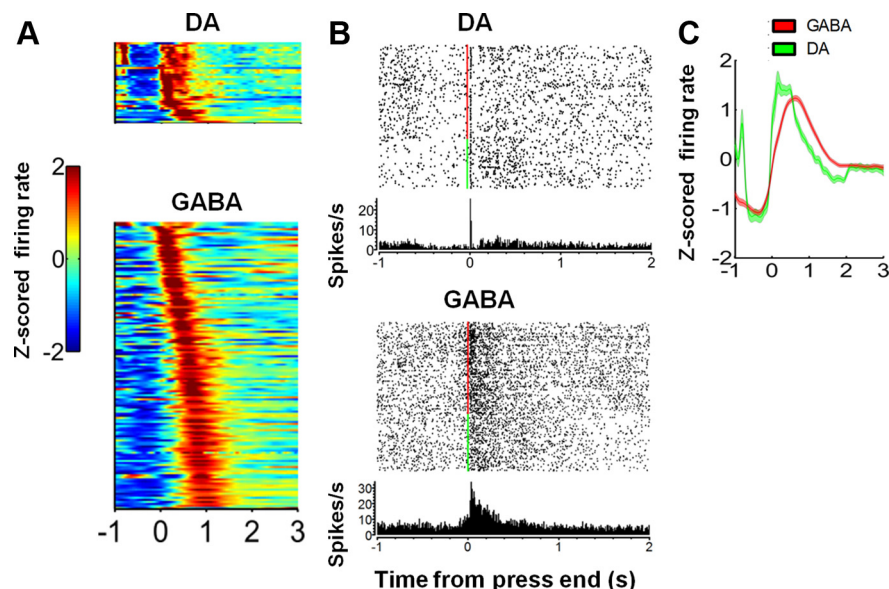
Figure 9. Press termination (upmodulated). **A**, Spike density function of GABA neurons ($n = 140$) and DA neurons ($n = 36$) sorted by the latency to the maximum value. Firing rates are z-score normalized. **B**, Raster plots showing representative upmodulated neurons with respect to the end of the press. Each row is a single trial from a temporal differentiation session. Note the increase of firing rate right after the end of the press regardless of the trial outcome (rewarded or not). Red markers indicate "Lever End Unrewarded"; green markers indicate "Lever End Rewarded." **C**, Population averages of all upmodulated "press termination" cells with shaded region representing SEM. Firing rates are z-score normalized.
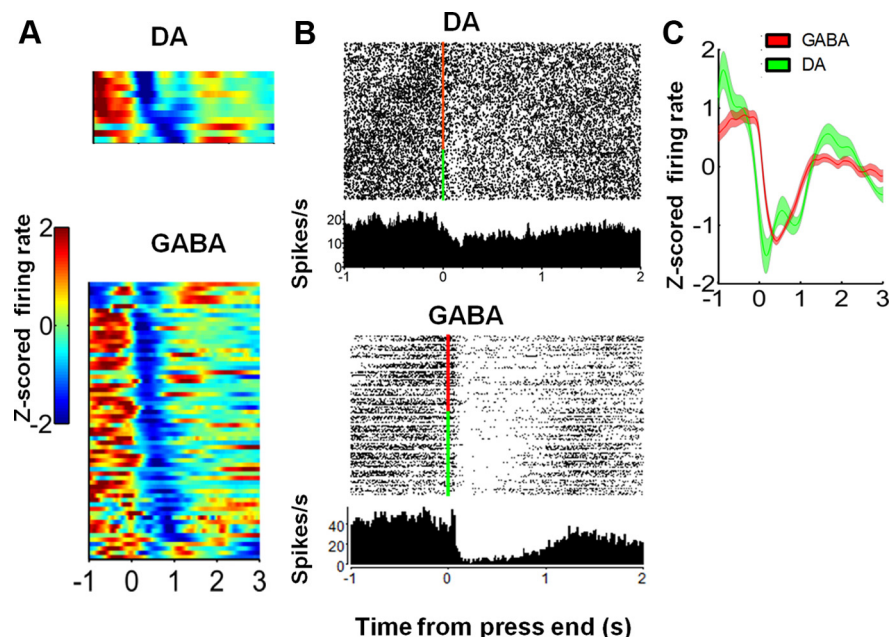

Figure 10. Press termination (downmodulated). **A**, Spike density function of GABA neurons ($n = 63$) and DA neurons ($n = 11$) sorted by the latency to the minimum value. Firing rates are z-score normalized. **B**, Raster plots showing representative downmodulated neurons with respect to the end of the press. Each row is a single trial from a temporal differentiation session. Note the decrease in firing rate right after the end of the press, regardless of the trial outcome (rewarded or not). Red markers indicate "Lever End Unrewarded"; green markers indicate "Lever End Rewarded." **C**, Population averages of downmodulated "press termination" cells with shaded region representing SEM. Firing rates are z-score normalized.

tion. These results suggest that the population activity in the recorded region is similar to the single unit activity, suggesting local domains of action-on and action-off neurons that are depolarized and hyperpolarized. To test this possibility, we examined the 25 channels with at least two single units encoding press duration. As shown in Figure 5E, of the 25 channels, 20 contained

neurons with the same direction of modulation during the lever press (i.e., neurons from that electrode are all upmodulated or all downmodulated). The remaining five channels had single units with both upmodulation and downmodulation. If we assume that neurons recorded by the same electrode wire are more likely to be spatially adjacent, then this analysis supports the hypothesis that local domains within the substantia nigra contain clusters of action-on or action-off neurons.

Action boundary

The activity of many neurons was modulated just before the initiation of the action (pressing the lever) and immediately following action termination (release of the lever) (Figs. 7–10). The modulation of their activity just preceding or following the lever press appears to mark the action boundaries. There is significant overlap between the action boundary neurons and the action-on and action-off neurons (Table 1). To take the most extreme example, the DA neuron in Figure 7B is classified as both "action initiation," "action termination," and action-off.

Neurons that increased firing just before the lever press are equally common among DA neurons and GABA neurons (22%, 19/86 of DA neurons; 25%, 145/580 of GABA neurons, $\chi^2 = 0.52, p > 0.1$, see Figs. 7, 16). On the other hand, cells that decreased firing before the press are more common among GABA neurons (34%, 197/580 of GABA neurons and 20%, 17/86 of DA neurons, $\chi^2 = 6.92, p < 0.01$, see Figs. 8, 16).

Neurons that increased firing just after the press are more common among DA neurons (42%, 36/86 of DA neurons; 24%, 140/580 of GABA neurons, $\chi^2 = 12.1, p < 0.0005$, see Figs. 9, 16). But cells that decreased firing after the lever press are more common among GABA neurons (13%, 11/86 of DA neurons; 11%, 63/580 of GABA neurons, $\chi^2 = 0.28, p > 0.5$, see Figs. 10, 16).

Decision confidence

When the press was too short, the animal rarely entered the food cup. Thus the probability of food cup entry (~1–2 s after the release of the lever) can indicate the animal's confidence about the completed action and the prospect of reward (Figs. 11, 12). The activity of some neurons immediately after the press could predict the probability of food cup entry (see Fig. 15B, ROC analysis, $p < 0.01$). By plotting the firing rate immediately after the press as a function of relative action duration (normalized to the criterion duration) on the last session of train-
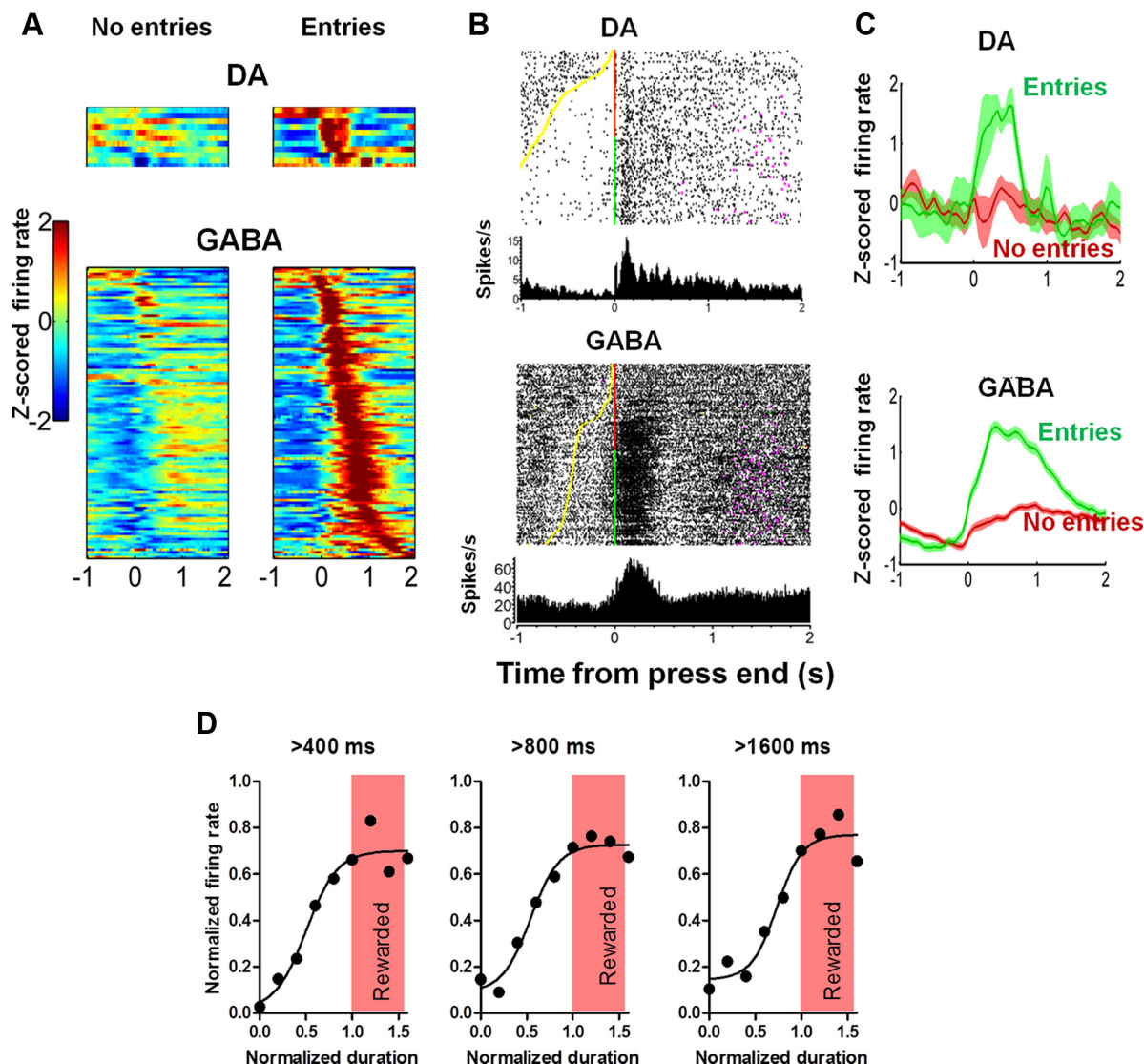
Figure 11. Neural activity encoding confidence (upmodulated). *A*, Spike density functions of DA (*n* = 11) and GABA (*n* = 112) neurons sorted by the latency to the maximum value (upmodulated). Each row represents activity from a single neuron from a temporal differentiation session. Firing rates are *z*-score normalized. *B*, Raster plots showing representative upmodulated neurons. Each row is a single trial from a temporal differentiation session. Yellow markers indicate the start of a trial; red markers indicate the end of unrewarded presses; green markers indicate the end of rewarded presses; purple markers indicate first food cup entry after press. The trials are sorted according to the duration of the lever press, starting with trials with the shortest action durations on top. Note the increase in firing rate immediately following the end of presses is modulated by the duration of the lever press just completed. Such activity occurs well before the detection of reward, and predicts the probability of head entries into the food cup (purple marker). *C*, Population averages of perievent spike density functions for upmodulated confidence-coding cells, with shaded regions representing SEM. Firing rates are *z*-score normalized. *D*, Firing rate as a function of the relative press duration. The *x*-axis shows press duration, normalized by the criterion duration, and the *y*-axis the normalized firing rate. The shaded region represents rewarded presses (1 being the criterion duration). The firing rates in each bin were normalized (baseline subtracted). This was done for the last day of each duration criterion. The curve is a sigmoid fitted to the data.

ing, we obtained a "neurometric" function for decision confidence (Figs. 11 *D*, 12 *D*).

Cells that increased firing with increasing decision confidence of receiving a reward are equally common among DA and GABA neurons (13%, 11/86 of DA neurons; 19%, 112/580 of GABA neurons, $\chi^2 = 2.1$, $p > 0.05$, see Figs. 11, 16), as are cells that are downmodulated by decision confidence (10%, 9/86 of DA neurons; 10%, 58/580 of GABA neurons, $\chi^2 = 0.02$, $p > 0.05$, see Figs. 12, 16).

The activity of these neurons is determined by the duration of the completed action. But instead of reflecting the absolute action duration, their firing rate reflects the relative duration of the press, i.e., how likely such a duration will be followed by a food reward. Our finding is in accord with previous work showing that

SNr activity is modulated by the likelihood of making a saccade to a particular target (Basso and Wurtz, 2002), yet for the first time we were able to relate the observed activity to the efficacy of the completed action.

One alternative explanation for these data is that the neurons simply modulate their activity in response to the sound of the food pellet, which was delivered following a successful trial. It is difficult to hear the sound of pellet delivery because of the noise produced by the lever retraction at the same time. And more importantly, neither the neural activity nor the food cup entry behavior reflected whether the reward was actually delivered. Were that the case, the unrewarded presses would have been followed by no significant changes in neural activity and no food cup entry while the rewarded presses would have been fol-
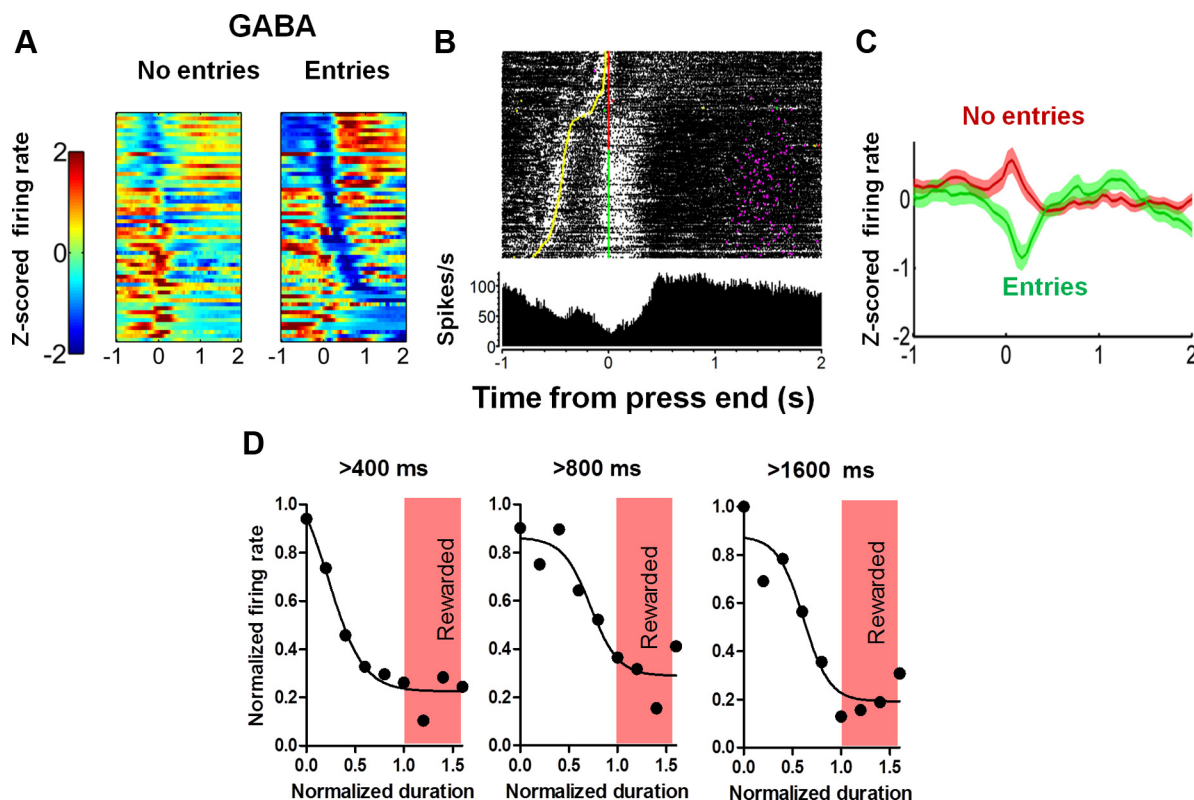
Figure 12. Neural activity encoding confidence (downmodulated). **A**, Spike density functions of GABA ($n = 58$) neurons sorted by the latency to the minimum value (downmodulated). Only GABA neurons are shown, because there were not enough DA neurons ($n = 9$) in this category. Each row represents activity from a single neuron from a temporal differentiation session. Firing rates are z-score normalized. **B**, Raster plots showing representative downmodulated neurons. Each row is a single trial from a temporal differentiation session. Yellow markers indicate the start of a trial; red markers indicate the end of unrewarded presses; green markers indicate the end of rewarded presses; purple markers indicate first food cup entry after press. The trials are sorted according to the duration of the lever press, starting with trials with the shortest action durations on top. Note the changes in firing rate immediately following the end of presses with longer durations, including rewarded presses. These firing rate changes occur well before the detection of reward. **C**, Population averages of perievent spike density functions for downmodulated confidence-coding cells. Firing rates are z-score normalized. **D**, Firing rate as a function of the relative press duration. The x-axis shows press duration, normalized by the criterion duration, and the y-axis the normalized firing rate. The shaded region represents rewarded presses (1 being the criterion duration). The firing rates in each bin were normalized (baseline subtracted). This was done for the last day of each duration criterion. The curve is a sigmoid fitted to the data.

lowed by significant modulation and food cup entry. But the confidence-related activity was found to be gradually increasing as a function of the previous lever press duration. It was clearly present following a press that was nearly long enough but did not result in a pellet delivery. As shown in Figures 11D and 12D, the firing rate of the "confidence" neurons is a sigmoid function of the normalized duration of the completed action. Since the sound of the pellet delivery is all or none, no psychometric function of sound detection would resemble the confidence function.

Reward

Although the food pellet is delivered as soon as the lever is released following a successful press, the mouse does not come into contact with the pellet until it enters the food cup, which is located ~5 cm away from the lever. To determine the response to reward collection, the firing rate of neurons at the first magazine entry after a press was calculated (Figs. 13, 14). For trials in which the first food cup entry occurred within 4 s of the end of the lever press, the 1 s window after the first magazine entry was used to measure the firing rate response. For other trials, the 1 s window after the mean latency to the first magazine entry of that session was used instead.

We used ROC analysis to quantify reward selectivity (Fig. 15C). Units whose AUC statistic is significant ($p < 0.01$) are

marked as either upmodulated or reward-on cells (AUC > 0.5), or downmodulated or reward-off cells (AUC < 0.5).

Reward-modulated units are equally common among DA and GABA neurons (13%, 11/86 of DA neurons; 14%, 82/580 of GABA neurons, $\chi^2 = 0.1$, $p > 0.05$), as are reward-off cells (21%, 18/86 of DA neurons; 15%, 89/580 of GABA neurons, $\chi^2 = 1.7$, $p > 0.05$).

Plasticity during learning

The proportion of GABA and DA cells modulated by each task-related event is shown in Figure 16A. With the exception of trial start, lever start, and lever end, there were no significant differences in the proportion of each cell type correlated with the task event. That is, the task-related modulation was surprisingly similar for DA and GABA neurons.

Recording from naive mice as they learned to press the lever, we were able to collect data on the change in neural activity over time during the learning of temporal differentiation. With training, the activity of most neurons reflected the temporal duration of each lever press. Such duration-related activity was not apparent when there was no press duration requirement, during initial training under continuous reinforcement, but rapidly emerged when mice had to hold down the lever for a minimum duration (Fig. 16B, 19%, 22/115 of CRF neurons; 63%, 112/179 of >400 ms neurons, $\chi^2 = 53.3$, $p < 0.0001$).
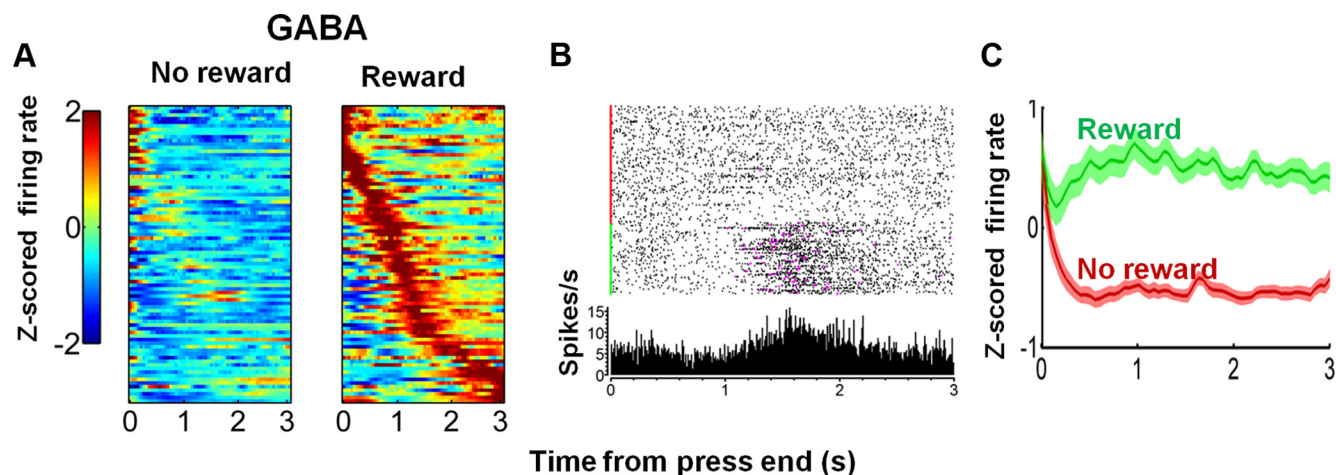
Figure 13. Neural activity encoding reward (upmodulated). *A*, Spike density functions and GABA (*n* = 82) neurons sorted by the latency to the maximum value (upmodulated). Only GABA neurons are shown here, as there were not enough DA neurons. Firing rates are *z*-score normalized. *B*, Raster plots showing representative upmodulated neurons. Each row is a single trial from a temporal differentiation session. Purple markers indicate first food cup entry after press; red markers indicate the end of unrewarded presses; green markers indicate the end of rewarded presses. Note the increase in firing rate at head entries after a rewarded press. *C*, Population averages of perievent spike density functions for upmodulated reward-coding related cells. Firing rates are *z*-score normalized.
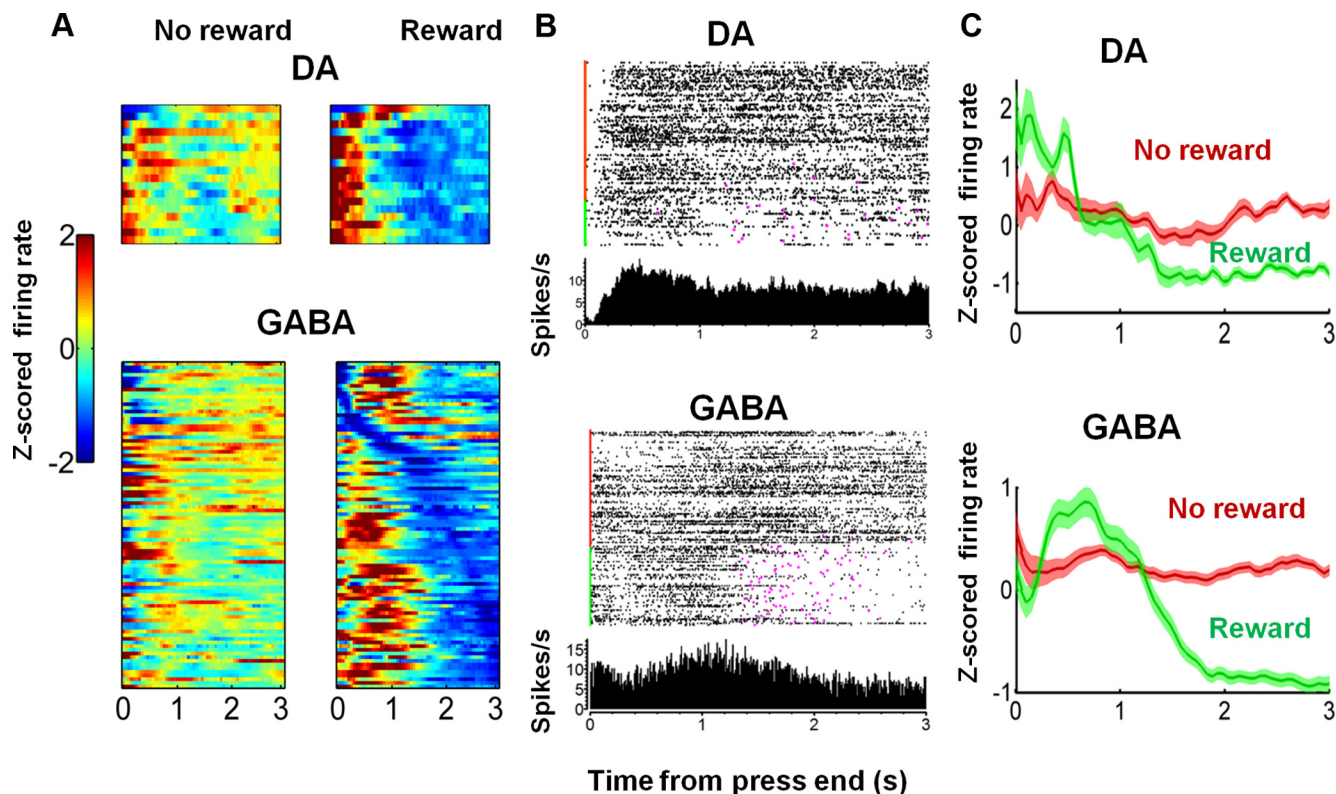

Figure 14. Neural activity encoding reward (downmodulated). *A*, Spike density functions of DA (*n* = 18) and GABA (*n* = 89) neurons sorted by the latency to the minimum value (downmodulated). Each row represents activity from a single neuron from a temporal differentiation session. Firing rates are *z*-score normalized. *B*, Raster plots showing representative downmodulated neurons. Each row is a single trial from a temporal differentiation session. Purple markers indicate first food cup entry after press; red markers indicate the end of unrewarded presses; green markers indicate the end of rewarded presses. Note the decrease in firing rate at head entries after a rewarded press. *C*, Population averages of perievent spike density functions for downmodulated reward-coding related cells, normalized by *z*-score, with shaded region representing SEM. Firing rates are *z*-score normalized.

Furthermore, confidence-encoding neurons increased as the duration requirement was increased (25%, 45/179 of >400 ms neurons; 44%, 73/166 of >1600 ms neurons, χ^2 = 13.6, $p <$ 0.0005). Similar plasticity during learning was observed with lever start neurons (39%, 45/115 of CRF neurons; 69%, 115/166 of CRF neurons, χ^2 = 25.2, $p <$ 0.0001) and lever end

neurons (23%, 26/115 of CRF neurons; 40%, 67/179 of CRF neurons, χ^2 = 9.67, $p <$ 0.002).

Discussion

Here, we used action duration as the operant to tag neural activity in the substantia nigra involved in action selection and timing.
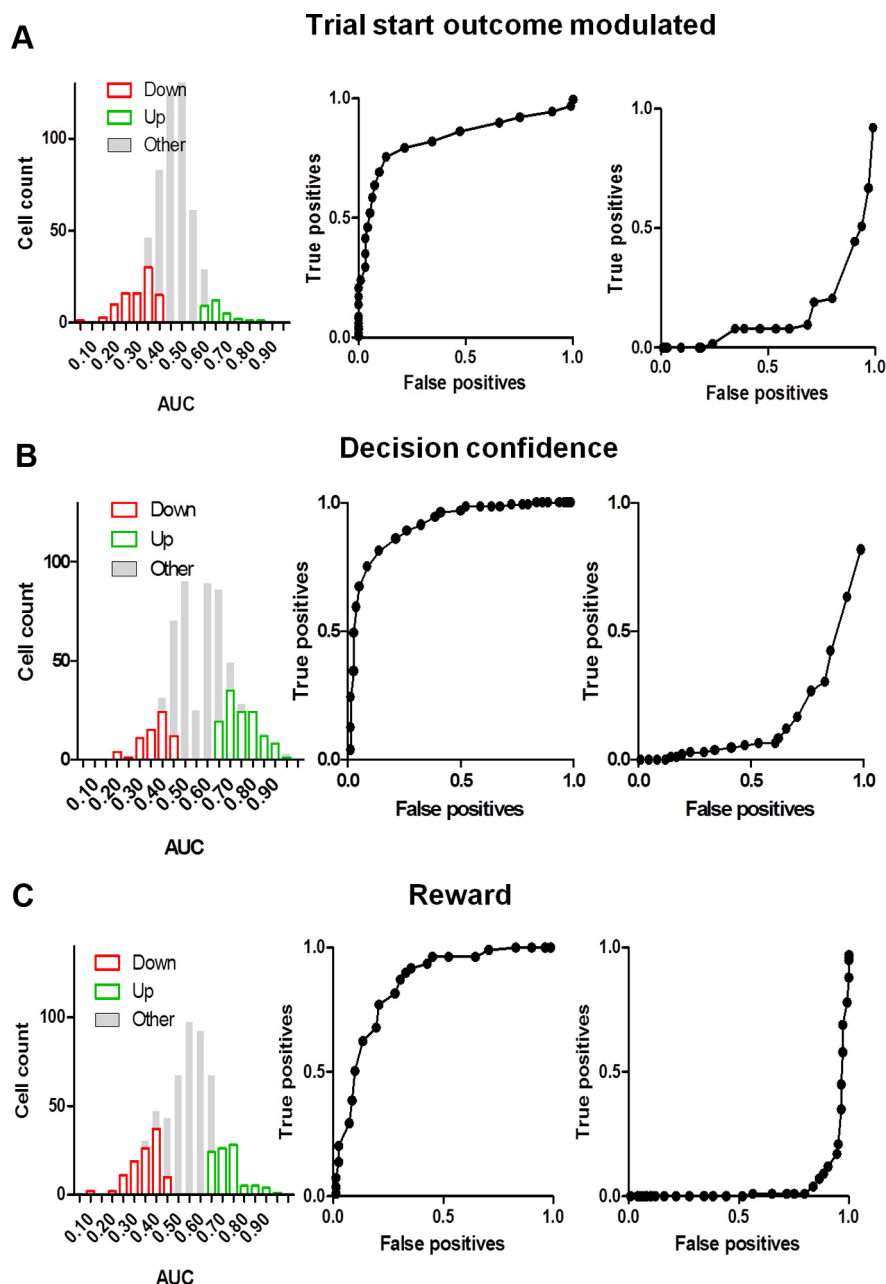
Figure 15. ROC analysis of reward detection and decision confidence. *A*, ROC analysis of trial start outcome-modulated neurons. From left to right: histogram of AUC values for all neurons. AUC values closer to either 0 or 1 indicate neurons with greater behaviorally discriminative power. Statistically significant AUC values (*p* < 0.01) are shown in color as "Down" or "Up," with statistically insignificant AUC values in gray. ROC curve for an "Up" neuron, whose firing rate increases following a reward (AUC = 0.83). ROC curve for a "Down" neuron, whose firing rate decreases following a previous success (AUC = 0.15). *B*, ROC analysis of confidence-modulated neurons. From left to right: histogram of AUC values for all neurons. ROC curve for an "Up" neuron, whose firing rate increases in anticipation of reward (AUC = 0.91). ROC curve for a "Down" neuron, whose firing rate decreases in anticipation of reward (AUC = 0.15). *C*, ROC analysis of reward detection. From left to right: histogram of AUC values for all neurons. ROC curve for a neuron ("Up") whose firing rate increases at the detection of reward (AUC = 0.85). ROC curve for a neuron ("Down") whose firing rate decreases at the detection of reward (AUC = 0.04).

We observed significant changes in nigral activity as mice learned to time their actions (Fig. 16). When no minimum duration was required, nigral activity rarely reflected press duration. Yet when a minimum action duration was required to earn rewards, behavior as well as neural activity adapted with training. The observed plasticity in the substantia nigra during learning has important implications. Skill learning is often thought to depend on the motor cortex (Shmuelof and Krakauer, 2011), although the refining of action timing, an essential component of skill learning, does not apparently require the primary motor cortex in mice (Yin, 2009). Results from our study clearly implicate the nigral output in the acquisition and performance of the learned action. Since the SNr GABA neurons can inhibit the motor thalamocortical circuit, our results do not necessarily rule out a role for the motor cortex. Rather, they underscore the need to consider the entire corticobasal ganglia network when studying motor learning (Ölveczky, 2011).

DA versus GABA neurons

The GABA neurons in the SNr inhibit downstream brainstem and midbrain structures (Hikosaka, 2007). The nearby DA neurons in the SNc, however, are part of the ascending modulatory projections to the cerebrum, particularly the striatum. While these neuronal populations project to distinct anatomical regions, they share common inputs, for example from the striatum (Yoshida and Precht, 1971; Gerfen and Wilson, 1996). Their activity during temporal differentiation can be quite similar in relation to a number of behavioral events (Fig. 16). Such similarity is probably due to common inputs, e.g., medium spiny neurons from the striatum project to both DA and GABA neurons in the nigra. If a common striatal input is responsible for the similarity in task-related modulation between DA and GABA neurons, then the striatum is also expected to show similar patterns of plasticity and correlation with behavior, a possibility that remains to be tested (Gerfen, 1992; Joel and Weiner, 2000).

Yet the differences between DA and GABA neurons are also striking. DA cells burst in response to the trial start signal (Fig. 3), whereas GABA cells rarely do (Fig. 16*A*). This "trial start" response, found in approximately half the DA neurons recorded, is similar to the well known phasic DA activity in response to reward predictors (Schultz, 1998). Unlike a pavlovian conditional stimulus, however, lever insertion is not a valid predictor of reward, since the reward outcome is contingent upon the production of the correct action (Yin et al., 2008). The observed activity is the same whether or not the animal earns any reward on that trial (Fig. 3). Rather than elicited by a reward predictor and modified by prediction errors, it is uniformly responsive to a discriminative stimulus indicating the possibility for action. Given the short latency of the observed responses (<50 ms), which allows very little time for the processing of the stimulus, our results may be more in accord with the possibility that the
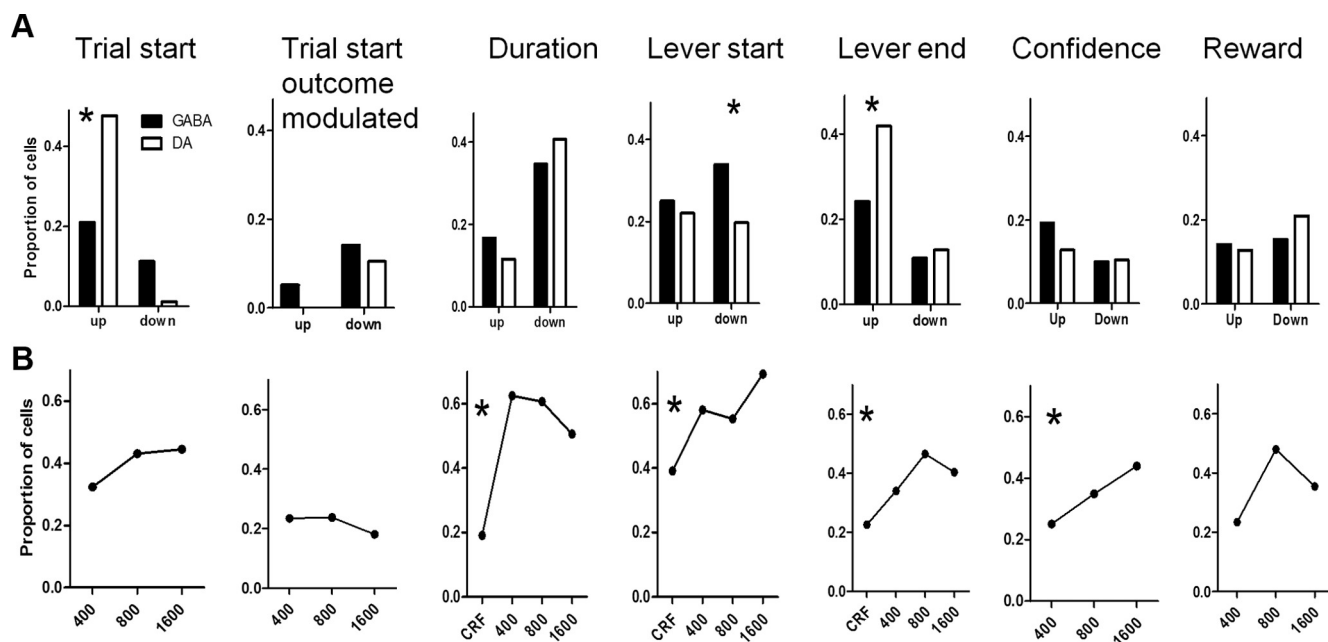
Figure 16. Summary of plasticity during training and proportion of neurons in each category. **A**, Proportion of GABA and DA cells modulated by each task-related event. Up, Upmodulated. Down, Downmodulated. *$p < 0.05$. **B**, Changes in the proportion of task-related neurons in the course of training. Upmodulated and downmodulated neurons are combined. CRF, Continuous reinforcement training with no minimum duration requirement; 400, >400 ms; 800, >800 ms; 1600, >1600 ms. *$p < 0.05$.

phasic activity of DA reflects sensory salience (Redgrave et al., 1999b). Direct projections from the superior colliculus, which shows short-latency activity following visual stimuli, could be responsible for evoking such responses (McHaffie et al., 2006).

Decision confidence

Because previous studies on decision confidence only manipulated sensory noise, without changing the feedback function linking the action and reward (Kepecs et al., 2008; Kiani and Shadlen, 2009), they tell us more about perceptual functions than about decision processes. Here, we show that neural activity immediately after the lever press can reflect confidence about the just completed action. Although the pellet is delivered into the food cup following a successful press, the mouse must enter the food cup to collect it. The probability of food cup entry, the consummatory behavior, is therefore a measure of subjective confidence about the efficacy of the instrumental action. The mouse rarely attempted to collect the food when the action duration was significantly shorter than the criterion duration (Fig. 11B). The "decision confidence" neurons (Figs. 11–12) are modulated by the learned instrumental contingency that defines how long the action duration should be, and their activity can predict the likelihood of food cup entries (Figs. 11D, 12D).

Opponent activity during action

A major finding here is the emergence of action-on and action-off neurons whose activity reflects the duration of the action. Although the heterogeneity among DA neurons is now widely recognized (Matsumoto and Hikosaka, 2009; Henny et al., 2012), we show for the first time that some DA neurons can show opponent activity during an action (Fig. 6). DA neurons are known to receive both excitatory and inhibitory projections from many brain regions, and send projections to cortical and striatal regions. Recent work suggests that SNc DA neurons vary in the strength of inhibitory input (Henny et al.,

2012), but it remains unknown whether the action-on and action-off DA neurons have distinct patterns of anatomical connectivity.

Most neurons that increase or decrease their activity for the duration of the action, however, are GABAergic. As the main output neurons of the basal ganglia, these neurons are known to inhibit the brainstem, tectum, and thalamus (Grillner et al., 2005; Hikosaka, 2007). Previous work has shown that pausing in the GABAergic output neurons is correlated with the initiation of voluntary saccades in monkeys (Hikosaka and Wurtz, 1983). We also observed pauses in nigral activity, especially in GABA neurons, both preceding and during the lever press. However, we observed increases in firing among action-on neurons (Fig. 5), whose activity appears to be a mirror image of that of the action-off neurons. These results raise two questions: (1) which inputs are responsible for the observed increases and decreases in activity during action, and (2) what is the functional significance of such bidirectional outputs? Let us consider these *seriatim*.

Classic work has established that the pause in SNr activity is a result of inhibition by the striatum (Yoshida and Precht, 1971; Hikosaka et al., 2000). Although we cannot rule out enhanced inhibition from the external globus pallidus, or inhibition from the collaterals of other GABA neurons in SNr (Tepper and Lee, 2007), the direct pathway from striatonigral neurons is the most likely source of inhibition for the action-off neurons. More difficult to answer, however, is the question of which inputs excite the action-on neurons. The two major possibilities are disinhibition and direct excitation. Disinhibition requires tonic inhibition from the external globus pallidus, which is lifted only during the action, as the relevant pallidal neurons pause. But it is not easy to reconcile tonic inhibition of nigral neurons with their high baseline firing rate. Direct excitation is more plausible given the known glutamatergic innervation of nigral neurons originating in the subthalamic nucleus (Bolam et al., 1993, 2000; Bevan et al.,

1996; Gerfen and Wilson, 1996; Nambu et al., 2002). Of course, direct excitation and indirect disinhibition are not mutually exclusive, and a combination of the two can be responsible for the observed increase in nigral activity during lever pressing.

If the action-on and action-off nigral neurons receive distinct sets of inputs, then neurons that receive excitatory input can become action on neurons, whereas those that receive inhibitory input can become action off neurons. But it is unclear how such exclusive anatomical connections can develop quickly through learning. Alternatively, each nigral GABA neuron can receive all three inputs: inhibitory inputs from the striatonigral pathway, disinhibitory inputs from the striatopallidal pathway, and excitatory inputs from the subthalamic nucleus (Hikosaka et al., 2000). If the net excitatory input exceeds the net inhibitory input, then the activity of the SNr neuron will increase, and vice versa should the inhibition exceed the excitation. The strengths of these inputs could be plastic, and, with learning, any GABA output neuron could be tuned to become either action-on or action-off neurons.

According to the standard model of the basal ganglia, the pause in nigral GABA neurons disinhibits the thalamus and motor initiation centers in the midbrain and brainstem (Hikosaka, 1989). However, here we found that two opponent channels of GABAergic outputs are sent simultaneously, one inhibiting and the other disinhibiting downstream structures. As we still do not know the destination of these two projections, it is difficult to understand their functional significance. In many cases, however, we observed increases and decreases in activity from a high baseline firing rate. This tonic level of nigral inhibitory output may be a "common-mode" signal sent to antagonistic downstream motor control systems, when there is no movement. Increases and decreases from the tonic output could maximize the action of one system while reducing the action of the antagonistic system, similar to the well known reciprocal inhibition at the level of spinal motor neurons innervating pairs of antagonistic muscles (Sherrington, 1906). Exactly where the opponent nigral outputs are directed, and how the downstream motor control systems coordinate the hundreds of muscles involved in any voluntary action, await future investigation.

References

Abeles M (1982) Quantification, smoothing, and confidence limits for single-units' histograms. J Neurosci Methods 5:317–325.

Andalman AS, Fee MS (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106:12518–12523.

Basso MA, Wurtz RH (2002) Neuronal activity in substantia nigra pars reticulata during target selection. J Neurosci 22:1883–1894.

Bevan MD, Smith AD, Bolam JP (1996) The substantia nigra as a site of synaptic integration of functionally diverse information arising from the ventral pallidum and the globus pallidus in the rat. Neuroscience 75:5–12.

Bolam JP, Smith Y, Ingham CA, von Krosigk M, Smith AD (1993) Convergence of synaptic terminals from the striatum and the globus pallidus onto single neurones in the substantia nigra and the entopeduncular nucleus. Prog Brain Res 99:73–88.

Bolam JP, Hanley JJ, Booth PA, Bevan MD (2000) Synaptic organisation of the basal ganglia. J Anat 196:527–542.

Brainard MS, Doupe AJ (2000) Auditory feedback in learning and maintenance of vocal behaviour. Nat Rev Neurosci 1:31–40.

Costa RM, Cohen D, Nicolelis MA (2004) Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr Biol 14:1124–1134.

Fan D, Rich D, Holtzman T, Ruther P, Dalley JW, Lopez A, Rossi MA, Barter

JW, Salas-Meza D, Herwik S, Holzhammer T, Morizio J, Yin HH (2011) A wireless multi-channel recording system for freely behaving mice and rats. PLoS One 6:e22033.

Gerfen CR (1992) The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. Annu Rev Neurosci 15:285–320.

Gerfen CR, Wilson CJ (1996) The basal ganglia. In: Handbook of chemical neuroanatomy (Swanson LW, Bjorklund A, Hokfelt T, eds), pp 371–468. Amsterdam: Elsevier.

Gold C, Henze DA, Koch C, Buzsáki G (2006) On the origin of the extracellular action potential waveform: A modeling study. J Neurophysiol 95:3113–3128.

Grace AA, Bunney BS (1983) Intracellular and extracellular electrophysiology of nigral dopaminergic neurons–1. Identification and characterization. Neuroscience 10:301–315.

Grillner S, Hellgren J, Ménard A, Saitoh K, Wikström MA (2005) Mechanisms for selection of basic motor programs - roles for the striatum and pallidum. Trends Neurosci 28:364–370.

Henny P, Brown MT, Northrop A, Faunes M, Ungless MA, Magill PJ, Bolam JP (2012) Structural correlates of heterogeneous in vivo activity of midbrain dopaminergic neurons. Nat Neurosci. Advance online publication. Retrieved February 12, 2012. doi: 10.1038/nn.3048.

Hikosaka O (1989) Role of basal ganglia in saccades. Rev Neurol 145:580–586.

Hikosaka O (2007) GABAergic output of the basal ganglia. Prog Brain Res 160:209–226.

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

Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol Rev 80:953–978.

Histed MH, Miller EK (2006) Microstimulation of frontal cortex can reorder a remembered spatial sequence. PLoS Biol 4:e134.

Hudzik TJ, Howell A, Georger M, Cross AJ (2000) Disruption of acquisition and performance of operant response-duration differentiation by unilateral nigrostriatal lesions. Behav Brain Res 114:65–77.

Joel D, Weiner I (2000) The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. Neuroscience 96:451–474.

Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. Nature 455:227–231.

Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. Science 324:759–764.

Macmillan NA, Creelman C (2004) Detection theory. Mahwah: Lawrence Erlbaum Associates.

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837–841.

McHaffie JG, Jiang H, May PJ, Coizet V, Overton PG, Stein BE, Redgrave P (2006) A direct projection from superior colliculus to substantia nigra pars compacta in the cat. Neuroscience 138:221–234.

Mita A, Mushiake H, Shima K, Matsuzaka Y, Tanji J (2009) Interval time coding by neurons in the presupplementary and supplementary motor areas. Nat Neurosci 12:502–507.

Mukamel R, Nir Y, Harel M, Arieli A, Malach R, Fried I (2011) Invariance of firing rate and field potential dynamics to stimulus modulation rate in human auditory cortex. Hum Brain Mapp 32:1181–1193.

Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. Neurosci Res 43:111–117.

Ölveczky BP (2011) Motoring ahead with rodents. Curr Opin Neurobiol 21:571–578.

Platt JR, Kuch DO, Bitgood SC (1973) Rats' lever-press durations as psychophysical judgements of time. J Exp Anal Behav 19:239–250.

Redgrave P, Prescott TJ, Gurney K (1999a) The basal ganglia: a vertebrate solution to the selection problem? Neuroscience 89:1009–1023.

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

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Sherrington CS (1906) The integrative action of the nervous system. New Haven: Yale UP.

Shimazaki H, Shinomoto S (2010) Kernel bandwidth optimization in spike rate estimation. J Comput Neurosci 29:171–182.

Shmuelof L, Krakauer JW (2011) Are we ready for a natural history of motor learning? Neuron 72:469–476.

Skinner BF (1938) The behavior of organisms. New York: Appleton-Century-Crofts.

Sober SJ, Wohlgemuth MJ, Brainard MS (2008) Central contributions to acoustic variation in birdsong. J Neurosci 28:10370–10379.

Sutton RS, Barto AG (1998) Reinforcement learning. Cambridge: MIT.

Tepper JM, Lee CR (2007) GABAergic control of substantia nigra dopaminergic neurons. Prog Brain Res 160:189–208.

Thorndike EL (1911) Animal intelligence: experimental studies. New York: Macmillan.

Wilson CJ (2004) Basal ganglia. In: The synaptic organization of the brain, Ed 5 (Shephard GM, ed). New York: Oxford UP.

Yin HH (2009) The role of the murine motor cortex in action duration and order. Front Integr Neurosci 3:23.

Yin HH (2010) The sensorimotor striatum is necessary for serial order learning. J Neurosci 30:14719–14723.

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464–476.

Yin HH, Ostlund SB, Balleine BW (2008) Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur J Neurosci 28:1437–1448.

Yin HH, Mulcare SP, Hilário MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM (2009) Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat Neurosci 12:333–341.

Yoshida M, Precht W (1971) Monosynaptic inhibition of neurons of the substantia nigra by caudato-nigral fibers. Brain Res 32:225–228.

Yu C, Gupta J, Yin HH (2010) The role of mediodorsal thalamus in temporal differentiation of reward-guided actions. Front Integr Neurosci 4. pii: 14.

Zeiler MD (1985) Pure timing in temporal differentiation. J Exp Anal Behav 43:183–193.