Multiple dynamic interactions from basal ganglia direct and indirect pathways mediate action selection

Hao Li[1], and Xin Jin[1,2,3,4*]

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

[2] Center for Motor Control and Disease, Key Laboratory of Brain Functional Genomics, East China Normal University, 3663 North Zhongshan Road, Shanghai 200062, China

[3] NYU–ECNU Institute of Brain and Cognitive Science, New York University Shanghai, 3663 North Zhongshan Road, Shanghai 200062, China

[4] Lead Contact

*Correspondence to: xjin@bio.ecnu.edu.cn

Summary

The basal ganglia are known to be essential for action selection. However, the functional role of basal ganglia direct and indirect pathways in action selection remains unresolved. Here by employing cell-type-specific neuronal recording and manipulation in mice trained in a choice task, we demonstrate that multiple dynamic interactions from the direct and indirect pathways control the action selection. While the direct pathway regulates the behavioral choice in a linear manner, the indirect pathway exerts a nonlinear inverted-U-shaped control over action selection, depending on the inputs and the network state. We propose a new center (direct) - surround (indirect) - context (indirect) "Triple-control" functional model of basal ganglia, which can replicate the physiological and behavioral experimental observations that cannot be simply explained by either the traditional "Go/No-go" or more recent "Co-activation" model. These findings have important implications on understanding the basal ganglia circuitry and action selection in health and disease.

In Brief

Using behavior analysis, in vivo electrophysiology, optogenetics and computational modeling in mice, Li and Jin unveiled the neuronal dynamics of basal ganglia direct and indirect pathways underlying action selection, and proposed a new “Triple-control” functional model of basal ganglia.

Introduction

Selecting the proper actions is essential for organism's survival and reproduction in the ever-changing environment (Gallistel, 1980). Numerous studies have implicated that the basal ganglia, a series of interconnected subcortical nuclei including the striatum and substantia nigra, play a primary role in action selection (Graybiel, 1998; Hikosaka et al., 1998; Jin and Costa, 2015; Mink, 2003; Redgrave et al., 1999). Indeed, a wide range of neurological and psychiatric disorders associated with the dysfunctional basal ganglia circuitry, including Parkinson's disease (Benecke et al., 1987), Huntington's disease (Phillips et al., 1995), Obsessive-compulsive disorder (OCD) (Graybiel and Rauch, 2000), are characterized by major deficits in action selection and movement control. Anatomically, commands for motor control are processed by basal ganglia through two major pathways, termed direct and indirect pathway, originating from striatal D1- and D2-expressing spiny projection neurons (D1-/D2-SPNs), respectively (Albin et al., 1989; DeLong, 1990). These two pathways collectively modulate substantia nigra pars reticulata (SNr) activity and the basal ganglia output, thus influence behavioral decisions. There are currently two major types of thinking on how the basal ganglia pathways work. An early classic theory has suggested that the basal ganglia direct and indirect pathways oppose each other to facilitate and inhibit action, respectively (the "Go/No-go" model) (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010). In contrast, a recent theory has proposed that direct pathway selects the desired action, while the indirect pathway inhibits other competing actions in order to highlight the targeted choice (the "Co-activation" model) (Cui et al., 2013; Hikosaka et al., 2000; Mink, 1996).

The two theories have essentially agreed upon the function of direct pathway being the positive driving force for initiating or facilitating the desired actions. Yet, the ideas about the indirect pathway function are largely controversial as either impeding the desired action in the "Go/No-go" model or inhibiting the competing actions in the "Co-activation" model. While the precise neuroanatomy on how the D2-SPNs control SNr through indirect pathway has yet to be mapped out at single-cell level to differentiate the two hypotheses, either theory has found its supports from behavioral and physiological observations. For instance, it has been found that stimulation of striatal direct and indirect pathways can bidirectionally regulate locomotion (Durieux et al., 2012; Kravitz et al., 2010), consistent with the traditional 'Go/No-go' model. On the other hand, in vivo electrophysiological and imaging experiments revealed that the striatal direct and indirect pathways are both activated during action initiation (Barbera et al., 2016; Cui et al., 2013; Geddes et al., 2018; Isomura et al., 2013; Jin et al., 2014; Klaus and Plenz, 2016; Markowitz et al., 2018; Nonomura et al., 2018), as the 'Co-activation' model predicted. Furthermore, physiological and optogenetic studies concerning complex behavior such as learned action sequences have further complicated the issue, and unveiled various neuronal subpopulations in both pathways are activated during the initiation, termination and switching of actions (Geddes et al., 2018; Jin and Costa, 2015; Jin et al., 2014; Tecuapetla et al., 2016). So far, how exactly the basal ganglia direct and indirect pathways work together to control action selection has been controversial and inconclusive, and the underlying circuit mechanism remains largely unclear (Calabresi et al., 2014).

Here we trained mice to perform an operant action selection task where they were required to select one out of two actions to achieve reward, based on self-monitored time intervals (Howard et al., 2017). By employing in vivo neuronal recording, we found that the net

output of two opponent SNr neuron populations is predictive of the behavioral choices. Through identifying striatal pathway-specific neuronal activity with optogenetic tagging, we found that there are neuronal populations in either the direct or indirect pathway that are activated during selecting one action and suppressed during another. Optogenetic inhibition, as well as selective ablation of direct pathway impairs action selection, and optogenetic excitation of direct pathway enhances current choice, confirming a role of direct pathway in facilitating desired actions. Furthermore, optogenetic inhibition of indirect pathway improves action selection and excitation of indirect pathway impairs behavioral choices, as predicted by the 'Go/No-go' model. However, selective ablation of indirect pathway impairs action selection, opposite from the behavioral effect of optogenetic inhibition and at odds with the 'Go/No-go' model, but consistent with the prediction from the 'Co-activation' model. To resolve these contradictions, we propose a new center (direct) - surround (indirect) - context (indirect) "Triple-control" functional model of basal ganglia pathways, in which there are two interacting indirect pathway subcircuits exerting opposite controls over the basal ganglia output. The new model can reproduce the neuronal and behavioral experimental results that cannot be simply explained by either the "Go/No-go" or the "Co-activation" model. Further systematic analyses from this new model suggested that the direct and indirect pathways modulate behavioral outputs in a linear and nonlinear manner, respectively. Notably, in the new 'Triple-control' model, the direct and indirect pathways can work together to dynamically control action selection and operate in a manner similar to 'Go/No-go' or 'Co-activation' model, depending on the activity level and the network state. These results revise our current understanding on how the basal ganglia control actions, and have important implications for a wide range of movement and psychiatric diseases where the dynamic balance

between the two pathways is compromised (Albin et al., 1989; Benecke et al., 1987; Calabresi et al., 2014; DeLong, 1990; Graybiel and Rauch, 2000; Mink, 1996; Phillips et al., 1995).

Results

Opponent SNr activities underlie action selection

To address the role of basal ganglia in action selection, we trained mice in a recently developed 2-8 s task in which they are required to choose the left versus right action based on self-monitored time intervals (Howard et al., 2017). Specifically, mice were put into an operant chamber with both left and right levers extended (Figure 1A, see Methods). For a given trial, both levers retract at trial initiation, and after either 2 s or 8 s (50% for each, randomly interleaved), both levers extend. The mouse has to judge the interval between lever retraction and extension as 2 s vs. 8 s and make a corresponding action choice by pressing the left vs. right lever, respectively (Figure 1A). The first lever press after lever extension was registered as the mouse's choice. The correct choice leads to sucrose delivery (10 µl) as reward, and any lever presses beyond the first press after lever extension yield no outcome. The animal only has one chance to select the correct choice and gets rewarded in a given trial. If the animal's very first press after levers extension is the wrong choice, then there's no reward, and the chance to get rewarded in this particular trial vanishes, or the trial is functionally "terminated" although both levers still available to press. The animal has no second chance to correct its wrong choice by pressing the correct lever after the wrong choice. During the 2s vs. 8s waiting period with lever retraction, the levers are not physically accessible to the animal. Even the animal is trying to approach to the lever during lever retraction, but no lever press will be generated (see Supplemental Video 1). A new trial starts at lever retraction again after a random inter-trial-

interval (ITI, 30 s on average; Figure 1A). Across 14 consecutive days of training, mice (n = 10) significantly increased the correct rate of choice from chance level to more than 90% (Figure 1B). In addition, the animals gradually shortened the choice latency and demonstrated a strong preference toward the left lever due to its association with the shorter waiting time (Figure S1A, B). As a result, during the longer-waiting 8 s trials the mouse initially moved toward the left lever, then crossing the midpoint between left and right levers at around 4 s, and stayed around the right lever afterwards (Howard et al., 2017) (Figure 1C; Supplemental Video 1). Note that, the mouse showed no stereotyped movement trajectories during the incorrect trials (Figure 1C). This emerged stereotyped movement trajectory in the 8-s trials thus provided us a unique opportunity for investigating the neural mechanisms underlying the internally-driven, dynamic action selection process.

The substantia nigra pars reticulata (SNr) is one of the major output nuclei of basal ganglia (Albin et al., 1989; DeLong, 1990; Hikosaka et al., 2000; Mink, 1996). To investigate how the basal ganglia contribute to the dynamic process of action selection, we began by recording the SNr neuronal activity in mice trained in the 2-8 s task (Figure 1D, Figure S1C, see Methods). It was found that a large proportion (211/261, 80.8%; recorded from n = 9 mice) of SNr neurons changed firing rate significantly during the correct 8-s trials as mice dynamically shifted the internal action selection from the left to the right (Figure 1E). The Z-score of the task-related neuronal firing rate, reflecting the firing activity changes related to baseline, was defined as Firing Rate Index (FRI, see Methods). We focus on the data analyses in the 8-s trials since the first 2-s of 8-s trials consists of the identical behavioral and neuronal profiles of the 2-s trials due to the task design (Figure 1C, Figure S1D-F). The task-related SNr neurons were categorized into four subtypes based on the dynamics of FRI in the correct 8-s trials: Type 1 - monotonic

decrease (Figure 1E, F, 102/211, 48.3%), Type 2 - monotonic increase (Figure 1E, G, 56/211, 26.5%), Type 3 - transient phasic increase (Figure 1E, H, 25/211, 11.9%) and Type 4 - transient phasic decrease (Figure 1E, I, 28/211, 13.3%). These four types of neuronal dynamics in SNr only appeared in the correct but not the incorrect trials (Figure 1F-I), nor on the day 1 of task training (Figure S1G-K), suggesting a tight correlation between the SNr neuronal dynamics and the behavioral performance. Here we show trial-by-trial firing activities of SNr example neurons in correct 8s trials from well-trained animals as follows. Although the time of initial approach to the left side varies across trials, trial-by-trial analysis showed that the firing activities are consistent across trials and the averaged activities faithfully reflect the dynamics of each trial, evident for all four types of neurons (Figure S2A-D). Specifically, the Type 1 and Type 2, but not the Type 3 and Type 4 neurons, exhibit firing changes co-varying with the action selection and these two types together consist in around 80% of all task-related SNr neuron population (Figure 1J, Figure S2A-D). There is no dramatic difference in dynamic subtypes and proportion between SNr neurons recorded in left and right hemispheres (Figure S3). Notably, for Type 1 neurons, the firing activities is much higher as animals selected left side at the correct 8s trials than the firing activities when animals selected left side at the incorrect 8s trials (Figure 1F, green squares). The same for Type 2 neurons, their firing activities are dramatically different when animals selected left side in the correct and incorrect trials (Figure 1G, green squares). Therefore, Type 1 and Type 2 dynamics cannot be simply explained by sensory or position-related neural activity. Furthermore, we compare the SNr neuron responses at rewarded and non-rewarded lever presses. The SNr neuron activities are aligned to lever press at 0 as shown below. For Type 1 SNr neurons, the firing activity at the rewarded left lever presses (defined as the left lever press in correct 2s trials) is much higher than the firing activity at the non-rewarded left

lever presses (defined as left lever presses in incorrect 8s trials and random left lever press during the inter-trial-interval). The firing activity difference can also be observed between the rewarded and non-rewarded right lever presses in Type 1 SNr neuron (Figure S1L). For Type 2 SNr neurons, although there's no difference between the rewarded and non-rewarded left lever presses, the firing activity at the rewarded right lever presses is higher than the firing activity at the non-rewarded right lever presses (Figure S1L). Again, given the same sensory inputs and spatial location for both rewarded and non-rewarded left presses, the difference between rewarded and non-rewarded lever presses indicates that the neural dynamics are action selection dependent, and not simply related to sensory or position information.

It has been suggested that SNr suppresses movements through the inhibition of downstream motor nuclei and releases action via disinhibition (Hikosaka et al., 2000; Mink, 1996). We thus ask whether the opponent neuronal dynamics in Type 1 and Type 2 SNr subpopulations mediate the dynamic shift of choice, by suppressing the competing selection of right vs. left action, respectively. Indeed, the SNr net output by subtracting Type 2 and Type 1 SNr neuronal dynamics (Figure 1K) is highly reminiscent of the animal's stereotyped movement trajectory during choice (Figure 1C). To further determine the relationship between the SNr net output and action selection, we tested the behavioral choice of the 2-8 s trained mice in a series of non-rewarded probe trials with novel intervals of 2.5, 3.2, 4, 5 and 6.3 s (see Methods). Consistent with what reported before (Howard et al., 2017), the probability of mice selecting the action associated with the long duration (8 s) gradually increases along with the time intervals of probe trials (Figure 1L). The resulting psychometric curve thus represents the animal's real-time action selection process during the 8-s trials. Further comparison between the psychometric curve and the SNr net output revealed a strong linear correlation (Figure 1M), indicating that the

SNr net output faithfully predicts momentary behavioral choice. Together, these results suggest that mice can learn to dynamically shift their choice based on internally-monitored time, and the opponent neuronal activities in SNr correlate with the action selection.

SNr neuronal dynamics reflect action selection but not simply time or value

In the 2-8 s task, the passage of time and expectation of reward both change simultaneously with the animal's internal choice. One may argue that the Type 1 and Type 2 neuronal dynamics observed in SNr during the 8-s trials might reflect the passage of time or value of expected reward rather than action selection. To differentiate these possibilities and specify the functional role of SNr activity, we presented mice previously trained in the 2-8 s task with random probe trials of 16-s interval (Figure 2A). In these 16-s probe trials which they have never experienced before during training, the animals sometimes wait on the right side and press the right lever, or shift back to the left side and press the left lever when the levers are extended at 16 s (Figure 2B, C). This arbitrary choice situation in the 16-s probe trials thus provides a special window to determine the functional relationship between SNr activity and behavioral choice. If the Type 1 and Type 2 SNr subpopulations encode information about time passage or expectation value, their neuronal activities would continue changing monotonically between 8 and 16 s. In contrast, if the Type 1 and Type 2 SNr subpopulations encode action selection, their neuronal activities would predict the behavioral choice and differentiate between the right vs. left action selection. Indeed, it was found that when the firing activity of Type 1 SNr neurons maintained below baseline from 8 to 16 s, the mice tended to select the right lever later (Figure 2D). However, when the firing activity reversed the decreasing tendency to increase, the mice chose the left lever instead (Figure 2D). A similar relationship between the neuronal activity and

behavior choice was also evident in Type 2 SNr neurons, albeit with opposite dynamics (Figure 2E). This is especially evident in the subtraction of Type 2 and Type 1 SNr neuronal dynamics, in which the SNr net output is strongly correlated with and predictive of behavioral choice (Figure 2F). These results thus suggested that the neuronal activities in SNr likely encode the ongoing action selection but not simply reflect time passage or reward value.

To further confirm this point, we recorded the firing activity from the same SNr neurons during both the 2-8 s control task (Figure 2G, 2s-left and 8s-right) and a modified version of 2-8 s task in which the contingency between action and interval is reversed (Figure 2J, 2s-right and 8s-left) on the same day (see Methods). It was found that the mice performed at around 80% correct in both tasks on the same day (Figure 2H, K, Figure S4A). Accordingly, the movement trajectories of the same mice in 8-s trials were reversed from left-then-right in the control task (Figure 2I) to right-then-left in the reversed task (Figure 2L). The left-lever preference during the ITI in the control task was also switched to right-lever preference in the reversed 2-8 s task (Figure S4B). Notably, the passage of time and expected value as well as other environmental factors are all identical in both versions of task, except that the animal's choice is now reversed from right to left for the 8-s trials (Figure 2H, I vs. Figure 2K, L). If Type 1 or Type 2 SNr neurons encode time or value, either neuronal population will exhibit the same neuronal dynamics in 8-s trials for both versions of task. On the other hand, if Type 1 and Type 2 SNr neurons encode action selection, their neuronal dynamics will reverse in the reversed version of 2-8 s task compared to the standard version. In fact, the Type 1 SNr neurons which showed monotonic decreasing dynamics in the control 2-8 s task (Figure 2M) reversed their neuronal dynamics to a monotonic increase in the reversed 2-8 s task (Figure 2P), consistent with the behavioral choice. The same reversal of neuronal dynamics was also observed in Type 2 SNr

neurons in the reversed version of standard task (Figure 2N, Q). The SNr net output by subtracting Type 2 and Type 1 SNr neuronal dynamics, which was tightly correlated with the action selection in the standard 2-8 s task (Figure 2O), is reversed and now predictive of the new behavioral choice in the reversed 2-8 s task (Figure 2R). Notably, Type 3 and Type 4 SNr neurons exhibiting transient change when mice switching between choices maintained the same neuronal dynamics in both tasks (Figure S4C-F). Together, these results therefore demonstrate that the output of basal ganglia reflects the dynamic action selection rather than simply time or value.

Distinct striatal direct vs. indirect pathway activity during action selection

The basal ganglia output is largely controlled by two major neural pathways, called 'direct' and 'indirect' pathway, originating from D1- vs. D2-expressing spiny projection neurons (D1- vs. D2-SPNs) in the striatum, respectively (Albin et al., 1989; DeLong, 1990; Hikosaka et al., 2000; Mink, 1996). We then decided to determine the neuronal dynamics in the striatum, specifically the neuronal activity in the direct and indirect pathways during action selection. We employed in vivo extracellular electrophysiology to record the neuronal activity in the dorsal striatum when mice perform the 2-8 s task, and classified putative SPNs based on the spike waveforms and firing properties (Geddes et al., 2018; Jin and Costa, 2010; Jin et al., 2014). Among all the SPNs recorded from the trained mice (n = 19), 341 out of 409 SPNs (83.4%) were defined as task-related neurons for showing significant firing changes during the 2-s and 8-s lever retraction period (Figure 3A, Figure S2E-H, Figure S5A). Similar to the various types of neuronal dynamics observed in SNr, task-related SPNs showed Type 1 (Figure 3B, monotonic decrease, 159/341, 46.6%), Type 2 (Figure 3C, monotonic increase, 103/341, 30.2%), Type 3

(Figure 3D, transient phasic increase, 49/341, 14.4%) and Type 4 (Figure 3E, transient phasic decrease, 30/341, 8.8%) activity profiles during the correct 8-s trials (Figure 3A, Figure S2E-H, Figure S5A). These neural dynamics were largely absent in SPNs on day 1 of training (Figure S5B-F). Also, SPNs recorded from left and right hemispheres showed similar proportions (Figure S6). These results indicate that the striatum, as one of the major input nuclei of basal ganglia, demonstrates the four types of neuronal dynamics similar with SNr during the dynamic process of action selection.

To further determine the neuronal activity in the direct and indirect pathways during action selection, we utilized an optogenetics-aided photo-tagging method (Geddes et al., 2018; Howard et al., 2017; Jin and Costa, 2010; Jin et al., 2014; Lima et al., 2009) to record and identify striatal D1- vs. D2-SPNs in freely behaving mice. Channelrhodopsin-2 (ChR2) was selectively expressed in D1- or D2-SPNs by injecting AAV-FLEX-ChR2 in the dorsal striatum of D1- and A2a-Cre mice, respectively (Geddes et al., 2018; Jin et al., 2014). In the end of each behavioral session with recording, optogenetic stimulation via an optic fiber attached to the electrode array was delivered to identify D1- vs. D2-SPNs through photo-tagging (Figure 3F, Figure S5G-J) (Geddes et al., 2018; Jin et al., 2014). Only those neurons exhibiting a very short latency ($\leq$ 6 ms) to light stimulation (Figure 3G-I) and showing identical spike waveforms ($R \geq 0.95$, Pearson correlation coefficient) between behavior and light-evoked response (Figure 3J, K) were identified as Cre-positive thus D1- or D2-SPNs (Geddes et al., 2018; Jin et al., 2014). Within all positively identified D1-SPNs (n = 92 from 6 mice) and D2-SPNs (n = 95 from 6 mice), 74 out of 92 (80.4%) D1-SPNs and 79 out of 95 (83.1%) D2-SPNs showed a significant change in firing rate during the correct 8-s trials. In addition, all four types of neuronal dynamics during action selection were found in both D1-SPNs (Figure 3L, M) and D2-SPNs (Figure 3N,

O), as observed in SNr. The Type 1 and Type 2 neuronal dynamics showing monotonic firing change (Figure 3M, O) were the predominant task-related subpopulations within either D1- (Figure 3L) or D2-SPNs (Figure 3N). Notably, the striatal D1-SPNs consist of significantly more Type 1 than Type 2 neurons (Figure 3L), while D2-SPNs show a similar proportion between the two Types (Figure 3N). These data thus suggest while neurons in both the striatal direct and indirect pathways encode information related to behavioral choice, the two pathways might reflect and contribute to distinct aspects of action selection.

Ablation of striatal direct vs. indirect pathway differently impaired action selection

Given the action-selection-related neuronal dynamics observed in striatum, we next asked whether the neural activity in striatum is necessary for learning and execution of action selection, and furthermore, what is the functional difference between the direct and indirect pathways. It has been reported that the NMDA receptors on striatal SPNs are critical for sequence learning (Geddes et al., 2018; Jin and Costa, 2010) and action selection (Howard et al., 2017). To further identify the functional role of NMDA receptors on D1- vs. D2-SPNs for action selection, we employed a genetic strategy to specifically delete NMDA receptors from D1- vs. D2-SPNs by crossing mice carrying a floxed NMDAR1 (NR1) allele with a dorsal striatum-dominant D1-cre line (Gong et al., 2007) and A2a-cre line (Geddes et al., 2018; Jin et al., 2014), respectively (referred to as D1-NR1 KO and D2-NR1 KO mice, respectively; see Methods). Both the D1-NR1 KO and D2-NR1 KO mice are significantly impaired in learning the 2-8 s task compared to their littermate controls (Figure 4A, B), suggesting that NMDA receptors on either D1- or D2-SPNs are critical for learning of proper action selection. In the end of two-week training, when given the probe trials with various intervals across 2 to 8 s, it was found that D1-NR1 KO mice

showed a systematic bias toward the lever associated with short interval and made deficient behavioral choice only in long interval trials (Figure 4C). In contrast, D2-NR1 KO mice showed impaired action selection across various probe trials of both short and long intervals (Figure 4D). These data suggest that while NMDA receptors on both D1- and D2-SPNs are required for action learning, the deletion of NMDA receptors in direct and indirect pathways impairs action selection in a different manner.

We then asked whether that neural activity in dorsal striatum is necessary for the proper execution of action selection after learning. We first conduct striatal inactivation in trained wildtype mice by bilateral intra-striatal infusion of muscimol (Figure 4E, see Methods). Striatal muscimol infusion significantly reduced the animal's overall performance in comparison with the pre- and post-saline injection controls (Figure 4F). When tested with probe trials, the psychometric curve indicated that the inactivation of striatum impairs action selection for the long trials (Figure 4G). These data thus suggested that the striatal neural activity is critical for appropriate execution of learned action selection.

To further elucidate the functional role of specific striatal pathways in action selection, we next employed a viral approach to bilaterally express diphtheria toxin receptors (AAV-FLEX-DTR-eGFP) in the dorsal striatum of trained D1- and A2a-Cre mice, followed by diphtheria toxin (DT) injections to selectively ablate D1- or D2-SPNs (Geddes et al., 2018) (Figure 4H, I, N; see Methods). Ablation of either D1- or D2-SPNs significantly impaired action selection and reduced the correct rate of choice (Figure 4J, O). Notably, the psychometric curve revealed that D1-SPNs ablation mice showed a selective impairment of choice in long interval trials (Figure 4K). In contrast, mice with D2-SPNs ablation exhibited choice deficits in both long

and short trials (Figure 4P). Consistent with the D1- and A2a-NR1 KO data, these results suggest that the direct and indirect pathways are both needed yet play distinct roles in action selection.

The classic 'Go/No-go' model of basal ganglia suggests the direct and indirect pathways work antagonistically to release and inhibit action, respectively (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010). On the other hand, more recent 'Co-activation' model of basal ganglia proposes that direct pathway initiates the selected action and at the same time, the indirect pathway inhibits the competing actions (Cui et al., 2013; Hikosaka et al., 2000; Mink, 1996). For visualization purpose, we diagram 'Go/No-go' and 'Co-activation' models as center-surround receptive field with D1-SPNs as the center and D2-SPNs as the surround (Figure 4S; Figure S7A, D). The "center-surround" layout is derived from the receptive field of neurons in the early visual system, as an intuitive analogy in describing the functional interaction among striatal pathways (Mink, 2003). The area of each region does not represent the amount of cells but mainly qualitative functional role (Figure 4S). While the direct pathway plays the similar role in both models (Figure S7B, E), the function of indirect pathway differs dramatically (Figure 4S). Lesion of the indirect pathway thus leads to contrast predictions on action selection from the two models (Figure S7C, F). Specifically, ablation of D2-SPNs would facilitate the action being selected through removing inhibition according to the Go/No-go model (Figure S7C) (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010), while blockage of indirect pathway would impair the action selection due to disinhibition of competing actions according to the Co-activation model (Figure S7F) (Cui et al., 2013; Hikosaka et al., 2000; Mink, 1996). Although our D1-SPNs ablation experiment indicates that direct pathway is required for action selection as suggested in both models (Figure 4J, K), the D2-SPNs ablation result favorably supports the Co-activation model over the Go/No-go model (Figure 4O, P, Figure S7F). In fact, close inspection

of the movement trajectories of D1-SPNs lesioned mice in the 8-s trials showed that compared to control mice (Figure 4L), they tend to stick on the left side more often with impaired right choice when lever extension at 8s (Figure 4M). In contrast, D2-SPNs lesioned mice demonstrated overall rather random movement trajectories, and the stereotyped left-then-right movement sequences were largely disrupted in comparison with the controls (Figure 4Q, R). These observations are mostly consistent with the idea of indirect pathway inhibiting competing actions in the Co-activation model (Figure 4S) and lesion of indirect pathway disrupts action selection for both the short and long trials (Figure 4P-R). Together, these data suggest that ablation of direct and indirect pathways both impair choice behavior but in a distinct manner due to their different roles in action selection.

Optogenetic manipulation of D1- vs. D2-SPNs distinctly regulates action selection

To further determine the specific function of direct vs. indirect pathway in action selection, we employed optogenetics to alter the D1- and D2-SPNs activity in vivo with high temporal precision and investigated its effects on the ongoing action selection process. Both the classic 'Go/No-go' (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010) and more recent 'Co-activation' (Cui et al., 2013; Hikosaka et al., 2000; Mink, 1996) models predict that activation of the direct pathway enhances the action selection (Figure S8A, E, I, K, O, Q), while inhibition of direct pathway reduces the correct choice (Figure S8B, F, L, R). To experimentally validate the models' predictions, AAV-FLEX-ChR2 was injected into the dorsal striatum of D1- or A2a-Cre mice and optic fibers were implanted bilaterally for in vivo optogenetic stimulation (Figure 5A, Figure S5K, L; see Methods) (Geddes et al., 2018; Jin et al., 2014). After mice learned the 2-8s task, 1-s pulse of constant light (wave length 473 nm) was delivered right before lever extension

in randomly chosen 50% of 2-s and 50% of 8-s trials (Figure 5B, C, see Methods). The correct rate of optogenetic stimulation trials is compared with the non-stimulation trials of the same animal as a within-subject design. We observed no significant change on the correct rate in 2-s trials, whereas the correct rate was significantly increased by optogenetic stimulation in 8-s trials (Figure 5D), indicating a facilitation effect on action selection by stimulating the D1-SPNs. We then sought to determine the effect of suppressing D1-SPN activity on action selection by viral expression of Halorhodopsin (AAV5-EF1a-DIO-eNpHR3.0-eYFP) in the dorsal striatum of D1-cre mice (Gradinaru et al., 2010). As expected, inhibiting D1-SPNs right before lever extension in trained mice reduced the correct rates in 8-s but not 2-s trials (Figure 5C, E), opposite to D1-SPN stimulation effects. These experimental data with bidirectional optogenetic manipulation suggest that the D1-SPN activity is positively correlated with the choice performance, consistent with the hypothesis of direct pathway facilitating the action selected in both the Go/No-go and Co-activation models (Figure S8K, L, Q, R).

Nevertheless, the two models have distinct views on the function of indirect pathway. While the classic 'Go/No-go' model suggests that the indirect pathway inhibits the selected action (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010), the 'Co-activation' model hypothesizes that the indirect pathway inhibits the competing actions instead (Cui et al., 2013; Hikosaka et al., 2000; Mink, 1996). These models thus provide contrasting predictions about the effect of activation of the indirect pathway on action selection, being decreased correct rate based on the Go/No-go model (Figure S8M) and increased correct rate from the Co-activation model (Figure S8S), respectively. We thus decided to test the distinct predictions from the two models by optogenetic manipulation of indirect pathway during action selection in the 2-8s task. ChR2 or Halorhodopsin (eNpHR3.0) was expressed in the dorsal striatum of A2a-cre mice for

bilaterally optogenetic activation or inhibition during behavior (Figure 5A, Figure S5L; see Methods). Notably, optogenetic excitation of D2-SPNs for 1s right before lever extension decreased the correct rate in both 2-s and 8-s trials (Figure 5F). In contrast, transient optogenetic inhibition of D2-SPNs before behavioral choice increased correct rates for both 2-s and 8-s trials (Figure 5G). These data suggest that opposite to the D1-SPN manipulation, optogenetic stimulation of D2-SPNs impairs action selection, while inhibition of D2-SPNs facilitates behavioral choice. These optogenetic results further unveil the distinct roles of direct vs. indirect pathway in action selection, and are in line with the predictions from the Go/No-go (Figure 5H, Figure S8M, N) but not the Co-activation model (Figure S8S, T).

A 'Triple-control' model of basal ganglia circuit for action selection

Our DT lesion experiments found that ablation of indirect pathway impairs action selection (Figure 4O, P), as predicted from the Co-activation but not Go/No-go model (Figure 4S), while the optogenetic results suggested that inhibition of D2-SPNs enhances behavioral choice (Figure 5G), a result in favor of the Go/No-go rather than Co-activation model (Figure 5H). We wonder whether these seemly discrepant effects are attributed to a more complex circuit mechanism involving in the indirect pathway different from either the Go/No-go or Co-activation model. To systematically investigate the cell type- and pathway-specific mechanisms underlying action selection, we firstly add Go/No-go and Co-activation models together to examine the whether the resulted combination model could explain the experimental observations (Figure S7G). The lesion of D1-SPNs in the combination model indeed selectively impaired choice in long interval trials (Figure S7H). However, the effect of D2-SPNs ablation in the combination model was neutralized due to the opposing contributions from Go/No-go and

Co-activation models respectively (Figure S7I). Based on these simulation results, none of the Go/No-go, Co-activation and combination models was able to fully capture the underlying mechanism of basal ganglia in action selection. Inspired by the data in current experiments, we decided to build a new computational model of the cortico-basal ganglia circuitry based on the realistic neuroanatomy (Aoki et al., 2019; Mailly et al., 2003a; Schmidt and Berke, 2017; Taverna et al., 2008) and empirical neuronal physiology during action selection (Figures 1-3).

Different from the dual control of action by direct vs. indirect pathway in either the Go/No-go or Co-activation model (Figure S7), our new model adds an additional layer of control derived from the indirect pathway, thus called 'Triple-control' model for action selection. The combination of Go/No-go or Co-activation models clearly failed to explain all the experimental results (Figure S7G-I), therefore in our model, the new layer of control is not a simple add-on but equipped with interaction with other layers. Specifically, the new model consists of one direct pathway and two indirect pathways defined as D2-SPN #1 and D2-SPN #2 two subpopulations, corresponding to the Co-activation and Go/No-go functional modules, respectively (Figure 6A, B). In addition, the indirect pathway D2-SPNs in the Co-activation module inhibits the indirect pathway D2-SPNs in the Go/No-go module through the well-known D2-SPN collaterals with the properties of short-term depression in the striatum (Gustafson et al., 2006; Schmidt and Berke, 2017; Taverna et al., 2008; Tecuapetla et al., 2007) (Figure 6A; see Methods), providing asymmetric modulation to D2-SPN subgroups and promoting Co-activation module as the dominant functional module at rest. In this 'Triple-control' basal ganglia model, striatal D1- and D2-SPNs associated with left and right actions receive excitatory inputs from corresponding cortical inputs (Figure 6A) to generate Type 1 and Type 2 neuronal dynamics (Figure S9A-D) (Lo and Wang, 2006). The D1- and D2-SPNs then regulate the SNr neuronal dynamics through

the direct and indirect pathways, respectively (Figure S9) (Albin et al., 1989; DeLong, 1990; Hikosaka et al., 2000; Mink, 1996). The net SNr output (Figure S9F, I), which controls the downstream brainstem and thalamic circuits necessary for action selection (Hikosaka, 2007; Lo and Wang, 2006; Redgrave et al., 1999), will determine the final behavioral choice (Figure S9G, J). The choice preference towards left lever over right lever was reflected in the direct pathway by the unevenly weighted connection strength from cortex to D1-SPN Left/Right, as well as the connection strength from D1-SPN Left/Right to SNr Left/Right neurons (see Methods). Our computational simulations showed that this 'Triple-control' network model could faithfully recapitulate the neuronal activity across the basal ganglia circuitry and predict the behavioral choice (Figure S9).

To dissect the functional role of direct vs. indirect pathway in action selection, we simulate the cell ablation experiments and examine the behavioral output in the 'Triple-control' basal ganglia model. Ablation of D1-SPNs in the network model (Figure S9E) modulates both Type 1 and Type 2 SNr dynamics but in different magnitude due to the biased striatal input to SNr left output and mutual inhibition between SNr left vs. right outputs (Figure S9F; see Methods). As a result, the lesion causes a downward shift in the net SNr output, especially evident at the late section of 8 s (Figure S9G). This change in net SNr output predicts a behavioral bias towards left choice as seen in the psychometric curve (Figure 6C), consistent with experimental results in mice with D1-SPNs ablation (Figure 4K). In contrast, ablation of D2-SPNs in the network model (Figure S9H), by removing the indirect pathways of both the Go/No-go and Co-activation modules, alters Type 1 and Type 2 SNr dynamics (Figure S9I) and change the net SNr output dramatically around 2s as well as 8s (Figure S9J). The model thus predicts behavioral choice deficits for both short and long trials during D2-SPNs ablation (Figure

6D), consistent with experimental observations (Figure 4P). Together, these data suggest that our new 'Triple-control' basal ganglia model, based on realistic neuroanatomy and empirical neuronal physiology, can perform action selection similar to the behavior of mice, and successfully replicate the pathway-specific lesion effects on choice.

We further simulate the neuronal and behavioral effects of optogenetic manipulation of D1- and D2-SPNs in the cortico-basal ganglia model. Consistent with the experimental results (Figure 5D, F), optogenetic stimulation of D1-SPNs facilitates the ongoing choice (Figure 6E), while optogenetic inhibition of D1-SPNs suppresses ongoing choice in the model (Figure 6G). In addition, optogenetic stimulation of D2-SPNs impairs the ongoing choice and causes switching (Figure 6F), while optogenetic inhibition of D2-SPNs facilitates ongoing choice, due to the now dominant Go/No-go module mediated by the short-term depression of D2 collaterals in the model (Figure 6H). Consistent with the experimental observations, the optogenetic inhibition effect is opposite from the D2-SPNs cell ablation in the model (Figure 6D).

We next investigate how the striatum influence SNr outputs. Since the collateral projection with STD in D2-SPNs is the key in our 'Triple-control' model to switch between Go/No-go and Co-activation modules, we first built a motif of indirect pathway with two D2-SPNs subgroups defined as D2-SPN #1 and D2-SPN #2 (Figure S10A). We tested this indirect pathway motif with monotonic neural dynamics observed in experiments meanwhile simulating the optogenetic activation at 1s and 7s (Figure S10D-I). The SNr therefore received more activation at 1s than at 7s (Figure S10J, K), suggesting that the D2-SPNs with short-term depression in collateral inhibition modulates SNr activities in a firing rate-dependent manner.

We next sought to test the model's predictions and experimentally investigate the distinctions in modulating SNr activities between the direct and indirect pathways during action

selection. In order to manipulate D1- or D2-SPNs and monitor SNr responses at the same time, we simultaneously implanted optogenetic fibers and recording array into striatum and SNr respectively on a single mouse (Figure S10L). While the mice performing the 2-8s task, optogenetic stimulation was delivered to activate either D1- or D2-SPNs. It was found that optogenetic activation of D1- or D2-SPNs caused both inhibition and excitation in Type 1 and Type 2 SNr neurons (Figure S10M-O). To further compare SNr activities responding to striatal activation at different time points during the lever retraction period, for a given trial, we activated D1-SPNs (or D2-SPNs) either at 1s or 7s after the lever retraction (Figure S10P-R). For direct pathway, the change of FRI in SNr activities caused by activation of D1-SPNs showed no significant difference between 1s and 7s (Figure S10P, S). For indirect pathway, activating D2-SPNs at 1s caused smaller activation of FRI than at 7s in Type 1 SNr neurons (Figure S10Q), whereas for Type 2 SNr neurons, activating D2-SPNs at 1s induced bigger FRI increase at 1s than at 7s (Figure S10R). Overall, activating D2-SPNs tended to bias the firing rate downward at 1s but upward at 7s in Type 1 SNr neurons, which was counteractive to the decreasing tendency of Type 1 SNr neuron (Figure S10T). In contrast, Type 2 SNr neurons showed higher FRI increase and smaller decrease in response to activating D2-SPNs at 1s than at 7s, which was opposing to the increasing dynamics of Type 2 SNr neurons (Figure S10T). This firing rate-dependent modulation on SNr activities through indirect pathway is consistent with the computational simulation (Figure S10J, K; Figure S11). Therefore, the underlying D2-SPNs collaterals might indeed be a key mechanism contributing to the modulation of SNr activity and action selection in vivo, as simulated in the ‘Triple-control’ model.

Taken together, our new ‘Triple-control’ basal ganglia model, based on realistic neuroanatomy and empirical neurophysiology, successfully reproduces both the lesion and

optogenetic data we collected during the animal experiments. It could thus potentially provide essential insights into the circuit mechanism of basal ganglia underlying action selection.

Linear and nonlinear control of action selection by direct vs. indirect pathway

To gain an overall picture of how basal ganglia control action selection, we run through the model with a wide continuous range of manipulation to mimic the effects from lesion to optogenetic inhibition and optogenetic activation (Figure 7A, B, E, F). The simulations of cell ablation and bidirectional optogenetic manipulations of D1-SPNs activity in the model reveal no significant effects at 2-s trials (Figure 7C), but a linear relationship between the neuronal activity in direct pathway and the behavioral performance of choice in 8 s trials (Figure 7D), as observed in animal experiments. It thus further confirms that direct pathway selects action and facilitates ongoing choice, consistent with the predictions from both the classic Go/No-go and recent Co-activation models (Figure 7I, Figure S12A-D) (Albin et al., 1989; DeLong, 1990; Hikosaka et al., 2000; Mink, 1996).

In contrast, manipulations of D2-SPNs activity from cell ablation to optogenetic inhibition and then optogenetic stimulation in the model demonstrate an inverted-U-shaped nonlinear relationship between the neuronal activity in indirect pathway and action selection, for both 2-s and 8-s trials (Figure 7G, H). Detailed analyses reveal that D2-SPNs ablation removes both the Co-activation and Go/No-go module in the indirect pathway and leaves SNr activity dictated by D1-SPN inputs. However, due to the inhibition from Co-activation to Go/No-go module in the indirect pathway via D2-SPN collaterals and short-term plasticity of these synapses (Figure S10A-C; see Methods), optogenetic manipulation of D2-SPNs differentially affects the D2-SPN subpopulations groups and promotes Go/No-go module to dominate the

basal ganglia network (Figure S10C-S). This dynamic switch of dominance between Co-activation and Go/No-go modules on the basal ganglia network gives rise to a nonlinear relationship between D2-SPNs manipulation and the behavioral outcome (Figure 7J).

Note that when the same inputs were applied to the Go/No-go or Co-activation model alone, the behavioral performance in either model exhibits linear negative (Figure S12E) or positive correlation (Figure S12F) with D2-SPNs activity, respectively. Both our experimental and modeling results thus indicate that different from either the Go/No-go or Co-activation model, the indirect pathway regulates action selection in a nonlinear manner, depending on the state of the network and D2-SPNs activity level. Besides collaterals within D2-SPNs, other collateral connections, for example connections between D1-SPNs or connections between D1- and D2-SPNs, could also contribute to the regulation of action selection (Taverna et al., 2008). We tested our 'Triple-control' model with adding additional collateral connections as D1→D1 (Figure S13A-C), D1→D2 (Figure S13D-F) and D2→D1 (Figure S13G-I), respectively. It was found while these additional collaterals further quantitatively regulate action selection, the general principle of linear vs. nonlinear modulation of action selection by direct and indirect pathways still qualitatively hold (Figure S13). Interestingly, our current 'Triple-control' model can also replicate the behavioral effects of optogenetic manipulation of nigrostriatal dopamine on behavioral choice (Howard et al., 2017), and further unveils an inverted-U-shaped relationship between striatal dopamine concentration change and action selection (Figure S14). Together, these results suggest that there are multiple levels of interactions from D1- and D2-SPNs to dynamically control SNr output, and the basal ganglia direct and indirect pathways distinctly control action selection in a linear and nonlinear manner, respectively.

Discussion

Here, by using an internally-driven 2-8s action selection task in mice, we investigated the function of basal ganglia direct and indirect pathways in mediating dynamic action selection. We found that the neuronal activities in SNr, the major output of basal ganglia, directly reflect animals' internal action selection process, other than simply time or value. It was also observed that the striatum, the main input of basal ganglia, shares the similar action selection-related neuronal dynamics with SNr and is needed for both learning and execution of proper action selection. Furthermore, the striatal direct and indirect pathways exhibit distinct neuronal activity and during manipulation, they have different functional effects on controlling action selection. Notably, the experimental observations on the physiology and function of direct and indirect pathways cannot be simply explained by either the traditional 'Go/No-go' model or the more recent 'Co-activation' model. We proposed a new 'Triple-control' functional model of basal ganglia, suggesting a critical role of dynamic interactions between different neuronal subpopulations within the indirect pathway for controlling basal ganglia output and behavior. In the model, a 'center (direct pathway) – surround (indirect pathway) – context (indirect pathway)' three layers of structure exerts dynamic control of action selection, depending on the input level and network state. This new model respects the realistic neuroanatomy, and can recapitulate and explain the essential in vivo electrophysiological and behavioral findings. It also provides a new perspective on understanding many behavioral phenomena involving in dopamine and basal ganglia circuitry in health and disease.

Our current 2-8s action selection task offers a unique opportunity to observe the animal's internal switch from one choice to another and monitor the underlying neuronal dynamics correspondingly. We observed two major types of monotonically-changing SNr neuronal dynamics during the internal choice switching, presumably one type associated with selecting

one action and another with selecting the competing action, respectively. The classic view on SNr activity is that it tonically inhibits the downstream motor nuclei and releases action via disinhibition (Albin et al., 1989; DeLong, 1990; Hikosaka and Wurtz, 1983; Wurtz and Hikosaka, 1986). The increased response in SNr, however, could potentially inhibit the competing actions or the movements toward the opposing direction through projections to the contralateral brain regions like superior colliculus (Jiang et al., 2003). Here, we found that two subpopulations in SNr showed opposite monotonic firing change during the left-then-right choice, and notably, their neuronal dynamics switched when the animals performed the reversed version of task which requires a right-then-left choice. It thus suggests that these SNr neurons are indeed associated with different action options during choice behavior, and actively adjust their firing rates to facilitate respective action selection. Given the opposite neuronal dynamics and functionally antagonistic nature of Type 1 and Type 2 SNr neurons, we defined the net output of basal ganglia by the subtracting the neuronal activity between the two SNr subpopulations and correlated it with the behavioral choice. The subtraction between Type 1 and Type 2 SNr neurons is the net output of two competing choices and indicates animals' choice in real time. Also, signals corresponding to left and right choices through direct/indirect pathways eventually converge to SNr (Albin et al., 1989; DeLong, 1990). The collateral inhibition within SNr (Brown et al., 2014; Mailly et al., 2003b) gives rise to the direct competition between different SNr functional subgroups. Therefore, the subtraction between Type 1 and Type 2 SNr neurons represents the outcome of competition between choices. Indeed, we found that the basal ganglia net output exhibited a tight correlation with the psychometric curve of behavioral choice, and faithfully represented a neural basis for the dynamic action selection process.

As one of the major input nuclei of basal ganglia, striatum influences SNr activity through direct/indirect pathways and undisputedly, plays an essential role in action selection (Ding and Gold, 2012; Geddes et al., 2018; Jin et al., 2014; Lauwereyns et al., 2002; Tai et al., 2012). By genetic manipulation and pharmacological inactivation, we showed that striatum is indispensable for both learning action selection and the proper performance of learned behavioral choice. The recording of neuronal activity in dorsal-central striatum during action selection further revealed that striatal spiny projection neurons share the similar types of neuronal dynamics as SNr. Through optogenetic-tagging in freely behaving mice, we further found that dorsal-central striatal SPNs in the direct and indirect pathways show distinct activity profile, with D1-SPNs representing a strong bias towards the preferred choice, while D2-SPNs encoding two choices equally.

Two prevailing models have been proposed to explain the functional distinction between D1- and D2-SPNs. The canonical model of the basal ganglia suggests that the D1- and D2-SPNs play antagonistic roles in controlling action as mediating “Go” and “No-go” signals, respectively (Albin et al., 1989; DeLong, 1990; Kravitz et al., 2010). A more recent model, however, implies that as D1-SPNs initiate an action, D2-SPNs co-activate with D1-SPNs to inhibit other competing actions (Cui et al., 2013; Hikosaka et al., 2000; Isomura et al., 2013; Jin et al., 2014; Mink, 1996). Essentially, these two models agree upon the functional role of D1-SPNs in releasing or facilitating the desired action, but contradict on the function of D2-SPNs on which targeted action of inhibiting. Here, our in vivo recording data indicate that both D1- and D2-SPNs share similar neuronal dynamics during action selection, and the neural activity alone is not sufficient to separate and determine whether “Go/No-go” or “Co-activation” model is supported (Cui et al., 2013; Isomura et al., 2013; Jin et al., 2014). To resolve the functional

distinction of the direct vs. indirect pathway, we applied a series of cell-type-specific manipulations on striatal D1- and D2-SPNs during action selection behavior. First, we generated mutant mice in which NMDA receptors are deleted from either striatal D1- or D2-SPNs (Geddes et al., 2018; Jin et al., 2014). Both the D1-NR1 KO and D2-NR1 KO mice showed learning deficits and behavioral choice impairments when tested with probe trials, suggesting that both D1- and D2-SPNs are necessary for learning and performing action selection. Notably, while the D1-NR1 KO mice are mostly impaired in the choice associated with 8s, a less-preferred option compared to 2s, the D2-NR1 KO mice are compromised in both 2s and 8s choice. Additional experiments with cell-type specific ablation further confirmed these results, consistent the distinct neuronal activity profile in these two pathways revealed during in vivo neuronal recording. While both the "Go/No-go" and "Co-activation" models predict the suppression of D1-SPNs activity leads to impaired action selection, supported by current KO and cell-ablation data, the manipulation experiments on D2-SPNs favor the "Co-activation" but not the "Go/No-go" model which the latter suggests D2-SPNs ablation would improve rather than impair action selection.

Next, we directly introduced transient bidirectional manipulations to D1- and D2-SPNs activity by optogenetics while mice performing the task. Our findings revealed that activation or inhibition of D1-SPNs increased and decreased the correct rate of choice respectively, suggesting a facilitating role of direct pathway in action selection, which again fits well with the "Go/No-go" as well as the "Co-activation" model. In contrast, optogenetic activation of D2-SPNs decreased the correct rate of choice, while inhibition of D2-SPNs promoted the correct choice. When stimulating D2-SPNs, animals are still able to press the lever and make a selection shortly after lever extension, therefore, the behavioral effect triggered by D2 stimulation is not simply due to

a general effect of decreased locomotion, but the altered action selection process. These results were supportive to the “Go/No-go” model but contradicted to the prediction of “Co-activation” theory, which the latter predicts that activation of D2-SPNs inhibits competing actions to facilitate desired choice, whereas inhibition of D2-SPNs releases competing actions and compromises the ongoing choice.

In summary, neither “Go/No-go” nor “Co-activation” models could fully explain the experimental results we found, particularly for experiments on D2-SPNs in the indirect pathway. Through computer simulation, we further demonstrated that a simple additive combination of “Go/No-go” nor “Co-activation” models by linear addition cannot reproduce all the experimental observations. To resolve these theoretical difficulties, we proposed a new center-surround-context “Triple-control” model of basal ganglia pathways for action selection. Specifically, two subpopulations of D2-SPNs in the indirect pathway function as “Co-activation” and “Go/No-go” modules respectively, and an activity-dependent inhibition from “Co-activation” to “Go/No-go” module mediates the dynamic switch between the dominant module depending on the inputs and network state. Due to the dominant “Co-activation” module in the default state, excessive inhibition of D2-SPNs or ablating the entire indirect pathway eliminates the promotive contribution and impairs action selection in the “Triple-control” model, consistent with the experimental observations. In contrast, transient increase of D2-SPNs firing activity during optogenetic stimulation introduces shift toward the “Go/No-go” dominance from the “Co-activation” module via firing-rate-dependent short-term depression of the inhibitory synapses between them, which amplifies the “No-go” signal and compromises action selection as experimentally found. In contrast, transient decrease of D2-SPNs firing activity during optogenetic inhibition results in disinhibition of “Go/No-go” module from inhibitory control of

“Co-activation” module, with an attenuated “No-go” signal which leads to better performance in choice. These results from our new “Triple-control” model thus suggest that the basal ganglia circuitry could be much more dynamic than previously thought, and it could employ a complex mechanism of functional module reconfiguration for context- or state-dependent flexible control. More importantly, our model further proposed that while direct pathway regulates action selection in a linear manner, the indirect pathway modulates action selection in a nonlinear inverted-U-shaped way depending on the inputs and the network state (Figure 7). Indeed, the amplitude of activities of D2 pathway is pivotal to the behavioral outcome (Meng et al., 2018). In other words, in certain conditions, activation of D2 pathway will facilitate the action as “start&go”; while too much activation of D2-SPNs will switch D2 pathway to “start&stop” mode (Meng et al., 2018), which is consistent with our proposed “nonlinear” control of D2 pathway over action selection. These results of various functional assemblies defied previous basal ganglia models in which either direct or indirect pathway has been treated as one uniform population and assigned with a single function in controlling action.

In the “Triple-control” model, we posited the collateral connections among striatal D2-SPNs and its short-term plasticity could serve as an operational mechanism for the dominant module switching. However, besides these well-known striatal local connections as one of the simplest possible mechanisms, other anatomical circuits within basal ganglia circuitry could potentially fulfill this functional role alone or additionally as well. For example, striatal D2-SPNs project to external global pallidus (GPe) through striatopallidal pathway, and meanwhile they receive arkypallidal projections from GPe to both the striatal SPNs and interneurons (Abdi et al., 2015; Fujiyama et al., 2016; Mallet et al., 2016). It is thus also possible that the dynamic interaction between “Co-activation” and “Go/No-go” modules is mediated through di-synaptic or

tri-synaptic modulation with GPe and/or striatal interneurons involved. Furthermore, in theory this dynamic interaction between “Co-activation” and “Go/No-go” modules can also occur outside striatum in the downstream nuclei including GPe and SNr, given their specific neuronal subpopulations receiving inputs from corresponding striatal D2-SPNs subgroups and proper collateral connections within the nuclei (Atherton et al., 2013; Cazorla et al., 2014; Fujiyama et al., 2011; Lee et al., 2020; Wu et al., 2000). Considering the crucial role of dopamine in basal ganglia circuitry, the new “Triple-control” model can also reproduce our previous experiments results on the effect of nigrostriatal dopamine on action selection (Howard et al., 2017). Importantly, it unveils that there is an inverted-U-shaped relationship between dopamine concentration change and action selection (Cools and D'Esposito, 2011). The model simulation suggests while moderate dopamine increase improves decision making, too much dopamine changes, either increase or decrease, dramatically impairs the choice behavior. These results might be able to explain some of behavioral observations involving in obscure decision making under the influence of addictive substances.

Our findings also have important implications in many neurological and psychiatric diseases. It was known that the loss of dopamine leads to hyperactivity of D2-SPNs and disruption of local D2-SPNs collaterals in Parkinson’s disease (Taverna et al., 2008; Wei and Wang, 2016). These alterations will not only break the balance of direct vs. indirect pathway, but also disrupt the multiple dynamic controls from the indirect pathway. The action selection will thus be largely problematic, even with L-DOPA treatment, which might restore the dopamine partially but not necessarily the altered basal ganglia circuitry and its circuit dynamics (Bastide et al., 2015). The current “Triple-control” model also provides some mechanistic insights into the inhibitory control deficits observed Schizophrenia (Taverna et al., 2008; Wei and Wang, 2016).

For instance, an increase in the density and occupancy of the striatal D2 receptors (D2R) has been frequently reported in schizophrenia patients (Abi-Dargham et al., 2000; Howes and Kapur, 2009; Laruelle et al., 1997; Wong et al., 1986). Many antipsychotic medications primarily aim to block the D2R (la Fougere et al., 2005; Lally and MacCabe, 2015; Yokoi et al., 2002), but the drug dose is the key to the treatment and severer adverse effects are associated with overdose of D2R antagonism (Levine and Ruha, 2012). In addition, prolonged exposure to antipsychotics often causes extrapyramidal symptoms, including Parkinsonian symptoms and tardive dyskinesia (Jarskog et al., 2007; Seeman, 2002). The dose-dependent effects when modulating D2R were also found in cognitive functions such as serial discrimination, in which relatively low and high dose of D2R agonist in striatum impair the performance in the discrimination task, while the intermediate dose of D2R agonist produces significant improvement (Cools and D'Esposito, 2011; Goldman-Rakic et al., 2000; Horst et al., 2019; Mattay et al., 2003). These observations thus further underscore the dynamic interplays and complexity of basal ganglia pathways in action control, as demonstrated in current study and the new triple-control functional model.

Acknowledgments: The authors would like to thank Tom Jessell and members of Jin lab for discussion and comments on the manuscript. This research was supported by grants from the NIH (R01NS083815), the Dystonia Medical Research Foundation and the McKnight Memory and Cognitive Disorders Award to X.J.

Author Contributions: X.J. conceived the project. X.J. and H.L. designed the experiments. H.L. performed the experiments and analyzed the data. X.J. supervises all aspects of the work. H.L. and X.J. wrote the manuscript.

Conflicts of Interest: None of the authors declare any conflict of interest, financial or otherwise.

Figure Legends:

Figure 1. The neuronal dynamics in SNr during the 2-8 s action selection task. (A) Schematic diagram for the design of 2-8 s task. (B) Correct rate for wild type mice across 14 days training (n = 10 mice, one-way repeated-measures ANOVA, significant effect of training days, $F_{13,117} = 32.54$, $p < 0.0001$). (C) Movement trajectory of an example mouse in correct (left panel) and incorrect (right panel) 8-s trials (gray line: trajectory of each trials; red/black line: the average trajectory). (D) Diagram of electrode array implanted into substantia nigra pars reticulata (SNr). (E) Firing Rate Index (FRI) of neuronal activity for all task-related SNr neurons in correct 8-s trials. The magnitude of FRI is color coded and the SNr neurons are classified as four different types based on the activity dynamics. (F-I) Averaged FRI for Type 1 (F, green squares indicating activities related to left choice), Type 2 (G, green squares indicating activities related to left choice), Type 3 (H), Type 4 (I) of SNr neurons in correct (red) and incorrect 8-s trials (gray). (J) The proportion of four types of SNr neurons. Type 1 and Type 2 are major types and significantly more than Type 3 and Type 4 (Z-test, $p < 0.05$). (K) Integrated SNr output defined as the subtraction of averaged FRI between Type 1 and Type 2 SNr neurons. (L) Averaged psychometric curve (n = 10 mice) of choice behavior. (M) The correlation between the Type 1 and Type 2 FRI subtraction and the behavioral choice ($R = 0.98$, $p < 0.0005$). Error bars denote s.e.m., same for below unless stated otherwise.

Figure 2. SNr neuronal dynamics reflect action selection but not interval time or reward value. (A) Task diagram of 2-8 s control task with 10% 16-s probe trials. (B) Percentage of behavioral choice in 2-s, 8-s and 16-s trials (blue: left choice; red: right choice) (n = 9 mice, paired t-test, $p < 0.05$). (C) Movement trajectory of an example mouse in 16-s trials (blue: left

choice; red: right choice). (D) Averaged SNr Type 1 FRI in 16-s trials (red: left choice; black: right choice). Firing rates from 8s to 16s (highlighted area) are compared between left and right choice (n = 26 neurons, two-way repeated-measures ANOVA, significant difference between left and right choices, $F_{1,25} = 6.646$, $p = 0.016$). (E) Averaged SNr Type 2 FRI in 16-s trials (red: left choice; black: right choice). Firing rates from 8s to 16s are compared between left and right choice (n = 16 neurons, two-way repeated-measures ANOVA, significant difference between left and right choices, $F_{1,15} = 5.785$, $p = 0.029$). (F) Subtraction of FRI for SNr Type 1 and Type 2 neurons in 16-s probe trials (red: left choice; black: right choice). (G) Task design of 2-8 s standard task. (H) Percentage of behavioral choice in 2-s and 8-s trials (blue: left choice; red: right choice) (n = 6 mice, paired t-test, $p < 0.05$). (I) Movement trajectory of an example mouse in 8-s trials (blue: left choice; red: right choice). (J) Task design of reversed 2-8 s task. (K) Percentage of behavioral choice in 2-s and 8-s trials in the reversed 2-8 s task (blue: left choice; red: right choice) (n = 6 mice, paired t-test, $p < 0.05$). (L) Movement trajectory of the same mouse as (I) in 8-s trials in the reversed 2-8 s task (blue: left choice; red: right choice). (M) Averaged FRI of the SNr Type 1 neurons in correct 8-s trials (n = 14 neurons). (N) Averaged FRI of the SNr Type 2 neurons in correct 8-s trials (n = 11 neurons). (O) Integrated SNr output as the subtraction of FRI for SNr Type 1 (M) and Type 2 neurons (N) in the standard 2-8 s task. (P) Averaged FRI of the same neurons as (M) in correct 8-s trials of the reversed 2-8 s task. (Q) Averaged FRI of the same neurons as (N) in correct 8-s trials of the reversed 2-8 s task. (R) Integrated SNr output as the subtraction of FRI for SNr Type 1 (P) and Type 2 neurons (Q) in the reversed 2-8 s task.

Figure 3. Neuronal activity of striatal D1- and D2-SPNs during action selection. (A) FRI of neuronal activity for all task-related SPNs in correct 8-s trials. SPNs were classified as Type 1 - 4. (B-E) Averaged FRI for Type 1 (B), Type 2 (C), Type 3 (D), Type 4 (E) of SPNs in correct (red) and incorrect 8-s trials (gray). (F) Diagram of simultaneous neuronal recording and optogenetic identification of D1- vs. D2-SPNs in dorsal striatum. (G) Top panel: Raster plot for a representative D1-SPN response to 100 ms optogenetic stimulation. Each row represents one trial and each black dot represents a spike. Bottom panel: Peristimulus time histogram (PETH) aligned to light onset at time zero. (H) PETH for the same neuron as shown in (G) with a finer time scale. (I) Distribution of light response latencies for D1- and D2-SPNs. (J) Action potential waveforms of the same neuron in (G) for spontaneous (black) and light-evoked (orange) spikes ($R = 0.998$, $P < 0.0001$, Pearson's correlation). (K) Principal component analysis (PCA) of action potential waveforms showing the overlapped clusters of spontaneous (black) and light-evoked (orange) spikes. (L) Proportion of D1-SPN subtypes. Type 1 neurons are significantly more than other three types of neurons in D1-SPNs (Z-test, $p < 0.05$). (M) Averaged FRI for Type 1 (blue) and Type 2 (red) D1-SPNs in correct 8-s trials. (N) Proportion of D2-SPN subtypes. (O) Averaged FRI for Type 1 (blue) and Type 2 (red) D2-SPNs in correct 8-s trials.

Figure 4. Selective genetic knockout and ablation of D1- or D2-SPNs distinctly alters action selection. (A) Correct rate of control (n = 11 mice) and D1-NR1 KO mice (n = 16) in 2-8 s task during 14 days training (two-way repeated-measures ANOVA, significant difference between control and KO mice, $F_{1,25} = 10.8$, $p = 0.003$). (B) Correct rate of control (n = 17) and D2-NR1 KO mice (n = 10) in 2-8 s task during 14 days training (two-way repeated-measures ANOVA, significant difference between control and KO mice, $F_{1,25} = 8.728$, $p = 0.007$). (C) The

psychometric curve for control (n = 11) and D1-NR1 KO mice (n = 16) (two-way repeated-measures ANOVA, significant difference between control and KO mice, $F_{1,25} = 12.27$, $p = 0.002$). (D) The psychometric curve for control (n = 17) and D2-NR1 KO mice (n = 10) (two-way repeated-measures ANOVA, significant difference between control and KO mice, $F_{1,25} = 9.64$, $p = 0.005$). (E) Schematic of muscimol infusion into the dorsal striatum in trained mice. (F) Correct rate for control (black: pre-muscimol, gray: post-muscimol) and mice with muscimol infusion (magenta) in dorsal striatum (n = 9 mice, paired t-test, $p < 0.01$). (G) The psychometric curve for control (n = 9 mice, black: pre-muscimol, gray: post-muscimol control) and mice with muscimol infusion (n = 9 mice, magenta) in dorsal striatum (two-way repeated-measures ANOVA, significant difference between control and muscimol infusion, $F_{2,16} = 11.74$, $p = 0.0007$). (H) Timeline for selective diphtheria toxin (DT) ablation experiments. (I) Schematic of diphtheria toxin receptor (DTR) virus (AAV-FLEX-DTR-GFP) injection in dorsal striatum of D1-Cre mice. (J) Correct rate for control (n = 9 mice) and mice with dorsal striatum D1-SPNs ablation (D1-DTR, n = 8 mice) (two-sample t-test, $p = 0.0016$). (K) The psychometric curve for control (n = 9 mice) and D1-SPNs ablation mice (n = 8 mice) (two-way repeated-measures ANOVA, main effect of ablation, $F_{1,15} = 1.84$, $p = 0.195$; interaction between trial intervals and ablation, $F_{6,90} = 4.14$, $p = 0.001$). (L) Movement trajectory of a control mouse in 8-s trials. (M) Movement trajectory of a D1-DTR mouse in 8-s trials. (N) Schematic of diphtheria toxin receptor (DTR) virus (AAV-FLEX-DTR-GFP) injection in dorsal striatum of A2a-Cre mice. (O) Correct rate for control (n = 8 mice) and mice with dorsal striatum D2-SPNs ablation (D2-DTR, n = 8 mice) (two-sample t-test, $p = 0.005$). (P) The psychometric curve for control (n = 9 mice) and D2-SPNs ablation mice (n = 8 mice) (two-way repeated-measures ANOVA, main effect of ablation, $F_{1,15} = 0.477$, $p = 0.5$; interaction between trial intervals and ablation, $F_{6,90} = 12.6$, p

< 0.001). (Q) Movement trajectory of a control mouse in 8-s trials. (R) Movement trajectory of a D2-DTR mouse in 8-s trials. (S) Schematic of center-surround receptive field diagram for Go/No-Go (left) and Co-activation (right) models. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection.

Figure 5. Optogenetic manipulation of D1- vs. D2-SPNs differently regulates action selection. (A) Schematic of optic fiber implantation for experimentally optogenetic excitation or inhibition of D1- or D2-SPNs in the dorsal striatum. (B, C) Schematic for optogenetic excitation (B) and inhibition (C) of D1-/D2-SPNs for 1 s right before lever extension in 2-8 s task. (D) Change of correct rate for optogenetic excitation of D1-SPNs in 2-s and 8-s trials (n = 11 mice, one-sample t-test, 2-s trials: $p = 0.248$; 8-s trials: $p < 0.05$). (E) Change of correct rate for optogenetic inhibition of D1-SPNs in 2-s and 8-s trials (n = 6 mice, one-sample t-test, 2-s trials: $p = 0.557$; 8-s trials: $p < 0.05$). (F) Change of correct rate for optogenetic excitation of D2-SPNs in 2-s and 8-s trials (n = 8 mice, one-sample t-test, 2-s trials: $p < 0.05$; 8-s trials: $p < 0.05$). (G) Change of correct rate for optogenetic inhibition of D2-SPNs in 2-s and 8-s trials (n = 5 mice, one-sample t-test, 2-s trials: $p < 0.05$; 8-s trials: $p < 0.05$). (H) Schematic of center-surround receptive field diagram for Go/No-Go (left) and Co-activation (right) models. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection.

Figure 6. A triple-control computational model of basal ganglia direct and indirect pathways for action selection. (A) Network structure of the cortico-basal ganglia model based on realistic anatomy and synaptic connectivity. (B) Schematic of center-surround-context receptive field diagram for 'Triple-control' model. '+' indicates facilitating effect to selection. '-'

indicates inhibitory effect to selection. (C) The psychometric curves of behavioral output in control (black) and D1-SPNs ablation condition (blue) in 'Triple-control' model ($n = 10$, two-way repeated-measures ANOVA, main effect of ablation, $F_{1,18} = 98.72$, $p < 0.0001$; interaction between trial intervals and ablation, $F_{6,108} = 7.799$, $p < 0.0001$). (D) The psychometric curves of behavioral output in control (black) and D2-SPNs ablation condition (red) in 'Triple-control' model ($n = 10$, two-way repeated-measures ANOVA, main effect of ablation, $F_{1,18} = 99.54$, $p < 0.0001$; interaction between trial intervals and ablation, $F_{6,108} = 177.6$, $p < 0.0001$). (E) Change of correct rate for optogenetic excitation of D1-SPNs in 2-s and 8-s trials ($n = 10$, one-sample t-test, 2-s trials: $p = 0.407$; 8-s trials: $p < 0.05$). (F) Change of correct rate for optogenetic excitation of D2-SPNs in 2-s and 8-s trials ($n = 10$, one-sample t-test, 2-s trials: $p < 0.05$; 8-s trials: $p < 0.05$). (G) Change of correct rate for optogenetic inhibition of D1-SPNs in 2-s and 8-s trials ($n = 10$, one-sample t-test, 2-s trials: $p = 0.28$; 8-s trials: $p < 0.05$). (H) Change of correct rate for optogenetic inhibition of D2-SPNs in 2-s and 8-s trials ($n = 10$, one-sample t-test, 2-s trials: $p < 0.05$; 8-s trials: $p < 0.05$).

Figure 7. Computational modeling reveals direct and indirect pathways regulating action selection in a distinct manner. (A) Schematic for manipulation of D1-SPNs in 'Triple-control' model. (B) Schematic of manipulation of D1-SPNs in the center-surround-context receptive field diagram for 'Triple-control' model. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection. (C) Correct rate change in 2s trials when manipulating D1-SPNs with different manipulation strengths ($n = 10$, one-way repeated-measures ANOVA, effect of manipulation strength, $F_{36,324} = 1.171$, $p = 0.238$). (D) Correct rate change in 8s trials when manipulating D1-SPNs with different manipulation strengths ($n = 10$, one-way repeated-

measures ANOVA, effect of manipulation strength, $F_{36,324} = 13.71$, $p < 0.0001$). (E) Schematic for optogenetic manipulation of D2-SPNs in ‘Triple-control’ model. (F) Schematic of manipulation of D2-SPNs in the center-surround-context receptive field diagram for ‘Triple-control’ model. ‘+’ indicates facilitating effect to selection. ‘-’ indicates inhibitory effect to selection. (G) Correct rate change in 2s trials when manipulating D2-SPNs with different manipulation strengths ($n = 10$, one-way repeated-measures ANOVA, effect of manipulation strength, $F_{36,324} = 59.13$, $p < 0.0001$). (H) Correct rate change in 8s trials when manipulating D2-SPNs with different manipulation strengths ($n = 10$, one-way repeated-measures ANOVA, effect of manipulation strength, $F_{36,324} = 40.75$, $p < 0.0001$). (I) Diagram of linear modulation of direct pathway. (J) Diagram of nonlinear modulation of indirect pathway.

METHODS

Animals

All experiments were approved by the Salk Institute Animal Care, and done in accordance with NIH guidelines for the Care and Use of Laboratory Animals. Experiments were performed on both male and female mice, at least two months old, housed on a 12-hour light/dark cycle. C57BL/6J mice were purchased from the Jackson Laboratory at 8 weeks of age and used as wildtype mice. BAC transgenic mice expressing cre recombinase under the control of the dopamine D1 receptor (referred as D1-cre, GENSAT: EY217; minimal labeling in cortex; mostly dorsal labeling in striatum) or the A2a receptor (referred as A2a-cre, GENSAT: KG139) promoter were obtained from MMRRC, and either crossed to C57BL/6 or Ai32 (012569) mice obtained from Jackson Laboratory (Cui et al., 2013; Geddes et al., 2018; Jin et al., 2014; Madisen et al., 2012; Tecuapetla et al., 2016). Striatal neuron-type-specific NMDAR1-knockout (referred as NR1-KO) and littermate controls were generated by crossing D1-cre and A2a-cre mice with NMDAR1-loxP (also denoted as Grin1 flox/flox in the Jackson Laboratory database) mice. The behavioral experiments using NR1-KO mice were performed on 8 to 12 weeks old D1/A2a-cre + / NMDAR1-loxP homozygous mice and their littermate controls, including D1/A2a-cre +, D1/A2a-cre + / NMDAR1-loxP heterozygous and NMDAR1-loxP homozygous mice. There was no difference between the three control groups, so the data were combined.

Behavior task

Mice were trained on a temporal bisection task in the operant chamber (21.6 cm L × 17.8 cm W × 12.7 cm H), which was isolated within a sound attenuating box (Med-Associates, St. Albans, VT). The food magazine was located in the middle of one wall, and two retractable

levers were located to the left and right side of the magazine. A house light was (3 W, 24 V) mounted on the opposite wall of the magazine. Sucrose solution (10 %) was delivered into a metal bowl in the magazine through a syringe pump. When a training session started, the house light was turned on and two levers were extended. After a random time interval (30 s on average), left and right levers were retracted and extended simultaneously. Mice were able to make a choice by pressing either left or right lever. Only the very first lever press after levers extension was registered as animals' choice. If the interval between the levers retraction and extension was 2s, then only the left lever press was active to trigger the sucrose reward; if the interval between the lever retraction and extension was 8s, then only the right lever press was active to trigger the sucrose reward (Howard et al., 2017). There was no punishment when mice made an unrewarded choice. 2s-trial and 8s-trial were randomized and interleaved by random inter-trial intervals (30 s on average). The mice were also trained in the reversed 2-8s task. In the reversed 2-8s task, if the interval between the levers retraction and extension was 2s, then only the right lever press was active to trigger the sucrose reward; if the interval between the lever retraction and extension was 8s, then only the left lever press was active to trigger the sucrose reward. Representative behavioral tracks were captured by EthoVision (Noldus).

Behavior training

Mice were placed on food restriction throughout the training, and fed daily after the training sessions with ~2.5 g of regular food to allow them to maintain a body weight of around 85 % of their baseline weight. Training started with continuous reinforcement (CRF), in which animals obtained a reinforcer after each lever press. The session began with the illumination of the house light and extension of either left or right lever, and ended with the retraction of the

lever and the offset of the house light. On the first day of CRF, mice received 5 reinforcers on left and right lever. On the second day of CRF, mice received 10 reinforcers on left and right lever. On the third day of CRF, mice received 15 reinforcers on left and right lever. The order of left lever CRF and right lever CRF on each day was randomized across all the CRF training days. After the training of CRF, animals started on the temporal bisection task (day 1). Mice were trained in the temporal bisection task for 14 consecutive days. On each day, there were 240 trials with 2s-trial and 8s-trial randomly intermixed at 50:50. After 14 days training, mice received an interval discrimination test, in which 20% of 2s/8s trials were replaced by probe trials. In probe trials, the levers retraction intervals were randomly selected from 2.3s, 3.2s, 4s, 5s and 6.3s. Neither choice in the probe trials was rewarded. Mice received 4 days of test, interleaved by training days without probes. The animals were trained daily without interruption and every day the training started approximately at same time (Howard et al., 2017). All timestamps of lever presses, magazine entries and licks for each animal were recorded with 10 ms resolution. The training chambers and procedures for NR1-KO mice and their littermate controls were exactly the same for C57BL/6J mice.

For the reversed task training, mice were trained in both the 2-8 s control task and reversed version of 2-8s task on the same day for at least 14 days. During each day, mice were trained in the 2-8s task first, and then mice were put back in the home cage for a 3-4 hours rest. After the rest period, the same mice were trained in the reversed 2-8s task. The order of these two tasks is fixed throughout the 14 days training.

Surgery

For in vivo electrophysiological data recording, each mouse was chronically implanted with an electrode array which consists of an array of 2 rows × 8 columns Platinum-coated

tungsten microwire electrodes (35 μm diameter) with 150 μm spacing between microwires in a row, and 250 μm spacing between 2 rows. The craniotomies were made at the following coordinates: 0.5 mm rostral to bregma and 1.5 mm laterally for dorsal striatum; 3.4 mm caudal to bregma and 1.0 mm laterally for SNr (Jin and Costa, 2010; Jin et al., 2014). During surgeries, the electrode arrays were gently lowered ~ 2.2 mm from the surface of the brain for dorsal striatum and ~ 4.3 mm for SNr, while simultaneously monitoring neural activity. Final placement of the electrodes was monitored online during the surgery based on the neural activity, and then confirmed histologically at the end of the experiment after perfusion with 10 % formalin, brain fixation in a solution of 30 % sucrose and 10 % formalin, followed by cryostat sectioning (coronal slices of 40 - 60 μm). For striatum recording, we implanted 11 mice in the left hemisphere and 8 mice in the right hemisphere. For the SNr recording, we implanted 5 mice in the left hemisphere and 4 mice in the right hemisphere.

For the cell type identification in striatum, the cre-inducible adeno-associated virus (AAV) vector carrying the gene encoding the light-activated cation channel channelrhodopsin-2 (ChR2) and a fluorescent reporter (DIO-ChR2-YFP/DIO-ChR2-mCherry) was stereotactic injected into the dorsal striatum of D1-Cre or A2a-Cre mice, enabling cell-type-specific expression of ChR2 in striatal D1-expressing or D2-expressing projection neurons (at exactly the same coordinates of electrode array implantation in striatum stated above). DIO-ChR2-YFP/DIO-ChR2-mCherry virus (1 μl, one site, 10^{12} titer) was injected through a micro-injection Hamilton syringe, with the whole injection taking ~10 min in total. The syringe needle was left in the position for 5–10 min after the injection and then slowly moved out. Following viral injections or for mice genetically expressing ChR2 under cre control (D1-Ai32, A2a-Ai32), electrode was implanted as previously described (Geddes et al., 2018; Jin et al., 2014). The electrode array was the same as used for

dorsal striatum recording, but with a guiding cannula attached (Innovative Neurophysiology) terminating ~300 μm above the electrode tips, and was implanted into the same site after virus injection, allowing for simultaneous electrophysiological recording and light stimulation. Following the implantation, a medal needle was inserted in the cannula and mice were placed in the home cage for 2 weeks, allowing both viral expression and surgery recovery, before further training and recording experiments.

For the optogenetic manipulation in striatum, we injected the AAV virus carrying the gene for coding ChR2 (DIO-ChR2-YFP/DIO-ChR2-mCherry) or Halorhodopsin (DIO-eNpHR3.0-eYFP). Virus was injected bilaterally at 0.5 mm rostral to bregma, 2 mm laterally and ~ 2.2 mm from the surface of the brain with 1 μl per site. 10 min after the virus injection, we bilaterally implanted optical fiber units in dorsal striatum to the same site as virus injection. An optical fiber unit was made by threading a 200-μm diameter, 0.37 NA optical fiber (Thor Labs) with epoxy resin into a plastic ferrule (Geddes et al., 2018; Howard et al., 2017). Optical fiber units were then cut and polished before the implantation.

For muscimol infusion in striatum, we bilaterally implanted cannulas (Plastics One, VA) in wildtype mice to the site at 0.5 mm rostral to bregma, 2 mm laterally and ~ 2.2 mm from the surface of the brain. After the implantation, cannulas were covered by dummy cannulas. Mice were placed in the home cage for 2 weeks, allowing surgery recovery, before further training and muscimol experiments.

For striatal neuron-type-specific ablation experiments, D1-cre and A2a-cre mice were stereotaxically injected with a cre-inducible adeno-associated virus carrying the diphtheria toxin receptor (Azim et al., 2014; Geddes et al., 2018) (AAV9-FLEX-DTR-GFP; Salk GT3 Core, CA). Virus was injected in eight different sites. We used two sets of AP/ML coordinates for each

hemisphere followed by two DV depths at each AP/ML site. The coordinates were +0.9 mm AP, ±1.6 mm ML, -2.2 and -3.0 mm DV and 0.0 mm AP, ±2.1 mm ML, -2.2 and -3.0 mm DV. A Hamilton syringe was used to inject 1 uL at the four -3.0 mm DV sites and another 0.5 uL at the four -2.2 mm DV sites for a total of 3 uL injected per hemisphere. Following each injection, the needle was left in place for ~5 minutes and then raised over ~5 minutes. This same protocol was used for each injection site.

Muscimol infusion

We daily trained wildtype mice with guide cannulas (Plastics One, VA) implanted until they achieved at least 80% correct rate for 3 consecutive days, we started muscimol infusion experiments. Muscimol was dissolved in saline before infusion (Sigma-Aldrich; 0.05 ug/ul). For the infusions, mice were briefly anesthetized with isoflurane and injection cannulas (Plastics One, VA) were bilaterally inserted into the guide cannulas, with the injection cannulas projecting 0.1 mm beyond the implanted guide cannulas. Each injection cannula was attached to an infusion pump (BASi, IN) via polyethylene tubing. Animals were bilaterally infused with 200 nL of liquid (saline or muscimol) followed by a five-minute waiting period before removal of the infusion cannulas. Mice were returned to their home cage and started in the behavioral task 30 minutes after infusion (Geddes et al., 2018). To estimate the effects of muscimol on choice, we repeated saline controls and muscimol infusions at least 3 times on a single mouse to gain enough probe trials for psychometric curve fitting.

DTR-mediated cell ablation

For striatal neuron-type-specific ablation experiments, D1-cre and A2a-cre were injected with AAV9-FLEX-DTR-GFP in striatum using the same coordinates described above. After three-week recovery, mice were food-restricted and, following completion of CRF, underwent training in the 2-8s task for two weeks. Immediately after day 14 of 2-8s task training, mice were randomly divided into control and treatment groups. Treatment mice were administered mice 1 ug of diphtheria toxin (DT) dissolved in 300 uL of phosphate buffered saline (PBS) via intraperitoneal (IP) injection on two consecutive days (Azim et al., 2014; Geddes et al., 2018), whereas control mice received IP injections of PBS. To allow for neuronal ablation, animals were stopped in behavioral training and placed back on food. Animals resumed 2-8s task training with probe trials 14 days after the first DT or PBS injection.

Neural recordings during the task

The mice with electrode array implanted were trained with exactly the same procedure as described above. When mice performed the 2-8s task with 80% correct rate for 3 consecutive days, we connected the array with recording cable and continued training until mice adapted to the mechanics of the recording cable and were able to maintain the correct rate greater than 80% (Howard et al., 2017).

Neural activity was recorded using the MAP system (Plexon Inc., TX). The spike activities were initially online sorted using a sorting algorithm (Plexon Inc., TX). Only spikes with a clearly identified waveforms and relatively high signal-to-noise ratio were used for further analysis. After the recording session, the spike activities were further sorted to isolate single units by offline sorting software (Plexon Inc., TX). Single units displayed a clear refractory period in the inter-spike interval histogram, with no spikes during the refractory period (larger

than 1.3 ms) (Geddes et al., 2018; Howard et al., 2017; Jin and Costa, 2010; Jin et al., 2014). All the timestamps of animal's behavioral events were recorded as TTL pulses which were generated by a Med-Associates interface board and sent to the MAP recording system through an A/D board (Texas Instrument Inc., TX). The animal's behavioral timestamps during the training session were synchronized and recorded together with the neural activity.

Neural dynamic analysis

The animal's behavior taking place during the lever retraction time period was critical to the choice to be made, so we focused on the analysis of the neural activity from levers retraction to levers extension. Neuronal firings aligned to lever retraction were averaged across trials in 20-ms bins, and smoothed by a Gaussian filter (Gaussian filter window size = 10, standard deviation = 5) to construct the peri-event histogram (PETH). The neurons showing significant firing changes during the lever retraction period were defined as task-related neurons (ANOVA); those showing no significant changes were defined as non-task-related neurons, which were not included in the further dynamic analysis.

During 2s trials, mice behaved exactly the same as they did during the 0s - 2s period in the rewarded 8s trials, so we mainly analyzed firing activities in 8s trials. To avoid confounding effect by the sensory responses triggered by the lever retraction, only neural activity from 1s to 8s following lever retraction were included (Howard et al., 2017). Then we calculated firing rate index (FRI) based on the PETH from 1s to 8s for each individual neuron as follows:

$$\mathrm{FRI} = \frac{\mathrm{PETH} - \mathrm{mean}(\mathrm{PETH})}{\mathrm{std}(\mathrm{PETH})}$$

We then used principal component analysis (PCA) and classification algorithm, a build-in toolbox in Matlab, to classify the task-related neurons based on types of dynamics. For striatum

and SNr, we used the same algorithm to classify neurons, and we found the same types of dynamics in striatum and SNr: Type 1, monotonic decreasing; Type 2, monotonic increasing; Type 3, peak at around 4s; Type 4, trough at around 4s.

Cell type classification

In dorsal striatum, we classified neurons as putative striatal projection neurons (SPNs) if they showed waveform trough half-width between 100 μs and 250 μs and the baseline firing rate less than 10 Hz. In substantia nigra pars reticulate (SNr), neurons with firing rate higher than 15 Hz were classified as putative SNr GABA neurons, which are most likely the SNr projection neurons, because the percentage of GABAergic interneurons in the SNr is rather small (Deniau JM, 2007; Jin and Costa, 2010).

To further identify the D1 and D2 SPNs in striatum, we utilized cre-loxp technique to exclusively express ChR2 on D1-SPNs or D2-SPNs by injecting the AAV-DIO-ChR2-YFP/ AAV-DIO-ChR2-mCherry virus into dorsal striatum or genetically express ChR2 by D1-Ai32 and A2a-Ai32. Optical stimulation on ChR2-expressed cells is able to directly evoke spiking activity with short latency (Geddes et al., 2018; Jin and Costa, 2010; Jin et al., 2014). Before the training session, we connected the recording cable to the electrode array for neuronal recording and inserted an optic fiber through the cannula attached to the array to conduct light into striatum for light stimulation. For better monitoring of the same cells stably during behavioral training and the later optogenetic identification, the optic fiber was well fixed to the array. After each training session, we delivered blue light stimulation through the optic fiber from a 473-nm laser (Laserglow Technologies) via a fiber-optic patch cord, and simultaneously recorded the neuronal responses, to testify the molecular identity of cells previously recorded during the behavioral

training. The stimulation pattern was 100-ms pulse width with 4s interval. The stimulation pattern was repeatedly delivered for 100 trials. We very carefully regulated the laser power to a relatively low level for each individual recording session which was strong enough to evoke reliable spikes in a small population of neurons recorded from certain electrodes, since high laser powers usually induced an electrical signal much larger and very different from the spike waveforms previously recorded in the same electrode, presumably resulting from synchronized activation of a large population of cells surrounding the electrode. For neuron identification in different sessions in the same mouse, substantial effort was made to optimize the position of optic fiber to identify those units recorded from different electrodes and that were not being able to be identified in the previous session. The final laser power used for reliable identification of D1/D2-SPNs was between 1.0 and 1.5 mW measured at the tip of the optical fiber (slightly varying for different mice and different sessions). Only those units showing very short (≤6 ms) response latency to light stimulation and exhibiting exactly the same spike waveforms ($R \geq 0.95$, Pearson's correlation coefficient) during the behavioral performance and light response were considered as direct light-activated and cre recombinase positive neurons thus D1-SPNs or D2-SPNs (Geddes et al., 2018; Howard et al., 2017; Jin and Costa, 2010). Strict criteria were employed to minimize the possibility of false positives (with the risk of increasing false negatives, and hence having to perform more recordings/mice to achieve the same number of neurons).

Optical stimulation during the task

For optogenetic manipulation experiments, mice were injected with AAV virus carrying were pre-trained in 2-8s task for two weeks and bilaterally implanted with optic fibers. After achieving a correct rate of 80%, stimulation trials began. D1-SPNs and D2-SPNs 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 constant for Halorhodopsin experiments). Rewards were delivered only at correct responses during 2 and 8 s trials. Within 50% of any type of trials, mice were optogenetically stimulated (or inhibited) for 1 s before lever extension (Howard et al., 2017). Mice only received stimulation (or inhibition) once per trial. Sessions with correct rate below 75% for control trials were excluded from further analysis.

Computational model

We constructed a neuronal network model, including cortico-basal ganglia circuitry, to simulate the behavioral effects of ablation and optogenetic manipulation on SPNs. Specifically, cortical information corresponding to left or right choice is sent to D1- and D2-SPNs associated with these two action options (Lo and Wang, 2006; Wang, 2002). One-way collateral inhibition is added between D2 SPNs subgroups. Signals from D1- and D2-SPNs eventually 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., 2003a), which could control the final motor output either through brainstem circuits or motor cortices (Aoki et al., 2019; 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.

Cortical neurons firing activities are defined as:

$$f_{left}^{cortex}(t) = k_{left}^{cortex} e^{-t_m \cdot t} + I_{noise}^{left}(t)$$

$$f_{right}^{cortex}(t) = k_{right}^{cortex} e^{-t_m \cdot t} + I_{noise}^{right}(t)$$

where $k_{left}^{cortex} = 2,\ k_{right}^{cortex} = -2$, $t_m = 0.4$ and $I_{noise}(t)$ is defined as Gaussian white noise (mean(I_{noise}^{left}) = 1, mean(I_{noise}^{right}) = 2, SD = 0.01).

Dopamine neuron firing activities is defined as:

$$f_{DA}(t) = k_{DA} e^{-t_{DA} \cdot t} + I_{noise}^{DA}(t)$$

where $k_{DA} = 3,\ t_{DA} = 0.4$ and $I_{noise}^{DA}(t)$ is defined as Gaussian white noise (mean (I_{noise}^{DA}) = 1, SD = 0.01).

Neuronal activities of D1-SPNs are defined as:

$$\tau \frac{d f_{left}^{D1}(t)}{dt} = w_0\left(E - f_{left}^{D1}(t)\right) + \widetilde{w}_{D1_left} f_{left}^{cortex}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{d f_{right}^{D1}(t)}{dt} = w_0\left(E - f_{right}^{D1}(t)\right) + \widetilde{w}_{D1_right} f_{right}^{cortex}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

where $w_0 = 1$, $E = 20$, $\widetilde{w}_{D1_left} = 3$, $\widetilde{w}_{D1_right} = 6$, $w_{D1} = 2$, $\tau = 0.1$, $I_{noise}(t)$ is defined as Gaussian white noise (mean ($I_{noise}(t)$) = 0, SD = 0.5).

Neuronal activities of D2-SPNs in “Co-activation” module (labelled as D2-SPN 1) are defined as:

$$\tau \frac{d f_{left1}^{D2}(t)}{dt} = w_0\left(E - f_{left1}^{D2}(t)\right) + \widetilde{w}_{D2_left} f_{left}^{cortex}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{d f_{right1}^{D2}(t)}{dt} = w_0\left(E - f_{right1}^{D2}(t)\right) + \widetilde{w}_{D2_right} f_{right}^{cortex}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

where $w_0 = 1$, $E = 21$, $\widetilde{w}_{D2_left} = 5$, $\widetilde{w}_{D2_right} = 5$, $w_{D2} = -0.3$, $\tau = 0.1$, $I_{noise}(t)$ is defined as Gaussian white noise (mean ($I_{noise}(t)$) = 0, SD = 0.5).

Neuronal activities of D2-SPNs in "Go/No-go" module (labelled as D2-SPN 2) are defined as:

$$\tau \frac{df_{left2}^{D2}(t)}{dt} = w_0\left(E - f_{left2}^{D2}(t)\right) + \widetilde{w}_{D2_left} f_{left}^{cortex}(t) + w_{D2_left} S_{left}(t) f_{right1}^{D2}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{df_{right2}^{D2}(t)}{dt} = w_0\left(E - f_{right2}^{D2}(t)\right) + \widetilde{w}_{D2_right} f_{right}^{cortex}(t) + w_{D2_right} S_{right}(t) f_{left1}^{D2}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

where $w_0 = 1$, $E = 21$, $\widetilde{w}_{D2_left} = 5$, $\widetilde{w}_{D2_right} = 5$, $w_{D2_left} = -0.7$, $w_{D2_right} = -0.5$, $w_{D2} = -0.3$, $\tau = 0.1$, $I_{noise}(t)$ is defined as Gaussian white noise (mean ($I_{noise}(t)$) = 0, SD = 0.5). $S_{left}(t)$ and $S_{right}(t)$ are short-term depression functions:

$$S_{left}(t) = 3/(1 + e^{0.3(f_{right1}^{D2}(t) - 15)})$$

$$S_{right}(t) = 3/(1 + e^{0.3(f_{left1}^{D2}(t) - 15)})$$

SNr neurons receive striatal inputs as well as the local inhibitory inputs from other SNr neurons. The SNr activities are defined as:

$$\tau_{SNr} \frac{df_{left}^{SNr}(t)}{dt} = w_0\left(E - f_{left}^{SNr}(t)\right) + \widetilde{w1}_{SNr_left} f_{left}^{D1}(t) + \widetilde{w2}_{SNr} f_{right1}^{D2}(t) + \widetilde{w3}_{SNr} f_{left2}^{D2}(t) + w_{left}^{SNr} f_{right}^{SNr}(t) + I_{noise}^{s}(t)$$

$$\tau_{SNr} \frac{df_{right}^{SNr}(t)}{dt} = w_0\left(E - f_{right}^{SNr}(t)\right) + \widetilde{w1}_{SNr_right} f_{right}^{D1}(t) + \widetilde{w2}_{SNr} f_{left1}^{D2}(t) + \widetilde{w3}_{SNr} f_{right2}^{D2}(t) + w_{right}^{SNr} f_{left}^{SNr}(t) + I_{noise}^{s}(t)$$

where $w_0 = 1$, $E = 40$, $\widetilde{w1}_{SNr_left} = -0.1$, $\widetilde{w1}_{SNr_right} = -0.105$, $\widetilde{w2}_{SNr} = 0.15$, $\widetilde{w3}_{SNr} = 0.07$, $w_{left}^{SNr} = -0.027$, $w_{right}^{SNr} = -0.01$, $\tau_{SNr} = 0.2$. $I_{noise}^{s}(t)$ is defined as Gaussian white noise (mean (I_{noise}^{s})= 0, SD = 0.3).

The time-dependent choice $C(t)$ is then determined by SNr outputs $f_{left}^{SNr}(t)$ and $f_{right}^{SNr}(t)$ as follows:

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

For optogenetic manipulation of striatal neurons, the stimulation pattern is defined as:

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

and for inhibition, the pattern is defined as:

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

where t_s is the onset of stimulation/inhibition, which lasts for 1 s. *amp* is defined as the strength of the optogenetic manipulation within the range of [1, 25].

To add D1-D1 collateral connections to the 'Triple-control' model, the neuronal activities of D1-SPNs are defined as:

$$\tau \frac{df_{left}^{D1}(t)}{dt} = w_0\left(E - f_{left}^{D1}(t)\right) + \widetilde{w}_{D1_{left}} f_{left}^{cortex}(t) + w_{D1_right} f_{right}^{D1}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{df_{right}^{D1}(t)}{dt} = w_0\left(E - f_{right}^{D1}(t)\right) + \widetilde{w}_{D1_{right}} f_{right}^{cortex}(t) + w_{D1_left} f_{left}^{D1}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

where $w_{D1_left} = -0.3$, $w_{D1_right} = -0.3$.

To add D1-D2 collateral connections to the 'Triple-control' model, the neuronal activities of D2-SPNs in "Go/No-go" module (labelled as D2-SPN 2) are defined as:

$$\tau \frac{df_{left2}^{D2}(t)}{dt} = w_0\left(E - f_{left2}^{D2}(t)\right) + \widetilde{w}_{D2_left} f_{left}^{cortex}(t) + w_{D2_{left}} S_{left}(t) f_{right1}^{D2}(t) + w_{D1_{left}} f_{left}^{D1}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{df_{right2}^{D2}(t)}{dt} = w_0\left(E - f_{right2}^{D2}(t)\right) + \widetilde{w}_{D2_right} f_{right}^{cortex}(t) + w_{D2_right} S_{right}(t) f_{left1}^{D2}(t) + w_{D1_right} f_{right}^{D1}(t) + w_{D2} f_{DA}(t) + I_{noise}(t)$$

where $w_{D1_left} = -0.3$, $w_{D1_right} = -0.3$.

To add D2-D1 collateral connections to the 'Triple-control' model, the neuronal activities of D1-SPNs are defined as:

$$\tau \frac{df_{left}^{D1}(t)}{dt} = w_0\left(E - f_{left}^{D1}(t)\right) + \widetilde{w}_{D1_{left}} f_{left}^{cortex}(t) + w_{D2_left} f_{left2}^{D2}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

$$\tau \frac{df_{right}^{D1}(t)}{dt} = w_0\left(E - f_{right}^{D1}(t)\right) + \widetilde{w}_{D1_{right}} f_{right}^{cortex}(t) + w_{D2_right} f_{right2}^{D2}(t) + w_{D1} f_{DA}(t) + I_{noise}(t)$$

where $w_{D2_left} = -0.3$, $w_{D2_right} = -0.3$.

All the modeling programs were coded in Matlab.

Psychometric curve fitting

Psychometric curves for behavioral data and for theoretical curves were fit using the following equation (Brunton et al., 2013; Howard et al., 2017):

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

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).

Statistical procedures

Statistics for the wild-type and KO mice learning data were performed on the basis of values for each mouse per day. One-way and two-way repeated measures ANOVA were used to investigate general main effects; and paired or unpaired t-test were used in all planned and post-hoc comparisons. Z-test was used for the comparison of neuron proportions (Sheskin, 2003). Statistics for the optogenetic data were performed on the basis of control and stimulated values for each mouse per stimulation condition. Statistical analyses were conducted in Matlab using the statistics toolbox (The MathWorks Inc., MA) and GraphPad Prism 7 (GraphPad Software Inc., CA). Results are presented as mean ± SEM for behavior readouts and the neuronal recording data. $p < 0.05$ was considered significant. All statistical details are located within the figure legends. The number of animals used in each experiment and the number of neurons are specified in the text and figure legend.

References:

Abdi, A., Mallet, N., Mohamed, F.Y., Sharott, A., Dodson, P.D., Nakamura, K.C., Suri, S., Avery, S.V., Larvin, J.T., Garas, F.N., *et al.* (2015). Prototypic and arkypallidal neurons in the dopamine-intact external globus pallidus. The Journal of neuroscience : the official journal of the Society for Neuroscience *35*, 6667-6688.

Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L.S., Weiss, R., Cooper, T.B., Mann, J.J., Van Heertum, R.L., *et al.* (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. Proceedings of the National Academy of Sciences of the United States of America *97*, 8104-8109.

Albin, R.L., Young, A.B., and Penney, J.B. (1989). The functional anatomy of basal ganglia disorders. Trends in neurosciences *12*, 366-375.

Aoki, S., Smith, J.B., Li, H., Yan, X., Igarashi, M., Coulon, P., Wickens, J.R., Ruigrok, T.J., and Jin, X. (2019). An open cortico-basal ganglia loop allows limbic control over motor output via the nigrothalamic pathway. eLife *8*.

Atherton, J.F., Menard, A., Urbain, N., and Bevan, M.D. (2013). Short-term depression of external globus pallidus-subthalamic nucleus synaptic transmission and implications for patterning subthalamic activity. The Journal of neuroscience : the official journal of the Society for Neuroscience *33*, 7130-7144.

Azim, E., Jiang, J., Alstermark, B., and Jessell, T.M. (2014). Skilled reaching relies on a V2a propriospinal internal copy circuit. Nature *508*, 357-363.

Barbera, G., Liang, B., Zhang, L., Gerfen, C.R., Culurciello, E., Chen, R., Li, Y., and Lin, D.T. (2016). Spatially Compact Neural Clusters in the Dorsal Striatum Encode Locomotion Relevant Information. Neuron *92*, 202-213.

Bastide, M.F., Meissner, W.G., Picconi, B., Fasano, S., Fernagut, P.O., Feyder, M., Francardo, V., Alcacer, C., Ding, Y., Brambilla, R., *et al.* (2015). Pathophysiology of L-dopa-induced motor and non-motor complications in Parkinson's disease. Progress in neurobiology *132*, 96-168.

Benecke, R., Rothwell, J.C., Dick, J.P., Day, B.L., and Marsden, C.D. (1987). Disturbance of sequential movements in patients with Parkinson's disease. Brain : a journal of neurology *110 (Pt 2)*, 361-379.

Brown, J., Pan, W.X., and Dudman, J.T. (2014). The inhibitory microcircuit of the substantia nigra provides feedback gain control of the basal ganglia output. eLife *3*, e02397.

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science *340*, 95-98.

Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., and Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: a critical reappraisal. Nature neuroscience *17*, 1022-1030.

Cazorla, M., de Carvalho, F.D., Chohan, M.O., Shegda, M., Chuhma, N., Rayport, S., Ahmari, S.E., Moore, H., and Kellendonk, C. (2014). Dopamine D2 receptors regulate the anatomical and functional balance of basal ganglia circuitry. Neuron *81*, 153-164.

Cools, R., and D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biological psychiatry *69*, e113-125.

Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. Nature *494*, 238-242.

DeLong, M.R. (1990). Primate models of movement disorders of basal ganglia origin. Trends in neurosciences *13*, 281-285.

Deniau JM, M.P., Maurice N, Charpier S. (2007). Gaba and the basal ganglia - from molecules to systems, Vol 160 (Elsevier).

Ding, L., and Gold, J.I. (2012). Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. Neuron *75*, 865-874.

Durieux, P.F., Schiffmann, S.N., and de Kerchove d'Exaerde, A. (2012). Differential regulation of motor control and response to dopaminergic drugs by D1R and D2R neurons in distinct dorsal striatum subregions. The EMBO journal *31*, 640-653.

Fujiyama, F., Nakano, T., Matsuda, W., Furuta, T., Udagawa, J., and Kaneko, T. (2016). A single-neuron tracing study of arkypallidal and prototypic neurons in healthy rats. Brain structure & function *221*, 4733-4740.

Fujiyama, F., Sohn, J., Nakano, T., Furuta, T., Nakamura, K.C., Matsuda, W., and Kaneko, T. (2011). Exclusive and common targets of neostriatofugal projections of rat striosome neurons: a single neuron-tracing study using a viral vector. Eur J Neurosci *33*, 668-677.

Gallistel, C.R. (1980). The Organization of Action: A New Synthesis.

Geddes, C.E., Li, H., and Jin, X. (2018). Optogenetic Editing Reveals the Hierarchical Organization of Learned Action Sequences. Cell *174*, 32-43 e15.

Goldman-Rakic, P.S., Muly III, E.C., and Williams, G.V. (2000). D receptors in prefrontal cells and circuits. Brain Research Reviews.

Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. The Journal of neuroscience : the official journal of the Society for Neuroscience *27*, 9817-9823.

Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K.R., and Deisseroth, K. (2010). Molecular and cellular approaches for diversifying and extending optogenetics. Cell *141*, 154-165.

Graybiel, A.M. (1998). The basal ganglia and chunking of action repertoires. Neurobiology of learning and memory *70*, 119-136.

Graybiel, A.M., and Rauch, S.L. (2000). Toward a neurobiology of obsessive-compulsive disorder. Neuron *28*, 343-347.

Gustafson, N., Gireesh-Dharmaraj, E., Czubayko, U., Blackwell, K.T., and Plenz, D. (2006). A comparative voltage and current-clamp analysis of feedback and feedforward synaptic transmission in the striatal microcircuit in vitro. Journal of neurophysiology *95*, 737-752.

Hikosaka, O. (2007). GABAergic output of the basal ganglia. Progress in brain research *160*, 209-226.

Hikosaka, O., Miyashita, K., Miyachi, S., Sakai, K., and Lu, X. (1998). Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. Neurobiology of learning and memory *70*, 137-149.

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

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

Horst, N.K., Jupp, B., Roberts, A.C., and Robbins, T.W. (2019). D2 receptors and cognitive flexibility in marmosets: tri-phasic dose–response effects of intra-striatal quinpirole on serial reversal performance. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *44*, 564-571.

Howard, C.D., Li, H., Geddes, C.E., and Jin, X. (2017). Dynamic Nigrostriatal Dopamine Biases Action Selection. Neuron *93*, 1436-1450 e1438.

Howes, O.D., and Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III—The Final Common Pathway. Schizophrenia Bulletin *35*, 549-562.

Isomura, Y., Takekawa, T., Harukuni, R., Handa, T., Aizawa, H., Takada, M., and Fukai, T. (2013). Reward-modulated motor information in identified striatum neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *33*, 10209-10220.

Jarskog, L.F., Miyamoto, S., and Lieberman, J.A. (2007). Schizophrenia: new pathological insights and therapies. Annual review of medicine *58*, 49-61.

Jiang, H., Stein, B.E., and McHaffie, J.G. (2003). Opposing basal ganglia processes shape midbrain visuomotor activity bilaterally. Nature *423*, 982-986.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457-462.

Jin, X., and Costa, R.M. (2015). Shaping action sequences in basal ganglia circuits. Current opinion in neurobiology *33*, 188-196.

Jin, X., Tecuapetla, F., and Costa, R.M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nature neuroscience *17*, 423-430.

Klaus, A., and Plenz, D. (2016). A Low-Correlation Resting State of the Striatum during Cortical Avalanches and Its Role in Movement Suppression. PLoS biology *14*, e1002582.

Kravitz, A.V., Freeze, B.S., Parker, P.R., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature *466*, 622-626.

la Fougere, C., Meisenzahl, E., Schmitt, G., Stauss, J., Frodl, T., Tatsch, K., Hahn, K., Moller, H.J., and Dresel, S. (2005). D2 receptor occupancy during high- and low-dose therapy with the atypical antipsychotic amisulpride: a 123I-iodobenzamide SPECT study. Journal of nuclear medicine : official publication, Society of Nuclear Medicine *46*, 1028-1033.

Lally, J., and MacCabe, J.H. (2015). Antipsychotic medication in schizophrenia: a review. British medical bulletin *114*, 169-179.

Laruelle, M., D'Souza, C.D., Baldwin, R.M., Abi-Dargham, A., Kanes, S.J., Fingado, C.L., Seibyl, J.P., Zoghbi, S.S., Bowers, M.B., Jatlow, P., *et al*. (1997). Imaging D2 receptor occupancy by endogenous dopamine in humans. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *17*, 162-174.

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

Lee, Y., Han, N.E., Kim, W., Kim, J.G., Lee, I.B., Choi, S.J., Chun, H., Seo, M., Lee, C.J., Koh, H.Y., *et al.* (2020). Dynamic Changes in the Bridging Collaterals of the Basal Ganglia Circuitry Control Stress-Related Behaviors in Mice. Molecules and cells *43*, 360-372.

Levine, M., and Ruha, A.M. (2012). Overdose of atypical antipsychotics: clinical presentation, mechanisms of toxicity and management. CNS drugs *26*, 601-611.

Lima, S.Q., Hromadka, T., Znamenskiy, P., and Zador, A.M. (2009). PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. PloS one *4*, e6099.

Lo, C.C., and Wang, X.J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nature neuroscience *9*, 956-963.

Madisen, L., Mao, T., Koch, H., Zhuo, J.-m., Berenyi, A., Fujisawa, S., Hsu, Y.-W.A., Garcia, A.J., Gu, X., and Zanella, S. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nature neuroscience *15*, 793-802.

Mailly, P., Charpier, S., Menetrey, A., and Deniau, J.-M. (2003a). Three-dimensional organization of the recurrent axon collateral network of the substantia nigra pars reticulata neurons in the rat. Journal of Neuroscience *23*, 5247-5257.

Mailly, P., Charpier, S., Menetrey, A., and Deniau, J.M. (2003b). Three-dimensional organization of the recurrent axon collateral network of the substantia nigra pars reticulata neurons in the rat. The Journal of neuroscience : the official journal of the Society for Neuroscience *23*, 5247-5257.

Mallet, N., Schmidt, R., Leventhal, D., Chen, F., Amer, N., Boraud, T., and Berke, J.D. (2016). Arkypallidal Cells Send a Stop Signal to Striatum. Neuron *89*, 308-316.

Markowitz, J.E., Gillis, W.F., Beron, C.C., Neufeld, S.Q., Robertson, K., Bhagat, N.D., Peterson, R.E., Peterson, E., Hyun, M., Linderman, S.W., *et al.* (2018). The Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. Cell *174*, 44-58 e17.

Mattay, V.S., Goldberg, T.E., Fera, F., Hariri, A.R., Tessitore, A., Egan, M.F., Kolachana, B., Callicott, J.H., and Weinberger, D.R. (2003). Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. Proceedings of the National Academy of Sciences of the United States of America *100*, 6186-6191.

Meng, C., Zhou, J., Papaneri, A., Peddada, T., Xu, K., and Cui, G. (2018). Spectrally Resolved Fiber Photometry for Multi-component Analysis of Brain Circuits. Neuron *98*, 707-717 e704.

Mink, J.W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. Progress in neurobiology *50*, 381-425.

Mink, J.W. (2003). The Basal Ganglia and involuntary movements: impaired inhibition of competing motor patterns. Archives of neurology *60*, 1365-1368.

Nonomura, S., Nishizawa, K., Sakai, Y., Kawaguchi, Y., Kato, S., Uchigashima, M., Watanabe, M., Yamanaka, K., Enomoto, K., Chiken, S., *et al*. (2018). Monitoring and Updating of Action Selection for Goal-Directed Behavior through the Striatal Direct and Indirect Pathways. Neuron *99*, 1302-1314 e1305.

Phillips, J.G., Chiu, E., Bradshaw, J.L., and Iansek, R. (1995). Impaired movement sequencing in patients with Huntington's disease: a kinematic analysis. Neuropsychologia *33*, 365-369.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? Neuroscience *89*, 1009-1023.

Schmidt, R., and Berke, J.D. (2017). A Pause-then-Cancel model of stopping: evidence from basal ganglia neurophysiology. Philosophical transactions of the Royal Society of London Series B, Biological sciences *372*.

Seeman, P. (2002). Atypical antipsychotics: mechanism of action. Canadian journal of psychiatry Revue canadienne de psychiatrie *47*, 27-38.

Sheskin, D.J. (2003). Handbook of parametric and nonparametric statistical procedures (Chapman and Hall/CRC).

Smith, Y., Bevan, M., Shink, E., and Bolam, J.P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. Neuroscience *86*, 353-387.

Tai, L.H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nature neuroscience *15*, 1281-1289.

Taverna, S., Ilijic, E., and Surmeier, D.J. (2008). Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of Parkinson's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 5504-5512.

Tecuapetla, F., Carrillo-Reid, L., Bargas, J., and Galarraga, E. (2007). Dopaminergic modulation of short-term synaptic plasticity at striatal inhibitory synapses. Proceedings of the National Academy of Sciences *104*, 10258-10263.

Tecuapetla, F., Jin, X., Lima, S.Q., and Costa, R.M. (2016). Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. Cell *166*, 703-715.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. Neuron *36*, 955-968.

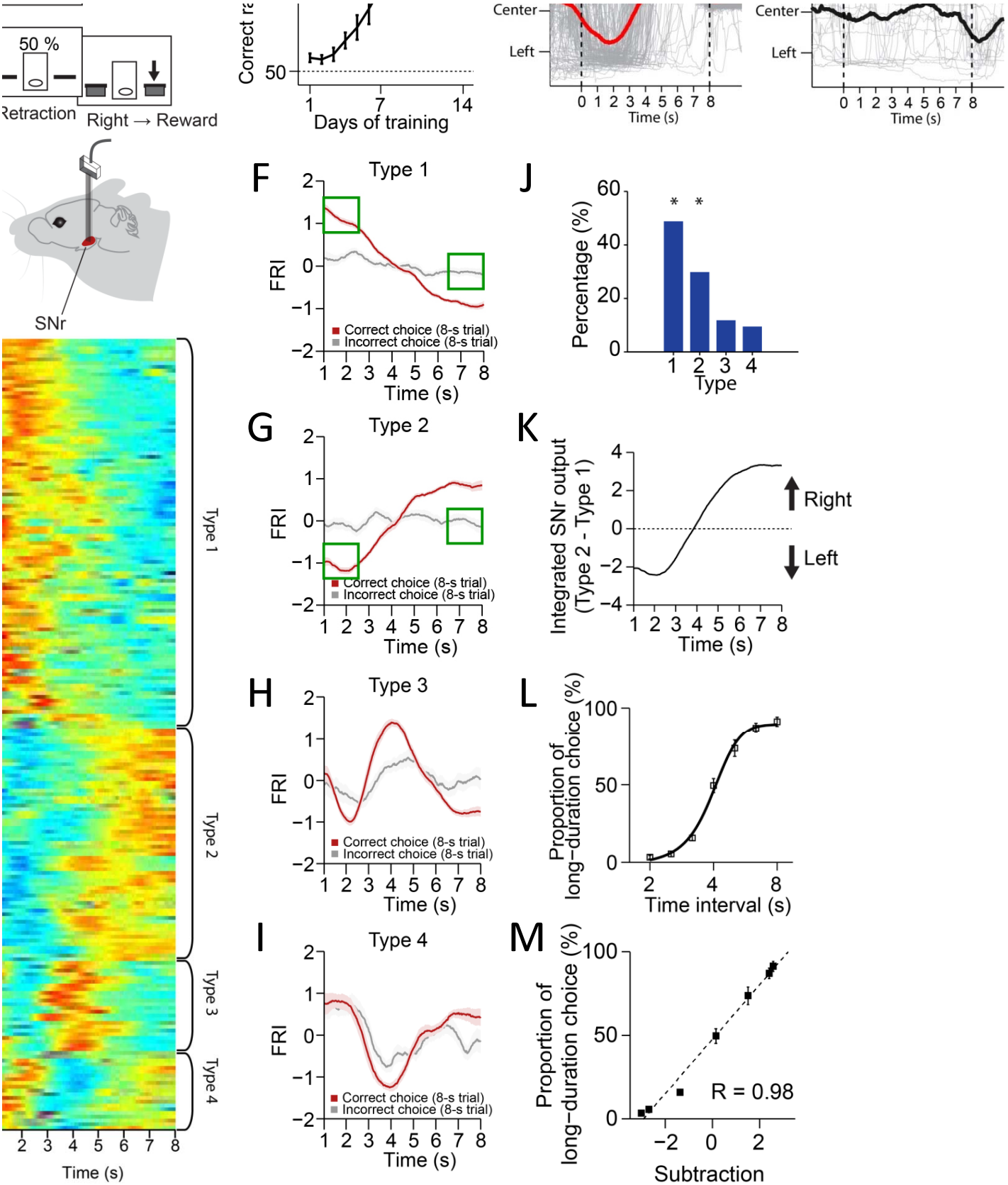
Wei, W., and Wang, X.J. (2016). Inhibitory Control in the Cortico-Basal Ganglia-Thalamocortical Loop: Complex Regulation and Interplay with Memory and Decision Processes. Neuron *92*, 1093-1105.

Wong, D.F., Wagner, H.N., Jr., Tune, L.E., Dannals, R.F., Pearlson, G.D., Links, J.M., Tamminga, C.A., Broussolle, E.P., Ravert, H.T., Wilson, A.A., *et al.* (1986). Positron emission tomography reveals elevated D2 dopamine receptors in drug-naive schizophrenics. Science *234*, 1558-1563.

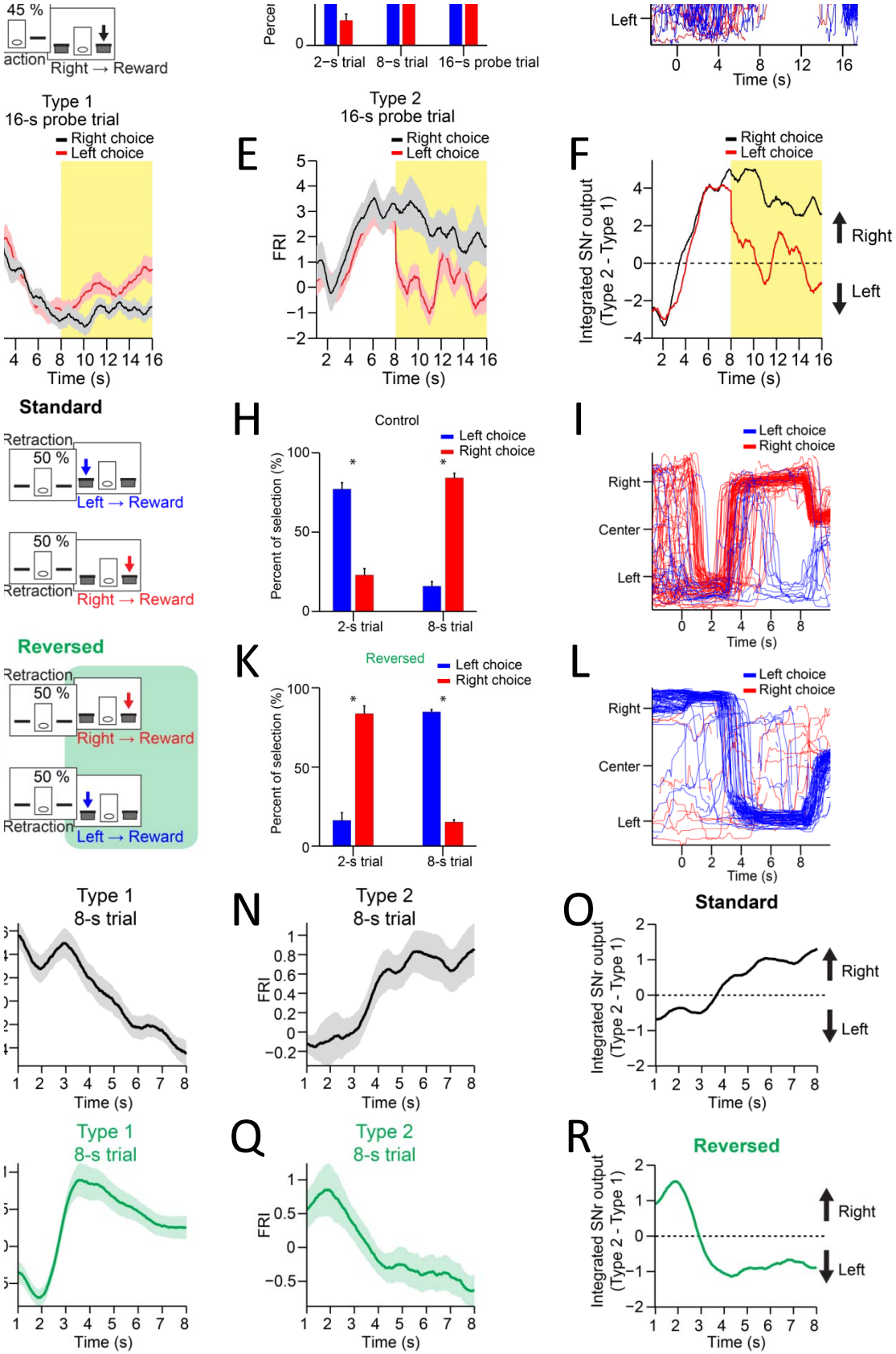
Wu, Y., Richard, S., and Parent, A. (2000). The organization of the striatal output system: a single-cell juxtacellular labeling study in the rat. Neurosci Res *38*, 49-62.

Wurtz, R.H., and Hikosaka, O. (1986). Role of the basal ganglia in the initiation of saccadic eye movements. Progress in brain research *64*, 175-190.

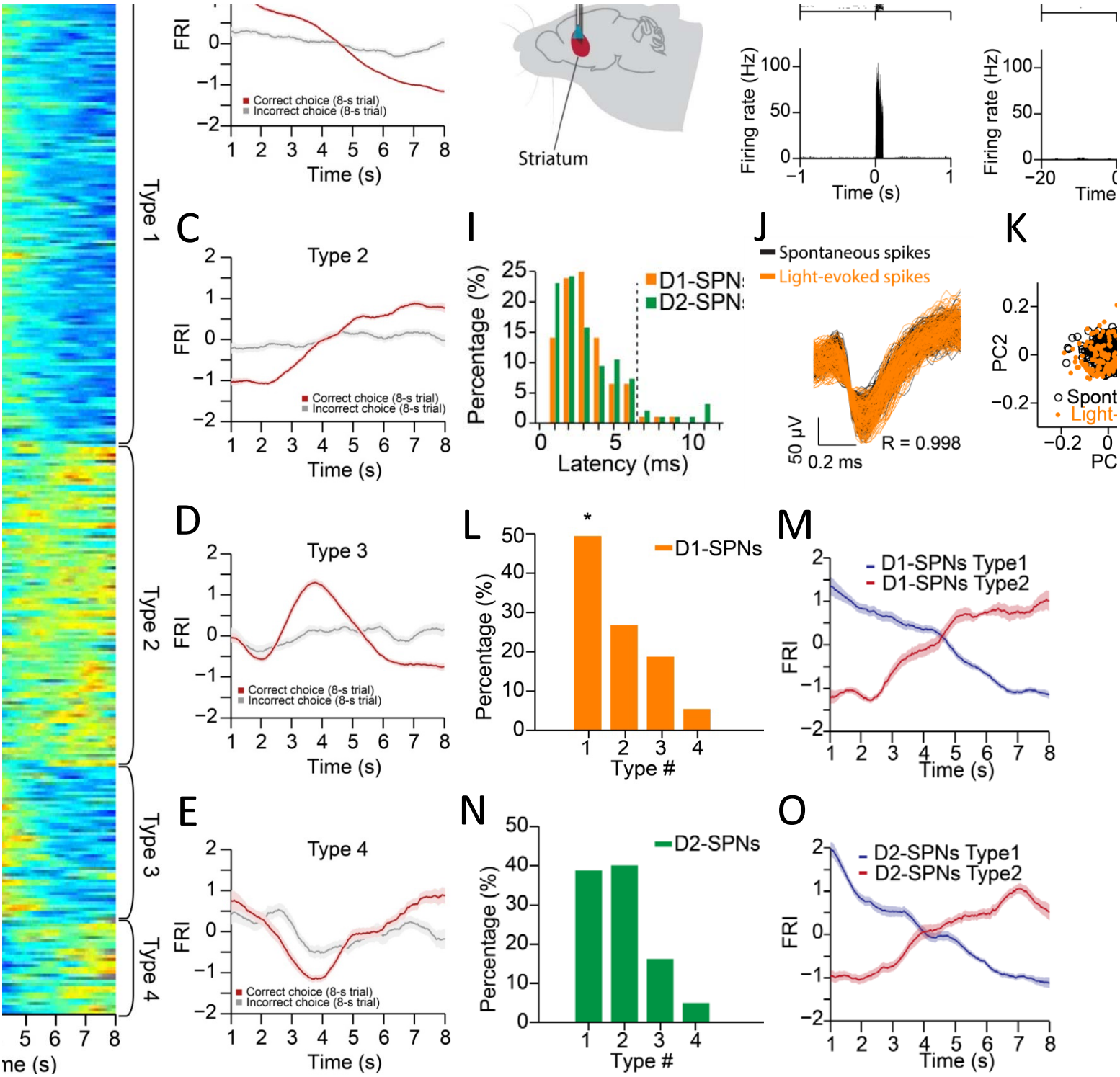
Yokoi, F., Gründer, G., Biziere, K., Stephane, M., Dogan, A.S., Dannals, R.F., Ravert, H., Suri, A., Bramer, S., and Wong, D.F. (2002). Dopamine D 2 and D 3 receptor occupancy in normal humans treated with the antipsychotic drug aripiprazole (OPC 14597): a study using positron emission tomography and [11 C] raclopride. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *27*, 248-259.



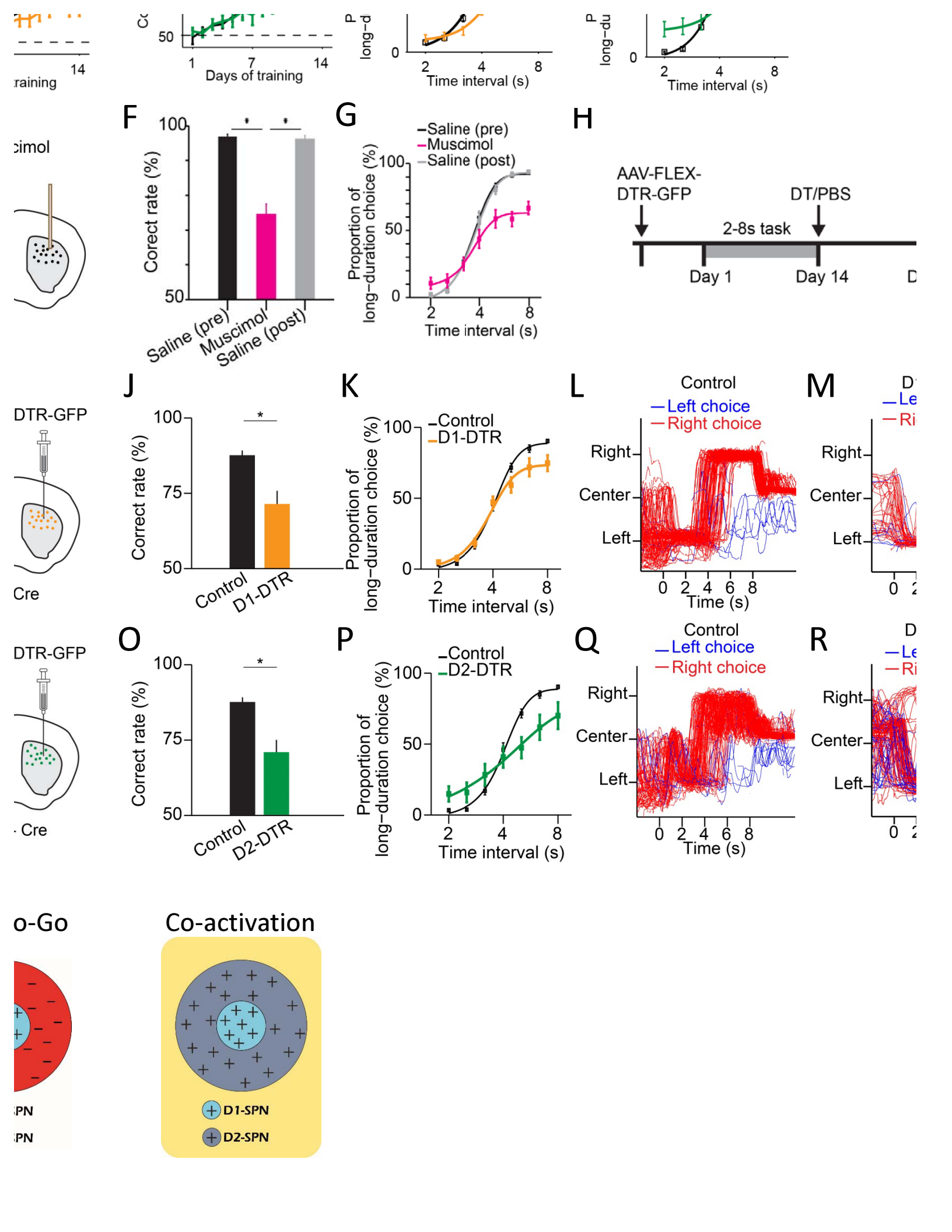
neuronal dynamics in SNr during the 2-8 s action selection task. (A) Schematic diagram for the design of 2-8 s task. ype mice across 14 days training (n = 10 mice, one-way repeated-measures ANOVA, significant effect of training days, F Movement trajectory of an example mouse in correct (left panel) and incorrect (right panel) 8-s trials (gray line: trajecto k line: the average trajectory). (D) Diagram of electrode array implanted into substantia nigra pars reticulata (SNr). (E) F neuronal activity for all task-related SNr neurons in correct 8-s trials. The magnitude of FRI is color coded and the SNr ur different types based on the activity dynamics. (F-I) Averaged FRI for Type 1 (F, green squares indicating activities re 2 (G, green squares indicating activities related to left choice), Type 3 (H), Type 4 (I) of SNr neurons in correct (red) and) The proportion of four types of SNr neurons. Type 1 and Type 2 are major types and significantly more than Type 3 an (K) Integrated SNr output defined as the subtraction of averaged FRI between Type 1 and Type 2 SNr neurons. (L) Aver



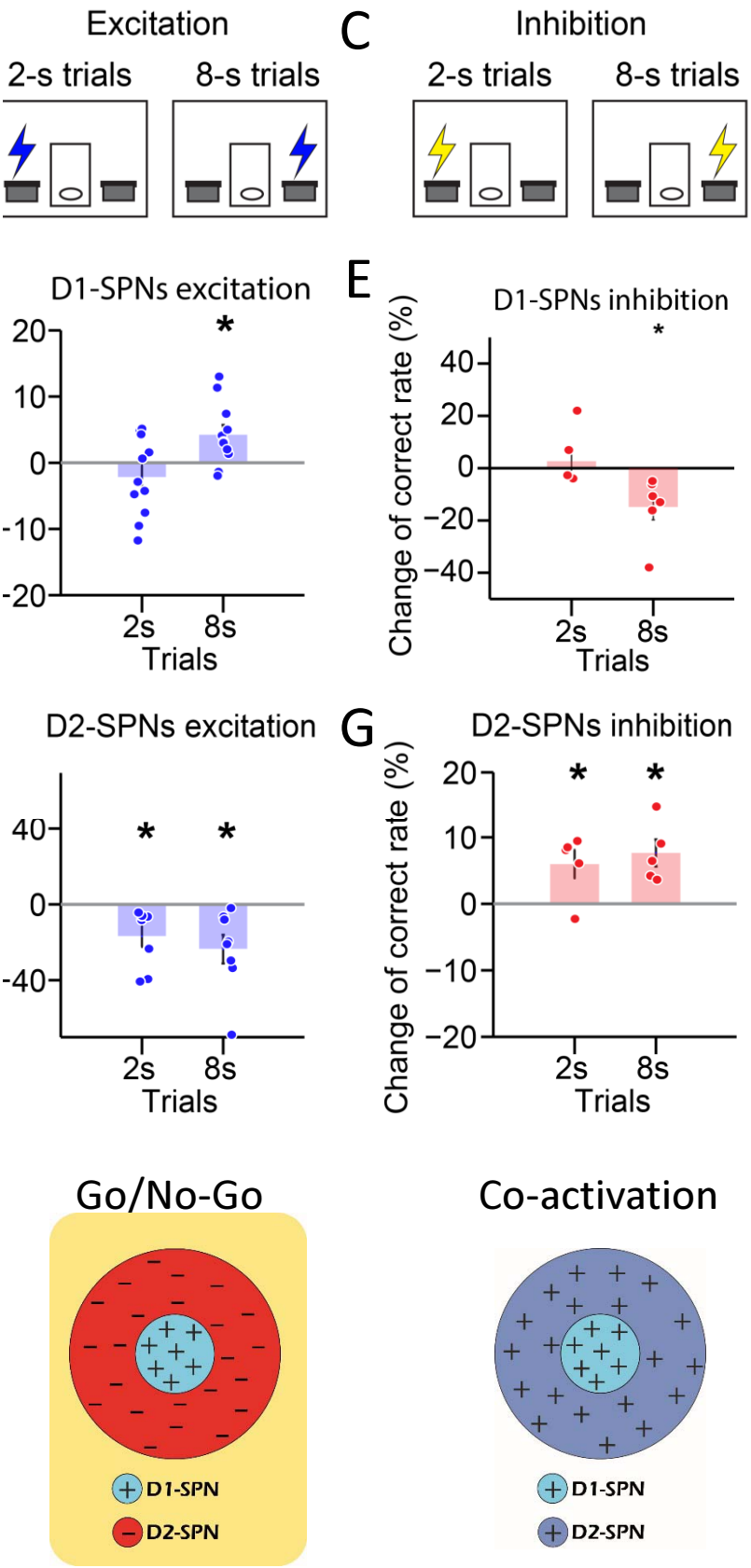
neuronal dynamics reflect action selection but not interval time or reward value. (A) Task diagram of 2-8 s control
e trials. (B) Percentage of behavioral choice in 2-s, 8-s and 16-s trials (blue: left choice; red: right choice) (n = 9 mice, p
Movement trajectory of an example mouse in 16-s trials (blue: left choice; red: right choice). (D) Averaged SNr Type 1 F
choice; black: right choice). Firing rates from 8s to 16s (highlighted area) are compared between left and right choice (n
way repeated-measures ANOVA, significant difference between left and right choices, $F_{1,25} = 6.646$, $p = 0.016$). (E) Avera
16-s trials (red: left choice; black: right choice). Firing rates from 8s to 16s are compared between left and right choice (
way repeated-measures ANOVA, significant difference between left and right choices, $F_{1,15} = 5.785$, $p = 0.029$). (F) Subtr
ype 1 and Type 2 neurons in 16-s probe trials (red: left choice; black: right choice). (G) Task design of 2-8 s standard task
behavioral choice in 2-s and 8-s trials (blue: left choice; red: right choice) (n = 6 mice, paired t-test, $p < 0.05$). (I) Moven
n example mouse in 8-s trials (blue: left choice; red: right choice). (J) Task design of reversed 2-8 s task. (K) Percentage



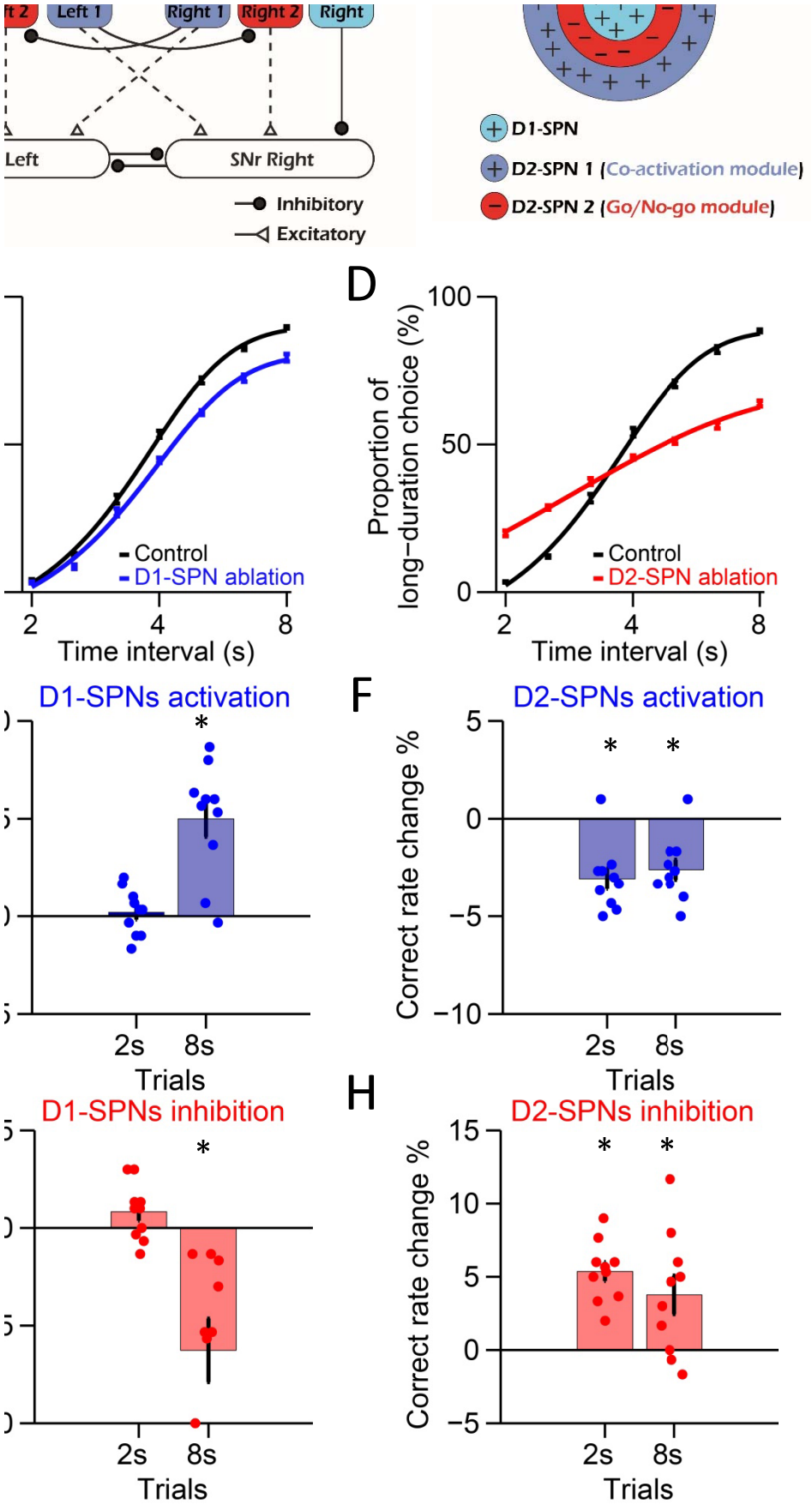
onal activity of striatal D1- and D2-SPNs during action selection. (A) FRI of neuronal activity for all task-related SPN were classified as Type 1 - 4. (B-E) Averaged FRI for Type 1 (B), Type 2 (C), Type 3 (D), Type 4 (E) of SPNs in correct (ls (gray). (F) Diagram of simultaneous neuronal recording and optogenetic identification of D1- vs. D2-SPNs in dorsal s r plot for a representative D1-SPN response to 100 ms optogenetic stimulation. Each row represents one trial and each b e. Bottom panel: Peristimulus time histogram (PETH) aligned to light onset at time zero. (H) PETH for the same neuron er time scale. (I) Distribution of light response latencies for D1- and D2-SPNs. (J) Action potential waveforms of the sam ous (black) and light-evoked (orange) spikes (R = 0.998, P < 0.0001, Pearson's correlation). (K) Principal component an potential waveforms showing the overlapped clusters of spontaneous (black) and light-evoked (orange) spikes. (L) Propo s. Type 1 neurons are significantly more than other three types of neurons in D1-SPNs (Z-test, p < 0.05). (M) Averaged I d Type 2 (red) D1-SPNs in correct 8-s trials. (N) Proportion of D2-SPN subtypes. (O) Averaged FRI for Type 1 (blue) an n correct 8-s trials.


Days of training
Time interval (s)
F
Corect rate (%)
Saline (pre)
Muscimol
Saline (post)
G
Proportion of long−duration choice (%)
H
AAV-FLEX-DTR-GFP
DT/PBS
2-8s task
Day 1
Day 14
J
Correct rate (%)
K
Control
D1-DTR
L
Control
Left choice
Right choice
Right
Center
Left
Time (s)
M
O
D2-DTR
P
Q
R
Cre
Co-activation
D1-SPN
D2-SPN

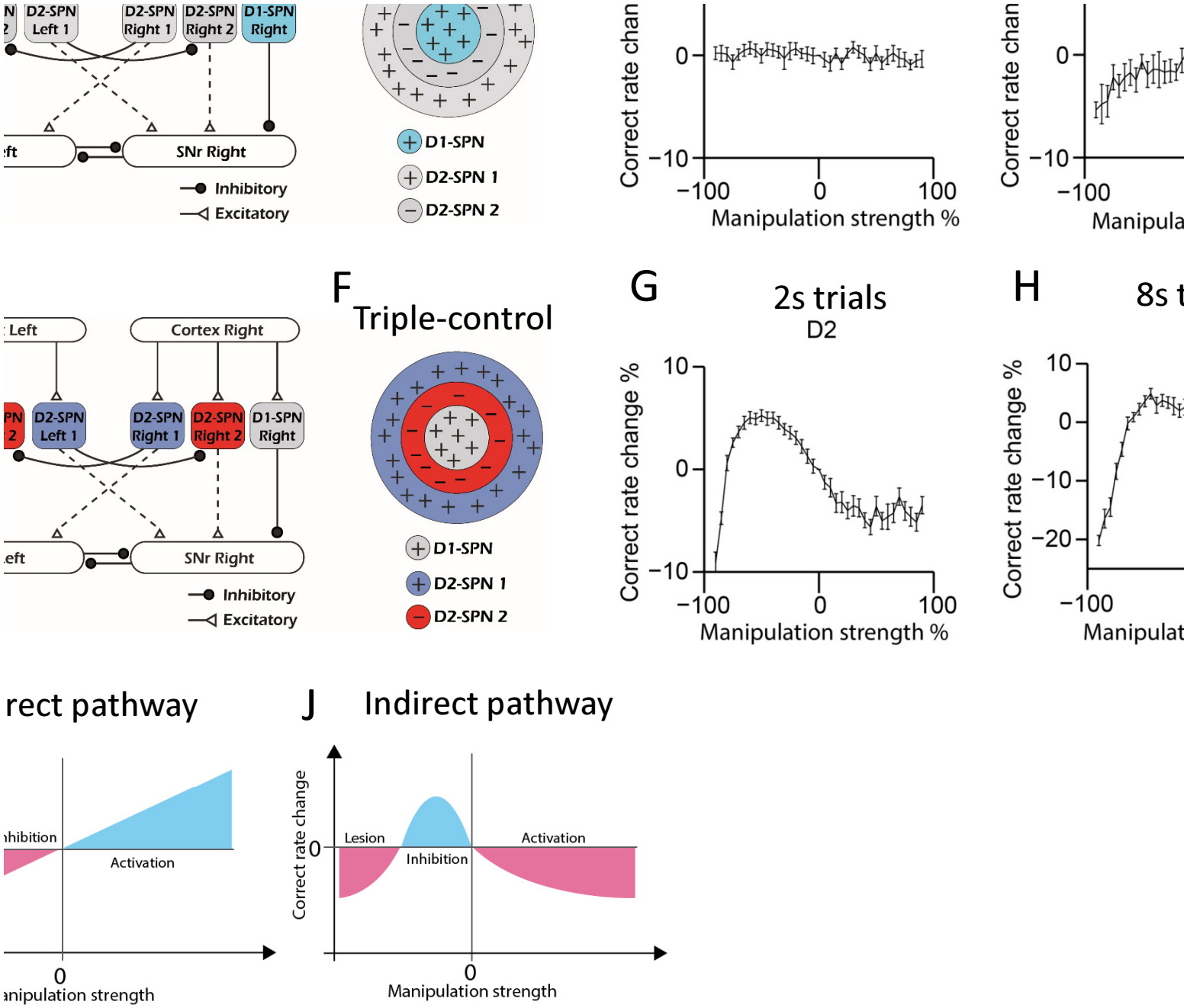
C) The psychometric curve for control (n = 11) and D1-NR1 KO mice (n = 16) (two-way repeated-measures ANOVA, sig
e between control and KO mice, $F_{1,25} = 12.27$, $p = 0.002$). (D) The psychometric curve for control (n = 17) and D2-NR1 I
two-way repeated-measures ANOVA, significant difference between control and KO mice, $F_{1,25} = 9.64$, $p = 0.005$). (E)
c of muscimol infusion into the dorsal striatum in trained mice. (F) Correct rate for control (black: pre-muscimol, gray: p
l) and mice with muscimol infusion (magenta) in dorsal striatum (n = 9 mice, paired t-test, $p < 0.01$). (G) The psychomet
ol (n = 9 mice, black: pre-muscimol, gray: post-muscimol control) and mice with muscimol infusion (n = 9 mice, magent
iatum (two-way repeated-measures ANOVA, significant difference between control and muscimol infusion, $F_{2,16} = 11.74$
(H) Timeline for selective diphtheria toxin (DT) ablation experiments. (I) Schematic of diphtheria toxin receptor (DTR) v
EX-DTR-GFP) injection in dorsal striatum of D1-Cre mice. (J) Correct rate for control (n = 9 mice) and mice with dorsa
D1-SPNs ablation (D1-DTR, n = 8 mice) (two-sample t-test, $p = 0.0016$). (K) The psychometric curve for control (n = 9 r
PNs ablation mice (n = 8 mice) (two-way repeated-measures ANOVA, main effect of ablation, $F_{1,15} = 1.84$, $p = 0.195$; int
rial intervals and ablation, F6,90 = 4.14, $p = 0.001$). (L) Movement trajectory of a control mouse in 8-s trials. (M) Mover
of a D1-DTR mouse in 8-s trials. (N) Schematic of diphtheria toxin receptor (DTR) virus (AAV-FLEX-DTR-GFP) injec
iatum of A2a-Cre mice. (O) Correct rate for control (n = 8 mice) and mice with dorsal striatum D2-SPNs ablation (D2-D
two-sample t-test, $p = 0.005$). (P) The psychometric curve for control (n = 9 mice) and D2-SPNs ablation mice (n = 8 mic
ated-measures ANOVA, main effect of ablation, $F_{1,15} = 0.477$, $p = 0.5$; interaction between trial intervals and ablation, F_6
0.001). (Q) Movement trajectory of a control mouse in 8-s trials. (R) Movement trajectory of a D2-DTR mouse in 8-s tri
c of center-surround receptive field diagram for Go/No-Go (left) and Co-activation (right) models. '+' indicates facilitati
on. '-' indicates inhibitory effect to selection.



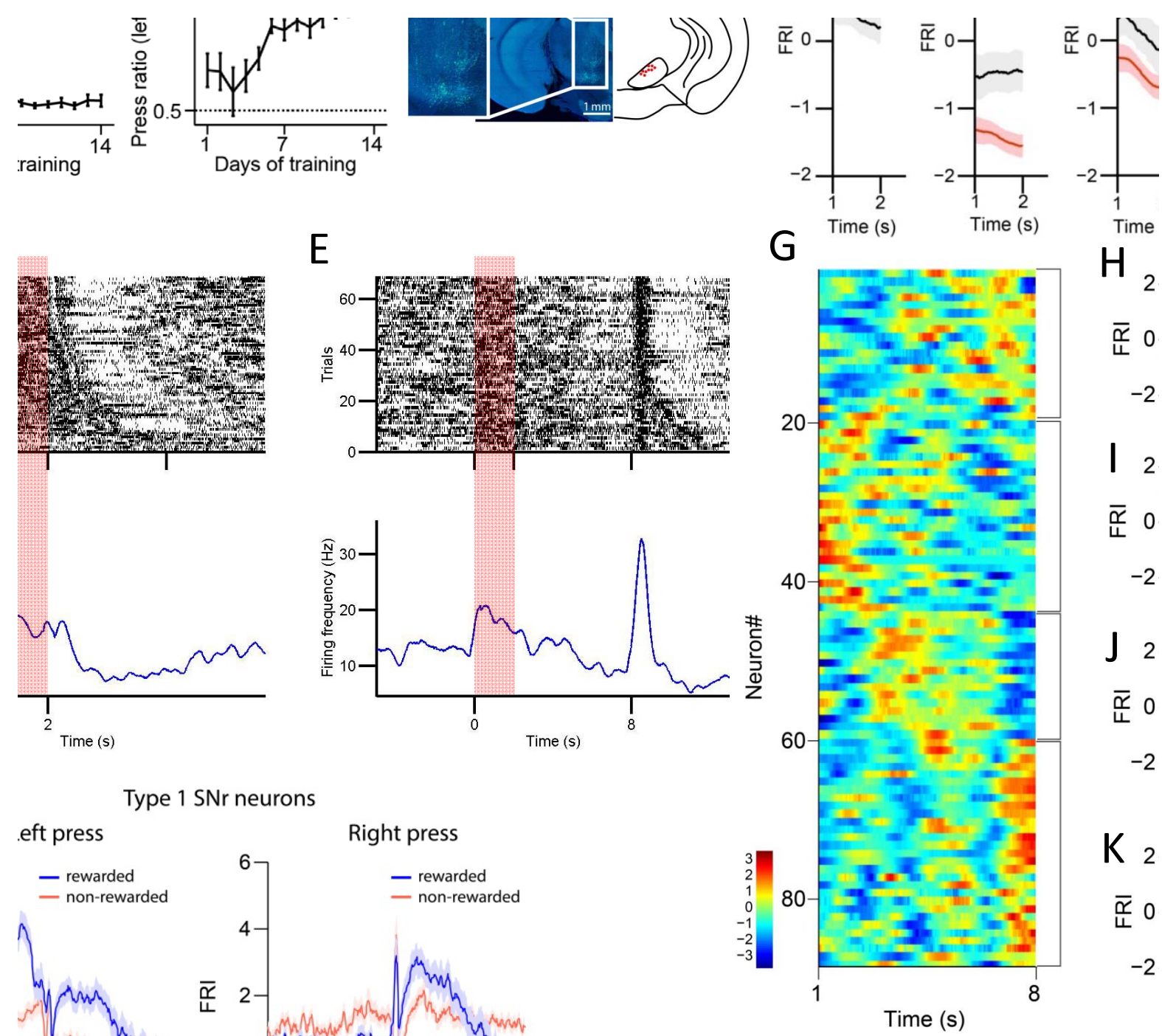
ptogenetic manipulation of D1- vs. D2-SPNs differently regulates action selection. (A) Schematic of optic fiber impl
lly optogenetic excitation or inhibition of D1- or D2-SPNs in the dorsal striatum. (B, C) Schematic for optogenetic excit
) of D1-/D2-SPNs for 1 s right before lever extension in 2-8 s task. (D) Change of correct rate for optogenetic excitation
rials (n = 11 mice, one-sample t-test, 2-s trials: p = 0.248; 8-s trials: p < 0.05). (E) Change of correct rate for optogenetic
2-s and 8-s trials (n = 6 mice, one-sample t-test, 2-s trials: p = 0.557; 8-s trials: p < 0.05). (F) Change of correct rate for

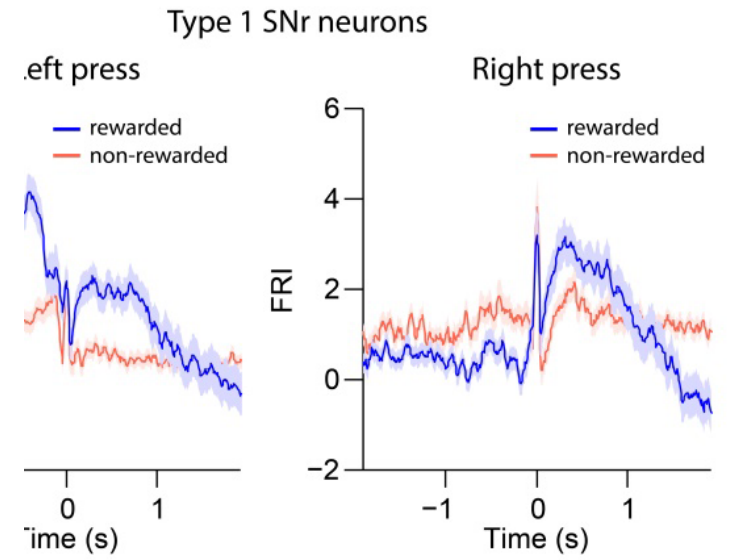


le-control computational model of basal ganglia direct and indirect pathways for action selection. (A) Network stru
nglia model based on realistic anatomy and synaptic connectivity. (B) Schematic of center-surround-context receptive fie
ol' model. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection. (C) The psychometric cu
ut in control (black) and D1-SPNs ablation condition (blue) in 'Triple-control' model (n = 10, two-way repeated-measure
blation, $F_{1,18} = 98.72$, $p < 0.0001$; interaction between trial intervals and ablation, $F_{6,108} = 7.799$, $p < 0.0001$). (D) The psy
oral output in control (black) and D2-SPNs ablation condition (red) in 'Triple-control' model (n = 10, two-way repeated-
effect of ablation, $F_{1,18} = 99.54$, $p < 0.0001$; interaction between trial intervals and ablation, $F_{6,108} = 177.6$, $p < 0.0001$). (E
ptogenetic excitation of D1-SPNs in 2-s and 8-s trials (n = 10, one-sample t-test, 2-s trials: p = 0.407; 8-s trials: p < 0.05



nputational modeling reveals direct and indirect pathways regulating action selection in a distinct manner. (A) Sch
of D1-SPNs in 'Triple-control' model. (B) Schematic of manipulation of D1-SPNs in the center-surround-context receptiv
'riple-control' model. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection. (C) Correct ra
manipulating D1-SPNs with different manipulation strengths (n = 10, one-way repeated-measures ANOVA, effect of ma
$_4$ = 1.171, p = 0.238). (D) Correct rate change in 8s trials when manipulating D1-SPNs with different manipulation streng
ted-measures ANOVA, effect of manipulation strength, $F_{36,324}$ = 13.71, p < 0.0001). (E) Schematic for optogenetic manip
'riple-control' model. (F) Schematic of manipulation of D2-SPNs in the center-surround-context receptive field diagram
. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection. (G) Correct rate change in 2s trial
D2-SPNs with different manipulation strengths (n = 10, one-way repeated-measures ANOVA, effect of manipulation stre
001). (H) Correct rate change in 8s trials when manipulating D2-SPNs with different manipulation strengths (n = 10, one-
ures ANOVA, effect of manipulation strength, $F_{36,324}$ = 40.75, p < 0.0001). (I) Diagram of linear relationship between the
way and correct rate of action selection. (J) Diagram of nonlinear relationship between the modulation of indirect pathwa
selection.





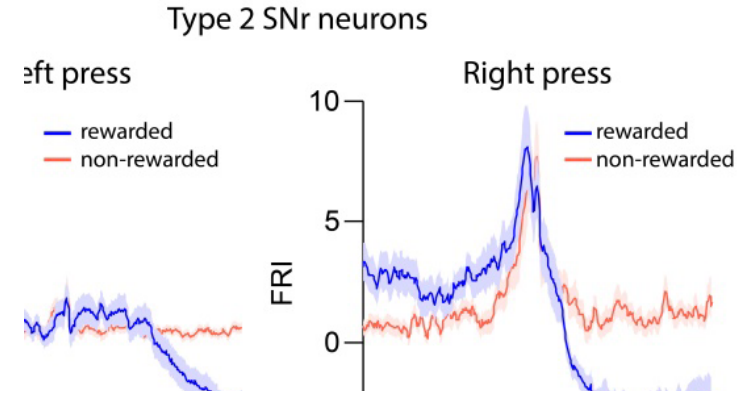
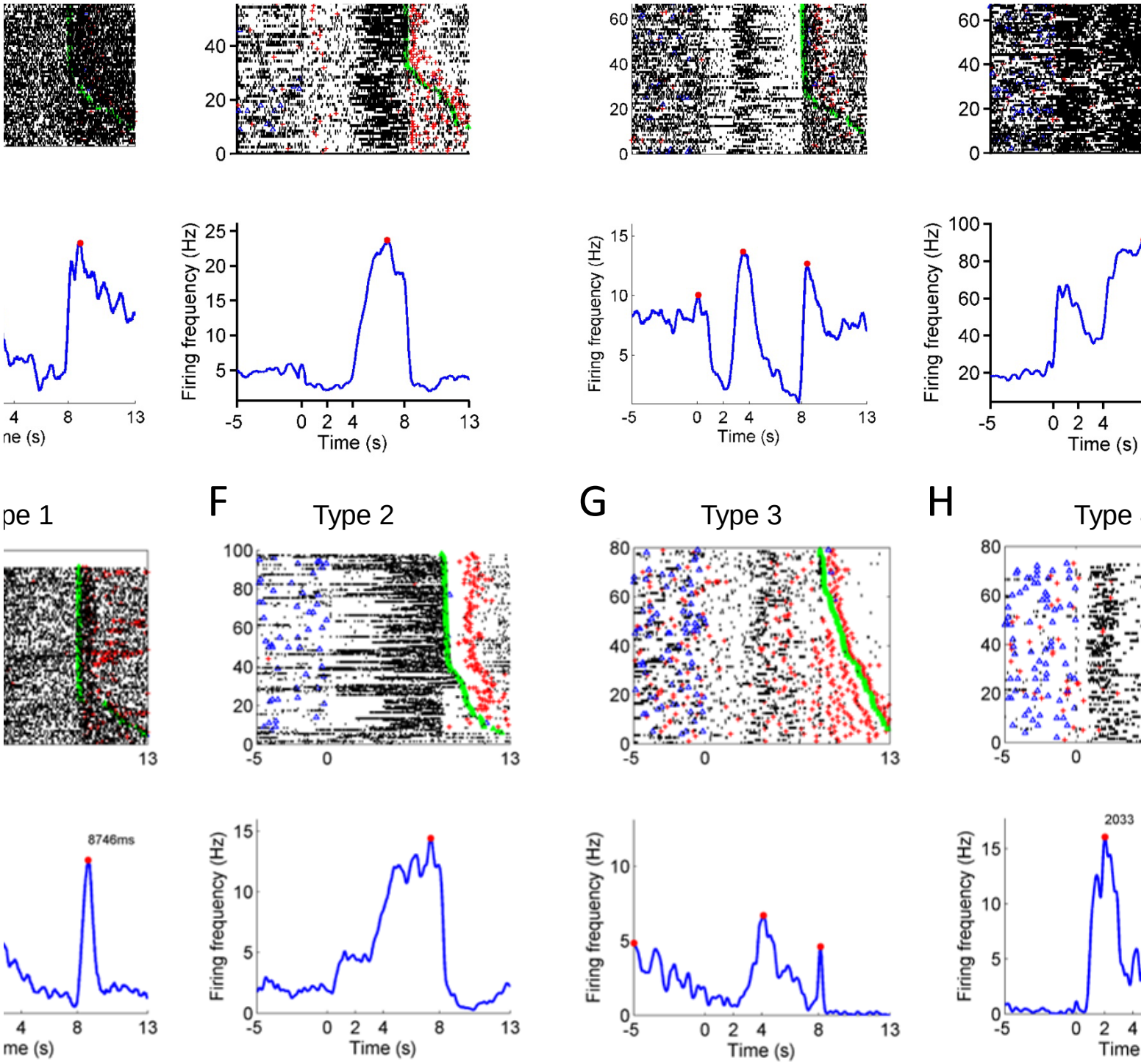
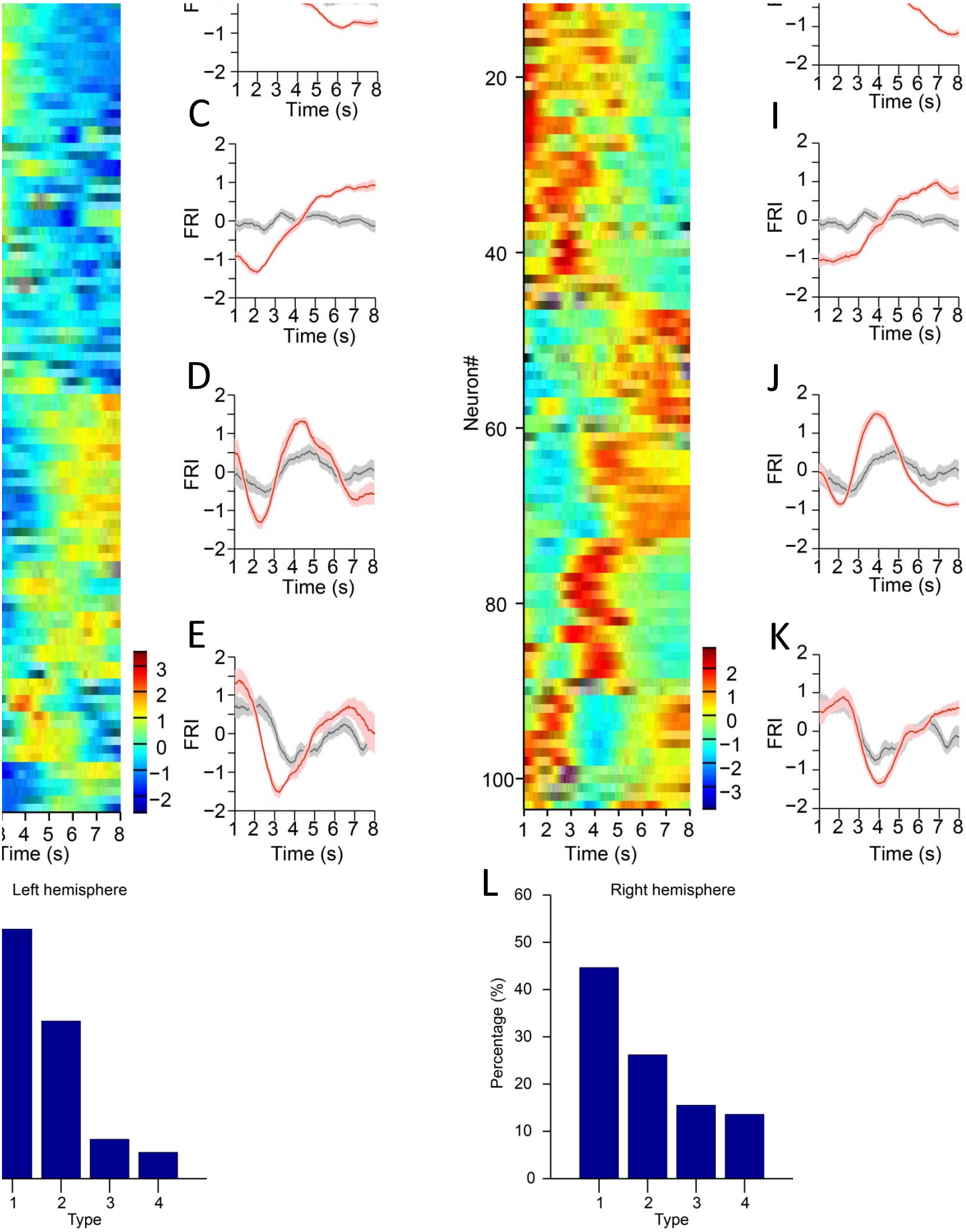


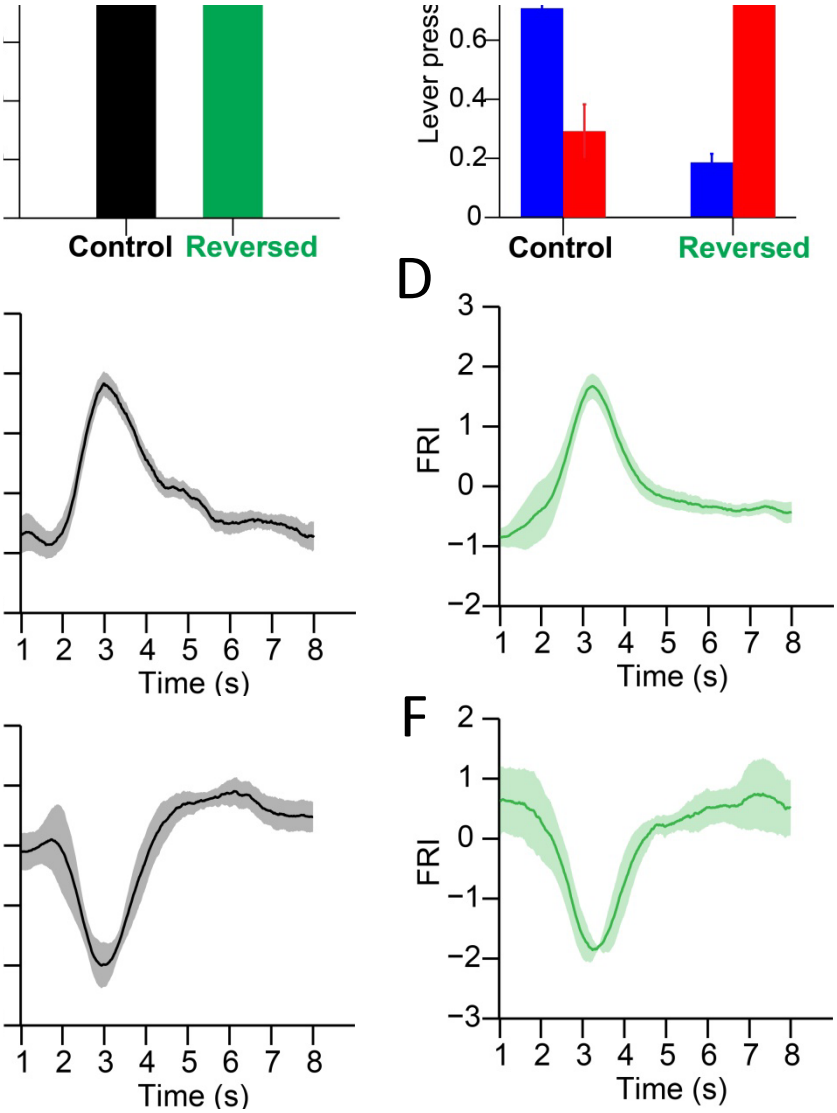
Figure S1. Behavioral performance across 14 days of tr
SNr neuronal recording on day 1. (A) Lever press latenc
extension for wild type mice across 14 days training (n = 1
way repeated-measures ANOVA, effect of training days, *F*
< 0.0001). (B) Lever press ratio for wild type mice across
training (n = 10 mice, one-way repeated-measures ANOVA
training days, $F_{13,117} = 6.472$, $p < 0.0001$). (C) Example of
array placement in SNr (left) and validation of array place
cohort of wildtype mice (right). Inset better demonstrates s
formed by the array implant. (D, E) Firing activities of an
neuron in correct 2-s (D) and 8-s trials (E) after 14 days tra
panels: raster plot of spikes across trials aligned to lever re
0. Bottom panel: PETH plot. Red shaded areas highlight th
period after the lever retraction in 2-s (D) and 8-s trials (E)
FRI for Type 1, Type 2, Type 3 and Type 4 of SNr neurons
and incorrect 2-s trials (gray). (G) Firing Rate Index (FRI)
activity for all task-related SNr neurons in correct 8-s trials



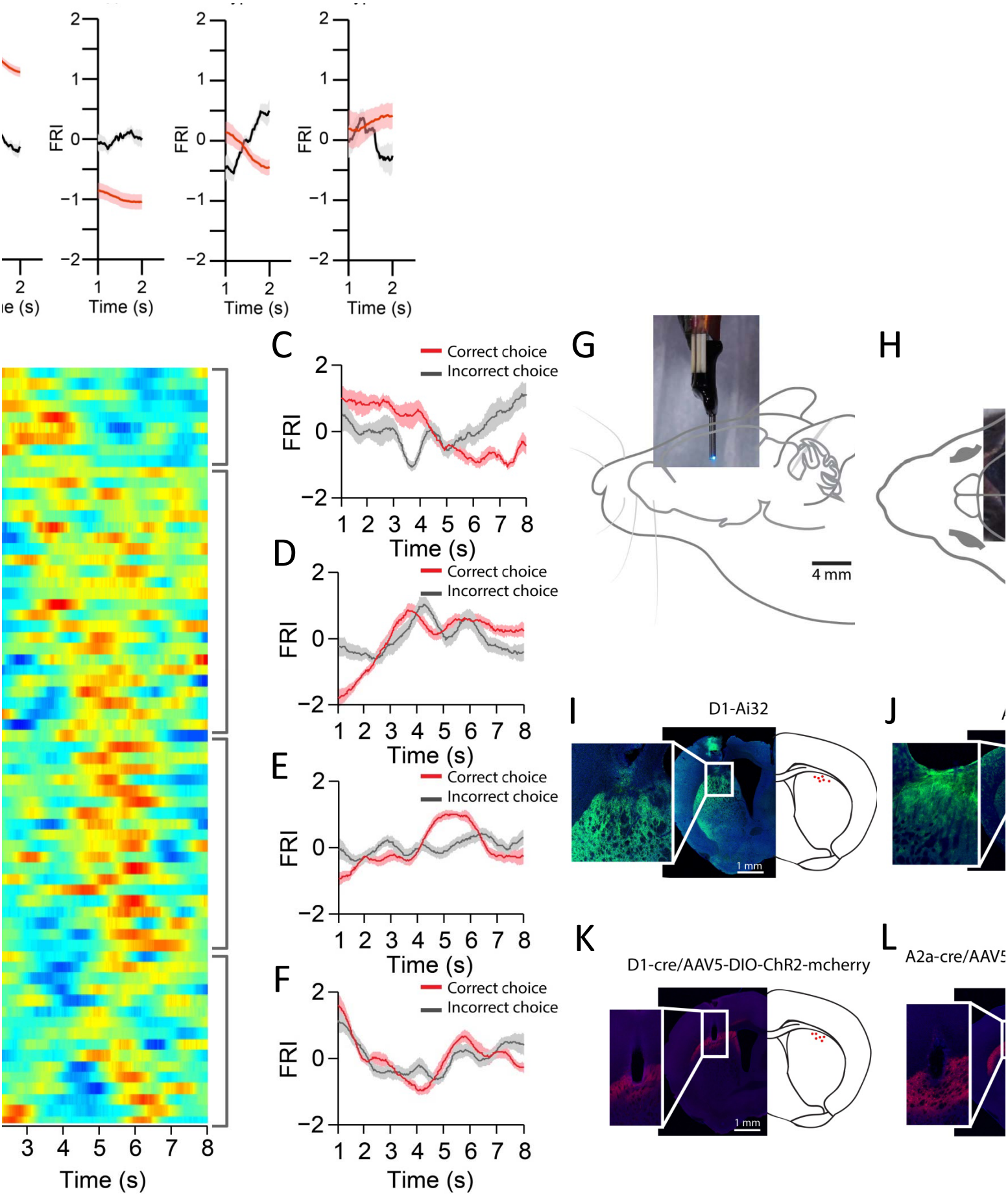
Examples of SNr neuron and SPN subtypes. (A) Firing activities of an example Type 1 SNr neuron in correct 8-s trials ning. Top panels: raster plot of spikes across trials aligned to lever retraction at time 0 (blue triangle: lever press; green tr cross: head entry). Bottom panel: PETH plot. (B) Firing activities of an example Type 2 SNr neuron in correct 8-s trials g. Top panels: raster plot of spikes across trials aligned to lever retraction at time 0. Bottom panel: PETH plot. (C) Firing an example Type 3 SNr neuron in correct 8-s trials after 14 days training. Top panels: raster plot of spikes across trials al action at time 0. Bottom panel: PETH plot. (D) Firing activities of an example Type 4 SNr neuron in correct 8-s trials aft g. Top panels: raster plot of spikes across trials aligned to lever retraction at time 0. Bottom panel: PETH plot. (E) Firing an example Type 1 SPN in correct 8-s trials after 14 days training. Top panels: raster plot of spikes across trials aligned t time 0 (blue triangle: lever press; green triangle: reward; red cross: head entry). Bottom panel: PETH plot. (F) Firing ac le Type 2 SPN in correct 8-s trials after 14 days training. Top panels: raster plot of spikes across trials aligned to lever re ottom panel: PETH plot. (G) Firing activities of an example Type 3 SPN in correct 8-s trials after 14 days training. Top p f spikes across trials aligned to lever retraction at time 0. Bottom panel: PETH plot. (H) Firing activities of an example T ect 8-s trials after 14 days training. Top panels: raster plot of spikes across trials aligned to lever retraction at time 0. Bot H plot.



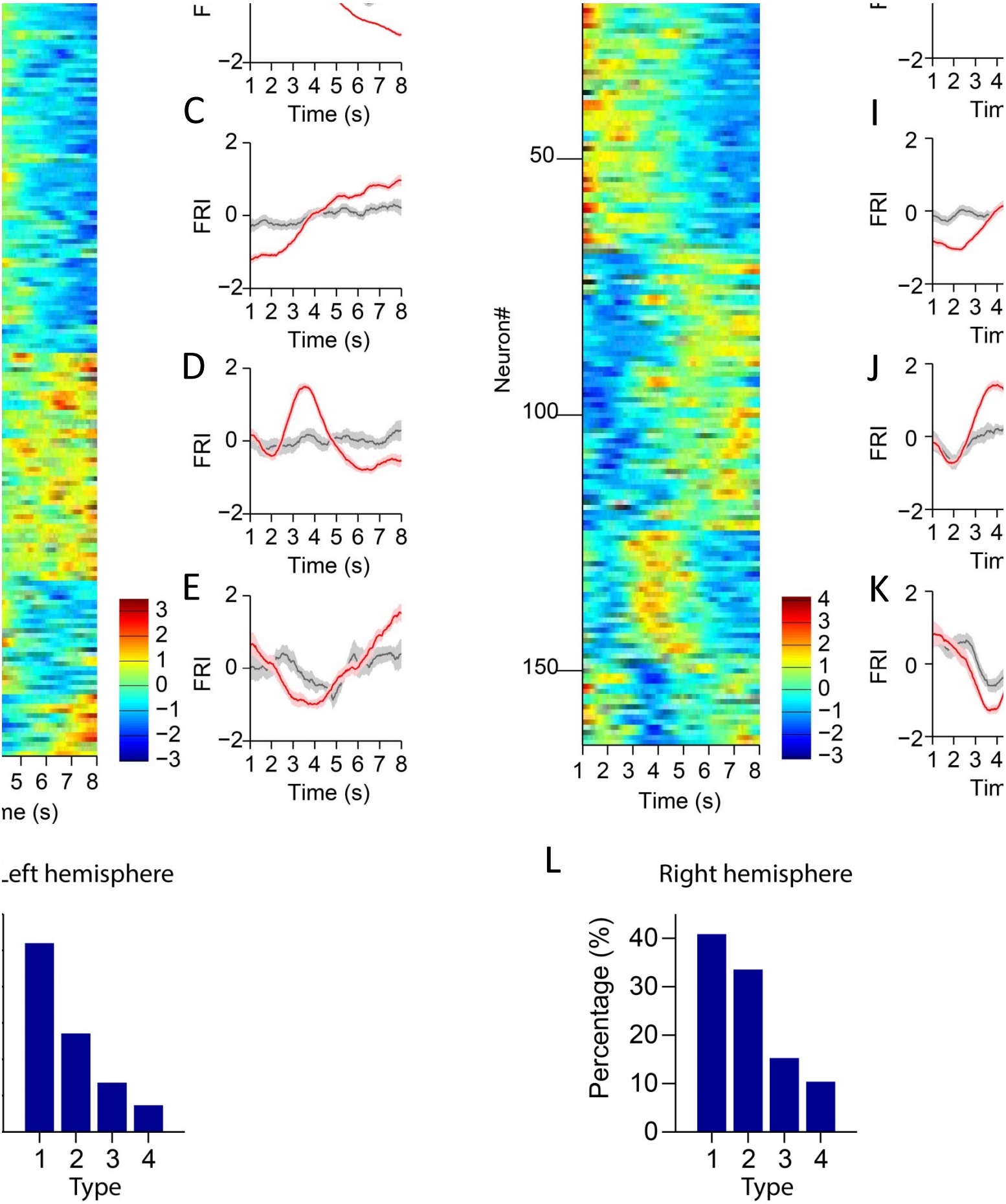
SNr neuron activities in left and right hemisphere. (A) Firing Rate Index (FRI) of neuronal activity for all task-related orrect 8-s trials recorded in the left hemisphere. The magnitude of FRI is color coded and the SNr neurons are classified es based on the activity dynamics. (B-E) Averaged FRI for Type 1 (B), Type 2 (C), Type 3 (D), Type 4 (E) of SNr neuron and incorrect 8-s trials (gray). (F) The proportion of four types of SNr neurons. Type 1 (57/108, 52.8%), Type 2 (36/108



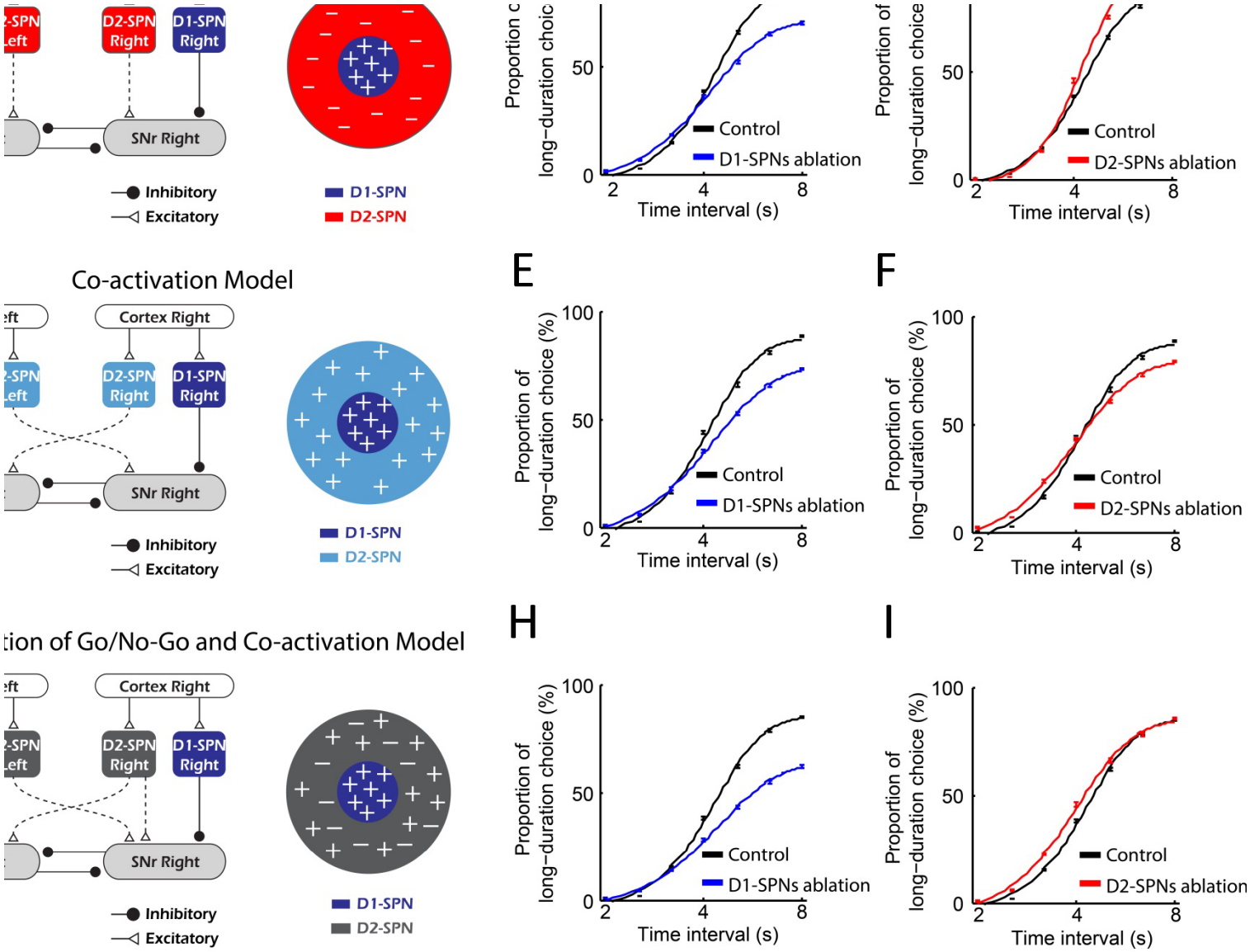
ehavioral statistics and neuronal dynamics of SNr neurons in the standard and reversed 2-8s tasks . (A) Correct rat
e both in the standard and reversed 2-8s tasks (n = 6 mice, paired t-test, $p = 0.33$). (B) Lever press ratios of the same grou
rd and reversed 2-8s tasks (n = 6 mice, paired t-test, $p < 0.05$). (C) Averaged FRI of the SNr Type 3 neurons in correct 8-s
s task. (D) Averaged FRI of the SNr Type 3 neurons in correct 8-s trials of the reversed 2-8s task. (E) Averaged FRI of the
orrect 8-s trials of the standard 2-8s task. (F) Averaged FRI of the SNr Type 4 neurons in correct 8-s trials of the reversed



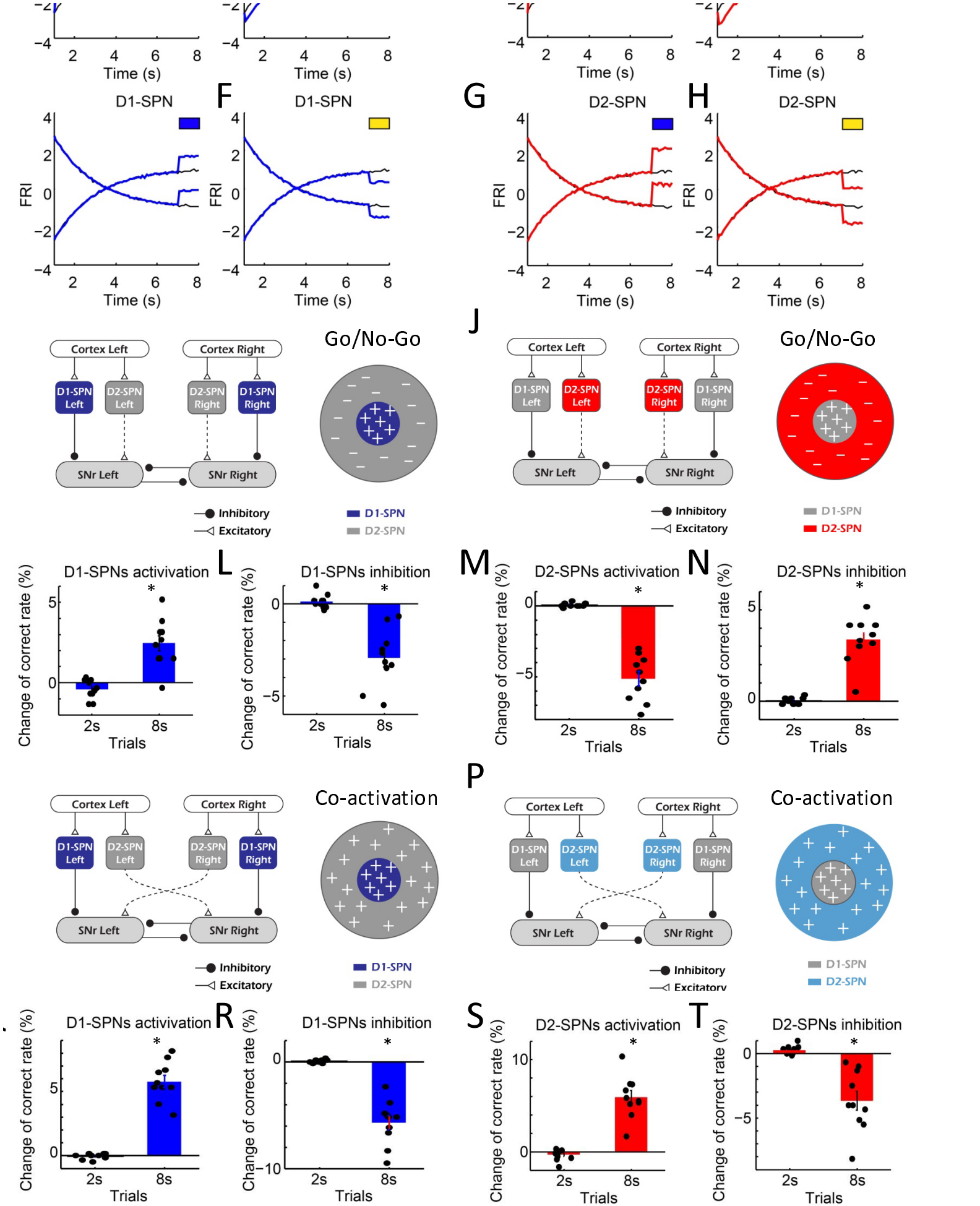
iatum neuronal recording on day 1 of training, recording array and optic fiber placement validation. (A) Averaged 3 and Type 4 of SPNs in correct (red) and incorrect 2-s trials (gray). (B) Firing Rate Index (FRI) of neuronal activity for correct 8-s trials on day 1 of training. The magnitude of FRI is color coded and the SPNs are categorized as four subgro amics. (C-F) Averaged FRI of SPNs in correct (red) and incorrect 8-s trials (gray). (G) Recording array affixed with a ca 1-Ai32 or A2a-Ai32 mice. Light emitted by optic fiber placed through the attached cannula is in close proximity to the ti



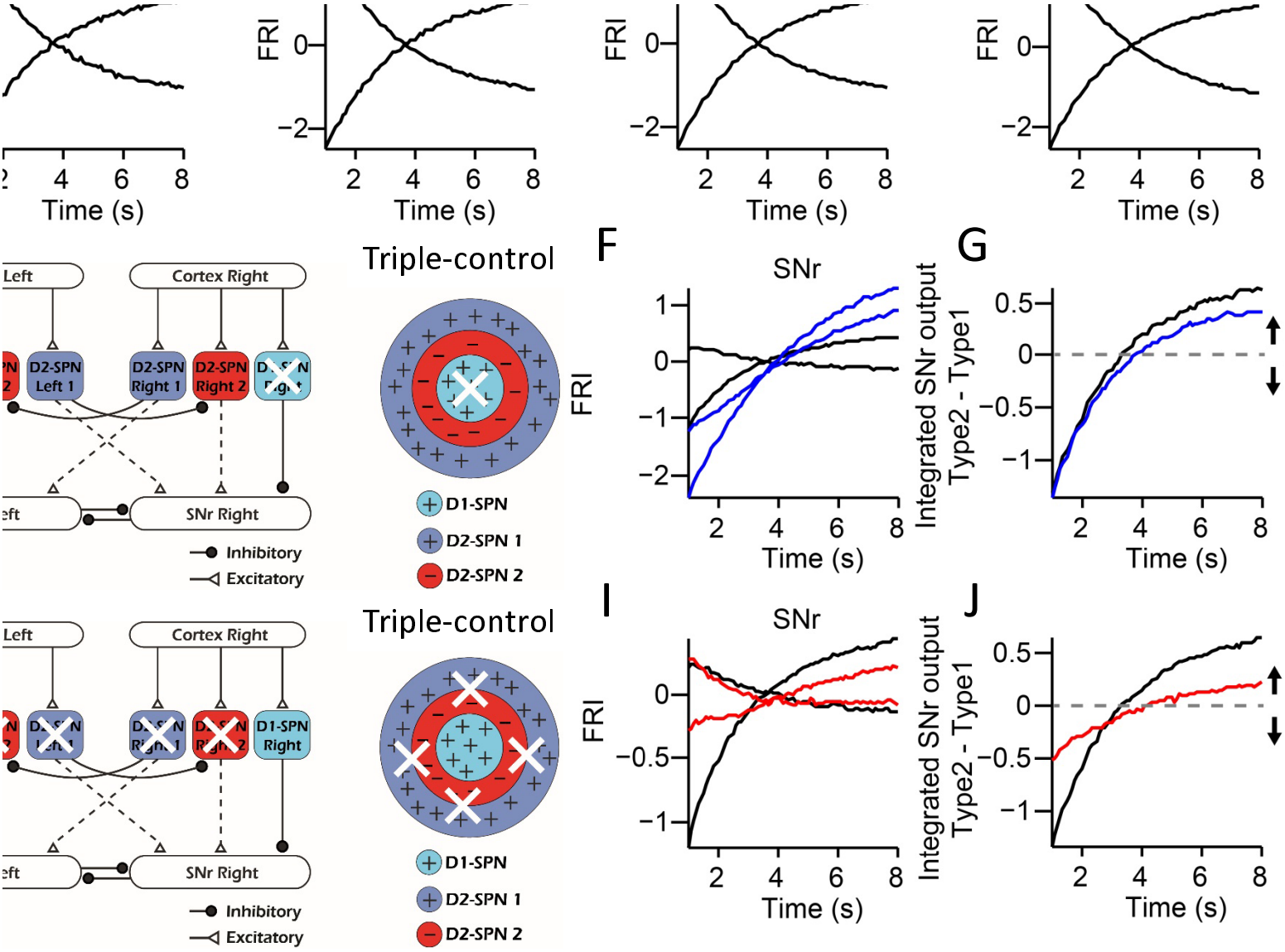
. **Striatal projection neuron activities in left and right hemisphere.** (A) Firing Rate Index (FRI) of neuronal activity fo
Ns in correct 8-s trials recorded in the left hemisphere. The magnitude of FRI is color coded and the SPNs are classified a
ypes based on the activity dynamics. (B-E) Averaged FRI for Type 1 (B), Type 2 (C), Type 3 (D), Type 4 (E) of SPNs in c
ect 8-s trials (gray). (F) The proportion of four types of SPNs. Type 1 (92/177, 52.0%), Type 2 (48/177, 27.1%), Type 3 (2
pe 4 (12/177, 7.2%). (G) Firing Rate Index (FRI) of neuronal activity for all task-related SPNs in correct 8-s trials recorr



nulation of lesion experiments in Go/No-Go, Co-activation and combination models. (A) Diagram of Go/No-Go mo
curves of behavior outputs simulated by Go/No-Go model in control (black) and D1-SPNs ablation condition (blue). (C)
curves of behavior outputs simulated by Go/No-Go model in control (black) and D2-SPNs ablation condition (red). (D) I
model. (E) The psychometric curves of behavior outputs simulated by Co-activation model in control (black) and D1-SP
e). (F) The psychometric curves of behavior outputs simulated by Co-activation model in control (black) and D2-SPNs al
). (G) Diagram of combination of Go/No-Go and Co-activation model. (H) The psychometric curves of behavior outputs
lel in control (black) and D1-SPNs ablation condition (blue). (I) The psychometric curves of behavior outputs simulated
ol (black) and D2-SPNs ablation condition (red).



imulation of optogenetic manipulation in Go/No-Go and Co-activation models. (A, B) Simulating optogenetic activa
) of D1-SPNs at 2s. Blue bar above indicates optogenetic activation. Yellow bar above indicates optogenetic inhibition. (
ptogenetic activation (C) and inhibition (D) of D2-SPNs at 2s. (E, F) Simulating optogenetic activation (E) and inhibitio



he neuronal activities in the "Triple-control" model and simulation of lesion experiments. (A) The simulated neuron
FRI for the cortical neurons in 8s trials. (B) The simulated neuronal dynamics quantified as FRI for the D1-SPN in 8s tr
uronal dynamics quantified as FRI for the D2-SPN 1 in 8s trials. (D) The simulated neuronal dynamics quantified as FRI
trials. (E) Schematic of selective ablation of D1-SPNs in the "Triple-control" model. (F) The model's Type 1 and Type 2
ition (black) and under D1-SPNs ablation (blue). (G) The subtraction of FRI between Type 1 and Type 2 SNr neurons in
s ablation condition (blue). (H) Schematic of selective ablation of D2-SPNs in the "Triple-control" model. (I) The model
FRI in control condition (black) and under D2-SPNs ablation (red). (J) The subtraction of FRI between Type 1 and Type
lack) and D2-SPNs ablation condition (red).

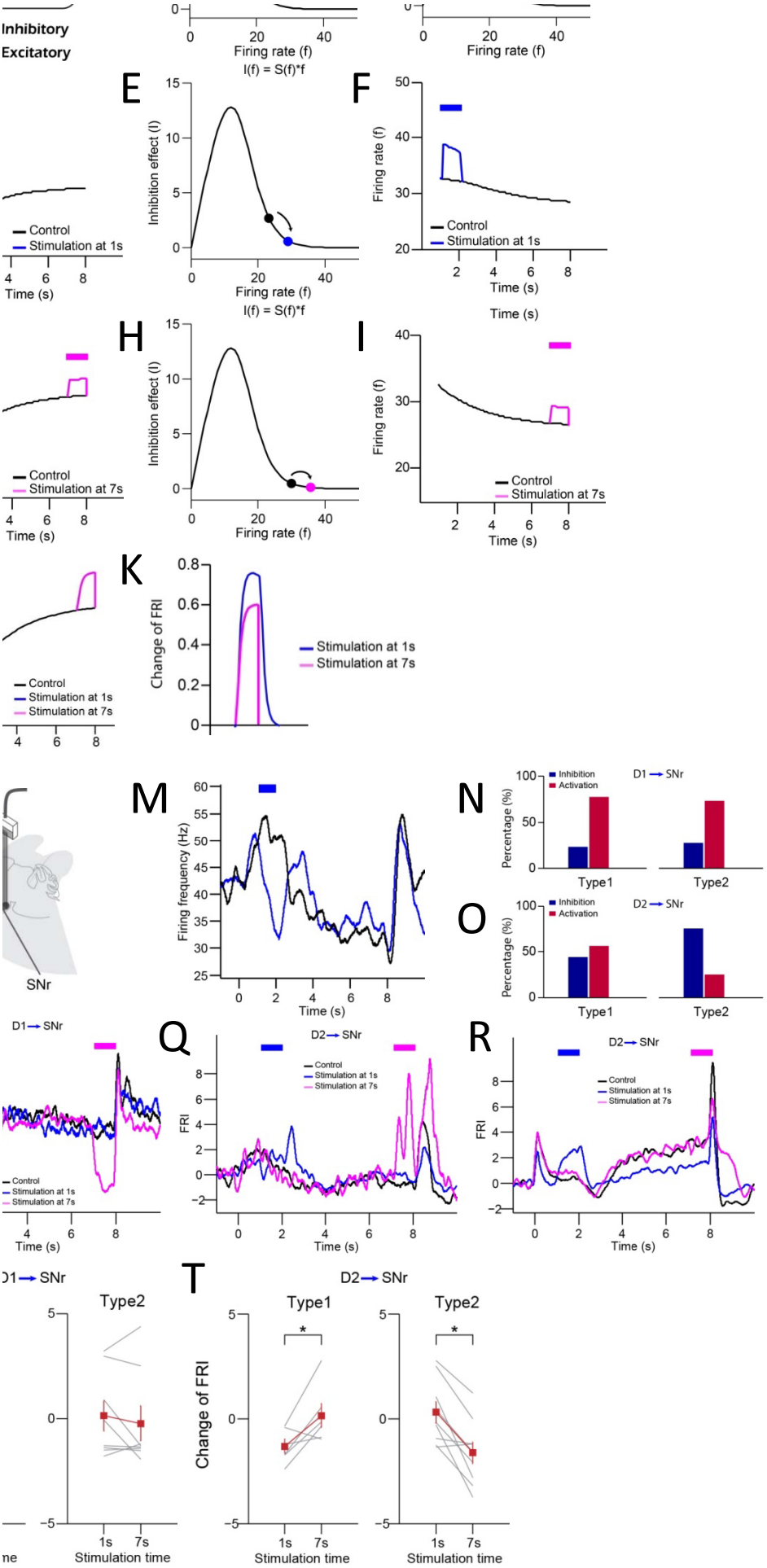
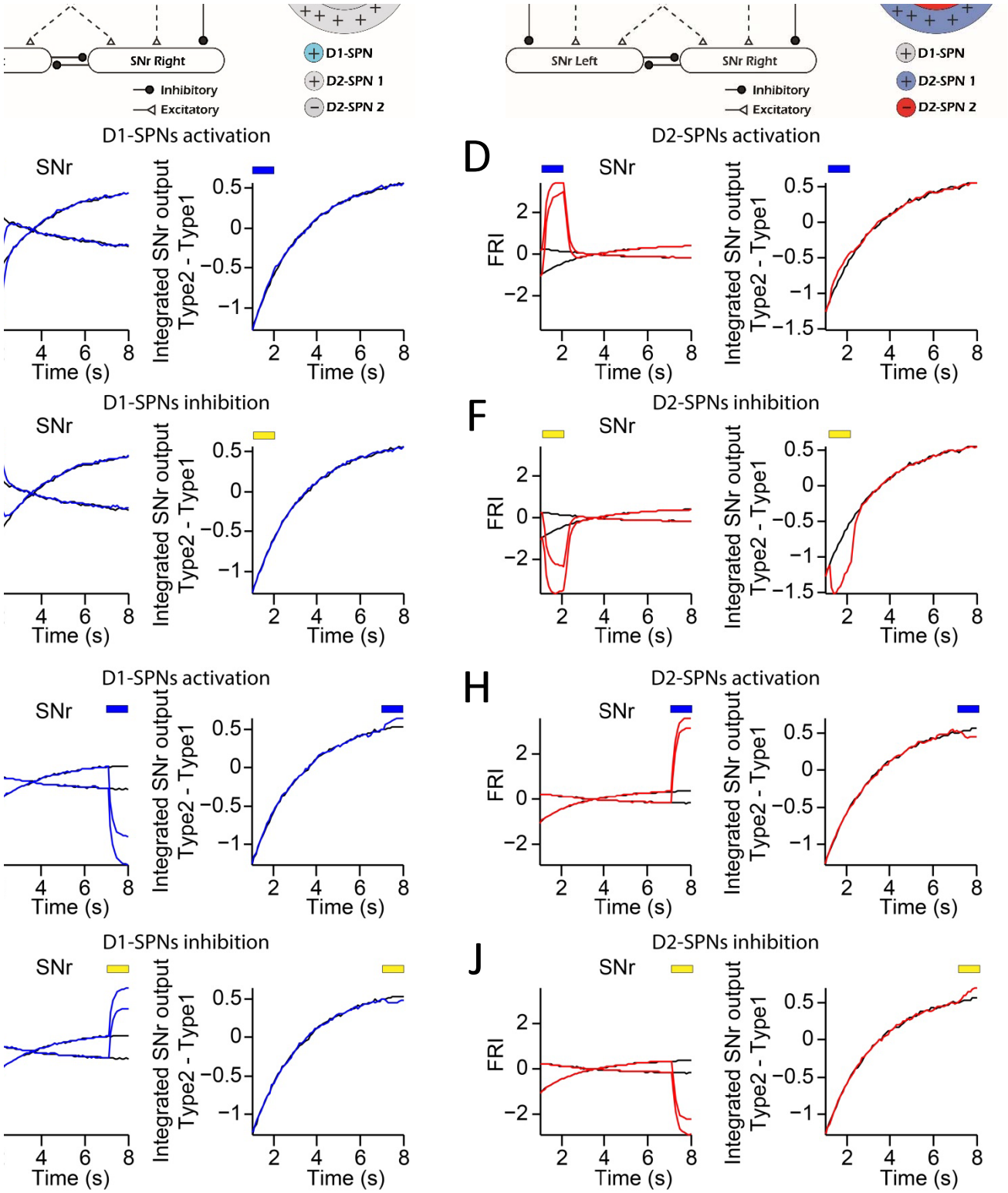
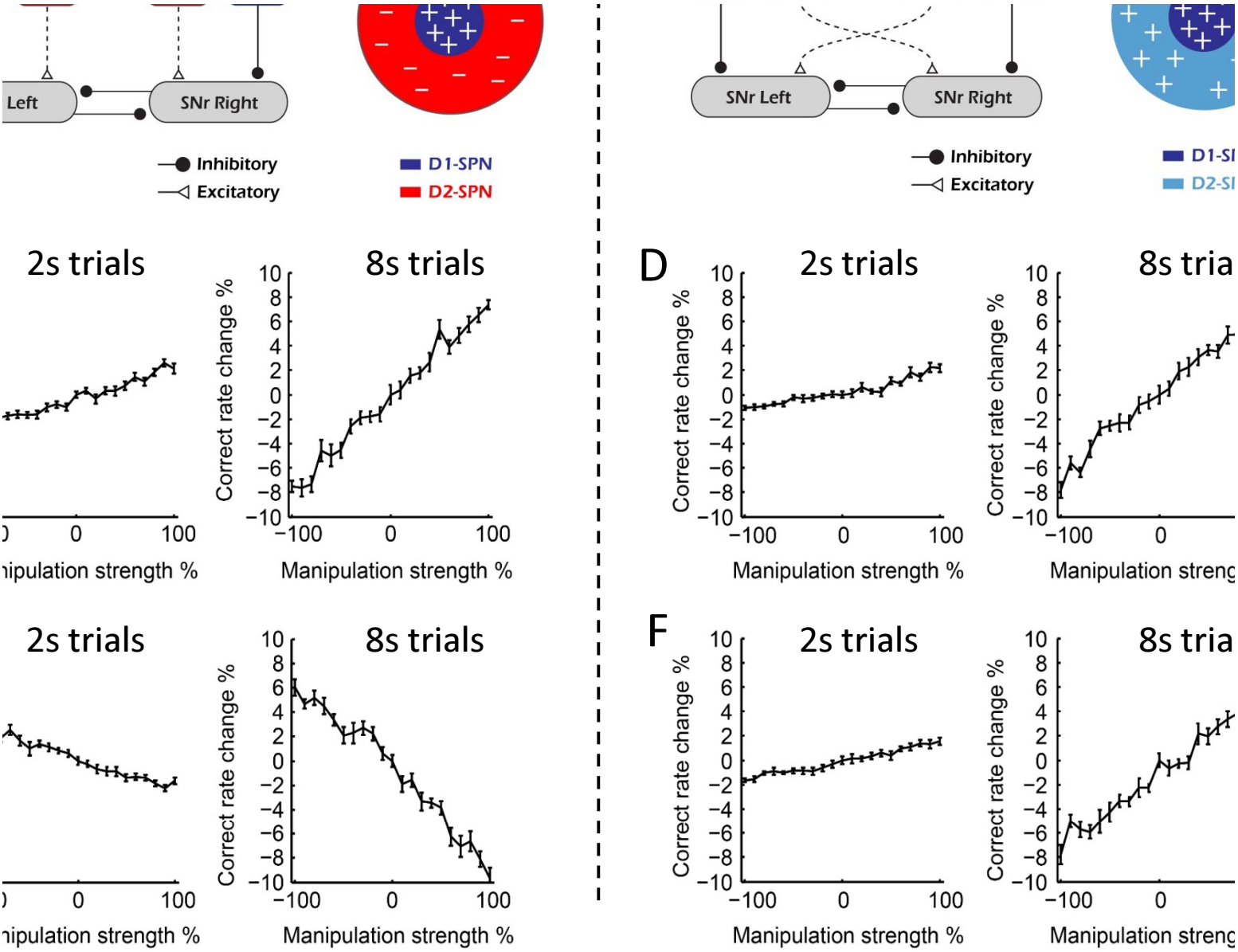


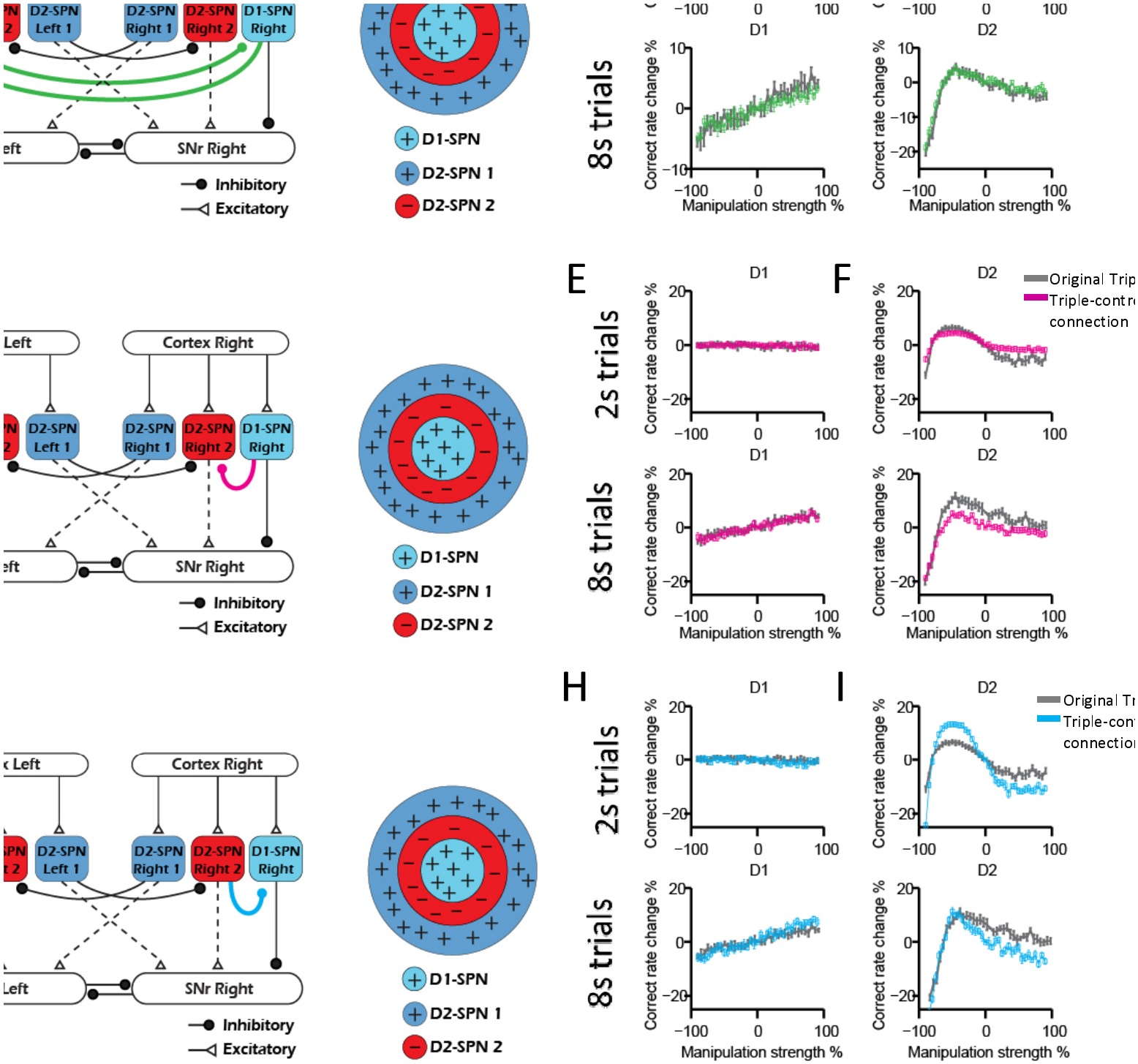
Figure S10. Optogenetic activation of D1- v differently regulates SNr activities in mode experiments. (A) A computational motif of in with collateral inhibitory synapse D2-SPN 1– collateral synapse between D2-SPNs exhibits depression. (B) Relationship between synapti SPN 1→ D2-SPN 2 and the firing rate of D2- inhibition effect of the collateral synapse betw (D) Activation of presynaptic D2-SPN 1 at 1s inhibition effect D2-SPN 1→ D2-SPN 2 syna activating D2-SPN 1 at 1s. (F) Activation of D (G) Activation of presynaptic D2-SPN 1 at 7s inhibition effect D2-SPN 1→ D2-SPN 2 syna activating D2-SPN 1 at 7s. (I) Activation of D (J) SNr neuron activities responding to activat at 1s (blue) and 7s (purple). (K) Comparison c SNr caused by activation of D2-SPNs at 1s an Schematic of simultaneous optogenetic excita D2-SPNs in the dorsal striatum and recording action selection. (M) Averaged neuronal activ example SNr Type 1 neuron responding to op activation of D1-SPNs at 1s during 8-s trials. percentage of SNr Type 1 (left) and Type 2 (ri showing excitation (blue) and inhibition (red) stimulating D1-SPNs. (O) The percentage of and Type 2 (right) neurons showing excitation inhibition (red) when stimulating D2-SPNs. (I neuronal activities of an example SNr Type 2 responding to optogenetic activation of D1-SI and 8s (purple) during 8-s trials. (Q, R) Avera activities of SNr Type 1 (Q) and Type 2 (R) n to optogenetic activation of D2-SPNs at 1s (b (purple) during 8-s trials. (S) Comparison of F SNr Type 1 (left) and Type 2 (right) neurons c optogenetic activation of D1-SPNs at 1s and 7 Comparison of FRI changes in SNr Type 1 (le (right) neurons caused by optogenetic activati at 1s and 7s (paired t-test, $p < 0.05$).



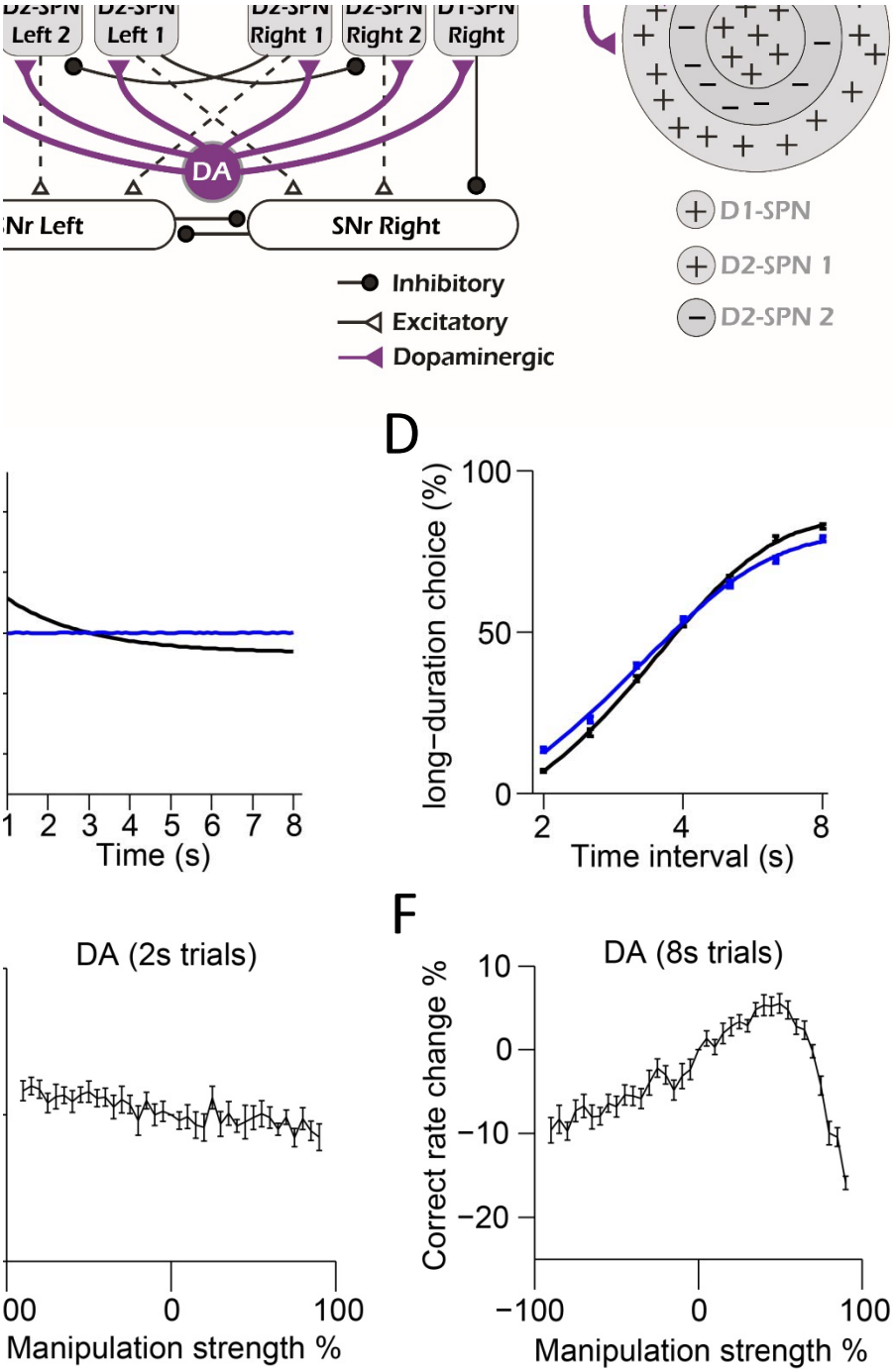
omputational modeling of optogenetic manipulation reveals that D1- vs. D2-SPNs differently regulates SNr output ol" model. (A, B) Schematic for optogenetic manipulation of D1-SPNs (A) and D2-SPNs (B) in the 'Triple-control' mod euronal dynamics of SNr Type 1/Type 2 (left panel) and integrated output (right panel) under control (black) and activati . (D) Modeling of neuronal dynamics of SNr Type 1/Type 2 (left panel) and integrated output (right panel) under control) of D2-SPNs at 2s. (E) Modeling of neuronal dynamics of SNr Type 1/Type 2 (left panel) and integrated output (right pa and inhibition (blue) of D1-SPNs at 2s. (F) Modeling of neuronal dynamics of SNr Type 1/Type 2 (left panel) and integ nder control (black) and inhibition (red) of D2-SPNs at 2s. (G) Modeling of neuronal dynamics of SNr Type 1/Type 2 (le ut (right panel) under control (black) and activation (blue) of D1-SPNs at 8s. (H) Modeling of neuronal dynamics of SN nd integrated output (right panel) under control (black) and activation (red) of D2-SPNs at 8s. (I) Modeling of neuronal pe 2 (left panel) and integrated output (right panel) under control (black) and inhibition (blue) of D1-SPNs at 8s. (J) Mod mics of SNr Type 1/Type 2 (left panel) and integrated output (right panel) under control (black) and inhibition (red) of D



Computational modeling of manipulation reveals that Go/No-Go and Co-activation model differently predicts the
A) Diagram of Go/No-Go model. (B) Diagram of Co-activation model. (C) Correct rate change in 2s (left panel) and 8s tr
manipulating D1-SPNs in Go/No-Go model with different manipulation strengths. (D) Correct rate change in 2s (left par
trials when manipulating D1-SPNs in Co-activation model with different manipulation strengths. (E) Correct rate change
s trials (right panel) when manipulating D2-SPNs in Go/No-Go model with different manipulation strengths. (F) Correct
) and 8s trials (right panel) trials when manipulating D2-SPNs in Co-activation model with different manipulation streng



omputational modeling reveals that the linear and nonlinear modulation of action selection by direct versus indire
nold with additional striatal collateral connections. (A) Schematic for 'Triple-control' model with D1-D1 collateral co
nange in 2s trials (upper panel) and 8s trials (bottom panel) when manipulating D1-SPNs with different manipulation stre
ted-measures ANOVA, effect of manipulation strength, 2s trials: $F_{40,369} = 1.328$, $p = 0.0945$; 8s trials: $F_{40,369} = 7.595$, $p <$
Triple-control' model with D1-D1 collateral connections. Gray lines: the same simulation results as shown in Figure 7(C,
nange in 2s trials (upper panel) and 8s trials (bottom panel) when manipulating D2-SPNs with different manipulation stre
ted-measures ANOVA, effect of manipulation strength, 2s trials: $F_{40,369} = 38.22$, $p < 0.0001$; 8s trials: $F_{40,369} = 34.29$, $p <$
Triple-control' model with D1-D1 collateral connections. Gray lines: the same simulation results as shown in Figure 7(D,
'Triple-control' model with D1-D2 collateral connections. (E) Correct rate change in 2s trials (upper panel) and 8s trials
nanipulating D1-SPNs with different manipulation strengths (n = 10, one-way repeated-measures ANOVA, effect of mani
als: $F_{40,369} = 0.9335$, $p = 0.5893$; 8s trials: $F_{40,369} = 8.778$, $p < 0.0001$). Purple lines: 'Triple-control' model with D1-D2 co
Gray lines: the same simulation results as shown in Figure 7(C, G). (F) Correct rate change in 2s trials (upper panel) and 8
) when manipulating D2-SPNs with different manipulation strengths (n = 10, one-way repeated-measures ANOVA, effect
strength, 2s trials: $F_{40,369} = 40.94$, $p < 0.0001$; 8s trials: $F_{40,369} = 26.61$, $p < 0.0001$). Purple lines: 'Triple-control' model wi
nections. Gray lines: the same simulation results as shown in Figure 7(D, H). (G) Schematic for 'Triple-control' model wi
nections. (H) Correct rate change in 2s trials (upper panel) and 8s trials (bottom panel) when manipulating D1-SPNs with



omputational modeling of dopaminergic modulation in the "Triple-control" model. (A) Diagram of Triple-control m modulation on SPNs. (B) Schematic of center-surround-context receptive field diagram with dopaminergic modulation a ' model. '+' indicates facilitating effect to selection. '-' indicates inhibitory effect to selection. (C) Simulation of two typ amics (black: decreasing dopamine; blue: constant dopamine with no change) in 8s trials. (D) Psychometric curves corre e dynamics (n = 10, two-way repeated-measures ANOVA, main effect of ablation, $F_{1,18}$ = 0.8743, p =0.362; interaction b blation, $F_{6,108}$ = 8.261, p < 0.0001). (E, F) Correct rate change in 2s (E) and 8s trials (F) trials when manipulating dopam with different manipulation strengths (n = 10, one-way repeated-measures ANOVA, effect of manipulation strength, 2-s 0001; 8-s trials: $F_{36,324}$ = 39.98, p < 0.0001).