



Published in final edited form as:

Neuron. 2017 March 22; 93(6): 1436–1450.e8. doi:10.1016/j.neuron.2017.02.029.

Dynamic nigrostriatal dopamine biases action selection

Christopher D. Howard^{1,3}, Hao Li^{1,3}, Claire E. Geddes^{1,2}, and Xin Jin^{1,4,*}

¹Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037

²Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093

Summary

Dopamine is thought to play a critical role in reinforcement learning and goal-directed behavior, but its function in action selection remains largely unknown. Here, we demonstrate that nigrostriatal dopamine biases ongoing action selection. When mice were trained to dynamically switch the action selected at different time points, changes in firing rate of nigrostriatal dopamine neurons, as well as dopamine signaling in the dorsal striatum, were found to be associated with action selection. This dopamine profile is specific to behavioral choice, scalable with interval duration, and doesn't reflect reward prediction error, timing, or value as single factors alone. Genetic deletion of NMDA receptors on dopamine or striatal neurons, or optogenetic manipulation of dopamine concentration, alters dopamine signaling and biases action selection. These results unveil a crucial role of nigrostriatal dopamine in integrating diverse information for regulating upcoming actions and have important implications for neurological disorders including Parkinson's disease and substance dependence.

Introduction

Selecting and executing motor behaviors are critical brain functions essential to survival. Specifically, choosing the appropriate action and performing it at the right time is crucial for the success of an organism (Gallistel, 1990). In mammals, several models have implicated the basal ganglia, a series of interconnected subcortical nuclei including the striatum and dopaminergic inputs from substantia nigra, as playing a primary role in action selection (Mink, 2003; Redgrave et al., 1999) and spatiotemporal sequencing of behaviors (Graybiel, 1998; Hikosaka et al., 1998; Jin and Costa, 2015). Indeed, a wide range of neurological and psychiatric disorders including Parkinson's disease, Obsessive-Compulsive Disorder (OCD),

*Correspondence: xjin@salk.edu.

³Co-first author

⁴Lead Contact

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Contributions X.J. conceived the project. X.J., H.L. and C.H. designed the experiments. C.H. conducted and analyzed the voltammetric recordings. H.L. conducted and analyzed the electrophysiological recordings. C.H. and H.L. conducted and analyzed the behavioral and optogenetic studies. C.G. conducted aspects of behavioral training and voltammetric recordings. H.L. built the neuronal network model and ran the simulations. X.J. supervised all aspects of the work. C.H., H.L., and X.J. wrote the paper.

and substance dependence are characterized by major deficits in action selection and movement control, underscoring the importance of basal ganglia circuits in decision-making and appropriate organization of behavior (Everitt and Robbins, 2005; Graybiel and Rauch, 2000; Mink, 2003). Furthermore, striatal neuronal activity has been shown to signal the initiation and termination of behavior (Jin and Costa, 2010; Jin et al., 2014; Jog et al., 1999), and bias the selection of future actions (Ding and Gold, 2012; Samejima et al., 2005), both of which are potentially influenced by interactions with dopaminergic inputs from the midbrain (Kim et al., 2015; Morris et al., 2006; Schultz, 2007).

While the function of the midbrain dopamine system in reward-associated learning has been widely studied (Berridge and Robinson, 1998; Bromberg-Martin et al., 2010; Schultz et al., 1997), the explicit role of dopamine in controlling actions remains poorly understood (Jin and Costa, 2010; Kim et al., 2015; Matsumoto and Hikosaka, 2009; Redgrave and Gurney, 2006; Redgrave et al., 2010). This is especially true for the nigrostriatal dopamine system, which undergoes degeneration in Parkinson's disease, resulting in severe deficits in self-initiated action selection and voluntary movement control (Mink, 2003). Emerging evidence suggests that, beyond the influential view of encoding reward prediction error associated with reinforcement learning (Schultz et al., 1997), phasic activity of nigrostriatal dopamine neurons could also reflect salience or value of sensory stimuli (Kim et al., 2015; Matsumoto and Hikosaka, 2009; Morris et al., 2006), and dopamine signaling may therefore modulate sensorimotor striatum accordingly to direct immediate actions. Moreover, recent work has indicated that dopamine signals may be more tightly associated with behavioral choice (Morris et al., 2006; Roesch et al., 2007) or action initiation (Jin and Costa, 2010; Phillips et al., 2003; Syed et al., 2016) than previously thought, implying an underappreciated role of dopamine in directly controlling actions in addition to its involvement in learning. However, little is known about how dopamine influences online action selection or the decision-making process. Here, we trained mice to perform an operant discriminative task where they are encouraged to internally switch the action selected at different time points to achieve reward. Fast-scan cyclic voltammetry was utilized to monitor subsecond dopamine changes in the dorsal striatum during the choice behavior. We found a unique dopamine profile in the striatum associated with internal action selection, which was further confirmed using *in vivo* electrophysiology with optogenetic identification of dopaminergic neurons. Accordingly, genetic or optogenetic manipulation of nigrostriatal dopamine signaling modulated both immediate and delayed choice, and biased action selection. A computational model of basal ganglia network reproduces the effects of optogenetic and genetic manipulation, and reveals that dopamine could influence choice through regulation of striatal direct and indirect pathways. The model further suggests that the dopamine dynamic, rather than the constant dopamine level, is crucial for optimizing action selection. These results uncover a novel role of nigrostriatal dopamine in modulating upcoming actions.

Results

Specific striatal dopamine signals are associated with action selection

To address dopamine's role in action selection, we designed a task where mice chose between two actions according to self-monitored time intervals (Church and Deluty, 1977)

(Figure 1A, see Methods). Specifically, in an operant chamber, trials started with simultaneous retraction of left and right levers, and after either 2 or 8 s, both levers extended. Left or right lever selection following extension yielded reward (10 μ l sucrose) during 2 or 8 s trials, respectively. Only the first response following lever extension could be rewarded, and any lever press during the inter-trial-interval periods or incorrect responses led to no outcome. Trials with 2 or 8 s retraction intervals were randomly interleaved with equal probability. For the remainder of the text we will refer to this paradigm as the 2–8 s task. Mice acquire this task quickly and reach a correct rate of more than 85% after two weeks training (Figure 1B; one-way repeated measures ANOVA, significant effect of training days, $F_{13, 208} = 77.51$, $P < 0.0001$). Following training, when tested on trials with lever extension at probe intervals spanning 2, 2.5, 3.2, 4, 5, 6.3 and 8 s, the probability of mice selecting the lever associated with 8 s gradually increases with longer probe intervals (Figure 1C; one-way repeated measures ANOVA, significant effect of time, $F_{6, 36} = 97.61$, $P < 0.0001$). The resulting psychometric curve suggests that choice behavior is dynamic across 8 s trials. Specifically, animals are initially biased toward the short-duration option (left action), neutral at the midpoint (4 s), and biased to the long-duration option (right action) at later time points (Figure 1D, see Supplementary Video 1 for an example 8 s trial). This task thus provides a unique opportunity for studying the dynamic process of action selection in a self-paced manner, as animals rely on internally-monitored time passage for updating of the response bias.

In order to directly assess the role of dopamine in implementation of this behavior, fast-scan cyclic voltammetry (FSCV) with chronically implanted carbon fiber microelectrodes (Clark et al., 2010; Oleson et al., 2014) was employed to monitor subsecond dopamine changes in the dorsal striatum during behavior (Figure 1E; Figure S1A–S1C). Dorsal striatal dopamine concentration increased after trial onset, and remained elevated during 2 s trials (Figure 1F and 1H; paired t-test, $t_{10} = 2.210$, $P < 0.05$). Interestingly, across 8 s trials, the initially elevated dopamine declined with time, continued to decrease and remained below baseline until lever extension (Figure 1G and 1I; one-way repeated measures ANOVA, significant effect of time; $F_{2, 20} = 12.48$, $P < 0.01$, significant Fisher's LSD *post hoc* tests). The vast majority of animals demonstrated this increasing-decreasing biphasic profile across the 8 s trials (11/13, 85%; 5 left hemisphere, 6 right hemisphere), though a monotonic increasing profile was seen in a minority of recordings (2/13, 15%; 1 left hemisphere, 1 right hemisphere; Figure S2A and S2B). Despite left and right orientation of these levers, we did not detect any difference in recordings between hemispheres in this task, indicating dopamine signaling might be generalized across the left and right striatum in this task.

In these recordings, dopamine concentrations are elevated initially, while mice prefer the left lever (Figure 1C and 1D) and these concentrations decrease as right lever preference increases. This raises the possibility that changes in dopamine are associated with choice behavior. If this is the case, behavioral output might be predicted based on changes in dopamine on a trial-by-trial basis. Therefore, we conducted trial-by-trial analysis (see Methods), and found that the behavioral outcomes of individual trials can be predicted based on dopamine signals alone at an accuracy significantly higher than chance level (Figure S3A–C, one sample t-test of difference from 50%, $t_{19} = 51.49$, $P < 0.0001$, $n = 20$, FSCV

trial-by-trial prediction 73.6% accuracy). These data suggest that changes in dopamine in the dorsal striatum are associated with choice behavior.

Firing activity of nigral dopamine neurons resembles striatal dopamine changes

Dopamine release in striatum is under the control of dopamine neuron somatic firing, as well as terminal modulation by local striatal circuitry (Cachope et al., 2012; Threlfell et al., 2012). To determine the activity of dopamine neurons during action selection, we employed *in vivo* electrophysiology to record and identify optogenetically-tagged nigral dopamine neurons during behavior. Mice expressing channelrhodopsin-2 (ChR2) selectively in dopamine neurons were chronically implanted with a multi-electrode array and optic fiber in substantia nigra pars compacta (SNc) for simultaneous neural recording and light stimulation (Cohen et al., 2012; Jin and Costa, 2010; Jin et al., 2014; Lima et al., 2009) (Figure 2A; see Methods; Figure S1I and S1J). Dopamine neurons were identified by demonstrating significant responses to light stimulation with a short latency of firing onset (Figure 2B and 2C; latency ≈ 6 ms), and exhibiting identical waveforms between spikes generated during behavior and those evoked by optogenetic stimulation ($R = 0.95$) (Figure 2D). According to these criteria, we successfully recorded and identified 23 SNc dopamine neurons during behavior (Figure 2E; Figure S2F–S2I).

Similar to the concentration changes observed with voltammetric recordings, many SNc dopamine neurons increased firing activity in the initial phase of 2 s (Figure 2F) and 8 s trials (Figure 2G), and the firing rate then gradually decreased and remained below baseline until lever extension in 8 s trials (Figure 2G). The majority (18 out of 23, 78%) of dopamine neurons demonstrated an increasing-decreasing biphasic profile across 8 s trials (Figure 2H–2J), though increasing profiles were noted in a small population (5 out of 23, 22%; Figure 2H–2J and Figure S2C). We used rather strict criteria for identifying dopamine neurons to avoid false positives, and slightly loosening the response latency restriction to include more neurons led to qualitatively similar results (Figure S2F–S2I). Consistent with FSCV recordings, trial-by-trial analysis reveals that dopamine neuron activity is also predictive of behavioral output (Figure S3D–S3F, electrophysiology trial-by-trial prediction 82.7% accuracy). These results suggest SNc dopamine neurons undergo a specific change in firing activity during action selection.

Dopamine signaling scales to time intervals and is predictive of choice outcome

One might ask whether this dopamine signal encodes a sensory cue-evoked response, or if it is related to specific internal state during action selection. To differentiate these possibilities, we trained mice on an action selection task with the same task structure but where the lever retraction intervals were doubled to 4 and 16 s, respectively (Figure 3A; referred to here as the 4–16 s task). Here, mice exhibited similar choice behavior under the scaled time intervals (Figure 3B; two-way repeated measures ANOVA, non-significant difference between groups, $F_{1, 10} = 0.3735$, $P > 0.05$; Figure S4D and S4E; paired t-test, $t_5 = 2.532$, $P < 0.05$). Interestingly, dopamine recorded during successful 16 s trials in the 4–16 s task exhibited a delayed initial peak, but similar biphasic increasing-decreasing profile to that observed in 8 s trials in the 2–8 s task (Figure 3C). Moreover, when dopamine recorded during these two tasks is normalized to fit into the same timescale, these two dopamine

profiles tend to overlap both in terms of profile (Figure 3D; two-way repeated measures ANOVA, non-significant effect of group $F_{1,15} = 0.3211$, $P > 0.05$; main effect of time $F_{159, 2385} = 17.31$, $P < 0.0001$; non-significant time \times group interaction $F_{159, 2385} = 0.6694$, $P > 0.05$) and magnitude (Figure 3F; unpaired t-test, $t_{15} = 0.4709$, $P > 0.05$ for early; $t_{15} = 0.6714$, $P > 0.05$ for late), revealing a high correlation with one and another (Figure 3E; Pearson's $R = 0.969$, $P < 0.0001$). These data suggests that the initial increase in dopamine in the 2–8 s task does not simply reflect a cue response, as it would appear at the same latency after trial initiation in the 4–16 s task if this signal was purely reflective of a cue response. To determine if this dopamine profile is consistent with changes in other factors, we plotted expected within-trial changes in reward prediction error (Schultz et al., 1997), value state (Hamid et al., 2016), and hazard rate (Janssen and Shadlen, 2005) across this task (Figure S4A–S4C, see Methods), and found that changes in dopamine concentration/activity reported here cannot be fully predicted under either of these models.

We next tested mice trained in the 2–8 s task with unrewarded probe trials lasting 16 s (Figure 3G). In this experiment, changes in dopamine were monitored to determine how dopamine changes are associated with behavioral choice given an interval longer than 8 s that was never experienced during training. Interestingly, when dopamine concentrations continued to decrease beyond 8 s, and remained below baseline until 16 s, the animals tended to choose the lever associated with long-duration intervals. In contrast, when dopamine tended to increase and returned to baseline levels before 16 s, the animals chose the lever associated with short-duration intervals (Figure 3H; two-way repeated measures ANOVA, non-significant effect of group $F_{1,6} = 1.174$, $P > 0.05$; main effect of time $F_{159,954} = 3.544$, $P < 0.0001$; significant time \times group interaction $F_{159,954} = 2.884$, $P < 0.0001$). Accordingly, the late, but not early, dopamine change predicted the behavioral choice for the 16 s probe trials (Figure 3I; paired t-test, $t_6 = 0.1471$, $P > 0.05$ for early; $t_6 = 2.864$, $P < 0.05$ for late). This result further suggests that dopamine signaling is associated with ongoing choice, as dopamine signaling differs depending on which lever is going to be selected. Together, these results indicate that the dopamine changes observed during the 2–8 s task might be particularly associated with the internal process of action selection.

Dopamine signaling does not simply reflect time interval or temporal discounting

In both the 2–8 s and 4–16 s tasks, changes in choice, time, and value occur simultaneously, though on different timescales. We next sought to parse the individual contribution of these task elements to the observed dopamine signal by systematically eliminating each factor and determining how dopamine signaling is altered. First, one may reason that dopamine changes recorded in the 2–8 s task reflect relative timing or correspond to delay discounting during long-duration trials. That is, once the trial has extended beyond the short-interval trial duration, a decrease in value state may occur (Hamid et al., 2016), which may contribute to the decrease in dopamine we detect during 2–8 s discrimination (Figure 1F and 1H). Therefore, we employed a task with identical 2 and 8 s lever retraction intervals, but where choice behavior was eliminated by removing the response requirement for reward delivery. Instead, sucrose was delivered immediately following lever extension regardless of response (Figure 4A, referred to as the 2–8 s Pavlovian task). Importantly, the same changes in time and value were present in this task, and temporal discounting across 8 s Pavlovian trials

mimics that in the 2–8 s task. A separate group of mice that had never experienced the 2–8 s task were trained on this Pavlovian task. The mice in this task demonstrated significantly lower lever press rates compared to mice trained in the 2–8 s task (Figure S4F; unpaired t-test, $t_9 = 2.077$, $P < 0.01$, $n = 7$ for 2–8 s task, $n = 4$ for 2–8 s Pavlovian task). We then recorded dopamine with FSCV during this task in these mice. Dopamine levels initially increase during these Pavlovian 8 s trials, similar to the original 2–8 s task (see Figure S5A for 2 s trials). However, dopamine levels then remain above baseline level throughout the remainder of the 8 s trials (Figure 4B; two-way repeated measures ANOVA, main effect of group $F_{1,13} = 29.17$, $P < 0.001$; main effect of time $F_{79,1027} = 4.538$, $P < 0.0001$; significant time \times group interaction $F_{79,1027} = 2.523$, $P < 0.0001$) and are statistically higher than dopamine detected in the 2–8 s task prior to lever extension (Figure 4C; unpaired t-test, $t_{13} = 0.5023$, $P > 0.05$ for early; $t_{13} = 3.887$, $P < 0.001$ for late). The failure of dopamine to decrease below baseline in this task suggests that dopamine detected in the 2–8 s task is not merely tracking changes in value state or simply encoding time intervals.

The 2–8 s Pavlovian task removed the influence of choice from the 2–8 s task. However, this task also removes the motoric component of lever pressing, which may also influence dopamine levels (Jin and Costa, 2010; Jin et al., 2014). Therefore, we designed a behavioral task where mice experienced both 2 and 8 s retraction intervals, but choice was forced by only rewarding lever pressing at one lever. Here, 2 and 8 s retraction trials were presented randomly, and following lever extension, the first press on the left lever always resulted in reward (Figure 4D, referred to here as the 2–8 s forced choice task). Mice trained in this task pressed exclusively on the rewarding lever (Figure S4G, paired t-test, $t_4 = 24.32$, $P < 0.0001$, $n = 5$). When dopamine is recorded in this task, an initial increase in dopamine is noted at lever retraction, which then returns to baseline just prior to lever extension. The initial release amplitude is comparable to the 2–8 s task (see Figure S5B for 2 s trials), but this dopamine profile never drops below baseline (Figure 4E; two-way repeated measures ANOVA, main effect of group $F_{1,14} = 10.22$, $P < 0.01$; main effect of time $F_{79,1106} = 8.932$, $P < 0.0001$, non-significant time \times group interaction $F_{79,1106} = 0.918$, $P > 0.05$) and is statistically different in the late phase of 8 s trials (Figure 4F; unpaired t-test, $t_{14} = 0.5319$, $P > 0.05$ for early; $t_{14} = 1.821$, $P < 0.05$ for late). These data further demonstrate that the dopamine profile in the 2–8 s task is likely related to dynamic choice behavior during 8 s trials, and is not merely a product of interval timing or changes in value (Howe et al., 2013; Hamid et al., 2016).

Since the dopamine profiles in the forced choice task do not demonstrate a significant time and group interaction (Figure 4E), one might argue that the observed dopamine signal is not associated with action selection *per se* but with action initiation or delayed responding. To explore this possibility, we trained a group of mice in a similar two-lever retraction/extension task design but where behavioral sessions were comprised exclusively of 8 s trials (Figure 4G). In this task, levers retracted for only 8 sec, similar to 8 s trials in the 2–8 s task, and reward was contingent on the first press occurring at the right lever following lever extension. However, 2 s trials were completely omitted and left lever presses never yielded any reward throughout training (thus referred to as the 8 s only task). The animals became highly efficient in this task and reached a correct rate of more than 95% in two weeks by almost exclusively lever pressing on the rewarded lever (Figure S4H; paired t-test, $t_6 =$

7.833, $P < 0.0001$, $n = 7$). Mice approached the active lever and displayed anticipatory behaviors during the 8 s interval, indicating that they are actively tracking the time passage (Figure S4I and S4J). Interestingly, dopamine signaling in the dorsal striatum during single-interval 8 s trials is very different from the dopamine trace observed in the 8 s trials in the 2–8 s task. Here, a small increase in dopamine during lever retraction was followed by sustained dopamine levels that slowly rise toward lever extension (Figure 4H; two-way repeated measures ANOVA, main effect of group $F_{1,15} = 8.060$, $P < 0.05$; main effect of time $F_{79,1185} = 3.622$, $P < 0.0001$; significant time \times group interaction $F_{79,1185} = 7.208$, $P < 0.0001$). Therefore, dopamine levels in the early and late phase of these 8 s trials are significantly different between the single-interval and two-interval task (Figure 3I; unpaired t-test, $t_{16} = 1.850$, $P < 0.05$ for early; $t_{16} = 3.715$, $P < 0.001$ for late). Given the same timing interval and action, it further supports that dopamine signaling observed in the 8 trials in the 2–8 s choice task is not simply reflecting interval timing or delayed responding.

Finally, to further separate the contribution of interval timing to dopamine dynamics, we designed an action selection task with the same temporal structure but where the behavioral choice depends on external sensory cues rather than interval timing. In this regard, timing would contribute to the dopamine dynamics to the same degree but play no role in directing action selection. To explore this, we employed a behavioral task where 2 and 8 s lever retraction intervals occurred, but 1 s prior to lever extension, one of two possible tones (3K or 10K Hz) indicated that left or right lever selection would yield reward, respectively (Figure 4J, referred to here as the 2–8 s tone task). Here, pressing the left lever following a 3K Hz tone (f_1) resulted in reward, whereas selection of the right lever or repeated responding at the left lever had no effect. Similarly, selection of the right lever following a 10K Hz tone (f_2) yielded reward, but repetitive responding on the right lever or selecting the left lever had no effect. This task maintains the element of choice, but does not require temporal discrimination in order to direct correct behavioral output. Mice can reach $> 80\%$ correct rate after two-week training (see Methods) and develop no significantly different preference to either lever in this task (Figure S4K; paired t-test, $t_3 = 1.630$, $P > 0.05$, $n = 4$), which may be expected as either lever had identical delay, reward possibility, and tone allotment. We recorded dopamine during this task in a separate group mice trained from naïve. Similar to the Pavlovian 2–8 s task, dopamine levels in 8 s trials of the 2–8 s tone task rise at lever retraction and remain elevated across the entire retraction interval regardless of which tone was presented (Figure 4K; two-way repeated measures ANOVA, main effect of group $F_{79,1422} = 7.466$, $P < 0.01$, main effect of time $F_{2,18} = 6.238$, $P < 0.0001$, significant time \times group interaction, $F_{158, 1422} = 2.519$, $P < 0.0001$; see Figure S5C for 2 s data). On average, dopamine levels differed significantly between those recorded during the 2–8 s task and the tone-guided task in the late phase of recordings (Figure 4L; two-way repeated measures ANOVA, significant time \times group interaction, $F_{2, 18} = 5.302$, $P < 0.05$, Fisher's LSD *post hoc* tests), with no significant difference between tone responses ($P > 0.05$, Fisher's LSD *post hoc* tests).

During tone presentation, a further increase in dopamine was present in several recordings (Figure 4K, and Figure S5D). While tones indicate a need to enact one of two behavioral choices in order to receive reward, one might argue that a choice related signal in this task may be more evident prior to lever selection, which occurred on average 5.3 ± 0.5 s

following tone onset. Therefore, we aligned voltammetric data to the rewarded lever press and compared the dopamine dynamics for left or right lever selection (Figure S5E). It was found that the difference in dopamine changes during left vs. right actions was significantly higher compared to the difference in dopamine response to tone *f1* vs. *f2* (Figure S5F, paired t-test, $t_{11} = 2.366$, $P < 0.05$), suggesting dopamine activity is more specific to action selection. In sum, the results of these control experiments indicate that dopamine dynamics detected in the 2–8 s task are not simply related to changes in value, time or uncertainty across 8 s trials, and are instead associated with the dynamic process of selecting actions during the 2–8 s task.

Selective genetic deletions of NMDA receptors disrupt dopamine signaling and alter behavior

N-methyl-D-aspartate (NMDA) receptors have been shown to be important for phasic firing of dopaminergic neurons as well as synaptic plasticity at dopamine neurons (Engblom et al., 2008; Nugent et al., 2007; Zweifel et al., 2008). To assess the necessity of these receptors on dopamine neurons during action selection, we crossed DAT-Cre mice with NMDAR1-loxP mice to generate mutant animals with selective deletion of functional NMDA receptors in dopamine neurons (referred to as DAT-NR1 KO mice). Although the DAT-NR1 KO mutants were relatively slower in acquiring the task compared to their littermate controls (DAT Cre, DAT Cre-NR1 +/- heterozygous and NR1 f/f mice, see Methods), they were able to improve their performance significantly across two weeks of training, though not to control levels (Figure 5A; two-way repeated measures ANOVA, significant effect of group, $F_{1,11} = 36.23$, $P < 0.0001$, significant Fisher's LSD *post hoc* tests). Although the reduced accuracy in the DAT-NR1 KO mice was evident in both 2 s and 8 s trials, the impaired performance was mostly attributable to incorrect choices during long-duration trials (Figure 5B; two-way repeated measures ANOVA, significant effect of group, $F_{1,10} = 7.68$, $P < 0.05$, significant Fisher's LSD *post hoc* tests). To determine how deletion of NMDA receptors on dopamine neurons affects dopamine signaling and impairs action selection, we implanted chronic FSCV electrodes in DAT-NR1 KO mice and recorded dopamine in the dorsal striatum during the action selection task. These animals showed weaker but similar initial elevation in dopamine concentration for 8 s trials, but the subsequent decrease in dopamine release was absent (Figure 5C; two-way repeated measures ANOVA, non-significant effect of group $F_{1,13} = 2.174$, $P > 0.05$, main effect of time $F_{79,1027} = 7.217$, $P < 0.0001$, effect of interaction $F_{79,1027} = 1.411$, $P < 0.05$ and Figure 5D; unpaired t-test, $t_{13} = 0.5493$, $P > 0.05$ for early; $t_{13} = 2.122$, $P < 0.05$ for late). Importantly, dips in dopamine are detectable in these KO animals (Figure S2D), suggesting they do not have a deficit in suppressing dopamine release *per se* but lack context-specific inhibitory inputs after training (Bonci and Malenka, 1999; Nugent et al., 2007). These data emphasize the importance of dopamine signaling in appropriate action selection and optimizing behavior, in which NMDA receptors on dopamine neurons play an essential role.

This idea was further supported by experiments using mutants with selective deletion of NMDA receptors in striatal projection neurons by crossing RGS9-Cre mice with NMDAR1-loxP mice (referred to here as RGS-NR1 KO mice) (Jin and Costa, 2010; Jin et al., 2014). Together with dopamine, NMDA receptors in the striatum have been shown to be critical for

action learning and corticostriatal plasticity (Calabresi et al., 1992; Jin and Costa, 2010; Surmeier et al., 2009). RGS-NR1 KO mice are completely impaired in the action selection task, and do not show any improvement in performance after two weeks of training (Figure 5E; two-way repeated measures ANOVA, significant effect of group, $F_{1,13} = 15.34$, $P < 0.01$, significant Fisher's LSD *post hoc* tests). Based on their performance during unrewarded probe trials, RGS-NR1 KO animals demonstrate no sign of learned action selection, with a constant bias toward the left lever irrespective of the duration of probe trials (Figure 5F; two-way repeated measures ANOVA, significant group \times time interaction, $F_{6,60} = 8.99$, $P < 0.0001$, significant Fisher's LSD *post hoc* tests). RGS-NR1 KO mice trained on this task were then implanted with chronic FSCV electrodes in dorsal striatum and dopamine changes were monitored during behavior. As predicted, striatal dopamine levels in RGS-NR1 KO mice showed no selection-associated dopamine release (Figure 5G; two-way repeated measures ANOVA, non-significant effect of group $F_{1,13} = 0.0008$, $P > 0.05$, main effect of time $F_{79,1027} = 3.713$, $P < 0.0001$; effect of interaction $F_{79,1027} = 3.642$, $P < 0.0001$; and Figure 5H; unpaired t-test, $t_{13} = 1.877$, $P < 0.05$ for early; $t_{13} = 1.613$, $P = 0.065$ for late). Importantly, phasic dopamine release can be reliably detected in these KO mice, typically to reward acquisition (Figure S2E), indicating these mice are capable of releasing dopamine and that impaired dopamine signaling is behavior-specific. These results further indicate a critical role for nigrostriatal dopamine and striatum for appropriate action selection.

Optogenetic manipulation of nigrostriatal dopamine biases action selection

We next sought to manipulate striatal dopamine levels during behavior with the expectation that changing striatal dopamine would interfere with the action selection process and alter behavioral choice. ChR2 was selectively expressed in dopamine neurons (Jin and Costa, 2010). Optogenetic stimulation of nigral dopamine neurons reliably induces dopamine neuron firing (Figure 6A), and evokes striatal dopamine release (Figure 6B). Next, simultaneous optogenetic stimulation and voltammetric monitoring of dopamine levels were conducted in order to manipulate and record dopamine in the same animal during action selection (Figure 6C; Figure S1E and S1F). As expected, optogenetically stimulating nigral dopamine neurons for 1 s at 1, 3, or 7 s after trial initiation promptly elevates striatal dopamine levels (Figure 6D–6F). To assess how this increase in dopamine affects immediate choice behavior, we stimulated dopamine at these time points, and then extended levers immediately after stimulation cessation at 2, 4, or 8 s (see Figure S6A and methods for experimental design). In other words, stimulation was fixed at 1 s before lever extension, and three different duration retractions (2, 4, 8 s) were tested. Driving increases in dopamine concentration biases the animal's choice immediately following stimulation toward the short-duration lever in 4 s unrewarded probe trials, though not in 2 or 8 s trials (Figure 6G; one sample t-test, $t_{10} = 3.122$, $P < 0.01$ for 4 s; $P > 0.05$ for 2 and 8 s). Further, this effect appears to be specific to driving high-frequency firing of dopamine neurons (Figure S6C–S6F). To confirm that these effects were specifically attributed to striatal dopamine changes, we performed an experiment in a separate group of mice with optical fibers targeting dopamine terminals in the striatum (Figure S1D). Optogenetically stimulating dopamine terminals for 1 s in the striatum resulted in qualitatively similar effects on immediate choice behavior to somatic stimulation (Figure 6H; one sample t-test, $t_2 = 4.304$, $P < 0.05$ for 4 s; P

> 0.05 for 2 and 8 s), indicating that specific modulation of nigrostriatal dopamine is sufficient to alter action selection.

Our simultaneous optogenetic and voltammetry results indicate that stimulation of dopamine cells can delay the tendency of dopamine concentrations to drop below baseline late in 8s trials. For instance, delivery of 1 s-duration light at 3 s significantly increases dopamine levels between 7s and 8s in the stimulated trials compared to non-stimulated trials (Figure 6I; paired t-test, $t_2 = 3.414$, $P < 0.05$ for 3 s; $t_2 = 3.774$, $P < 0.05$ for 7 s; $P > 0.05$ for 1 s), suggesting that optogenetic stimulation at early time points might also alter choice behavior in a delayed manner. Therefore, we next stimulated dopamine neurons at different time points during 8 s trials to determine how stimulation proximity to lever extension affected choice. Specifically, dopamine stimulation occurred for 1s across each second from 0 to 7s during 8 s trials (see Figure S6B and methods for experimental design). Here, driving dopamine release systematically biased animals toward the short-duration lever on 8 s trials, which was especially evident for the stimulation delivered at middle time points (Figure 6J; one sample t-test, $P < 0.05$ for 2 to 6 s). We next sought to inhibit dopamine neurons using this same experimental design with the expectation that decreasing dopamine may bias choice in the opposite direction. Selective inhibition of dopamine activity by archaerhodopsin (Chow et al., 2010) increased long-duration lever choice in 8s trials, but this effect was only observed when the inhibition was delivered immediately after trial initiation (Figure 6K; one sample t-test, $t_6 = 2.273$, $P < 0.05$ for 0 s), coincident with the early phase increase in dopamine concentration. Taken together, these data indicate that manipulation of nigrostriatal dopamine is sufficient to modulate behavioral choice bidirectionally, and to bias the animal's online action selection in both an immediate and delayed manner.

Dopamine biases action selection through modifying basal ganglia activity

To further explore the circuit mechanisms underlying dopamine biasing action selection, we constructed a neuronal network model of the cortico-basal ganglia circuitry (see Methods), wherein nigrostriatal dopamine acts on striatal D1- (direct pathway) and D2-expressing (indirect pathway) spiny projection neurons (SPNs) to regulate behavioral output (Hikosaka et al., 2000; Jin et al., 2014; Mink, 2003) (Figure 7A). Specifically, cortical information corresponding to left or right choice are sent to dorsal striatal D1- and D2-subpopulations associated these two action options (Lo and Wang, 2006). Signals from these striatal subpopulations then converge to two separate SNr populations (Hikosaka et al., 2000; Jin et al., 2014; Mink, 2003), which drive left or the right choice, respectively (Lo and Wang, 2006). Behavioral output is determined by the dominant action between the mutually inhibiting left and right SNr populations (Mailly et al., 2003), which could control the final motor output either through brainstem circuits or motor cortices (Hikosaka, 2007; Lo and Wang, 2006; Redgrave et al., 1999). Therefore, dopamine could potentially bias action selection through modulation of striatal direct and indirect pathways, thus influencing SNr activity and behavioral output.

Using this neuronal network model, we simulated both the immediate and delayed effects of optogenetic manipulation of dopamine on choice behavior. The changes of dopamine

concentration in the model are based on experimentally observed dopamine dynamics (Figure 7B). Given the same cortical inputs (Lo and Wang, 2006), neural activity in striatal D1-/D2-SPNs and SNr neurons are further modulated due to the optogenetic manipulation of dopamine release (Figure 7C–7J). The model simulation revealed that optogenetic stimulation of dopamine biases the animals toward left choice only during 4 s but not 2 or 8 s trials (Figure 7K), faithfully recapitulating the experimentally observed immediate effects of optogenetic stimulation (Figure 6G and 6H). In addition, computational simulation using the same set of parameters was able to reproduce the effects of stimulation at each earlier time point on choice in the 8 s trials under both optogenetic stimulation (Figure 7L) and inhibition (Figure 7M). Notably, the model also successfully predicted the action selection impairments observed in DAT-NR1 KO mice (Figure S7A–S7J). These results thus suggest that nigrostriatal dopamine could bias action selection through a circuitry mechanism by modulating neuronal activity in the basal ganglia direct and indirect pathways.

To further explore the significance of dopamine dynamics in choice behavior, we manipulated dopamine levels in this model to determine how basal ganglia activity and behavioral choice were altered. First, by removing the dopamine dynamic observed during 8 s trials and instead constantly maintaining dopamine at baseline level during the 8s retraction period (Figure 8A), action selection is blunted and choice behavior at short- and long-duration trials becomes less selective (Figure 8B; two-way repeated measures ANOVA, significant effect of group, $F_{1, 18} = 34.05$, $P < 0.0001$, significant Fisher's LSD *post hoc* tests). This suboptimal behavior is driven by altered neuronal responses in basal ganglia (Figure S7K–S7M). Furthermore, the model simulation suggested that dopamine at different concentrations devoid of any changes will generally bias behavioral choice toward one or another option, depending on absolute concentration (Figure 8C and D). Elevated dopamine generally biased the behavior toward short-duration choice, whereas suppressed dopamine biased the behavior toward long-duration choice (Figure 8D, for corresponding basal ganglia neuronal activity see Figure S7N–S7P). Together with the lack of dopamine dynamics observed in the RGS-NR1 mice where action selection was largely compromised (Figure 5E–5H), these results suggest that dopamine dynamics are critical for optimizing action selection and behavioral choice.

Discussion

Under different experimental settings, nigrostriatal dopamine neurons have been suggested to encode the salience of stimuli (Matsumoto and Hikosaka, 2009), the value of cues in stimulus-response behavior (Kim et al., 2015; Morris et al., 2006), the initiation and termination of self-paced action sequences (Jin and Costa, 2010), and might even directly reflect movement kinematics (Barter et al., 2015). In the current study, mice were trained to dynamically select two alternative actions according to internally-monitored temporal information. This particular experimental design allowed us to observe dopamine dynamics throughout a dynamic decision process, as mice switched action preference based on internally-monitored passage of time. When initially employing a preferred option, dopamine levels are significantly elevated above baseline. However, when mice move away from their initial choice and switch to the other option, instead of a further increase, we note a gradual decrease in dopamine concentration, which remains below baseline. This

dopamine profile can't be easily attributed to within-trial reward prediction error (Schultz et al., 1997) or value state (Hamid et al., 2016) across this task (Figure S4A–S4C).

While the current behavioral paradigm nicely allows us to observe and perturb the neural dynamics underlying the internal processes of action selection and preference switch, an argument one might have is that the timing aspect of the task is a potential confounding factor. Previously, dopamine has been implicated in timing behaviors (Buhusi and Meck, 2005; Oleson et al., 2014), mostly based on results of pharmacological experiments with global manipulation, though it remains unclear whether these effects resulted from altering time perception or because of disrupted action selection. A recent study by Soares et al. (2016) utilized a temporal discrimination task to determine the contribution of midbrain dopamine to interval duration judgment. Similar to the current study, the authors found that the population dopamine neuron activity, reflected by calcium signals recorded through optic fibers, is associated with trial-by-trial variability in behavioral performance, and that optogenetic stimulation of dopamine neurons alters psychometrics of behavioral choice (Soares et al., 2016). Interestingly, both studies have found that driving dopamine activity increases, rather than decreases short-duration selection (Figure 6 and Figure S8) (Soares et al., 2016), at odds with the so-called dopamine clock hypothesis in the timing literature that predicts increased dopamine release would speed up the internal clock (Buhusi and Meck, 2005; Simen and Matell, 2016). The authors also report changes in dopamine are tightly associated with behavioral choice in their task, highlighting a potential role for dopamine in action selection (Soares et al., 2016). Somewhat contrary to these results, however, the authors further use a computational model to rule out the contribution of action bias and conclude dopamine neurons reflect and directly control the subjective estimation of time (Soares et al., 2016).

Much evidence suggests that the nigrostriatal dopamine profile we observed in the current study is highly unlikely to be a pure timing signal. First, if dopamine is purely tracking time, one may expect a monotonic change in dopamine across 8 s trials. However, the biphasic nature of this dopamine dynamic was evident in both voltammetric and electrophysiological recording data. In addition, using four control experiments with altered elements of choice and an identical time intervals, we note robust differences in dopamine signaling between choice and non-choice or Pavlovian settings (Figure 4A – 4L). In these experiments, mice appear to faithfully anticipate the interval (Figure S4J), though dopamine concentrations remain elevated across 8 s trials, in contrast with a biphasic change in dopamine observed when discrimination and choice are involved (Figure 4B, 4E, 4H, and 4K). Additionally, dopamine signaling during discrimination is predictive of behavioral outcome (Figure S3). Furthermore, when given probe trials of 16 s in the 2–8 s task (Figure 3G – 3I), or when mice use a separate discriminative sensory modality to cue correct behavioral choice (Figure 4J – 4L, Figure S5), dopamine signaling clearly follows dynamic choice but not interval timing. This point was further confirmed in the behavior and dopamine dynamics seen in the DAT-NR1 KO mice (Figure 5A–D). Finally, taking advantage of optogenetics, we performed an additional experiment to directly test whether the nigrostriatal dopamine exerts its effect on action selection at the sensory input or at the decision-making level. The logic was this: if optogenetic stimulation of dopamine effects purely at the sensory input (increase or decrease time/value) but not decision-making level, it will cause a clear horizontal shift in the

psychometric curve. To avoid a time/state-dependent change in the optogenetic effect and ensure all the test trials experience the same manipulation, we delivered the same stimulation (1 s constant light) at a fixed timepoint (time 1s since trials begin) to examine its effect on all following probe trials spanning 2 to 8 s. Theoretical psychometric curves were constructed with varying parameters associated with different aspects of decision and were compared to the behavioral effects from the optogenetic experiments (see Methods). The optogenetic results are inconsistent with predicted changes in the psychometric curve under pure manipulation of time (Figure S8). Therefore, while the dopamine system has been linked to timing behaviors under certain experimental settings (Buhusi and Meck, 2005; Oleson et al., 2014), neither our physiological measurements nor our behavioral manipulations can be fully explained as dopamine signaling the time intervals alone.

Instead of encoding time or value, we posit that the nigrostriatal dopamine might integrate information across multiple factors essential to decision processing, and send this integrative signal to downstream brain regions including dorsal striatum for modulating action selection (Figure 7). Indeed, recent studies with optical imaging have revealed that dopamine sends both reward- and movement-related information to the striatum, likely by different dopaminergic axons (Howe and Dombeck, 2016; Parker et al., 2016). Importantly, it has been further demonstrated that the movement-related signaling in dopaminergic axons is not only associated with but also capable of directly triggering movements, emphasizing a motor nature of the nigrostriatal dopamine system (Barter et al., 2015; Howe and Dombeck, 2016; Jin and Costa, 2010). Together with the current study, these results support functionally heterogeneous roles of dopamine signaling during behavior (Hamid et al., 2016; Kim et al., 2015; Schultz, 2007), acting at both fast and slow timescales (Hamid et al., 2016; Schultz, 2007). They also underscore a critical function of the nigrostriatal dopamine system in controlling actions (Barter et al., 2015; Jin and Costa, 2010; Kim et al., 2015; Matsumoto and Hikosaka, 2009; Panigrahi et al., 2015; Redgrave and Gurney, 2006; Redgrave et al., 2010), besides the well-known role it plays in learning (Schultz, 2007; Schultz et al., 1997).

Our optogenetic experiments suggest a bidirectional effect of nigrostriatal dopamine on learned action selection. Furthermore, dopamine stimulation can immediately bias action selection and this effect was stronger under uncertain situations (4 s probe trials), but less so for more certain conditions (2 and 8 s trials), consistent with the notion that well-learned actions are less reliant on dopamine (Wickens et al., 2007). Importantly, this modulatory role of dopamine in biasing action selection does not necessarily imply dopamine directly selects actions *per se*. Dopamine stimulation does not elicit a stereotypic behavioral choice, as behavioral outcome is highly dependent on the animal's state and when optogenetic stimulation occurs (Figure 6J and 6K). It is likely that nigrostriatal dopamine biases action selection through its interactions with downstream brain structures, notably the striatum. Indeed, based on a cortico-basal ganglia neuronal network model, our computational simulations have demonstrated that both the optogenetic stimulation and inhibition effects of dopamine on action selection could be accounted for by alteration of basal ganglia output through modulation of striatal D1- and D2-SPNs (Jin et al., 2014) (Figure 7). This possibility is indeed supported by *in vivo* electrophysiological recordings in both primate and rodents. Striatal neuronal activity has been found to encode action value (Lau and Glimcher, 2008; Samejima et al., 2005), which could be used to bias the selection of future

actions through regulation of SNr activity (Ding and Gold, 2012; Hikosaka et al., 2000; Jin et al., 2014; Lauwereyns et al., 2002; Tai et al., 2012). Our findings favor the idea that nigrostriatal dopamine plays an important modulatory role in action selection as the striatum cues up available options. Notably, recent studies have found that cue-evoked phasic release in the mesolimbic dopamine system is strongly contingent on correct action initiation, not only on reward prediction (Syed et al., 2016), suggesting that even the mesolimbic dopamine system may be more closely tied to actions than previously thought (Niv et al., 2007; Phillips et al., 2003; Stopper et al., 2014; Syed et al., 2016). Importantly, our model simulations indicate that the dopamine dynamic might be critical for appropriate action selection as constantly maintained dopamine at any concentration does not fully optimize behavioral output (Figure 8C and D). This result suggests that updating of dopamine levels are a crucial underlying factor during the dynamic process of selecting upcoming actions or while switching among behavioral states. The capability of selection is largely compromised when dopamine is “clamped” to a constant high or low level as may be expected during pharmacological manipulation or in certain neurological diseases. The current findings thus unveil an important role of nigrostriatal dopamine in action selection, and have far-reaching implications for understanding various neurological and psychiatric disorders in humans from Parkinson’s disease, Obsessive-Compulsive Disorder (OCD) to substance dependence (Everitt and Robbins, 2005; Graybiel and Rauch, 2000; Mink, 2003; Redgrave et al., 2010).

STAR★Methods

Contact for Reagent and Resource Sharing

Further information and requests for reagents may be directed to, and will be fulfilled by the Lead Contact, Xin Jin (xjin@salk.edu).

Experimental Model and Subject Details

Mice—All procedures were approved by the Institutional Animal Care and Use Committee at the Salk Institute for Biological Studies, and were conducted in accordance with the National Institute of Health’s Guide for the Care and Use of Laboratory Animals. Male and female mice (2–6 months of age) with C57BL/6 background were group housed and kept on a 12 hour light/dark cycle (lights on at 6:00 am) with *ad libitum* access to food and water. In total, 173 mice were used in this study. Specifically, 81 mice were implanted with FSCV hardware, and we successfully collected data from 32 animals (see below under Fast Scan Cyclic Voltammetry for inclusion criteria). Additionally, 23 mice were implanted with optical fibers for optogenetic experiments, though three were excluded for failing to reach performance criterion (described below under Optogenetic Experiments). A total of 20 mice were implanted with multielectrode arrays, and we successfully identified 23 dopamine neurons in 5 of these animals (inclusion criteria below under *In Vivo* Electrophysiology). Finally, 49 mice were used in behavioral experiments. Dopamine neuron-specific NMDAR1-knockout mice were generated by crossing *DAT-cre* mice with *NMDAR1-loxP* mice (Engblom et al., 2008; Zweifel et al., 2008), and striatal-specific NMDAR1-knockout mice were generated by crossing *RGS9-cre* mice with *NMDAR1-loxP* mice as formerly described (Dang et al., 2006; Jin and Costa, 2010). Mice with viral expression of channelrhodopsin-2 were *DAT-cre* mice with injection of AAV to drive opsin expression.

Mice with genetic expression of channelrhodopsin-2 or archaerhodopsin in dopamine neurons were bred by crossing *DAT-cre* with Ai32 (ChR2 (H134R)-EYFP, Figure S1) or Ai35 (RCL-Arch/GFP) mice, respectively (Jackson Laboratory) (Madisen et al., 2012).

Method Details

Behavioral training—Behavioral training took place in operant chambers (Med Associates) which were placed inside a sound attenuating box. Behavioral chambers were 21.6 cm L × 17.8 cm W × 12.7 cm H and were comprised of a central food magazine with retractable levers on both sides. A house light (3 W, 24 V) was centered near the ceiling on the opposite side of the food magazine (Jin and Costa, 2010; Jin et al., 2014). Sucrose solution (10%, 10 μ L) was delivered into a bowl in the food magazine by a syringe pump. Operant chambers were computer controlled and all behavioral programs were custom written. Mice were food restricted for at least 24 hours prior to training and were maintained at ~85% free-feeding weight by providing ~2.5 g regular chow per day per mouse.

Behavioral training began with 60 minute continuous reinforcement (CRF) training, where mice were presented with either the left or the right lever which resulted in one reward each time the lever was pressed. Sessions began with illumination of the house light and extension of either the right or left lever. The first day, mice received 5 rewards on each lever. The second day, mice received 10 rewards on each lever. The third day, mice received 15 rewards on each lever. Lever presentation order was counterbalanced and following completion of the task levers were retracted and the house light was turned off.

Following the completion of CRF, mice were trained on an adopted version of a temporal bisection task (Figure 1A), where they were trained to discriminate two time intervals and respond on either the left or right lever accordingly (Church and Deluty, 1977). Sessions began with illumination of the house light and extension of both the left and the right levers. Following a random interval which ranged from 30–40 s (35 s on average), two levers were retracted for either 2 or 8 s (50 % chance of either interval, random order). Both levers were then extended and the first response on the left lever following a 2 s retraction interval or the first response on the right lever following an 8 s retraction interval resulted in reward delivery. For a subgroup of experiments, the contingency between time interval and lever pressing was reversed, i.e. 2 and 8 s trials were rewarded following right or left lever responses, respectively. Only the first press on the correct lever following lever extension could result in reward, that is, mice were not rewarded with any further lever pressing after making an incorrect choice or by continuous responding on the correct lever. The inter-trial-interval period with both levers extended thus served as an extinction test for reading out the animal's net preference. Behavioral sessions were terminated after 180 min or following 160 rewards, whichever happened first. Mice that were included in voltammetric or electrophysiological studies were trained to 75% accuracy before implanting recording electrodes. Following surgery, mice were trained back to 75% accuracy before physiological recordings took place. For all tasks, mice that needed to be attached to cables for stimulation or recording were trained with cables attached through the post-surgery period to allow better habituation to the weight of the headstage.

Mice previously trained on 2 and 8 s discrimination were later trained on 4 and 16 s discrimination (Figure 3A). These trials were in the same structure and contingency as the 2 and 8 s task, the only difference being that lever retraction intervals were doubled.

Mice previously trained on 2 and 8 s discrimination were later tested using unrewarded 16 s probe trials (Figure 3G). Here, trials were identical to those described above except that 10% of trials were comprised of 16 s retraction intervals which remain unrewarded irrespective of the animal's choice. The other 90% of trials were 2 and 8 s retraction intervals (45% each) with maintained normal action-reward contingency. These trials were presented in a random order.

A separate group of naïve mice that had not been trained on the 2–8 s task were utilized for the 2–8 s Pavlovian task. Here, mice were initially trained on CRF, as described above. Next, mice were trained on a task with 2 and 8 s lever retraction intervals, but where reward was delivered at lever extension regardless of response (Figure 4A). Sessions began with illumination of the house light and extension of both the left and the right levers. Following a random interval which ranged from 30–40 s (35 s on average), two levers were retracted for either 2 or 8 s (50 % chance of either interval, random order). Both levers were then extended and a sucrose reward was delivered.

Mice previously trained on the 2–8 s Pavlovian task were then trained in the 2–8 s forced choice task. This task had an identical structure as the 2–8 s task, but mice only received rewards by selecting the left lever following retraction periods (Figure 4D). Following a random interval which ranged from 30–40 s (35 s on average), two levers were retracted for either 2 or 8 s (50 % chance of either interval, random order). Both levers were then extended and selection of the left lever resulted in delivery of a sucrose reward. Only the first left press following extension yielded reward and right lever presses and repetitive pressing at the left lever had no effect.

For the task with 100% 8 s trials, the task structure was exactly the same as 2 and 8 s bisection, but 2 s trials were completely omitted (Figure 4G). Mice therefore only experienced 8 s trials the whole session and received rewards on the right lever following 8 s lever retractions. Left lever presses and repetitive responding on the right lever were unrewarded. Mice were trained on this task until they achieved an accuracy of ~95%. Voltammetric recordings in this task were collected from one subgroup of mice previously trained on 2 and 8 s discrimination, as well as a separate group of naïve animals trained only on 8 s trials. Data from these two groups were consistent and were therefore combined.

A separate group of naïve mice that had not been trained previously were used in the 2–8 s tone task. Here, a task with identical structure to the 2–8 s task was used, but where two different tones indicated appropriate behavioral response (left or right lever selection, Figure 4J). Sessions began with illumination of the house light and extension of both the left and the right levers. Following a random interval which ranged from 30–40 s (35 s on average), two levers were retracted for either 2 or 8 s (50 % chance of either interval, random order). Then, one second prior to lever extension, either a 3,000 or 10,000 Hz tone was played. Following lever extension in trials with a 3,000 Hz tone, a left lever press yielded a single

sucrose reward whereas repetitive responding or right lever presses yielded no response. Conversely, following lever extension in trials with a 10,000 Hz tone, a right lever press yielded a single sucrose reward whereas repetitive responding or left lever presses yielded no response. Only the first response following lever extension could result in reward. Mice were trained 2–3 weeks to reach ~70% correct rate before recordings occurred. One group of mice was trained prior to voltammetric implantation and another was trained previous to implantation. Recordings from these groups did not differ significantly and were combined. To increase numbers of rewarded trials, mice could earn up to 250 rewards during recordings.

Learning curves for both wildtype and KO animals were determined by first training mice on CRF as described above. Once they met CRF training criterion, mice were trained daily on 2 and 8 s lever retraction discrimination with sessions lasting 180 min or until mice received 160 rewards. The training lasted 14 days. Accuracy was defined as percentage of correct trials. Psychometric curves were constructed using unrewarded probe trials after these 14 days of training. Here, mice were presented with unrewarded probe trials with retraction intervals of 2.5, 3.2, 4, 5, or 6.3 s in duration. These probe trial durations represent evenly spaced intervals between 2 and 8 s on a logarithmic scale, as temporal bisection is thought to fall at the logarithmic, rather than arithmetic, mean of the intervals tested (Church and Deluty, 1977). These probe trials were randomly presented and comprised 30% of all trials, with the remaining 70% of trials being 2 or 8 s in duration (35% each). Mice were tested the first, third, fifth and seventh day following 14 day training. On the second, fourth and sixth day following training, mice were retrained on the 2 and 8 s task without probes. All probe trials were not rewarded irrespective of the animal's choice.

Both wildtype and KO mice used for physiological experiments were trained for two weeks on 2 and 8 s discrimination as described above prior to implantation of recording electrodes. Following recovery, mice were again food restricted and retrained to ~75% accuracy, or, in the case of RGS-NR1 KO mice, for one additional week prior to recording sessions.

Surgical procedures—Mice were anesthetized using either ketamine (100 mg/kg) and xylazine (5 mg/kg) or isoflurane (4% induction; 1–2% sustained) and were placed in a stereotactic frame. For voltammetry experiments, custom fabricated carbon fiber voltammetry electrodes were inserted bilaterally into dorsal striatum (+ 0.8 AP, ± 1.5–2.0 ML, – 2.5 DV) and were affixed using dental cement (Clark et al., 2010). Chloridized silver reference electrodes were implanted ipsilaterally to their associated recording electrode, and recording and reference electrodes were attached to a connector that extended from the dental cement allowing connection to a voltammetric headstage (Clark et al., 2010) (Figure S1B).

For electrophysiological identification of dopamine neurons, we utilized electrode arrays (Innovative Neurophysiology Inc.) of 16 tungsten contacts (2 × 8) that were 35 µm in diameter. Electrodes were spaced 150 µm apart in the same row and 200 µm apart between two rows. Total length of electrodes was 5 mm. Each array had a cannula attached (300 µm from the electrode tips) allowing for insertion of fiber optics capable of delivering laser light (Jin and Costa, 2010) (Figure S1I). Another method was also employed in some animals by

directly attaching an optic fiber to the electrode array (Figure S1J). Arrays targeting substantia nigra pars compacta ($-3.1\sim 3.4$ AP, ± 1.0 ML, $-3.8\sim 4.1$ DV) were unilaterally and incrementally lowered into substantia nigra pars compacta (SNc) and laser stimulation (473 nm) occurred to elicit firing in putative dopamine neurons. Silver grounding wire was attached to skull screws. Following optimization of placement, the array was affixed using dental cement (Jin and Costa, 2010; Jin et al., 2014).

For optogenetic stimulation experiments, we utilized both *DAT-Ai32* mice and *DAT-cre* mice with viral expression of ChR2 (Figure S1G and S1H). No significant differences were detected between these groups, so data from these mice was combined. To express ChR2 selectively in SNc dopamine neurons, 1 μ l of concentrated adeno-associated virus encoding cre-inducible Channelrhodopsin-2 (AAV9-DIO-ChR2-EYFP; University of Pennsylvania Vector Core) was injected bilaterally into substantia nigra pars compacta ($-3.1\sim 3.4$ AP, ± 1.5 ML, -3.8 DV) of *DAT-cre* mice. Then, immediately following viral injection, custom fabricated fiber optics were implanted above substantia nigra pars compacta ($-3.1\sim 3.4$ AP, ± 1.5 ML, -3.6 DV, Figure S1D and S1F). These were affixed using dental cement with ferrules allowing attachment to a commutator extending above the cement. To determine if behavioral effects were attributable to nigrostriatal dopamine, a separate group of *DAT-Ai32* mice had fiber optics implanted into striatum ($+0.8$ AP, ± 1.5 ML, -2.3 DV). For optogenetics inhibition experiments, we used *DAT-Ai35* mice with fiber optics chronically implanted above substantia nigra pars compacta ($-3.1\sim 3.4$ AP, ± 1.5 ML, -3.6 DV) in the same manner as the optogenetic stimulation experiments.

For simultaneous optogenetic and voltammetric recordings, chronically implantable carbon fiber voltammetric electrodes and references were unilaterally implanted in striatum ($+0.8$ AP, $\pm 1.5\sim 2.0$ ML, -2.5 DV for carbon fiber probe). Optic fibers were then implanted bilaterally above substantia nigra pars compacta ($-3.1\sim 3.4$ AP, $\pm 1.0\sim 1.5$ ML, -3.6 DV), and were incrementally lowered until robust dopamine release was detected in dorsal striatum. Fiber optics and voltammetry headstage were then affixed using dental cement (Figure S1E). Following surgery, mice received analgesia that consisted of either Ibuprofen dissolved in water (100 mg/ml) or buprenorphine (1 mg/kg). For all surgical procedures, animals were allowed to fully recover for at least two weeks before behavioral training commenced.

Optogenetic experiments—For optogenetic experiments, mice were first trained in 2 and 8 s interval discrimination task for two weeks and surgically implanted with optic fibers. After achieving a success rate of 75% with fiber optic cables attached, stimulation trials began. The trials were comprised of 2, 4, or 8 s retraction intervals (40%, 20%, and 40%, respectively) presented in random order. Dopamine neurons were stimulated or inhibited bilaterally in 50% of trials using a single pulse of light (Laserglow, 473 nm, 5 mW, 1 s constant for ChR2 experiments; Laserglow, 532 nm 10 mW, 1 s or 8 s constant for Arch experiments). Rewards were delivered only at correct responses during 2 and 8 s trials. We used three experimental protocols, one to assess the immediate effects of dopamine stimulation, one to determine delayed effects of modulating dopamine, and one to construct a psychometric curve. To determine the immediate effect of dopamine stimulation, mice were presented with 2 s (40%), 4 s (20%), and 8 s (40%) trials in random order. Responses

during 4 s trials were unrewarded irrespective of left or right choice. Within 50% of any type of trials, mice were optogenetically stimulated for 1 s before lever extension (Figure S6A). For the delayed stimulation protocol, mice were presented with 2 s (50%) and 8 s (50%) trials in random order. In 50% of 8 s trials, mice were optogenetically stimulated for 1 s occurring randomly at 0–7 s (Figure S6B). Mice only received stimulation once per stimulation trial. For psychometric curves, mice were presented with trials spanning 2, 2.5, 3.2, 4, 5, 6.3, and 8 s in duration. On 50% of trials, stimulation occurred for 1 s constant at 1 s following lever retraction (for stimulation) or 8 s constant beginning at lever retraction (for inhibition). Mice only received rewards during successful 2 or 8 s trials, and probe trials occurred randomly during 35% of trials. For all optogenetic experiments, sessions with correct rate below 75% for control trials were excluded from further analysis.

Fast-Scan Cyclic Voltammetry (FSCV)—Dopamine was recorded during behavior using fast-scan cyclic voltammetry, which consisted of a triangular waveform (–0.4 to 1.3 V and back at 400 V/s) applied to the tip of carbon fiber voltammetric electrodes at 10 Hz (Bucher and Wightman, 2015; Howard et al., 2013). Potentiostats and headstages were custom fabricated by the Department of Psychiatry & Behavioral Sciences at the University of Washington, and voltammetry was controlled with Tarheel software (ESA Biosciences Inc.) (Clark et al., 2010). Prior to each recording session, a triangular waveform was applied to the recording electrode at a frequency of 60 Hz for at least one hour. The cycling frequency was then reduced to 10 Hz for 30 min or until stable background was detected prior to initiation of the recording session. During this ‘cycling’ period, mice remained in their homecage. Voltammetric electrodes were validated for functionality and for data inclusion using two criteria. First, only electrodes with a background size above 200 nA that were stable (background signal:noise above 250:1) were included. Additionally, a salient event prior to each experiment (i.e. tapping on homecage) and/or the opening of the sound proofing box at the end of each experiment yielded a dopamine signal with a cyclic voltammogram consistent with dopamine recorded *in vitro* ($R = 0.90$). Principal component analysis (Keithley et al., 2009) was used to isolate dopamine changes from changes in current attributed to pH or electrode drift. Training sets for principal component analysis were constructed using *in vitro* recordings of dopamine and pH changes collected using glass electrodes in a flow cell apparatus. Unique changes in background current were collected for each electrode used on the same day that recordings were taken. Therefore, training sets for dopamine and pH changes were consistent across recordings, whereas training sets for background drift were unique for each recording. Current recorded was converted to concentration using background size and an equation collected from glass-sealed carbon fiber electrodes *in vitro*. Voltammetric analysis was conducted using HDCV software which was kindly provided by the Department of Chemistry at the University of North Carolina at Chapel Hill (Bucher et al., 2013).

In vivo electrophysiology—Dopamine neurons were recorded and identified as previously described (Jin and Costa, 2010; Jin et al., 2014). Briefly, neural activity was recorded using the MAP system (Plexon Inc., TX). The spike activities were initially online sorted with a build-in algorithm (Plexon Inc., TX). Only spikes with stereotypical waveforms clearly distinguished from noise and relatively high signal-to-noise ratio were

tagged and saved for further analysis. After the recording session, the recorded spikes were further isolated into individual units by an offline sorting software (Offline Sorter, Plexon Inc.). Each individual unit displayed a clear refractory period in the inter-spike interval histogram, with no spikes during the refractory period (larger than 1.3 ms). All the timestamps of animal's behavioral events during performing the task were converted into TTL pulses by a Med-Associates interface board and recorded by the MAP recording system through an A/D board (Texas Instrument Inc., TX). The animal's behavioral timestamps were simultaneously recorded together with the neural activity.

To identify the dopamine neurons in SNc, we utilized *DAT-Ai32* mice which express ChR2 selectively on dopamine neurons. Before the recording session, we connected the recording cable to the electrode array for neuronal recording, and inserted an optic fiber through the cannula attached the array to target SNc. The tip of the fiber was ~200 μm away from the tips of the electrodes and the optic fiber was firmly fixed to the array for the duration of each recording session. For each training session with recording, blue laser stimulation was delivered through the optic fiber from a 473-nm laser (Laserglow Technologies) via a fiber-optic patch cord, and the neuronal responses were simultaneously recorded. The stimulation patterns included 100-ms constant light and 5 or 50 Hz (10-ms pulse width, 5 or 50 pulses in 1s) (Figure S6C–S6F). The inter-stimulation interval was 4 s and the stimulation pattern was repeated for 50 ~ 100 times. The laser power was adjusted carefully to a relatively low level (~ 1.0 – 1.5 mW) to evoke reliable spikes from individual neurons, since high laser powers frequently caused large electrical response generated by simultaneously activated large neuronal population, which differ greatly from spike waveforms recorded from single cells in the same electrode. Only those units 1) showing very short (~ 6 ms) response latency to light stimulation, 2) exhibiting exactly the same spike waveforms ($R = 0.95$, Pearson's correlation coefficient) during the behavioral performance and light response were identified as DAT-cre positive thus dopaminergic neurons (Cohen et al., 2012; Jin and Costa, 2010; Jin et al., 2014). Thus the rather strict criteria were used to avoid any potential false positive. The results and conclusions were qualitatively consistent if a longer response latency was considered (~ 11 ms, Figure S2F–S2I).

Computational model—We constructed a neuronal network model, including cortico-basal ganglia circuitry and dopamine signaling, to simulate the behavioral effects of optogenetic manipulation on dopamine. Specifically, cortical information corresponding to left or right choice is sent to the dorsal striatal D1- and D2-subpopulations associated with these two action options (Lo and Wang, 2006; Wang, 2002). Signals from D1- and D2-subpopulations then converge to two separate SNr populations through distinct pathways (Hikosaka et al., 2000; Jin et al., 2014; Mink, 2003), and exert opposing effects on SNr activity (Smith et al., 1998). Behavioral output is then determined by the dominant activity between the mutually inhibiting left and right SNr populations (Mailly et al., 2003), which could control the final motor output either through brainstem circuits or motor cortices (Hikosaka, 2007; Lo and Wang, 2006; Redgrave et al., 1999). Here for simplicity, other basal ganglia nuclei such as globus pallidus and subthalamic nucleus are not included in the model. The increase in dopamine concentration activates striatal D1 neurons, while inhibiting striatal D2 neurons, whereas a decrease in dopamine has an opposite effect on

these two populations (Gerfen and Surmeier, 2011). Therefore, dopamine could bias action selection by modulating striatal direct and indirect pathways, which in turn alter the SNr activity and eventually behavioral output. The left and right choice related cortical activity is defined as:

$$\begin{aligned} f_{left}^{cortex}(t) &= k_{left}^{cortex}(t - t_m) + I_{noise}(t) \\ f_{right}^{cortex}(t) &= k_{right}^{cortex}(t - t_m) + I_{noise}(t) \end{aligned}$$

where $k_{left}^{cortex} = -0.02$, $k_{right}^{cortex} = 0.02$, $t_m = 2$ and $I_{noise}(t)$ is defined as Gaussian white noise (mean = 0, SD = 0.1). In the following section, if not specifically defined, all the noise is referred to as the same Gaussian white noise.

Striatal dopamine concentration $f_{DA}(t)$ is defined as:

$$\begin{aligned} \tau_{DA} \frac{df_{DA}(t)}{dt} &= w_0(E - f_{DA}(t)) + w_{DA} I_{DA}(t) + I_{noise}(t) \\ I_{DA}(t) &= k_{DA}(t - t_m) \end{aligned}$$

where $w_0 = 1$, $E = 0$, $w_{DA} = 2$, $\tau_{DA} = 2$. Rate of dopamine decrease is defined as $k_{DA} = -0.06$.

Neuronal activity of striatal D1-/D2-spiny projection neurons are defined as:

$$\begin{aligned} \tau_{D1} \frac{df_{left}^{D1}(t)}{dt} &= w_0(E - f_{left}^{D1}(t)) + \tilde{w}_{D1} f_{left}^{cortex}(t) + w_{D1} f_{DA}(t) + I_{noise}(t) \\ \tau_{D1} \frac{df_{right}^{D1}(t)}{dt} &= w_0(E - f_{right}^{D1}(t)) + \tilde{w}_{D1} f_{right}^{cortex}(t) + w_{D1} f_{DA}(t) + I_{noise}(t) \\ \tau_{D2} \frac{df_{left}^{D2}(t)}{dt} &= w_0(E - f_{left}^{D2}(t)) + \tilde{w}_{D2} f_{left}^{cortex}(t) + w_{D2} f_{DA}(t) + I_{noise}(t) \\ \tau_{D2} \frac{df_{right}^{D2}(t)}{dt} &= w_0(E - f_{right}^{D2}(t)) + \tilde{w}_{D2} f_{right}^{cortex}(t) + w_{D2} f_{DA}(t) + I_{noise}(t) \end{aligned}$$

where $w_0 = 1$, $E = 0$, $\tilde{w}_{D1} = 1$, $w_{D1} = 2$, $\tilde{w}_{D2} = 1$, $w_{D2} = -2$, $\tau_{D1} = 0.2$, $\tau_{D2} = 0.2$.

SNr neurons receive striatal inputs as well as the local inhibitory inputs from other SNr neurons. Also based on the dominant lever press preference on left action (Figure S4D), we assume that the short-duration selection is more robust than the long-duration choice, thus to mimic the preference, the systematic noise in SNr ($I_{noise}^s(t)$) for short-duration selection is defined less than the long-duration selection. Then SNr activity is thus defined as:

$$\begin{aligned} \tau_{SNr} \frac{df_{left}^{SNr}(t)}{dt} &= w_0(E - f_{left}^{SNr}(t)) - \tilde{w}_{SNr} f_{left}^{D1}(t) + \tilde{w}_{SNr} f_{right}^{D2}(t) + w_{left}^{SNr} f_{right}^{SNr}(t) + I_{noise}^s(t) \\ \tau_{SNr} \frac{df_{right}^{SNr}(t)}{dt} &= w_0(E - f_{right}^{SNr}(t)) - \tilde{w}_{SNr} f_{right}^{D1}(t) + \tilde{w}_{SNr} f_{left}^{D2}(t) + w_{right}^{SNr} f_{left}^{SNr}(t) + I_{noise}^s(t) \end{aligned}$$

where $w_0 = 1$, $E = 0$, $\tilde{w}_{SNr} = 1$, $w_{left}^{SNr} = -0.1$, $w_{right}^{SNr} = -2$, $\tau_{SNr} = 0.5$. The time-dependent choice $S(t)$ is then determined by SNr outputs $f_{left}^{SNr}(t)$ and $f_{right}^{SNr}(t)$ as follows:

$$S(t) = \begin{cases} \text{left choice (short - duration choice),} & f_{left}^{SNr}(t) - f_{right}^{SNr}(t) < 0 \\ \text{right choice (long - duration choice),} & f_{left}^{SNr}(t) - f_{right}^{SNr}(t) \geq 0 \end{cases}$$

For optogenetic stimulation, the stimulation pattern is defined as:

$$F_{stim}(t) = \begin{cases} 1, & t_s \leq t \leq t_s+1 \\ 0, & t < t_s \text{ or } t > t_s+1 \end{cases}$$

and for inhibition, the pattern is defined as:

$$F_{inhibit}(t) = \begin{cases} -1, & t_s \leq t \leq t_s+1 \\ 0, & t < t_s \text{ or } t > t_s+1 \end{cases}$$

where t_s is the onset of stimulation/inhibition, which lasts for 1 s. The dopamine increases evoked by stimulation negatively correlate with the current dopamine concentration, that is when the current dopamine concentration is relatively high, then the optogenetically evoked increase is weaker than the one evoked when the current concentration is relatively low. Oppositely, the optogenetic inhibition only effectively depresses dopamine concentration when the current dopamine concentration is higher than baseline. Then, dopamine concentrations under optogenetic manipulation are defined as:

$$\begin{aligned} \tau_{DA} \frac{df_{DA}(t)}{dt} &= w_0(E - f_{DA}(t)) + w_{DA} I_{DA}^{stim}(t) + I_{noise}(t) \\ I_{DA}^{stim}(t) &= k_{DA}(t - t_m) + w_{stim} F_{stim}(t) e^{-(f_{DA}(t) + f_0)} \\ \tau_{DA} \frac{df_{DA}(t)}{dt} &= w_0(E - f_{DA}(t)) + w_{DA} I_{DA}^{inhibit}(t) + I_{noise}(t) \\ I_{DA}^{inhibit}(t) &= k_{DA}(t - t_m) + w_{inhibit} F_{inhibit}(t) e^{-t \cdot f_1} \end{aligned}$$

where $w_0 = 1$, $E = 0$, $w_{DA} = 2$, $\tau_{DA} = 2$, $k_{DA} = -0.06$, $w_{stim} = 3.5$, $w_{inhibit} = 10$, $f_0 = 2$, $f_1 = 0.15$.

For the simulation of DAT-NR1 KO mice, the dopamine concentration is defined as:

$$\begin{aligned} \tau_{DA} \frac{df_{DA}^{KO}(t)}{dt} &= w_0(E - f_{DA}^{KO}(t)) + w_{DA} I_{DA}^{KO}(t) + I_{noise}(t) \\ I_{DA}^{KO}(t) &= \begin{cases} k_{DA}^{KO}(t - t_m), & t \leq 3 \\ 0, & t > 3 \end{cases} \end{aligned}$$

where $w_0 = 1$, $E = 0$, $w_{DA} = 2$, $\tau_{DA} = 2$, $k_{DA}^{KO} = -0.04$.

In the constant dopamine simulation, dopamine concentration is defined as constant values (-0.05 , -0.025 , 0 , 0.025 , 0.05) combined with Gaussian white noise (mean = 0, SD = 0.03) during the 8s period. All the modeling programs were coded in Matlab.

Trial-by-trial analysis—The electrophysiological recording data and fast-scan voltammetry data are analyzed trial by trial to predict the behavioral choice. For the electrophysiological recording on a single mouse, N dopamine neurons are identified and recorded. n neurons ($n < N$) are randomly selected from these N dopamine neurons for further analysis. For these n neurons, the action potentials from 7s to 8s after lever retraction in the 8s trials are binned with 10ms-window. Then for a single trial, activity of n neurons are transformed into 100 n -dimensional vectors. Half of the rewarded and non-rewarded trials are randomly selected as training data set. All the n -dimensional vectors from those training data set (X) are included in the principal component analysis (PCA). n -dimensional vectors are projected onto PC plane by a coefficient matrix A (linear transformation)

$$PC_s = X \cdot A$$

The vectors corresponding to rewarded and non-rewarded trials are clustered as two isolated groups on the PC plane. The centroids of clusters are calculated by K-means. Then the n -dimensional vectors from the other half of the rewarded and non-rewarded trials (\bar{X}) are defined as testing data set, and they are also projected onto PC plane by

$$\overline{PC}_s = \bar{X} \cdot A$$

The Euclidean distance between \overline{PC}_s of a testing trial and the centroids of the two clusters is calculated ($D_{rewarded}$, $D_{non-rewarded}$). The behavioral choice in the testing trial is predicted as

$$choice = \begin{cases} \text{left choice (short - duration choice),} & D_{non-rewarded} < D_{rewarded} \\ \text{right choice (long - duration choice),} & D_{rewarded} < D_{non-rewarded} \end{cases}$$

The prediction accuracy is determined by the percentage of correctly predicted trials out of the total testing trials.

For the fast-scan voltammetry data, the dopamine concentration trace from 7s to 8s in the 8s trials are recorded with 100ms temporal resolution, thus the dopamine trace within this 1s period is defined as 10-dimensional vector. The same prediction approach is implemented as described above.

Psychometric curve simulation—Psychometric curves for behavioral data and for theoretical curves were fit using the following equation (Brunton et al., 2013):

$$y = a + \frac{b}{1 + e^{-\frac{c-x}{d}}}$$

where a is the percentage of long-lever selection during short duration trials, b is the difference between a and the percentage of long-lever selection during long duration trials, c is the x-intercept where long-duration selection equals 0.5, and d is the rate of increase or

decrease in the curve (slope). These can be interpreted as change in overall choice, long-duration choice, time, and sensitivity, respectively (Brunton et al., 2013). We systematically increased and decreased each of these terms to generate Figure S8B–F.

Hazard rate and value state—Hazard rate (Janssen and Shadlen, 2005) was calculated from a bimodal distribution $B(t)$. $B(t)$ was defined as the sum of two non-overlapping Gaussian distributions:

$$B(t) = \frac{G_1(t) + G_2(t)}{2}$$

where

$$G_1(t) = \begin{cases} \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(t-\mu_1)^2}{2\sigma_1^2}}, & \text{for } t \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

$$G_2(t) = \begin{cases} \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(t-\mu_2)^2}{2\sigma_2^2}}, & \text{for } t > 4 \\ 0, & \text{otherwise} \end{cases}$$

$$\sigma_1 = 0.3, \mu_1 = 2, \sigma_2 = 1, \mu_2 = 8$$

The anticipation should satisfy the conditional probability that an event will occur given that it has not yet occurred, which is termed as hazard rate (Janssen and Shadlen, 2005). Thus, the hazard rate is defined as the ratio between the probability that an event will occur at time t and the probability that it has not yet occurred:

$$H(t) = \frac{B(t)}{1 - F(t)}$$

where $F(t)$ is the cumulative distribution

$$F(t) = \int_0^t B(x) dx$$

For state value in 8s trials, we simplified the reinforcement-learning model described in Hamid et al. (2016) and limited it to within-trial, rather than across-trial analysis. The exponential function to replicate the state value $V(t)$ is defined as :

$$V(t) = \begin{cases} e^{(t+1)*0.07}, & 0 < t \leq 2 \\ e^{t*0.035}, & 2 < t \leq 8 \\ 0, & \text{other} \end{cases}$$

Data analysis—For voltammetry recordings, dopamine concentration changes in all the trials were aligned to lever retraction. The successful 2s and 8s trials were averaged respectively to demonstrate dopamine changes during action selection for each individual animal. Dopamine concentration changes were then averaged across mice. For all the 8s

trials, including 8s trials in 2–8 s task and 8s-only task, dopamine concentration was averaged within 0.5–1.5 s and 7–8 s following lever retraction. Time periods of 0.5–1.5 s and 7–8 s were defined as ‘early’ and ‘late’ phases, respectively. For analysis of dopamine concentration in 4–16 s task, 2–3 s and 14–15 s were defined as ‘early’ and ‘late’, respectively. For 16 s probe trials in the 2–8 s task, ‘early’ changes were defined as peak dopamine concentration within the first 1.5 s of trials, and ‘late’ changes were defined as 14–15 s. Dopamine traces shown in Figures 1H, 3C, 4B, 4E, 4H, 4K, 5C, 5G, 6D, 6E, and 6F represent rewarded trials. Differences in dopamine concentration between f_1 and f_2 tones (Figure S5) were defined as the absolute value of the difference in dopamine concentrations during a 1 s window following tone onset. Similarly, differences in dopamine concentration between left and right rewarded presses were defined as the absolute value of the difference in dopamine concentrations during a 1 s window prior to rewarded lever presses. These values were then compared across mice (Figure S5F).

For electrophysiological data analysis during 2–8 s task, neuronal firing was aligned to lever retraction and averaged across trials in 20-ms bins, and then smoothed by a Matlab build-in Gaussian filter (Gaussian filter window size = 50, standard deviation = 20) to construct the peri-event histogram (PETH). During 2s trials, mice behaved exactly the same as they do during the 0–2s period in the rewarded 8s trials, so we focused on the analysis of firing activities in 8s trials. To avoid confounding effects by sensory responses triggered by lever retraction, neural activity from 1s to 8s following lever retraction were included. We calculated Z-score based on the PETH from 1s to 8s for each individual neuron as follows:

$$Z - score = \frac{PETH - \text{mean}(PETH)}{\text{std}(PETH)}$$

We then used principal component analysis (PCA) and a classification algorithm (Matlab) to classify these dopamine neurons as increasing or decreasing.

For dopamine neuron identification, neuronal firings were aligned to laser onset and averaged across stimulation trials in 1-ms bins, and then smoothed by a Gaussian filter (Gaussian filter window size = 5, standard deviation = 3) to construct a light-evoked PETH. The baseline firing was defined as distributions of the PETH from –1000 to 0 ms before laser onset. The neuronal firing latency to light stimulation was defined as the beginning of significant firing rate increase after light onset. The threshold for significance test was defined as mean (baseline firing) + 1.96 × std (baseline firing) (Jin and Costa, 2010; Jin et al., 2014). A significant increase in firing rate was defined by at least 5 consecutive bins showing a significant firing rate increase above the threshold. Data analyses were conducted in Matlab with custom-written programs (MathWorks). The linear and fitting was conducted in Prism (Version 6.02, GraphPad) and Matlab (MathWorks). Representative behavioral tracking was done using EthoVision (Noldus).

Quantification and Statistical Analysis

Statistics—Changes in dopamine concentration were compared to baseline using a one-way repeated measures ANOVA with post hoc multiple comparison tests. Changes

dopamine ‘early’ and ‘late’ in 8 s trials were compared with a t-test. Changes in learning data, psychometric curves, and 16 s probe trial dopamine responses were compared using a two-way ANOVA with repeated measures for both factors and post hoc Fisher’s LSD multiple comparison tests for significant interactions. Changes in lever selection in optogenetics experiments and modeled data were compared using one sample t-test. Dopamine levels compared in simultaneous voltammetry and optogenetics experiments, as well as unrewarded lever press rates, were compared using a paired t-test. Statistical analysis was performed in Prism (Version 6.02, GraphPad) and Matlab (MathWorks).

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
AAV9-DIO-ChR2-EYFP	University of Pennsylvania Vector Core	Cat#AV-9-20298P
Experimental Models: Organisms/Strains		
C57BL/6 mice	Jackson Laboratory	Stock# 000664
DAT-cre mice	Jackson Laboratory	Stock# 020080
NR1-flox mice	Jackson Laboratory	Stock# 005246
RGS9-cre mice	Dang et al. 2006	N/A
Ai32(RCL-ChR2(H134R)/EYFP) mice	Jackson Laboratory	Stock# 024109
Ai35(RCL-Arch/GFP) mice	Jackson Laboratory	Stock# 012735
Software and Algorithms		
HDCV	Department of Chemistry at the University of North Carolina at Chapel Hill; Bucher et al. 2013	N/A
Matlab	The Mathworks Inc.	R2013a
Offline Sorter	Plexon Inc.	Version 3.3.3
OmniPlex	Plexon Inc.	Version 1.4.5
GraphPad Prism	GraphPad Software Inc.	Version 6.02
Med-PC	Med Associates Inc.	Cat#SOF-735
Tarheel	ESA Biosciences Inc.	N/A
EthoVision	Noldus Information Technology Inc.	Version 8.5
Other		
Med Associates operant chamber	Med Associates Inc.	Cat#MED-307W-D1
Electrode array	Innovative Neurophysiology Inc.	N/A
Voltammetric electrode	Clark et al. 2010	N/A

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Ed Callaway, Martyn Goulding, Tom Jessell, Chris Kintner, Terry Sejnowski and members of Jin lab for discussion and comments on the manuscript, as well as Erik Oleson, Yolanda Mateo, Scott Ng-Evans, and Kendra Bunner for technical discussion on FSCV. This research is supported by grants from the US

National Institutes of Health (R01NS083815, R01AG047669 and P30NS072031), the Dana Foundation, Ellison Medical Foundation and Whitehall Foundation to X.J.

References

- Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH. Beyond reward prediction errors: the role of dopamine in movement kinematics. *Front Integr Neurosci.* 2015; 9:39. [PubMed: 26074791]
- Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev.* 1998; 28:309–369. [PubMed: 9858756]
- Bonci A, Malenka RC. Properties and plasticity of excitatory synapses on dopaminergic and GABAergic cells in the ventral tegmental area. *J Neurosci.* 1999; 19:3723–3730. [PubMed: 10234004]
- Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron.* 2010; 68:815–834. [PubMed: 21144997]
- Brunton BW, Botvinick MM, Brody CD. Rats and humans can optimally accumulate evidence for decision-making. *Science.* 2013; 340:95–98. [PubMed: 23559254]
- Bucher ES, Brooks K, Verber MD, Keithley RB, Owesson-White C, Carroll S, Takmakov P, McKinney CJ, Wightman RM. Flexible software platform for fast-scan cyclic voltammetry data acquisition and analysis. *Anal Chem.* 2013; 85:10344–10353. [PubMed: 24083898]
- Bucher ES, Wightman RM. *Electrochemical Analysis of Neurotransmitters. Annual review of analytical chemistry.* 2015; 8:239–261.
- Buhusi CV, Meck WH. What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci.* 2005; 6:755–765. [PubMed: 16163383]
- Cachope R, Mateo Y, Mathur BN, Irving J, Wang HL, Morales M, Lovinger DM, Cheer JF. Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. *Cell reports.* 2012; 2:33–41. [PubMed: 22840394]
- Calabresi P, Pisani A, Mercuri NB, Bernardi G. Long-term Potentiation in the Striatum is Unmasked by Removing the Voltage-dependent Magnesium Block of NMDA Receptor Channels. *Eur J Neurosci.* 1992; 4:929–935. [PubMed: 12106428]
- Chow BY, Han X, Dobry AS, Qian X, Chuong AS, Li M, Henninger MA, Belfort GM, Lin Y, Monahan PE, et al. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature.* 2010; 463:98–102. [PubMed: 20054397]
- Church RM, Deluty MZ. Bisection of temporal intervals. *J Exp Psychol Anim Behav Process.* 1977; 3:216–228. [PubMed: 881613]
- Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, et al. Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. *Nature methods.* 2010; 7:126–129. [PubMed: 20037591]
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* 2012; 482:85–88. [PubMed: 22258508]
- Dang MT, Yokoi F, Yin HH, Lovinger DM, Wang Y, Li Y. Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc Natl Acad Sci U S A.* 2006; 103:15254–15259. [PubMed: 17015831]
- Ding L, Gold JJ. Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron.* 2012; 75:865–874. [PubMed: 22958826]
- Engblom D, Bilbao A, Sanchis-Segura C, Dahan L, Perreau-Lenz S, Balland B, Parkitna JR, Lujan R, Halbout B, Mameli M, et al. Glutamate receptors on dopamine neurons control the persistence of cocaine seeking. *Neuron.* 2008; 59:497–508. [PubMed: 18701074]
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci.* 2005; 8:1481–1489. [PubMed: 16251991]
- Gallistel, CR. *The organization of learning.* Cambridge, Mass.: MIT Press; 1990.
- Gerfen CR, Surmeier DJ. Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci.* 2011; 34:441–466. [PubMed: 21469956]

- Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem.* 1998; 70:119–136. [PubMed: 9753592]
- Graybiel AM, Rauch SL. Toward a neurobiology of obsessive-compulsive disorder. *Neuron.* 2000; 28:343–347. [PubMed: 11144344]
- Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD. Mesolimbic dopamine signals the value of work. *Nat Neurosci.* 2016; 19:117–126. [PubMed: 26595651]
- Hikosaka O. GABAergic output of the basal ganglia. *Prog Brain Res.* 2007; 160:209–226. [PubMed: 17499116]
- Hikosaka O, Miyashita K, Miyachi S, Sakai K, Lu X. Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. *Neurobiol Learn Mem.* 1998; 70:137–149. [PubMed: 9753593]
- Hikosaka O, Takikawa Y, Kawagoe R. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev.* 2000; 80:953–978. [PubMed: 10893428]
- Howard CD, Daberkow DP, Ramsson ES, Keefe KA, Garris PA. Methamphetamine-induced neurotoxicity disrupts naturally occurring phasic dopamine signaling. *Eur J Neurosci.* 2013; 38:2078–2088. [PubMed: 23574406]
- Howe MW, Dombeck DA. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature.* 2016; 535:505–510. [PubMed: 27398617]
- Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature.* 2013; 500:575–579. [PubMed: 23913271]
- Janssen P, Shadlen MN. A representation of the hazard rate of elapsed time in macaque area LIP. *Nat Neurosci.* 2005; 8:234–241. [PubMed: 15657597]
- Jin X, Costa RM. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature.* 2010; 466:457–462. [PubMed: 20651684]
- Jin X, Costa RM. Shaping action sequences in basal ganglia circuits. *Curr Opin Neurobiol.* 2015; 33:188–196. [PubMed: 26189204]
- Jin X, Tecuapetla F, Costa RM. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat Neurosci.* 2014; 17:423–430. [PubMed: 24464039]
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science.* 1999; 286:1745–1749. [PubMed: 10576743]
- Keithley RB, Heien ML, Wightman RM. Multivariate concentration determination using principal component regression with residual analysis. *Trends in analytical chemistry : TRAC.* 2009; 28:1127–1136. [PubMed: 20160977]
- Kim HF, Ghazizadeh A, Hikosaka O. Dopamine Neurons Encoding Long-Term Memory of Object Value for Habitual Behavior. *Cell.* 2015; 163:1165–1175. [PubMed: 26590420]
- Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. *Neuron.* 2008; 58:451–463. [PubMed: 18466754]
- Lauwereyns J, Watanabe K, Coe B, Hikosaka O. A neural correlate of response bias in monkey caudate nucleus. *Nature.* 2002; 418:413–417. [PubMed: 12140557]
- Lima SQ, Hromadka T, Znamenskiy P, Zador AM. PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One.* 2009; 4:e6099. [PubMed: 19584920]
- Lo CC, Wang XJ. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci.* 2006; 9:956–963. [PubMed: 16767089]
- Madisen L, Mao T, Koch H, Zhuo JM, Berenyi A, Fujisawa S, Hsu YW, Garcia AJ 3rd, Gu X, Zanella S, et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nat Neurosci.* 2012; 15:793–802. [PubMed: 22446880]
- Mailly P, Charpier S, Menetrey A, Deniau JM. Three-dimensional organization of the recurrent axon collateral network of the substantia nigra pars reticulata neurons in the rat. *J Neurosci.* 2003; 23:5247–5257. [PubMed: 12832549]

- Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
- Mink JW. The Basal Ganglia and involuntary movements: impaired inhibition of competing motor patterns. *Arch Neurol*. 2003; 60:1365–1368. [PubMed: 14568805]
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci*. 2006; 9:1057–1063. [PubMed: 16862149]
- Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 2007; 191:507–520. [PubMed: 17031711]
- Nugent FS, Penick EC, Kauer JA. Opioids block long-term potentiation of inhibitory synapses. *Nature*. 2007; 446:1086–1090. [PubMed: 17460674]
- Oleson EB, Cachope R, Fitoussi A, Tsutsui K, Wu S, Gallegos JA, Cheer JF. Cannabinoid receptor activation shifts temporally engendered patterns of dopamine release. *Neuropsychopharmacology*. 2014; 39:1441–1452. [PubMed: 24345819]
- Panigrahi B, Martin KA, Li Y, Graves AR, Vollmer A, Olson L, Mensh BD, Karpova AY, Dudman JT. Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell*. 2015; 162:1418–1430. [PubMed: 26359992]
- Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat Neurosci*. 2016; 19:845–854. [PubMed: 27110917]
- Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM. Subsecond dopamine release promotes cocaine seeking. *Nature*. 2003; 422:614–618. [PubMed: 12687000]
- Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci*. 2006; 7:967–975. [PubMed: 17115078]
- Redgrave P, Prescott TJ, Gurney K. The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*. 1999; 89:1009–1023. [PubMed: 10362291]
- Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y, DeLong MR, Obeso JA. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat Rev Neurosci*. 2010; 11:760–772. [PubMed: 20944662]
- Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*. 2007; 10:1615–1624. [PubMed: 18026098]
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science*. 2005; 310:1337–1340. [PubMed: 16311337]
- Schultz W. Multiple dopamine functions at different time courses. *Annu Rev Neurosci*. 2007; 30:259–288. [PubMed: 17600522]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
- Simen P, Matell M. Why does time seem to fly when we're having fun? *Science*. 2016; 354:1231–1232. [PubMed: 27940832]
- Smith Y, Bevan MD, Shink E, Bolam JP. Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*. 1998; 86:353–387. [PubMed: 9881853]
- Soares S, Atallah BV, Paton JJ. Midbrain dopamine neurons control judgment of time. *Science*. 2016; 354:1273–1277. [PubMed: 27940870]
- Stopper CM, Tse MT, Montes DR, Wiedman CR, Floresco SB. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*. 2014; 84:177–189. [PubMed: 25220811]
- Surmeier DJ, Plotkin J, Shen W. Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. *Curr Opin Neurobiol*. 2009; 19:621–628. [PubMed: 19896832]
- Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME. Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nat Neurosci*. 2016; 19:34–36. [PubMed: 26642087]
- Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci*. 2012; 15:1281–1289. [PubMed: 22902719]

- Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron*. 2012; 75:58–64. [PubMed: 22794260]
- Wang XJ. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*. 2002; 36:955–968. [PubMed: 12467598]
- Wickens JR, Horvitz JC, Costa RM, Killcross S. Dopaminergic mechanisms in actions and habits. *J Neurosci*. 2007; 27:8181–8183. [PubMed: 17670964]
- Zweifel LS, Argilli E, Bonci A, Palmiter RD. Role of NMDA receptors in dopamine neurons for plasticity and addictive behaviors. *Neuron*. 2008; 59:486–496. [PubMed: 18701073]

Highlights

- Nigrostriatal dopamine signaling is associated with ongoing action selection
- Dopamine signaling is necessary for appropriate action selection
- Optogenetic manipulations can bidirectionally modulate online action selection
- Modeling suggests dopamine could bias choice by modifying striatal activity

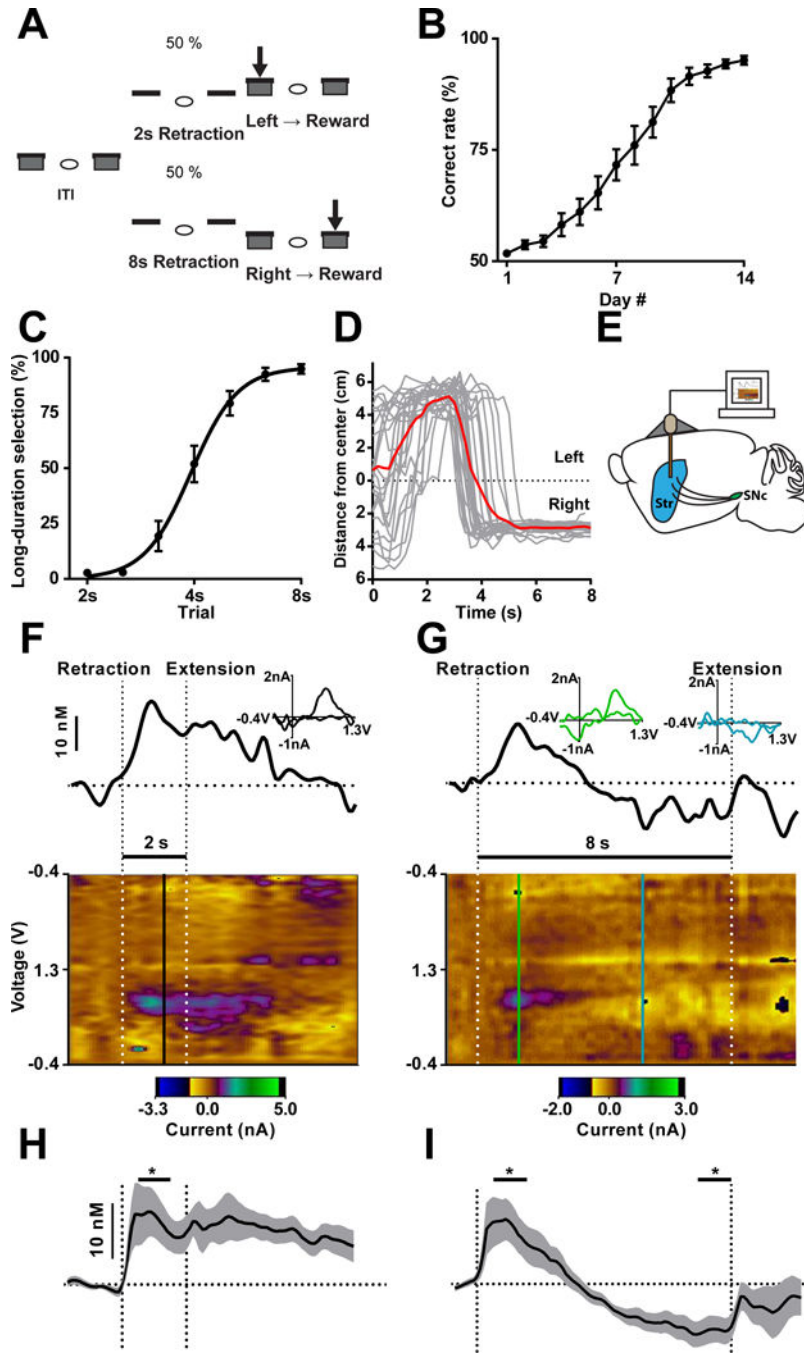


Figure 1. Nigrostriatal dopamine signaling associated with action selection
 (A) Task design of 2–8 s task. (B) Mice improve performance of the action selection task over two weeks training. $n = 17$ mice. $P < 0.0001$; one-way repeated measures ANOVA. (C) Psychometric curve of interval discrimination. $n = 7$ mice. $P < 0.0001$; one-way repeated measures ANOVA. (D) Behavioral tracking of a representative mouse during 8 s trials. Tracks for individual trials shown in grey and average shown in red. (E) Cartoon of carbon fiber microelectrodes implanted bilaterally into dorsal striatum (Str). (F and G) Representative real-time changes in dopamine concentration in dorsal striatum during

correctly performed 2 s (F) and 8 s (G) trials. Thick black line shows change in dopamine concentration over time aligned to lever retraction/extension (vertical dotted lines) above a pseudocolor plot, which illustrates the change in current recorded at each point in the voltage sweep across time. Baseline is denoted by horizontal dotted line. INSET shows a cyclic voltammogram collected at the vertical lines on the pseudocolor plot identifying the current recorded as either positive (black in F, green in G) or negative (blue in G) changes in dopamine. (H) Average correct 2 s trials show increases in dopamine at lever retraction relative to baseline. $n = 11$ mice. $P < 0.05$; paired t-test. (I) Dopamine increases in the early phase (0.5–1.5 s) and decreases below baseline in the late phase (7–8 s) during 8 s trials. $n = 11$ mice. $P < 0.01$; one-way repeated measures ANOVA with Fisher's LSD *post hoc* tests. Data are shown as mean \pm SEM. * $P < 0.05$. Same for below unless stated otherwise.

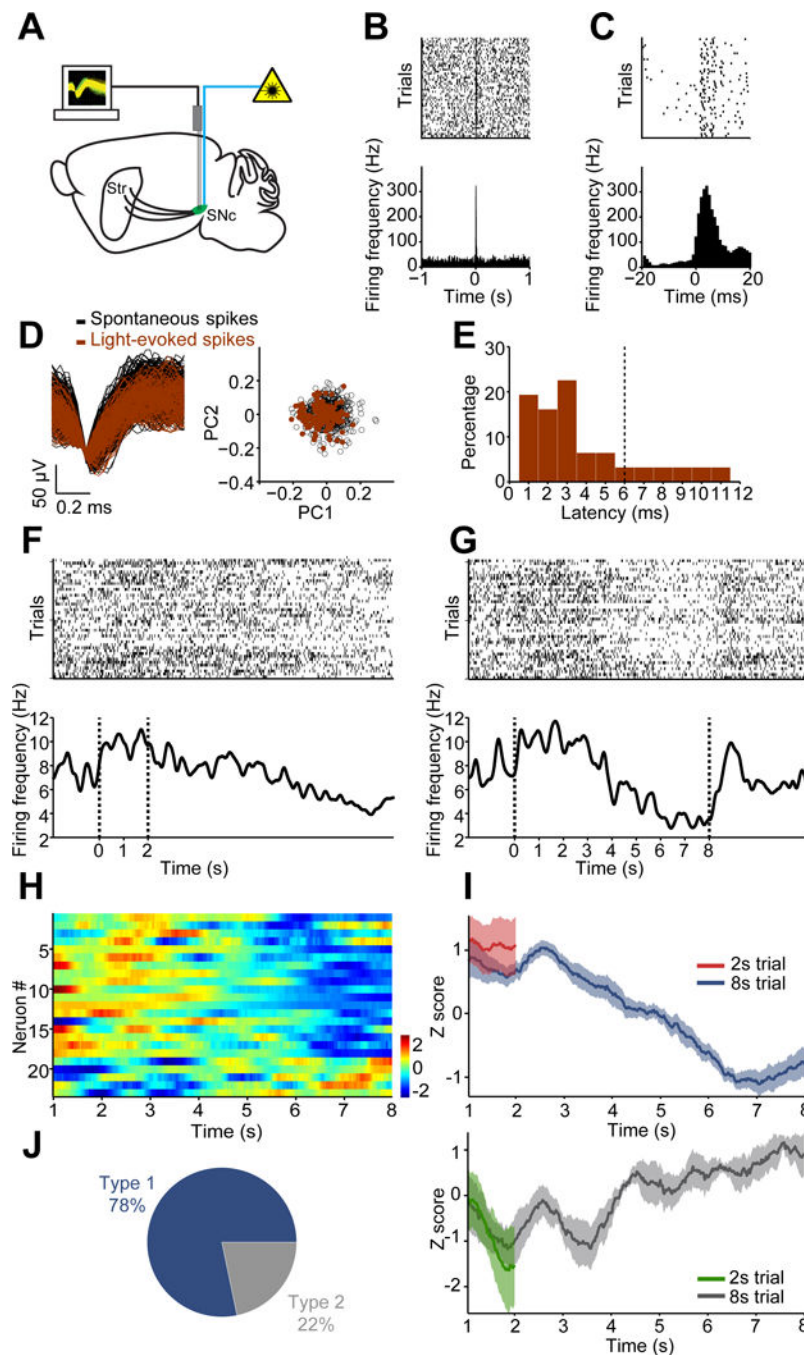


Figure 2. SNc Dopamine neurons display action selection-specific firing activity

(A) Diagram of simultaneous neuronal recording and optogenetic stimulation in substantia nigra pars compacta (SNc). (B) Top panel: Raster plot for a representative dopamine neuron response to 100 ms optogenetic stimulation. Each row represents one trial and black ticks represent spikes. Bottom panel: Averaged firing rate aligned to light onset at 0. (C) Raster plot and averaged firing rate for the same neuron as shown in (B) with a finer time scale. (D) Left panel: Waveforms from the same neuron in (B) for spontaneous (black) and light-evoked (red) spikes ($R = 0.998$, $P < 0.0001$, Pearson's correlation). Right panel: Principal

component analysis (PCA) of spontaneous and light-evoked waveforms shows the overlapped clustering of spontaneous (black) and light-evoked (red) spikes. (E) Histogram showing latency of optogenetically-evoked neuronal response. The vertical line (6ms) shows the criterion considered for inclusion. (F and G) Raster plots (top) and firing rate (bottom) for a representative dopamine neuron during 2 s (F) and 8 s (G) trials. (H) Z-score of neuronal activity for all positively identified dopamine neurons in 8 s trials. (I) Averaged neuronal activity of decreasing (Type 1) and increasing (Type 2) neurons during the 2 s trials and 8 s trials. (J) Proportion of Type 1 and Type 2 dopamine neurons.

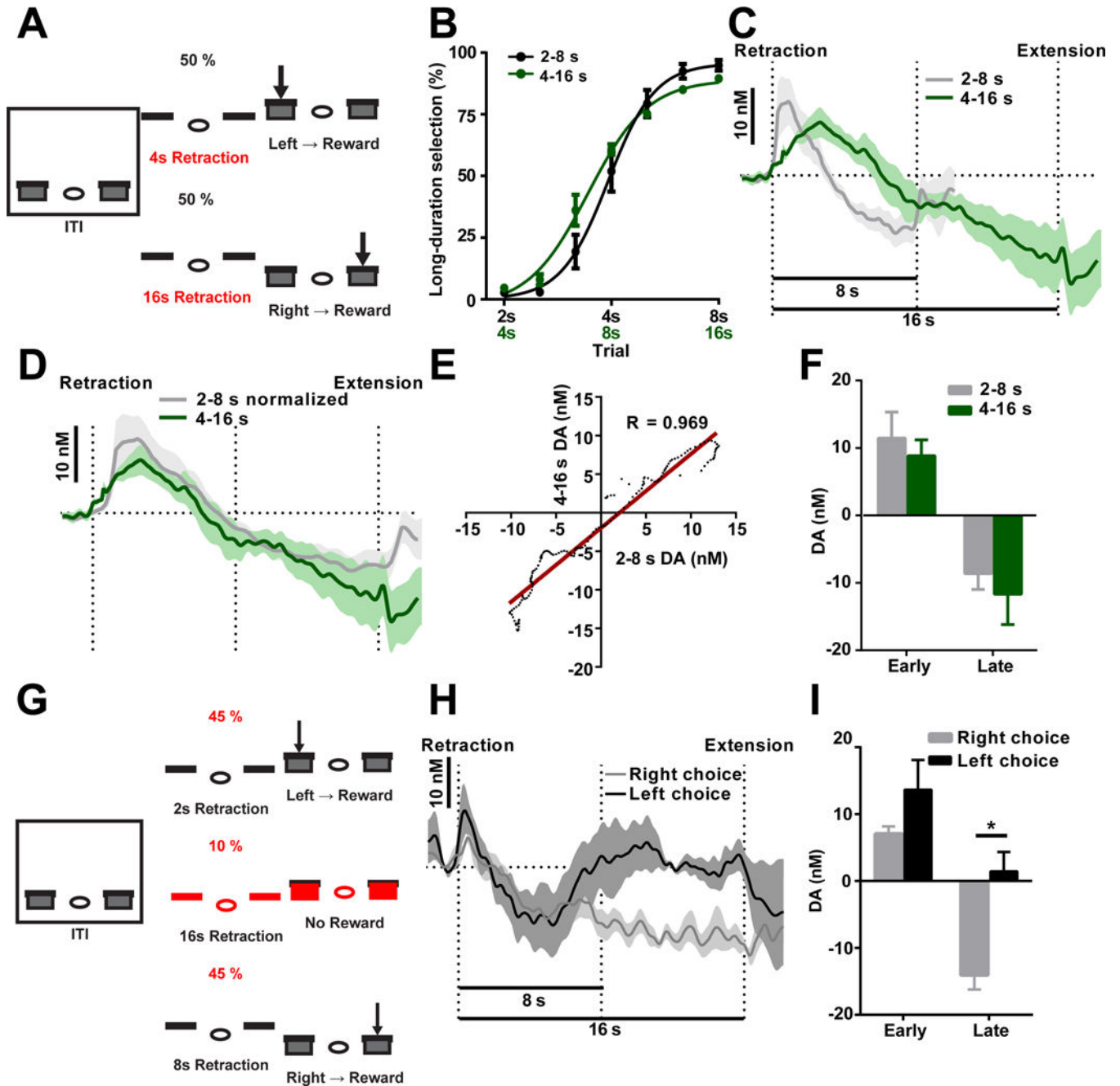


Figure 3. Dopamine changes are scalable to time intervals and are related to action selection
 (A) Task design for the 4–16 s task. (B) Psychometric curves for mice trained on 2–8 s task (black) and 4–16 s task (green). $n = 7$ for 2–8 s, $n = 4$ for 4–16 s. $P > 0.05$; two-way repeated measures ANOVA. (C) Dopamine recorded during the 16 s trials in the 4–16 s. (D and E) Dopamine recorded in both tasks are normalized to fit the same timescale (D), and they are significantly correlated (E; $R = 0.969$, $P < 0.0001$). (F) Dopamine magnitude is similar during the early (unpaired t-test, $P > 0.05$, $n = 6$ for 4–16 s, $n = 11$ for 2–8 s) and late phases (unpaired t-test, $P > 0.05$, $n = 6$ for 4–16 s, $n = 11$ for 2–8 s). (G) Task diagram for 16 s probe trials. (H) Dopamine recorded in 16 s probe trials (black: short-duration lever

selection; grey: long-duration lever selection). (I) Dopamine signaling in the early phase (paired t-test, $P > 0.05$, $n = 4$) and late phase (paired t-test, $P > 0.05$, $n = 4$).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

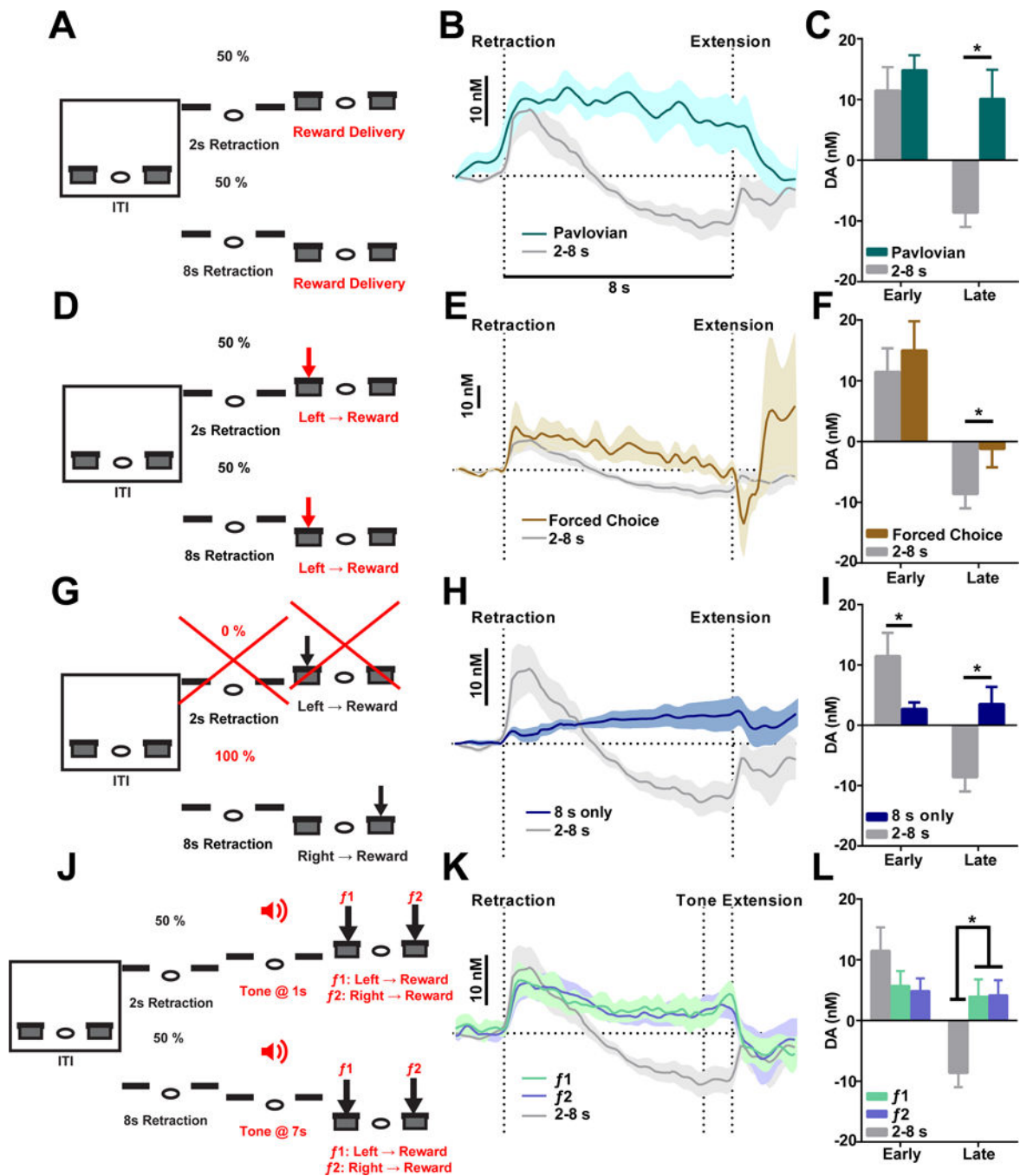


Figure 4. Dopamine changes in 2–8 s task are not simply tracking relative timing, changes in value, or uncertainty
 (A) Task design for 2–8 s Pavlovian task. (B) Recordings in the Pavlovian task are shown in teal and recordings from the 2–8 s task are shown in grey. (C) Dopamine magnitude between the 2–8 s Pavlovian and 2–8 s task during the early phase and late phase. $P > 0.05$ for early; $P < 0.001$ for late; unpaired t-test. (D) Task design for 2–8 s forced choice task. (E) Recordings in the forced choice task are shown in brown and recordings from the 2–8 s task are shown in grey. (F) Dopamine magnitude between the forced choice task and 2–8 s task during the early phase and late phase. $P > 0.05$ for early, $P < 0.05$ for late; unpaired t-test. (G) Task design for 2–8 s tone task with 0% and 100% trials. (H) Recordings in the tone task are shown in blue and recordings from the 2–8 s task are shown in grey. (I) Dopamine magnitude between the tone task and 2–8 s task during the early phase and late phase. $P < 0.05$ for early, $P < 0.05$ for late; unpaired t-test. (J) Task design for 2–8 s tone task with f1 and f2 tones. (K) Recordings in the tone task are shown in green and recordings from the 2–8 s task are shown in grey. (L) Dopamine magnitude between the tone task and 2–8 s task during the early phase and late phase. $P < 0.05$ for early, $P < 0.05$ for late; unpaired t-test.

(G) Task design for 8 s only task. (H) Recordings in the 8 s only task are shown in blue and recordings from the 2–8 s task are shown in grey. (I) Dopamine magnitude between the 8 s only task and 2–8 s task during the early phase and late phase. $P < 0.05$ for early, $P < 0.001$ for late; unpaired t-test. (J) Task design for 2–8 s tone task. (K) Dopamine recorded during 8 s trials with presentation of f_1 (green) and f_2 (purple). (L) Dopamine magnitude in the tone task and 2–8 s task during the early phase and late phase. $P > 0.05$ for early, $P < 0.05$ for late; two-way repeated measures ANOVA, significant interaction between time and groups, with significant LSD *post hoc* tests.

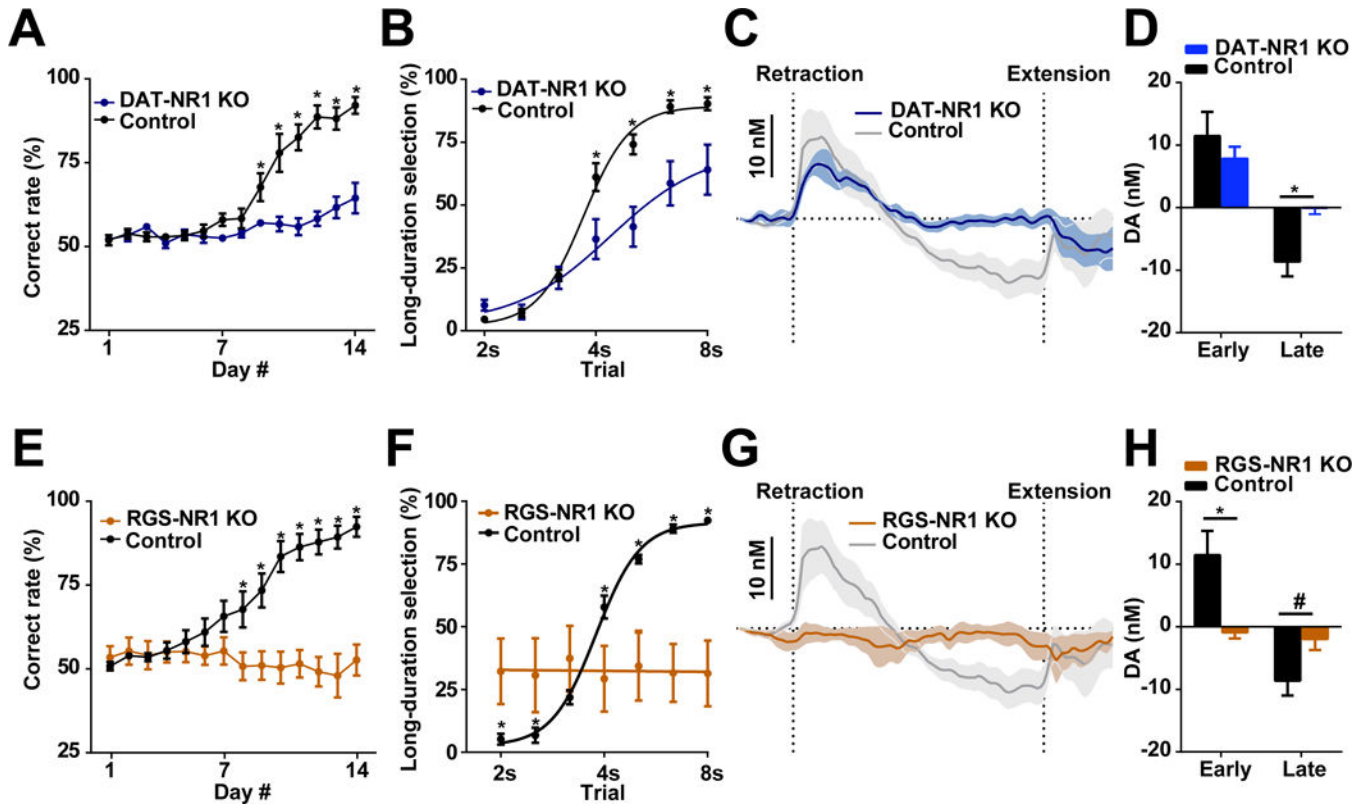


Figure 5. Selectively deleting NMDA receptors in either dopamine or striatal projection neurons disrupts dopamine signaling and alters action selection

(A) Performance of DAT-NR1 KO mice during 14 days training. $n = 7$ DAT-NR1 KO, $n = 6$ control. $P < 0.0001$; two-way repeated measures ANOVA, significant effect of group; * significant Fisher's LSD *post hoc* tests. (B) Psychometric curve of DAT-NR1 KO and control mice. $n = 7$ DAT-NR1 KO, $n = 6$ control. $P < 0.05$; two-way repeated measures ANOVA, significant effect of group; * significant Fisher's LSD *post hoc* tests. (C) Dopamine recorded in DAT-NR1 KO mice. $n = 4$ DAT-NR1 KO, $n = 11$ control. (D) Dopamine magnitude during early and late phase. $n = 4$ DAT-NR1 KO, $n = 11$ control. $P > 0.05$ for early, $P < 0.05$ for late; unpaired t-test. (E) Performance of RGS-NR1 KO mice during 14 days training. $n = 6$ RGS-NR1 KO, $n = 9$ control. $P < 0.01$; two-way repeated measures ANOVA, significant effect of group, * significant Fisher's LSD *post hoc* tests. (F) Psychometric curve of RGS-NR1 KO mice. $n = 6$ RGS-NR1 KO, $n = 9$ control. $P < 0.0001$; two-way repeated measures ANOVA, significant group \times time interaction, * significant Fisher's LSD *post hoc* tests. (G) Dopamine recorded in RGS-NR1 KO mice. $n = 4$ RGS-NR1 KO, $n = 11$ control. (H) Dopamine magnitude during early and late phase. $n = 4$ RGS-NR1 KO, $n = 11$ control. $P < 0.05$ for early; $P = 0.065$ for late; unpaired t-test. Dopamine traces are shown as mean \pm SEM. * $P < 0.05$, # $P < 0.1$.

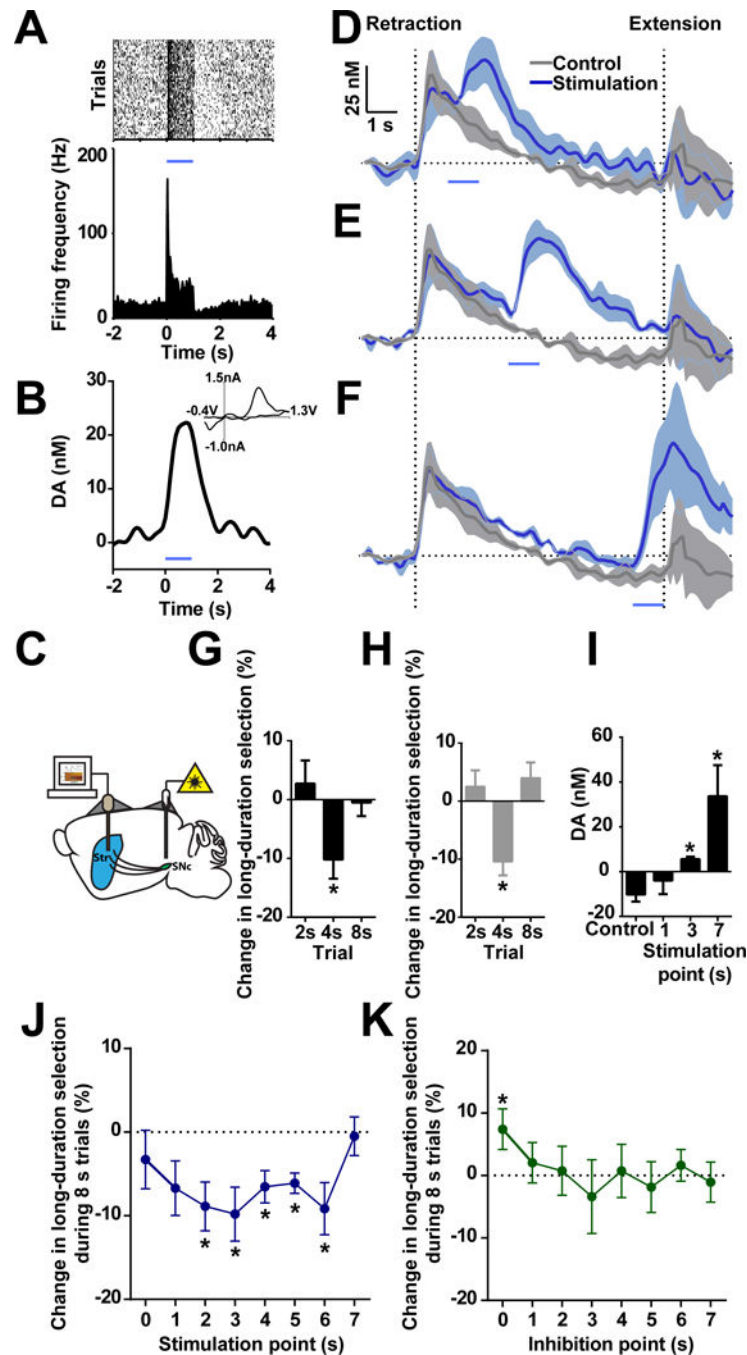


Figure 6. Optogenetic manipulations of dopamine signaling bias action selection

(A) Optogenetic stimulation of nigral dopamine neuron for 1 s drives firing activity *in vivo*. Stimulation is shown as blue line for current and following panels. (B) 1 s optogenetic stimulation evokes robust dopamine release. INSET shows current versus voltage identifying the recorded analyte as dopamine. (C) Fiber microelectrode placement in dorsal striatum (Str) and bilateral fiber optic placement in substantia nigra pars compacta (SNc). (D–F) Simultaneous optogenetic stimulation and FSCV recording during optical stimulation at 1 (D), 3 (E), and 7 s (F) in 8 s trials. (G) Immediate behavioral effects under dopamine neuron

stimulation. $n = 11$. $P < 0.01$ in 4 s trials; $P > 0.05$ in 2 or 8 s trials; one sample t-test. (H) Immediate behavioral effects while stimulating dopamine terminals in striatum. $n = 3$. $P < 0.05$ in 4 s trials; $P > 0.05$ in 2 or 8 s trials; one sample t-test. (I) Dopamine recorded with optogenetic stimulation at 1, 3 or 7 s. $n = 3$. $P > 0.05$ for 1 s; $P < 0.05$ for 3 and 7 s; paired t-test. (J) Change in long-duration selection in 8 s trials during optogenetic stimulation of dopamine at various time points. $n = 11$. one sample t-test, $*P < 0.05$. (K) Change in long-duration selection in the 8 s trials during optogenetic inhibition of dopamine at various time points. $n = 7$. one sample t-test, $*P < 0.05$. Dopamine traces are shown as mean \pm SEM.

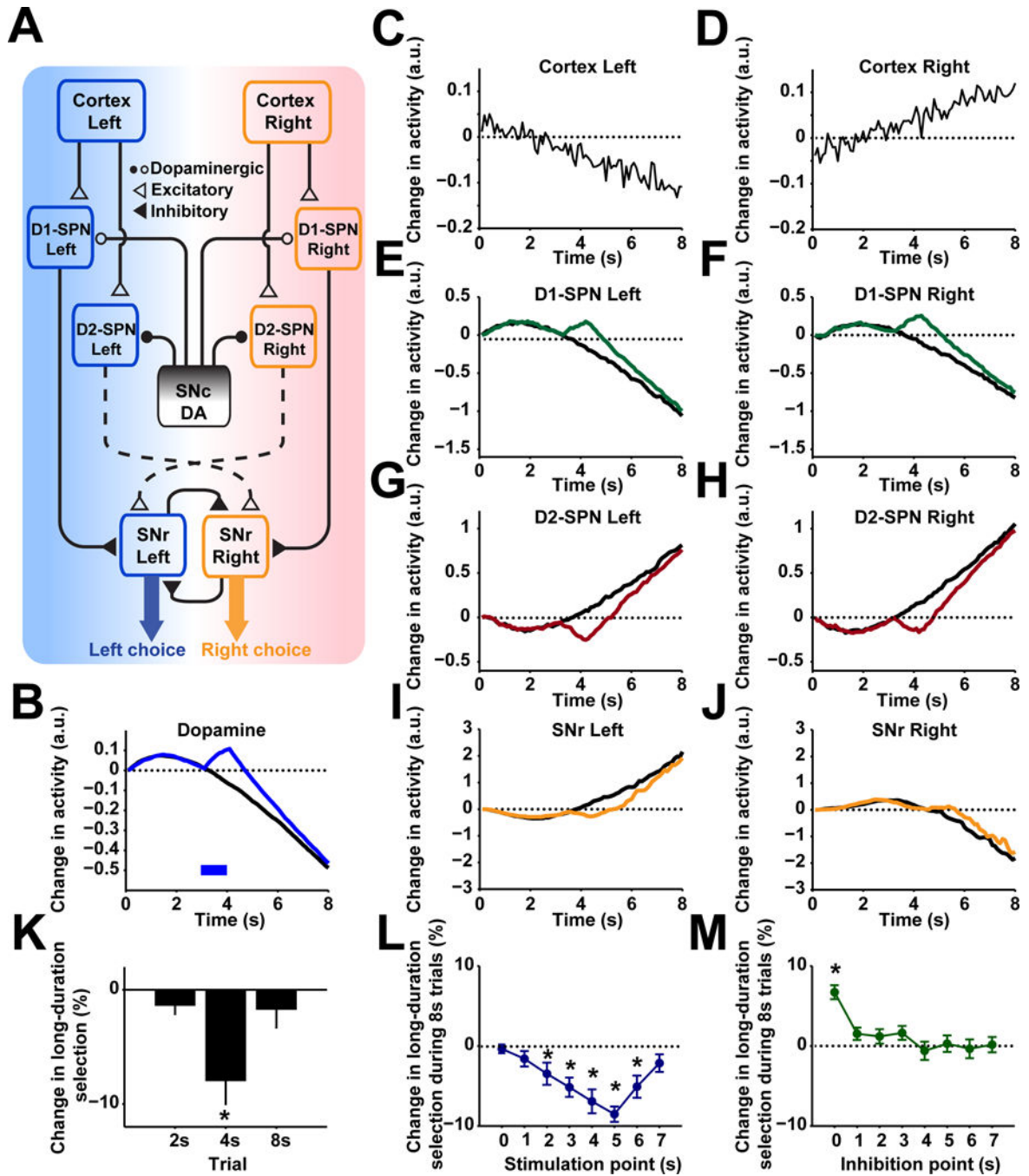


Figure 7. Computational model of dopamine biasing action selection through basal ganglia circuitry

(A) Network structure of the cortico-basal ganglia action selection model. For all panels, ‘left’ and ‘right’ refers to activity encoding the left or right choice. (B) Modeled changes in dopamine during control (black) and stimulation trials (blue) across 8 s trials with stimulation occurring at 3 s following lever retraction. (C and D) Modeled cortical inputs to striatal populations that encode information about left or right choice. (E and F) Modeled changes in D1 spiny projection neurons (SPNs) populations encoding either left or right choice under control (black) and stimulation (green). (G and H) Modeled changes in D2

SPNs encoding either left or right choice under control (black) and stimulation (red). (I and J) Modeled changes in SNr populations encoding left or right choice under control (black) and stimulation (orange). (K) Predicted changes in behavioral output based on SNr activity during probe trials of 2, 4, or 8 s duration where 1-s stimulation occurs at 1, 3, or 7 s, respectively. $n = 10$. one sample t-test, $*P < 0.05$. (L) Predicted changes in behavioral output during 8 s trials with stimulation occurring at each second across the 8 s interval. $n = 10$. one sample t-test, $*P < 0.05$. (M) Predicted changes in behavioral output during 8 s trials with inhibition occurring at each second across the 8 s interval. $n = 10$. one sample t-test, $*P < 0.05$. Data show mean \pm SEM of 10 modeled subjects.

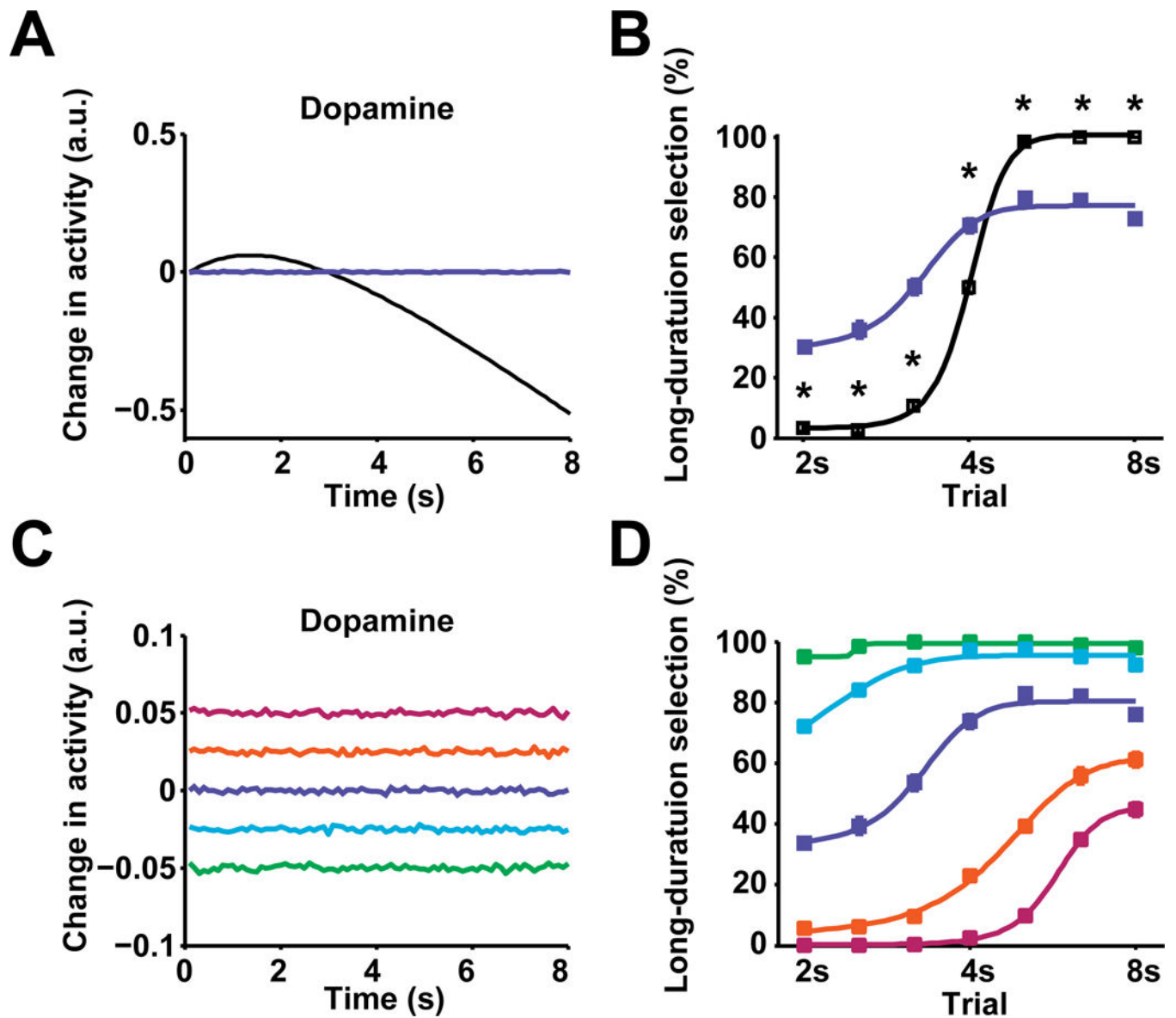


Figure 8. Dynamic dopamine is crucial for optimizing action selection

(A) Modeled dynamic (black) and constant (blue) dopamine in 8 s trials. (B) Model output of psychometric curve for behavioral choice under dynamic (black, $n = 10$) and constant dopamine (blue, $n = 10$; $P < 0.0001$; two-way repeated measures ANOVA, significant effect of group, * significant Fisher's LSD *post hoc* tests.). (C) Modeled dopamine at five different constant levels (-0.05 , -0.025 , 0 , 0.025 , 0.05) in 8 s trials. (D) Model output of psychometric curves for behavioral choice under constant dopamine of corresponding levels. Psychometric curves for different dopamine levels are color coded the same as (C). $n = 10$ modeled subjects.