Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors

Nick G. Hollon[1], **Elora W. Williams**[1], **Christopher D. Howard**[1,2], **Hao Li**[1], **Tavish I. Traut**[1], **Xin Jin**[1,3,4,5,*]

[1]Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

[2]Current address: Neuroscience Department, Oberlin College, Oberlin, OH 44074, USA

[3]Center for Motor Control and Disease, Key Laboratory of Brain Functional Genomics, East China Normal University, Shanghai 200062, China

[4]NYU–ECNU Institute of Brain and Cognitive Science, New York University Shanghai, Shanghai 200062, China

[5]Lead Contact

SUMMARY

Dopamine has been suggested to encode cue-reward prediction errors during Pavlovian conditioning, signaling discrepancies between actual versus expected reward predicted by the cues[1–5]. While this theory has been widely applied to reinforcement learning concerning instrumental actions, whether dopamine represents action-outcome prediction errors and how it controls sequential behavior remain largely unknown. Indeed, the vast majority of previous studies examining dopamine responses primarily have used discrete reward-predictive stimuli[1–15], whether Pavlovian conditioned stimuli for which no action is required to earn reward, or explicit discriminative stimuli that essentially instruct an animal how and when to respond for reward. Here, by training mice to perform optogenetic intracranial self-stimulation, we examined how self-initiated goal-directed behavior influences nigrostriatal dopamine transmission during single as well as sequential instrumental actions, in behavioral contexts with minimal overt changes in the animal's external environment. We found that dopamine release evoked by direct optogenetic stimulation was dramatically reduced when delivered as the consequence of the animal's own action, relative to non-contingent passive stimulation. This dopamine suppression generalized to food rewards, was specific to the reinforced action, was temporally restricted to counteract

the expected outcome, and exhibited sequence-selectivity consistent with hierarchical control of sequential behavior. Together these findings demonstrate that nigrostriatal dopamine signals sequence-specific prediction errors in action-outcome associations, with fundamental implications for reinforcement learning and instrumental behavior in health and disease.

eTOC BLURB

Dopamine signals prediction errors in cued-reward contexts. Here, Hollon et al. show that nigrostriatal dopamine is inhibited when a reward outcome is the expected consequence of self-initiated action. These action-outcome prediction errors suppress even optogenetically stimulated dopamine and exhibit hierarchical control during action sequences.

Keywords

Nigrostriatal dopamine; action-outcome; prediction error; action sequence; efference copy

RESULTS

Suppression of optogenetically stimulated nigrostriatal dopamine by goal-directed action

Mice expressing channelrhodopsin-2 selectively in their dopamine neurons[16,17] (see Methods) were implanted with a fiber optic over the substantia nigra pars compacta (SNc) for optogenetic stimulation[18] and a carbon-fiber microelectrode[19] in the ipsilateral dorsal striatum to record nigrostriatal dopamine transmission using fast-scan cyclic voltammetry (FSCV; Figures 1A and S1A–B). The mice were trained in a free-operant optogenetic intracranial self-stimulation (opto-ICSS) task (Figure 1B), in which they learned to press a continuously reinforced "Active" lever to optogenetically stimulate their own dopamine neurons (50 Hz for 1 s) and rarely pressed the non-reinforced "Inactive" lever yielding no outcome (Figure 1C). Therefore, consistent with other recent reports[20–23], selective stimulation of SNc dopamine neurons is sufficient to reinforce novel actions.

To examine the extent to which this behavior is indeed goal-directed, a subset of mice underwent a contingency degradation test[24–27]. During this test phase, stimulation was decoupled from the lever-pressing action and instead delivered non-contingently at a rate yoked to that animal's own stimulation rate from the preceding self-stimulation phase (Methods). The mice significantly reduced their performance rate (Figures 1D and S1C–D), indicating that they readily learned that their action was no longer required to earn stimulation. This demonstrates that nigrostriatal dopamine neuron self-stimulation under this simple fixed-ratio schedule of continuous reinforcement (CRF) is sensitive to changes in the action-outcome contingency, which is an established operational hallmark of goal-directed behavior[28,29].

To investigate whether goal-directed action affects the nigrostriatal dopamine response to the consequence of that action, we used FSCV to record subsecond dopamine transmission in behaving mice during a session that included two phases: In the Self-Stimulation phase, as in prior opto-ICSS training, mice earned optogenetic stimulation for each Active lever press (Figure 1B). In the subsequent Passive Playback phase, the levers were retracted, and

the mice received non-contingent stimulations, with timestamps identical to the stimulations that each individual had self-administered in its Self-Stimulation phase. Thus, in this entirely within-subject design, we recorded at the same striatal location with the same chronically implanted electrode, with each animal yoked to its own performance, receiving the same temporal sequence of stimulations across both phases of the session, delivered to the same site within the SNc using identical optogenetic stimulation parameters to directly depolarize these nigrostriatal dopamine neurons (Figures 1A and 1E–F).

We observed a remarkably robust difference between the amplitude of Self-Stimulated dopamine release and the significantly greater amplitude evoked by the non-contingent Passive Playback stimulation (Figure 1F–I). All individual mice (9/9, 100%) exhibited less dopamine release when evoked as the consequence of their own action; this difference was significant at the individual level in 7 of 9 mice (Ps < 0.0001) and was a trend in the same direction for the remaining 2 mice (Ps = 0.0623 and 0.0825). The latency for the onset of the dopamine response to exceed baseline was significantly longer for Self-Stimulation than for Playback (Figure S1E), whereas the latency to the respective peaks did not significantly differ (Figure S1F), consistent with the transient suppression observed in the Self-vs.-Playback Difference trace (Figure 1H).

Although the free-operant opto-ICSS task was designed to minimize discrete external cues, it nevertheless is possible that the offset of a previous stimulation essentially could serve as a stimulus that might elicit the next lever-pressing response. However, when we isolated the initiation of lever-pressing bouts using an inter-stimulation interval (ISI) criterion of at least 10 s since the previous stimulation, this subset of stimulations still showed a significant difference between Self-Stimulated and non-contingent Playback-evoked dopamine release (Figure S1G–I). Self-stimulations following shorter ISIs evoked less dopamine release than did those with long ISIs (Figure S1J–K), consistent with a recent investigation of mesolimbic dopamine[30]. However, a similar effect of ISI was observed for the non-contingent Playback stimulations, such that the net Self-vs.-Playback stimulation difference was not significantly related to the ISI (Figure S1J–M), highlighting the importance of temporally matched stimulations in the current experimental design. The Self-vs.-Playback difference also was stable across early and late stimulations within this recorded session (Figure S1N–Q), consistent with the mice being well-trained by the time of this recording (mean ± SEM = 11.1 ± 1.2 prior training days).

Using fiber photometry to record the red fluorescent dopamine sensor $rGRAB_{DA1h}$[31] (Figures 1J and S1R–U), we recapitulated the suppression of self-stimulated dopamine release in the dorsomedial striatum (DMS; Figure 1K–M) as observed in our FSCV recordings. This effect generalized to natural reward outcomes, as the $rGRAB_{DA1h}$ response was lower when retrieving self-administered sucrose pellets than when retrieving non-contingent 'Playback' pellets (Figure 1N–Q). We also found similar suppression of the optogenetically evoked $rGRAB_{DA1h}$ response in the dorsolateral striatum (DLS; Figure S1V–Y), suggesting that this main effect generalizes across these recording modalities, types of reinforcer, and striatal subregions. Collectively, these findings indicate that reward-evoked dopamine release is lower when it is the expected outcome of self-initiated, goal-directed actions.

Nigrostriatal dopamine signals action-outcome prediction errors

The reward prediction error theory implies decreased dopamine responses to expected versus unexpected outcomes[1–5,32–34]. Nevertheless, the relative difference we observed does not alone resolve whether dopamine release is in fact inhibited by the animal's action. To address this question, we recorded additional opto-ICSS FSCV sessions in which a random 20% of Active lever presses did not yield stimulation, instead causing a 5-s timeout period during which no further stimulation could be earned (Figure 2A). During these Omission Probes, there was a clear dip in dopamine below baseline levels (Figure 2B–C), consistent with a neurochemical instantiation of a negative prediction error[15]. Indeed, the timecourse for this Omission Probe dip was remarkably similar to the digital subtraction ("Difference Trace") of the Self-Stimulated dopamine response minus the Passive Playback response (Figure 1H; overlaid in Figure 2D), and there was a significant correlation between the amplitude of the Omission Probe dip and the Self-vs.-Playback Difference (Figure S2A). This Omission Probe dip was not merely an artifact of FSCV background subtraction[35], where reuptake during the stimulation-free timeout period might follow an elevated baseline from several preceding stimulations. Rather, a significant dip below baseline was still prominent for the subset of Omission Probes with a minimum latency of at least 5 s since the previous stimulation, whereas no such decrease was detected at the equivalent time points from the Playback phase (Figure S2B–C). Furthermore, additional lever presses during an ongoing stimulation augmented the suppression of Self-Stimulated dopamine release, and similarly, additional presses during an Omission Probe timeout period prolonged the duration of the dip below baseline (Figure S2D–G). *In vivo* extracellular electrophysiological recording further revealed reduced somatic firing in an optogenetically-identified SNc dopamine neuron in response to action-evoked optogenetic Self-Stimulation relative to non-contingent Passive Playback stimulation (Figure S2H–M). Collectively, these results demonstrate that the action indeed causes inhibition of dopamine transmission.

It recently has been reported that some dopamine neurons transiently reduce their firing rate during certain types of spontaneous movement[8,36,37]. We therefore considered the possibility that the action-induced suppression observed in our recordings may be a generalized inhibition following any lever pressing action, regardless of whether that action is associated with a particular reinforcing outcome. However, we found no such inhibition in the instances when the animal pressed the Inactive lever, which had never been reinforced throughout training (Figure 2D–E), indicating that the action-induced inhibition of dopamine release is specific to the typically reinforced action and conveys a bona fide prediction-error signal mediated by expectation.

To examine the temporal specificity of this action-induced suppression, we recorded Delay Probe sessions in which 20% of Active lever presses instead resulted in stimulation that was delayed by 5 s (Figure 2F). The initial 5 s of this delay period was equivalent to the timeout period of the Omission Probes, and we again observed a dip in dopamine below baseline (Figure 2G–H). When the probe stimulation was finally delivered at the end of the delay period, there now was a high amplitude of dopamine release that did not differ from the corresponding Playback stimulations (Figure 2I–J). Because these Delay Probes were randomly interleaved throughout the Self-Stimulation phase, this indicates that there

is not a global suppression of dopamine neuron excitability throughout the whole context of the Self-Stimulation phase. Rather, this action-induced inhibition is precisely timed to counteract the expected consequence of that action, namely the immediate stimulation that is its typical outcome.

We further determined the nature of this inhibition in Magnitude Probe sessions, where 20% of Active lever presses yielded 5 s of stimulation rather than the standard 1-s stimulation used throughout training (Figure 2K). These increased Magnitude Probes indeed evoked much greater dopamine release, as expected for longer-duration stimulation (Figure 2L–M). Closer examination of the time course of these dopamine responses also revealed a transient suppression during the self-stimulated Magnitude Probes that was restricted to the first second or so following stimulation onset, but no longer differed from Probe Playback by the end of the 5-s probe stimulation. This brief inhibition also was borne out by the transient dips and similar overall time courses in the Difference traces for both the Magnitude Probes and the standard 1-s stimulations, comparing each type of Self-Stimulation to their respective Playback stimulations (Figure 2N–O). This again highlights the timing and duration specificity of the action-induced suppression, and suggests that there is not a global inhibition of dopamine throughout the Self-Stimulation context. Together, these data suggest that nigrostriatal dopamine can encode a reward prediction error signal for individual goal-directed action and its expected outcome.

Sequence-specific suppression of nigrostriatal dopamine release

In real life, goals are seldom achieved by a single action but instead mostly through a series of actions organized in spatiotemporal sequences[38–41]. Having established that the observed prediction error-like suppression of nigrostriatal dopamine is temporally restricted and specific to an action associated with a reinforcing outcome, we next turned to the question of whether such regulation of dopamine transmission reflects hierarchical control over learned action sequences[38,41,42]. To this end, we trained a separate cohort of mice to perform a spatiotemporally heterogeneous action sequence, pressing the Left and then Right lever (LR) to earn optogenetic nigrostriatal dopamine neuron self-stimulation (See Methods; Figures 3A and S3A–B). As mice increased the number of stimulations earned across days of training (Figure 3B), their behavior exhibited several indications of successfully learning this LR action sequence: They increased both their probability of correctly completing a sequence by transitioning to a Right lever press following each Left lever press, and their probability of reinitiating with a Left lever press following each stimulation (Figure 3C). Their duration to complete these LR sequences was shorter than the post-reinforcer reinitiation latency (Figure 3D), and the proportion of correct LR sequences increased relative to other non-reinforced press pairs (Figure 3E). The total presses per sequence and the number of consecutive presses on either lever both decreased throughout training, collectively contributing to an increase in overall efficiency (Figure S3C–F). Therefore, rather than simply associating the reinforcing outcome with the most proximal action at the Right lever, the animals' behavior suggested that they indeed concatenated the distinct action elements into chunked action sequences. Furthermore, the mice significantly reduced their LR sequence performance during a contingency degradation test (Figure S3G), indicating that these chunked action sequences also were goal-directed.

We then recorded nigrostriatal dopamine transmission with FSCV in these sequence-trained mice using the same within-subject manipulation comparing Self-Stimulation versus Passive Playback-evoked dopamine responses. We again found a robust suppression of the Self-Stimulated dopamine response (Figures 3F–I and S3H–I), recapitulating the main result from the single-lever CRF cohort (Figure 1E–I). We also examined dopamine transmission aligned to the completion of other combinations of non-reinforced press pairs (Figure 3J–K), essentially the multi-press analogues of the non-reinforced Inactive Lever presses from the single-lever CRF cohort (Figure 2D–E). Importantly, there was no significant inhibition to any combination of non-reinforced press pairs, in stark contrast to the strong suppression of dopamine revealed in the Difference between LR Self-Stimulation versus Playback (Figure 3J–K). Therefore, analogous to the action-specificity observed in the single-lever CRF cohort, these results indicate that this inhibition was specific to the learned action sequence associated with the expected reward outcome.

Differential regulation of dopamine by individual actions within learned sequence

Beyond this sequence-type specificity, we further examined the question of whether dopamine transmission might reflect regulation at the level of individual action elements or instead at a higher sequence level in a hierarchy of behavioral control. For example, if regulated with each action element, we might expect similar inhibition for each individual Left and Right lever press, and summation of each to the full inhibition at outcome delivery. Alternatively, since animals chunked these action elements into fully concatenated action sequences, we might expect the action-induced inhibition of dopamine to begin at sequence initiation and persist throughout performance of this chunked action sequence. The results were inconsistent with either of these hypotheses, instead exhibiting a distinct form of sequence-specificity consistent with hierarchical control[41,43]. Initiating Left lever presses did not cause any inhibition of dopamine, instead revealing a slight, albeit non-significant increase in dopamine release (Figures 4A–B and S4). Similarly, additional recording sessions with probe stimulations delivered on 20% of initiating Left presses revealed no inhibition of dopamine evoked by these Left Probes (Figure 4C–H). Instead, these Left Probes actually evoked significantly greater dopamine release than their Playback (Figure 4F). At the individual animal level, 5 out of 8 mice (62.5%) showed a significant dopamine increase, and none showed a significant suppression. Outside of LR sequences, single Right lever presses (see Methods) did not result in inhibition of dopamine, in stark contrast with the full inhibition of Self-Stimulated versus Playback-evoked dopamine for correct LR sequences (Figure 4I–J). The dopamine response to Probe stimulations for these isolated Right presses also did not differ from their Playback, exhibiting no significant inhibition (Figure 4K–P). Importantly, these presumably unexpected Right Probe stimulations evoked significantly greater dopamine release than did the standard Self-Stimulations, which resulted from the same proximal action of pressing the identical Right lever to complete a LR sequence (Figure 4K–L). This same action at the Right lever therefore reveals highly distinct regulation of dopamine dynamics depending on the action's membership within the learned sequence or not. These results indicate that the dopaminergic prediction errors are selective to the learned action sequence and reflect sequence-level hierarchical control over instrumental behavior.

DISCUSSION

Our brains constantly generate predictions about the world around us[44,45], particularly regarding the expected consequences of environmental cues or our own actions[46–49]. Indeed, the effects of such expectations have long been recognized when examining the phasic activity of midbrain dopamine neurons following reward-predictive stimuli[1]. Here, we have demonstrated that nigrostriatal dopamine transmission to reinforcing outcomes is strongly suppressed when this outcome is the expected consequence of the animal's own action. This inhibition of outcome-evoked dopamine following self-initiated actions parallels commonly observed reward prediction errors in explicit stimulus-outcome and stimulus-response behavioral contexts. The current results therefore expand this phenomenon to include action-outcome prediction errors that support instrumental associations underlying self-initiated goal-directed behavior. This action-outcome prediction error was specific to the typically reinforced action, temporally restricted to counteract the expected consequence of that action, and exhibited sequence selectivity consistent with a high level of hierarchical control over chunked action sequences. The prediction errors signaled by dopamine transmission therefore reflect not only expectations associated with Pavlovian cues or behavioral responses to such discrete stimuli, but also the expected outcomes of self-initiated instrumental actions and sequences. Compelling behavioral and neural evidence for action chunking also is well established[39–41,43,50–55], but the mechanisms subserving such sequence learning remain poorly understood. While the exact role of nigrostriatal dopamine throughout sequence acquisition requires further direct investigation, the current results demonstrate that the performance of well-learned action sequences entails distinct dopamine dynamics for actions within these sequences. That nigrostriatal dopamine transmits specific action-outcome prediction errors and exhibits sequence-dependent hierarchical regulation provides critical new insight into these important neuromodulatory dynamics in goal-directed behavioral control, an under-examined domain of instrumental action beyond spontaneous movement of unknown purpose and responding to reward-predictive cues.

Several aspects of our opto-ICSS experimental design conferred distinct advantages for examining the regulation of nigrostriatal dopamine dynamics in goal-directed behavior. The current study used an entirely within-subject design and direct optogenetic excitation to selectively stimulate dopamine neurons and record dopamine transmission at identical locations within a given animal, in contrast to previous ICSS studies that used non-selective electrical stimulation of the midbrain and compared dopamine release between trained versus naïve animals[56,57] or did not include temporally matched non-contingent playback[30,58,59]. Although selective optogenetic stimulation lacks the specific sensory features such as flavor that typically define the identity of natural reward outcomes[23,29,55,60–62], the direct intracranial delivery permitted precise temporal control over outcome receipt across the matched session phases. This obviated any potential complications that might arise in traditional procedures with natural rewards regarding the timing of when the animal detected and retrieved the outcome, particularly during the non-contingent Playback phase. Direct optogenetic stimulation also bypasses afferent circuitry representing the natural reward itself or its associated sensory features, permitting the current focus on regulation of dopamine by specific action-associated expectancies.

Nevertheless, we also observed a similar suppression of dopamine when mice made consummatory actions to retrieve self-administered sucrose pellets, suggesting that this phenomenon generalizes to natural reward as well. These features of the current design collectively yielded results consistent with nigrostriatal dopamine transmitting an action-outcome prediction error signal.

Although direct optogenetic stimulation indeed approaches an essentially identity-less outcome[63], this outcome delivery does coincide with sensory feedback during the action, such as somatosensory contact or auditory feedback from pressing the lever. However, these sensory reafferents are comparable for inactive lever presses or other non-reinforced action sequences, and therefore cannot account for the selective suppression of dopamine evoked as the consequence of reinforced actions (Figures 2D–E and 3J–K). Other modalities such as visual or proprioceptive feedback admittedly would differ between the spatially segregated levers or between session phases depending on the animal's spatial location and posture, but nevertheless remain direct consequences of the animal's own action rather than experimenter-controlled external cues. Indeed, the distinct regulation of dopamine to the same action depending on sequence membership (Figure 4I–P) provides clear evidence that the observed suppression was due to specific action-outcome expectancies rather than consequent sensory feedback from this proximal action. Whereas the suppression of outcome-evoked dopamine release is therefore unlikely accounted for by different sensory features between the session phases, this action-induced suppression may instead share important commonalities with efference copy (or corollary discharge) phenomena widely observed in numerous other sensorimotor systems throughout the nervous systems of many different species[46–49]. Indeed, the current results provide evidence that a learned, sequence-level efference copy can suppress the neurochemical consequence of the complete action sequence, distinct from the regulation by individual action elements. These findings align with the recent demonstration of dopaminergic prediction errors for evaluating sequential sensorimotor control relative to internal performance templates[64], and are broadly consistent with the prominent role proposed for efference copies in striatal-dependent learning[65,66].

The current study's FSCV recordings targeted mainly the DMS, which is widely implicated in goal-directed instrumental behavior[28,29,67–69]. Despite the preponderance of evidence that the DMS plays critical roles in the acquisition and performance of goal-directed action, we do note that one study found that lesions of DLS rather than DMS impaired sequence learning[70]. Therefore, natural next questions include whether regulation of dopamine dynamics differs across distinct striatal subregions, how these dynamics evolve throughout learning, and how each might causally contribute to learning and performance. Recent work found an attenuation of mesolimbic dopamine release in the nucleus accumbens core within a session of self-paced opto-ICSS of ventral tegmental area dopamine neurons, albeit without comparison to temporally matched non-contingent playback stimulation[30]. Together, this finding and the present study extend earlier work reporting suppression of both mesolimbic[56] and nigrostriatal dopamine[57] evoked by non-selective electrical self-stimulation in trained animals versus non-contingent playback in naïve animals. Further, in a discrete-trial, cued task variant, Covey and Cheer[30] also found an attenuation of optogenetically stimulated dopamine release and a concomitant increase in cue-evoked release, consistent with classic reward prediction errors in natural reward contexts[1]. Indeed,

another recent study found predominant prediction-error responses in dopamine axonal activity throughout much of the ventral striatum, DMS, and DLS in a cued discrimination task for natural reward[71]. In that study, a notable difference in the DLS was a lack of dips below baseline despite similarly suppressive effects of reward expectation across regions. In our photometry recordings with a fluorescent dopamine sensor, we also observed similar suppression of self-stimulated dopamine in both the DMS and DLS during single-lever CRF, but did not examine negative prediction errors to reward omission in these sessions. Based on these collective findings, we therefore would predict that most effects observed within the DMS in the current study would be largely similar in the accumbens core[30] and DLS, although we also might not expect negative prediction errors to cause dips below baseline in the DLS[71]. In contrast, the predictions are perhaps less clear for aspects of the accumbens shell and the caudal-most tail of the striatum, where distinct and surprising dopamine dynamics have been revealed particularly in aversive domains[72–74]. Overall, potential heterogeneity of dopamine signaling across striatal subregions remains an important topic of investigation.

Uncovering the circuit mechanisms responsible for this dopaminergic action-outcome prediction error also remains an important open question for future research. The current results constrain candidate mechanisms to those with fairly rapid onset, transient duration, and sufficiently strong inhibition to suppress or shunt even direct optogenetic depolarization. Nigrostriatal dopamine neurons receive monosynaptic inputs from all basal ganglia nuclei[75–77], the majority of which are predominantly inhibitory GABAergic projections[78–80]. Striatal, pallidal, and nigral basal ganglia nuclei contain many cells exhibiting prominent activity related to action sequence initiation, termination, and transitions[39–41], as well as action-outcome value information[81–86] that may converge and contribute to these dopamine neuron computations. Striatal patch (striosome) compartment neurons are compelling striatonigral candidates, given their dense anatomical innervation and strong inhibition of nigral dopamine neurons[75,80,87–89], although the striatal neurons in the surrounding matrix compartment could potentially contribute as well[80,87]. The rostromedial tegmental nucleus is another major GABAergic input, inverting lateral habenula signals that often resemble negative prediction errors[90–92]. Recent investigations of circuitry regulating prediction-error computations by adjacent mesolimbic dopamine neurons in cued-reward contexts has suggested that distinct afferents provide dissociable information[93–95], though there also may be a high degree of redundancy and mixed selectivity distributed across these inputs[96]. Although subpopulations within several of these nuclei also exhibit prediction error-like activity which dopamine neurons might passively relay[90,96–100], the action-induced suppression of dopamine evoked by direct optogenetic stimulation in the current experiments further implies computation within dopamine neurons themselves, akin to mesolimbic dopamine neurons in Pavlovian contexts[4,5,96,101]. The observation that an identified SNc dopamine neuron exhibits reduced spiking to self- vs. passive optogenetic stimulation (Figure S2H–M) provides initial evidence that somatodendritic inhibition likely contributes to this action-induced inhibition. Nevertheless, a variety of axonal and terminal mechanisms ultimately regulating dopamine release within the striatum also merit further functional investigation, including local GABAergic, cholinergic, and neuropeptide regulation[102–105]. Finally, given the prominent role proposed

for efference copy signals in striatal-dependent learning[65,66], corticostriatal projections, particularly from premotor regions involved in both action initiation[106] and efference copy signal generation[48], could be distal upstream sources contributing action-outcome information for these dopaminergic prediction error computations, whether via multisynaptic pathways to the midbrain or striatal terminal regulation.

Dopaminergic prediction errors are thought to convey a teaching signal that is critical for multiple forms of associative learning across the corticostriatal topography[28,29], spanning both classical Pavlovian stimulus-outcome conditioning[20,107–110] and the formation of stimulus-response habits[12,52,111–114]. The action-outcome prediction errors with sequence-specific hierarchical regulation observed in the current work likely reveal fundamental computations supporting instrumental learning. Such action-outcome prediction errors may underlie the assignment of credit to antecedent actions, thereby updating action values and policies[34,115]. Subtractive inhibition by reward expectation minimizes new learning to fully predicted outcomes, permits performance to stabilize, and supports extinction following reward omission[4,34,109]. Dysregulation of dopamine dynamics and disrupting the suppressive effects of action-outcome expectations in particular may contribute to the development of compulsive behavior characteristic of addiction[116,117] or impulse control disorders that are common side effects of dopamine replacement therapies for Parkinson's disease[118]. Related to this notion of credit assignment, action-outcome prediction errors also may be fundamental to the attribution of agency when outcomes are under instrumental control[65,115]. Efference copies attenuate the expected sensory consequences of self-generated actions, permitting organisms to distinguish outcomes resulting from their own action versus external causes[46–49]. Deficits in efference copy signaling and hierarchical predictive processing more broadly may cause perturbed regulation of dopaminergic prediction errors implicated in agency-related delusions and hallucinations comprising core positive symptoms of schizophrenia and psychosis[119–128]. Recent studies have suggested that dynamic nigrostriatal dopamine might regulate ongoing actions[36,39,129,130] and bias online action selection[131]. The current results revealed that nigrostriatal dopamine can encode action-outcome prediction errors critical for action learning. Together they underscore the importance of dopamine for action selection at short as well as long timescales, and have important implications in many neurological disorders such as Parkinson's disease, schizophrenia, and addiction.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Xin Jin (xjin@bio.ecnu.edu.cn).

Materials Availability—This study did not generate new unique reagents.

Data and Code Availability—Data have been deposited at Zenodo: https://doi.org/10.5281/zenodo.5501733 and are publicly available as of the date of publication. The DOI is listed in the Key Resources Table. Any additional information required to reanalyze the data reported in this paper is available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals—All procedures were approved by the Institutional Animal Care and Use Committee at the Salk Institute for Biological Studies and were conducted in accordance with the National Institute of Health's Guide for the Care and Use of Laboratory Animals. Experiments were performed using male and female mice, at least two months old, group-housed (2–5 mice / cage) on a 12 hr light/dark cycle (lights on at 6:00 am). *DAT-cre* mice[16] (Jackson Laboratory # 020080) were either crossed with the Ai32 line[17] (*RCL-ChR2(H134R)-EYFP*, Jackson Laboratory # 024109) or injected with cre-dependent AAV in the SNc to selectively express channelrhodopsin-2 in their dopamine neurons.

METHOD DETAILS

Surgical Procedures—Mice were anesthetized with isoflurane (3% induction, 0.5–1.5% sustained), their head shaved, and they were placed in a stereotaxic frame. The scalp was swabbed with 70% isopropyl alcohol and a povidine-iodine solution, and given a subcutaneous injection of bupivicaine (2 mg/kg) for local anesthesia. After a midline incision and leveling the skull, skulls were dried and coated with OptiBond adhesive and/or implanted with skull screws. Craniotomies were drilled over the dorsal striatum (+ 0.5–0.8 mm AP, 1.5 mm ML from bregma) for the voltammetric working electrode, the substantia nigra pars compacta (SNc: −3.1–3.3 mm AP, 1.3 mm ML) for the fiber optic(s), and an arbitrary distal site for the Ag/AgCl reference electrode. For *DAT-cre* mice not already crossed with the *Ai32* line, 300 nl of AAV5-EF1a-DIO-ChR2(H134R)-mCherry (UNC vector core) was injected into the SNc (4.1 mm ventral from dura; 100 nl/min), and the injection needle was left in place for 5 min before being slowing withdrawn[131]. For all FSCV mice, the Ag/AgCl reference was inserted under the skull and cemented in place, and a carbon-fiber microelectrode[19] was lowered into the striatum (2.3–2.5 mm DV from dura) while applying a voltammetric waveform (see FSCV section) at 60 Hz for 10–15 min, and then at 10 Hz until the background had stabilized. A fiber optic[18,131] (200 μm core) was lowered targeting the ipsilateral SNc (3.8–4.1 mm DV). *DAT-cre x Ai32* mice received 1-s 50-Hz optical stimulation while striatal dopamine was recorded with FSCV to ensure electrode functionality and fiber placement. Mice subsequently trained in the Left-Right sequence task cohort (see Behavioral Training) also were implanted with a fiber optic over the contralateral SNc for bilateral stimulation. All implants were cemented to the skull along with a connector from the reference and working electrodes for later attachment to the FSCV head-mounted amplifier (headstage). For fiber photometry recordings, *DATcre x Ai32* mice were injected unilaterally with 300 nl of red fluorescent dopamine sensor[31] AAV1-hSyn-rGRAB$_{DA1h}$ (Addgene plasmid # 140557 was a gift from Yulong Li, packaged in the Salk GT3 viral vector core) at two depths in either the DMS (+ 0.7 mm AP, 1.5 mm ML from bregma, −2.6 and −2.3 mm DV from dura) or the DLS (+ 0.2 mm AP, 2.5 mm ML from bregma, −3.3 and −3.0 mm DV from dura), implanted with a fiber optic (200 μm core, black ceramic ferrule; Neurophotometrics or RWD) at the same site (−2.4 mm DV for DMS, −3.1 mm DV for DLS), and implanted unilaterally with a fiber optic targeting the ipsilateral SNc as above (−3.3 mm AP, 1.3 mm ML, −3.9 mm DV). For electrophysiological identification of dopamine neurons, *DATcre x Ai32* mice were implanted unilaterally in the SNc with an electrode array (Innovative Neurophysiology) with 16 tungsten contacts

(2 × 8), 35 μm in diameter, spaced 150 μm apart within rows and 200 μm apart between rows. The array had a fiber optic directly attached, positioned ~300 μm from the electrode tips, to permit coupling to the laser for stimulation delivery[39,131]. The silver grounding wire was attached to a skull screw, and the array was affixed with dental cement. Mice received buprenorphine (1 mg/kg, s.c.) for analgesia and dexamethasone (2.5 mg/kg, s.c.) or ibuprofen in their drinking water for post-operative anti-inflammatory treatment, recovered in a clean home cage on a heating pad, were monitored daily for at least 3 days, and allowed to recover for at least 10 days before beginning behavioral training.

Behavioral Training—Behavioral training was conducted in standard operant chambers (Med Associates) inside sound attenuating box, as previously described[41,131]. Mice were connected to the fiber optic patch cable from the laser (LaserGlow; 473 nm, ~5 mW measured before each session) and placed in the operant chamber, and optogenetic intracranial self-stimulation (opto-ICSS) sessions began with the insertion of two levers and the onset of a central house light on the opposite wall. The levers remained extended and the house light remained on for the duration of the 60 min sessions.

<u>Continuous reinforcement cohort:</u> Each press on the designated Active lever resulted in 1 s of optical stimulation (50 Hz, 10 ms pulse width) on a continuous reinforcement (CRF) schedule, other than additional presses during an ongoing stimulation train, which were recorded but had no consequence. Presses on the other, Inactive lever also were recorded but had no consequence. The sides of the Active and Inactive levers were counterbalanced relative to both the operant chamber and implanted hemisphere across mice, and remained fixed across training days for a given animal. Once FSCV mice reliably made at least 100 Active lever presses per session for 3 consecutive days, they also were connected to a voltammetry headstage before each session to allow habituation to behaving with this additional tethering. If a mouse failed to interact with the levers during its first 3 days of training, it was placed on food restriction overnight and a sucrose pellet was placed on the lever during its next behavioral session to encourage exploration. Once mice were reliably pressing the Active lever, they remained on *ad libitum* access to food and water in their home cages for all subsequent behavioral training and FSCV recordings. Mice were trained for at least 3 days while tethered to the FSCV headstage and meeting the behavioral criteria of at least 100 Active lever presses before FSCV recordings commenced (mean ± SEM = 11.1 ± 1.2 training days).

Fiber photometry and electrophysiology mice were trained as described above prior to their respective opto-ICSS recordings. DMS-implanted fiber photometry mice also were food restricted to 85% of their free-feeding baseline weight and trained to press a lever for sucrose pellets (20 mg, Bio-Serv) on a CRF schedule. In these sucrose pellet CRF sessions, only a single Active lever was present, which had been the Inactive lever from their opto-ICSS sessions. These sessions lasted 60 min or for a maximum of 60 sucrose pellet reinforcers, whichever came first, and fiber photometry recordings commenced on the tenth training session.

<u>Left-Right sequence cohort:</u> Mice in the Left-Right (LR) sequence cohort were initially trained on single-lever CRF opto-ICSS. For this cohort's CRF training, only one lever was

extended in each of two 30-min blocks per session (order counterbalanced across mice), and presses in both left- and right-lever blocks yielded the same 1-s, 50-Hz stimulation. To expedite this initial training stage, all mice in this cohort were food restricted prior to their first session, and were maintained at 85% their free-feeding baseline weight with ~2.5 g of standard lab chow per mouse in their home cage after the daily training sessions. Once mice made at least 100 presses in each block for 3 consecutive days, they were returned to *ad libitum* food access in their home cage, CRF training continued until they again met this 100-press criterion for another 3 days, and they then began training on the LR sequence task.

In LR sequence session, both levers were inserted at the start of the session and remained extended for the duration of the 60 min sessions. To receive stimulation (1 s, 50 Hz), mice now had to press the Left and then Right lever. No other combination of lever press pairs (Left-Left, Right-Right, or Right-Left) was reinforced with stimulation. After reaching the behavioral criterion of receiving at least 100 stimulations per session for 3 consecutive days, mice were habituated to tethering with the FSCV headstage, and received further training while tethered until they again met this 100-stimulation 3-day criterion and FSCV recording sessions commenced (mean ± SEM = 56.9 ± 7.6 training days). A subset of animals was trained under the same procedures to instead perform the Right-Left sequence as a spatial control, but we refer to the LR sequence throughout for simplicity. The hemisphere of the implanted FSCV recording electrode also was counterbalanced relative to this sequence direction across mice.

Contingency degradation: A contingency degradation test session began with 30 min of standard opto-ICSS (CRF for the CRF cohort, LR sequence task for the LR cohort). In the subsequent 30-min contingency degradation test phase, the levers remained extended, but stimulation was decoupled from task performance and instead was delivered regardless of whether the mice pressed any levers[24–27]. For each mouse, the timing of these non-contingent stimulations during the test phase was matched to the time stamps of stimulations earned during that animal's preceding opto-ICSS phase in the first half of the session, ensuring that the stimulation rate and distribution of inter-stimulation intervals were yoked within-subject to a given animal's own opto-ICSS performance.

Fast-Scan Cyclic Voltammetry (FSCV)—Striatal dopamine was recorded with *in vivo* FSCV in behaving animals as previously described[6,19,131]. Briefly, voltammetric waveform application consisted of holding the potential at the carbon-fiber electrode at −0.4 V relative to the Ag/AgCl reference between scans, and ramping to +1.3 V and then back to −0.4V at 400 V/s for each scan. Prior to the initial FSCV recording during opto-ICSS performance, this voltammetric waveform was applied at 60 Hz for at least one hour while mice were in a 'cycling chamber' outside the operant box, then at 10 Hz until the background current had stabilized. Mice then received experimenter-delivered optical stimulations (1 s, 50 Hz) to ensure electrode functionality.

For opto-ICSS sessions with FSCV recordings, electrodes were first cycled at 60 Hz for ~40 min and then at 10 Hz for at least 20 min until background current equilibration and throughout the opto-ICSS behavioral session. Mice received a series of 3 experimenter-delivered stimulations before and after the session to validate electrode functionality the day

of each recording and for generating voltammetric training sets (see Statistical Analyses). The opto-ICSS session began at least 5 min after the final pre-session stimulation. The first half of each FSCV session consisted of a standard opto-ICSS phase (CRF for the CRF cohort, LR sequence task for the LR cohort) that was identical to the previous behavioral training sessions. At the conclusion of this active Self-Stimulation phase, the both levers retracted and the house light turned off for a 5 min interim period, followed by a Passive Playback phase in which mice received non-contingent stimulations with the same timing and stimulation parameters (1 s at 50 Hz) as in the active Self-Stimulation phase. The timing of these non-contingent Passive Playback stimulations was matched to the time stamps of stimulations earned during a given animal's preceding Self-Stimulation phase, again ensuring that the stimulation rate and distribution of inter-stimulation intervals were identical across both the active and passive phases for a given animal. Mice in the LR sequence cohort also performed another 30 min of active LR sequence opto-ICSS following the Passive Playback phase to permit assessment of possible temporal order effects (Figure S3H–I).

Mice also underwent additional FSCV recordings in several types of probe sessions, including Omission, Delay, and Magnitude Probes for the CRF cohort, and Left and Right Lever Probes for the LR cohort. These FSCV sessions consisted of the same basic protocol described above, with active Self-Stimulation and non-contingent Passive Playback yoked within-subject. In Omission Probe sessions, 20% of presses on the typically Active lever did not yield stimulation, and instead caused a 5-s timeout period during which no further stimulation could be earned. This timeout period was not explicitly cued with any overt stimulus, other than the absence of the typical stimulation delivery. In Delay Probe sessions, 20% of presses on the Active lever resulted in stimulation that was delayed by 5 s. As for the Omission Probe timeout period, no further stimulation could be earned during this delay period. In Magnitude Probe sessions, 20% of Active lever presses yielded an increased magnitude of stimulation (5 s at 50 Hz). For the LR sequence cohort, the single-press probe sessions consisted of probe stimulations delivered on a random subset of first lever presses after previous reinforcement, in addition to continuous reinforcement for LR sequences as usual. For the Left Probe sessions, the next left lever press following the last reinforcement was stimulated with 20% probability. Due to the lower probability of an additional right lever press following a reinforcement, a right lever press following the last reinforcement was stimulated with 50% probability to collect enough probes for data analyses in the Right Probe sessions. Probe sessions were recorded at least 2 days apart, with standard opto-ICSS behavioral training sessions performed on the intervening days to allow return to baseline performance.

Fiber Photometry—The red fluorescent dopamine sensor[31] rGRAB$_{DA1h}$ was recorded using fiber photometry (FP3002, Neurophotometrics) controlled via Bonsai[132]. LEDs delivering two excitation wavelengths (560 nm for detection of dopamine and 415 nm for a dopamine-independent control[31], light intensity ~50 μW each at the tip of the patch cord) were interleaved at 40 Hz throughout recording sessions. Fluorescence emission was focused onto a CMOS sensor for detection with a region of interest drawn around the end of the connected patch cable. Opto-ICSS recording sessions consisted of a 15-min Self-

Stimulation phase (CRF), a 3-min interim, and then temporally matched Passive Playback stimulations. For sucrose pellet CRF sessions, the Self-administration phase lasted 15 min or terminated after a maximum of 30 pellets, whichever came first, and the Playback phase was recorded 24–48 hours later to control for satiety, again using the same pellet-delivery timestamps as that animal's Self phase.

***In Vivo* Electrophysiology**—SNc dopamine neurons were recorded and identified as previously described[39,131]. Briefly, neural activity was recorded using the MAP system (Plexon), and spike activities first were sorted online with a build-in algorithm. Only spikes with stereotypical waveforms distinguishable from noise and high signal-to-noise ratio were saved for further analysis. Behavioral training and recording sessions were conducted as described above for the opto-ICSS CRF cohort. After recording the opto-ICSS session with active Self-Stimulation and Passive Playback phases, the recorded spikes were further isolated into individual units using offline sorting software (Offline Sorter, Plexon). Each individual unit displayed a clear refractory period in the inter-spike interval histogram, with no spikes during the refractory period (larger than 1.3ms). To identify laser-evoked responses, neuronal firing was aligned to stimulation onset and averaged across stimulations in 1-ms bins, and baseline was defined by averaging neuronal firing in the 1 s preceding stimulation onset. The latency to respond to stimulation was defined as the as the time to significant firing rate increase, with a threshold defined as > 99% of baseline activity (3 standard deviations). Only units with short response latency (< 10 ms) from stimulation onset and high correlation between spontaneous and laser-evoked spike waveforms (r > 0.95) were considered cre-positive, optogenetically identified dopamine neurons[39,131].

Histology—Mice were anesthetized with ketamine (100 mg/kg, i.p.) and xylazine (10 mg/kg, i.p.), and the FSCV recording site was marked by passing a 70 μA current through the electrode for 20 s. Mice were transcardially perfused with 0.01 M phosphate-buffered saline (PBS) and then 4% paraformaldehyde (PFA) in PBS. Brains were removed, post-fixed in PFA at 4° for 24 hr, and then stored at 4° in a solution of 30% sucrose in 0.1 M phosphate buffer until ready for cryosectioning. Tissue was sectioned at 50 μm thickness on a freezing microtome, and striatal and SNc sections were mounted onto glass slides and coverslipped with AquaPoly mounting media containing DAPI (1:1000). Some sections also were processed for immunohistochemistry as previously described[41,87]. Briefly, sections were washed 3 times for 15 min each in tris-buffered saline (TBS), and incubated for 1 hr in blocking solution containing 3% normal horse serum and 0.25% Triton-X 100 in TBS. Tissue was incubated for 48 hr in primary antibody against tyrosine hydroxylase (anti-TH, raised in rabbit, 1:1000, Abcam) and green fluorescent protein (anti-GFP, raised in chicken, 1:1000, Novus Biologicals) in this blocking solution at 4°, washed twice for 15 min in TBS and then for 30 min in the blocking solution, and then incubated for 3 hr in secondary antibody (anti-Chicken AlexaFluor 488 and anti-Rabbit Cy3, each 1:250, Jackson ImmunoResearch) in blocking solution. Finally, sections were washed 3 times for 15 min in TBS, mounted onto slides, and coverslipped with DAPI mounting media as above. Sections were imaged on a Zeiss LSM 710 confocal microscope with 10x and 20x objectives. All included FSCV animals were confirmed to have electrode placement in the dorsal striatum and fiber optics targeting the SNc.

QUANTIFICATION AND STATISTICAL ANALYSIS

FSCV data were low-pass filtered at 2 kHz, aligned to each lever press and/or stimulation onset, and background-subtracted using the mean voltammetric current in the 1 s prior to each aligned event of interest. Dopamine responses were isolated using chemometric principal component analysis with training sets consisting of cyclic voltammograms for dopamine, pH, and electrode drift[131,133,134]. Electrode-specific training sets were used for each animal and represented additional inclusion criteria for a given electrode, but similar results were obtained when reanalyzing data with a standardized training set across animals[135]. Changes in dopamine concentration were estimated based on average post-implantation electrode sensitivity[19]. Analysis of fiber photometry data consisted of fitting the 415 nm control channel with a biexponential decay to account for photobleaching across the session, linearly scaling that fit to the 560 nm dopamine-dependent channel using robust regression, and dividing the 560 nm data by this scaled fit. Peri-event change in fluorescence (F/F) then was calculated by subtracting the 250-ms baseline period preceding the stimulation or pellet delivery.

Mean changes in dopamine concentration summarized in bar graphs throughout the opto-ICSS results analyzed time periods spanning 0.5–1.5 s following the aligned event onset. The summary bar graph for the sucrose pellet photometry data quantifies the $rGRAB_{DA1h}$ sensor response in the 1 s following reward retrieval. Analysis of the FSCV Magnitude Probes also included a late time point at 4.5–5.5 s after Probe onset, as did supplementary analysis of Omission Probes with versus without additional presses during the timeout period. For the LR sequence cohort, analysis of non-reinforced press pairs was restricted to pairs with short inter-press intervals (IPI < 5 s), consistent with the short duration of most LR sequences. Analysis of the non-reinforced single Left and Right lever presses was restricted to the first press following previous reinforcement, to match the press that could receive probe stimulation in the corresponding single-press probe sessions. Analysis of non-reinforced Right lever presses and Right Probe stimulations was restricted to those where the animal did not first approach the Left lever, as determined by examination of the video, to ensure that the right presses analyzed were individual actions and not part of a LR sequence. Analysis of LR sequences aligned to approach initiation entailed first identifying this time of approach initiation from the video[55], and the subsequent intervals from approach initiation to Left press and from the Left to Right press were scaled to each interval's median duration to normalize over time and then concatenated[55,136]. Statistical analyses of behavioral and recording data consisted of t tests and repeated-measures ANOVAs with post-hoc tests corrected for multiple comparisons as indicated throughout the corresponding figure legends. Stimulation-evoked dopamine traces also were analyzed with Difference Traces that digitally subtracted the Passive Playback response from the Self-Stimulation response for each pair of matched stimulations. Dopamine trace time courses following event onset were analyzed with permutation tests (10,000 random shuffles) with a cluster-based correction for multiple comparisons over time[137,138]. For electrophysiological data analysis, neuronal firing was aligned to stimulation onset, averaged within each session phase, and smoothed with a Gaussian filter (window size = 50 ms, standard deviation = 10) to construct peri-event time histograms for Self-Stimulation and Passive Playback responses. Statistical analyses were performed in Prism (GraphPad) and Matlab (MathWorks).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Schultz W, Dayan P, and Montague PR (1997). A Neural Substrate of Prediction and Reward. Science 275, 1593–1599. [PubMed: 9054347]

2. Fiorillo CD, Tobler PN, and Schultz W (2003). Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. Science 299, 1898–1902. [PubMed: 12649484]

3. Cohen JY, Haesler S, Vong L, Lowell BB, and Uchida N (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482, 85–88. [PubMed: 22258508]

4. Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, and Uchida N (2015). Arithmetic and local circuitry underlying dopamine prediction errors. Nature 525, 243–246. [PubMed: 26322583]

5. Eshel N, Tian J, Bukwich M, and Uchida N (2016). Dopamine neurons share common response function for reward prediction error. Nat. Neurosci 19, 479–486. [PubMed: 26854803]

6. Hollon NG, Arnold MM, Gan JO, Walton ME, and Phillips PEM (2014). Dopamine-associated cached values are not sufficient as the basis for action selection. Proc. Natl. Acad. Sci 111, 18357–18362. [PubMed: 25489094]

7. Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, and Witten IB (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat. Neurosci 19, 845–854. [PubMed: 27110917]

8. Coddington LT, and Dudman JT (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nat. Neurosci 21, 1563–1573. [PubMed: 30323275]

9. Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, et al. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature 570, 509–513. [PubMed: 31142844]

10. Kremer Y, Flakowski J, Rohner C, and Lüscher C (2020). Context-Dependent Multiplexing by Individual VTA Dopamine Neurons. J. Neurosci 40, 7489–7509. [PubMed: 32859713]

11. Matsumoto M, and Hikosaka O (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459, 837–841. [PubMed: 19448610]

12. Kim HF, Ghazizadeh A, and Hikosaka O (2015). Dopamine Neurons Encoding Long-Term Memory of Object Value for Habitual Behavior. Cell 163, 1165–1175. [PubMed: 26590420]

13. Morris G, Arkadir D, Nevet A, Vaadia E, and Bergman H (2004). Coincident but Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons. Neuron 43, 133–143. [PubMed: 15233923]

14. Roesch MR, Calu DJ, and Schoenbaum G (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci 10, 1615–1624. [PubMed: 18026098]

15. Hart AS, Rutledge RB, Glimcher PW, and Phillips PEM (2014). Phasic Dopamine Release in the Rat Nucleus Accumbens Symmetrically Encodes a Reward Prediction Error Term. J. Neurosci 34, 698–704. [PubMed: 24431428]

16. Zhuang X, Masson J, Gingrich JA, Rayport S, and Hen R (2005). Targeted gene expression in dopamine and serotonin neurons of the mouse brain. J. Neurosci. Methods 143, 27–32. [PubMed: 15763133]

17. Madisen L, Mao T, Koch H, Zhuo J, Berenyi A, Fujisawa S, Hsu Y-WA, Garcia AJ, Gu X, Zanella S, et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci 15, 793–802. [PubMed: 22446880]

18. Sparta DR, Stamatakis AM, Phillips JL, Hovelsø N, van Zessen R, and Stuber GD (2012). Construction of implantable optical fibers for long-term optogenetic manipulation of neural circuits. Nat. Protoc 7, 12–23.

19. Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, et al. (2010). Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nat. Methods 7, 126–129. [PubMed: 20037591]

20. Saunders BT, Richard JM, Margolis EB, and Janak PH (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat. Neurosci 21, 1072–1083. [PubMed: 30038277]

21. Rossi MA, Sukharnikova T, Hayrapetyan VY, Yang L, and Yin HH (2013). Operant Self-Stimulation of Dopamine Neurons in the Substantia Nigra. PLoS One 8, e65799. [PubMed: 23755282]

22. Ilango A, Kesner AJ, Keller KL, Stuber GD, Bonci A, and Ikemoto S (2014). Similar Roles of Substantia Nigra and Ventral Tegmental Dopamine Neurons in Reward and Aversion. J. Neurosci 34, 817–822. [PubMed: 24431440]

23. Keiflin R, Pribut HJ, Shah NB, and Janak PH (2019). Ventral Tegmental Dopamine Neurons Participate in Reward Identity Predictions. Curr. Biol 29, 93–103.e3. [PubMed: 30581025]

24. Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, et al. (2011). Recombinase-Driver Rat Lines: Tools, Techniques, and Optogenetic Application to Dopamine-Mediated Reinforcement. Neuron 72, 721–733. [PubMed: 22153370]

25. Koralek AC, Jin X, Long II JD, Costa RM, and Carmena JM (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483, 331–335. [PubMed: 22388818]

26. Clancy KB, Koralek AC, Costa RM, Feldman DE, and Carmena JM (2014). Volitional modulation of optically recorded calcium signals during neuroprosthetic learning. Nat. Neurosci 17, 107–109.

27. Neely RM, Koralek AC, Athalye VR, Costa RM, and Carmena JM (2018). Volitional Modulation of Primary Visual Cortex Activity Requires the Basal Ganglia. Neuron 97, 1356–1368.e4. [PubMed: 29503189]

28. Yin HH, Ostlund SB, and Balleine BW (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur. J. Neurosci 28, 1437–1448. [PubMed: 18793321]

29. Balleine BW (2019). The Meaning of Behavior: Discriminating Reflex and Volition in the Brain. Neuron 104, 47–62. [PubMed: 31600515]

30. Covey DP, and Cheer JF (2019). Accumbal Dopamine Release Tracks the Expectation of Dopamine Neuron-Mediated Reinforcement. Cell Rep. 27, 481–490.e3. [PubMed: 30970251]

31. Sun F, Zhou J, Dai B, Qian T, Zeng J, Li X, Zhuo Y, Zhang Y, Wang Y, Qian C, et al. (2020). Next-generation GRAB sensors for monitoring dopaminergic activity in vivo. Nat. Methods 17, 1156–1166. [PubMed: 33087905]

32. Houk JC, Adams JL, and Barto AG (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In Models of Information Processing in the Basal Ganglia, Houk JC, Davis JL, and Beiser DG, eds. (MIT Press), pp. 249–270.

33. Montague PR, Dayan P, and Sejnowski TJ (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci 16, 1936–1947. [PubMed: 8774460]

34. Sutton RS, and Barto AG (2018). Reinforcement Learning: An Introduction. 2nd ed. (MIT Press).

35. Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, and Berke JD (2016). Mesolimbic dopamine signals the value of work. Nat. Neurosci 19, 117–126. [PubMed: 26595651]

36. da Silva JA, Tecuapetla F, Paixão V, and Costa RM (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. Nature 554, 244–248. [PubMed: 29420469]

37. Dodson PD, Dreyer JK, Jennings KA, Syed ECJ, Wade-Martins R, Cragg SJ, Bolam JP, and Magill PJ (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. Proc. Natl. Acad. Sci 113, E2180–E2188. [PubMed: 27001837]

38. Gallistel CR (1980). The Organization of Action (Lawrence Erlbaum Associates).

39. Jin X, and Costa RM (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature 466, 457–462. [PubMed: 20651684]

40. Jin X, Tecuapetla F, and Costa RM (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nat. Neurosci 17, 423–430. [PubMed: 24464039]

41. Geddes CE, Li H, and Jin X (2018). Optogenetic Editing Reveals the Hierarchical Organization of Learned Action Sequences. Cell 174, 32–43.e15. [PubMed: 29958111]

42. Lashley KS (1951). The problem of serial order in behavior. In Cerebral Mechanisms in Behavior: The Hixon Symposium, Jeffress LA, ed. (Wiley), pp. 112–136.

43. Jin X, and Costa RM (2015). Shaping action sequences in basal ganglia circuits. Curr. Opin. Neurobiol 33, 188–196. [PubMed: 26189204]

44. Rao RPN, and Ballard DH (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci 2, 79–87. [PubMed: 10195184]

45. Keller GB, and Mrsic-Flogel TD (2018). Predictive Processing: A Canonical Cortical Computation. Neuron 100, 424–435. [PubMed: 30359606]

46. Wolpert DM, Ghahramani Z, and Jordan MI (1995). An internal model for sensorimotor integration. Science (80-.). 269, 1880–1882.

47. Crapse TB, and Sommer MA (2008). Corollary discharge across the animal kingdom. Nat. Rev. Neurosci 9, 587–600. [PubMed: 18641666]

48. Schneider DM, Sundararajan J, and Mooney R (2018). A cortical filter that learns to suppress the acoustic consequences of movement. Nature 561, 391–395. [PubMed: 30209396]

49. Wurtz RH (2018). Corollary Discharge Contributions to Perceptual Continuity Across Saccades. Annu. Rev. Vis. Sci 4, 215–237. [PubMed: 30222532]

50. Graybiel AM (1998). The Basal Ganglia and Chunking of Action Repertoires. Neurobiol. Learn. Mem 70, 119–136. [PubMed: 9753592]

51. 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. Neurobiol. Learn. Mem 70, 137–149. [PubMed: 9753593]

52. Matsumoto N, Hanakawa T, Maki S, Graybiel AM, and Kimura M (1999). Nigrostriatal Dopamine System in Learning to Perform Sequential Motor Tasks in a Predictive Manner. J. Neurophysiol 82, 978–998. [PubMed: 10444692]

53. Barnes TD, Kubota Y, Hu D, Jin DZ, and Graybiel AM (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature 437, 1158–1161. [PubMed: 16237445]

54. Wassum KM, Ostlund SB, and Maidment NT (2012). Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. Biol. Psychiatry 71, 846–854. [PubMed: 22305286]

55. Collins AL, Greenfield VY, Bye JK, Linker KE, Wang AS, and Wassum KM (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. Sci. Rep 6. [PubMed: 28442741]

56. Garris PA, Kilpatrick M, Bunin MA, Michael D, Walker QD, and Wightman RM (1999). Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature 398, 67–69. [PubMed: 10078530]

57. Kilpatrick MR, Rooney MB, Michael DJ, and Wightman RM (2000). Extracellular dopamine dynamics in rat caudate–putamen during experimenter-delivered and intracranial self-stimulation. Neuroscience 96, 697–706. [PubMed: 10727788]

58. Owesson-White CA, Cheer JF, Beyene M, Carelli RM, and Wightman RM (2008). Dynamic changes in accumbens dopamine correlate with learning during intracranial self-stimulation. Proc. Natl. Acad. Sci 105, 11957–11962. [PubMed: 18689678]

59. Rodeberg NT, Johnson JA, Bucher ES, and Wightman RM (2016). Dopamine Dynamics during Continuous Intracranial Self-Stimulation: Effect of Waveform on Fast-Scan Cyclic Voltammetry Data. ACS Chem. Neurosci 7, 1508–1518. [PubMed: 27548680]

60. Kruse JM, Overmier JB, Konz WA, and Rokke E (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. Learn. Motiv 14, 165–181.

61. Corbit LH, and Janak PH (2007). Inactivation of the Lateral But Not Medial Dorsal Striatum Eliminates the Excitatory Impact of Pavlovian Stimuli on Instrumental Responding. J. Neurosci 27, 13977–13981. [PubMed: 18094235]

62. Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, and Schoenbaum G (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. Neuron 95, 1395–1405.e3. [PubMed: 28910622]

63. Wise RA (2002). Brain Reward Circuitry: Insights from Unsensed Incentives. Neuron 36, 229–240. [PubMed: 12383779]

64. Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, and Goldberg JH (2016). Dopamine neurons encode performance error in singing birds. Science 354, 1278–1282. [PubMed: 27940871]

65. Redgrave P, and Gurney K (2006). The short-latency dopamine signal: a role in discovering novel actions? Nat. Rev. Neurosci 7, 967–975. [PubMed: 17115078]

66. Fee MS (2014). The role of efference copy in striatal learning. Curr. Opin. Neurobiol 25, 194–200. [PubMed: 24566242]

67. Yin HH, Ostlund SB, Knowlton BJ, and Balleine BW (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci 22, 513–523. [PubMed: 16045504]

68. Gremel CM, and Costa RM (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun 4, 2264. [PubMed: 23921250]

69. Matamales M, McGovern AE, Mi JD, Mazzone SB, Balleine BW, and Bertran-Gonzalez J (2020). Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. Science 367, 549–555. [PubMed: 32001651]

70. Yin HH (2010). The Sensorimotor Striatum Is Necessary for Serial Order Learning. J. Neurosci 30, 14719–23. [PubMed: 21048130]

71. Tsutsui-Kimura I, Matsumoto H, Akiti K, Yamada MM, Uchida N, and Watabe-Uchida M (2020). Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. Elife 9, e62390. [PubMed: 33345774]

72. de Jong JW, Afjei SA, Pollak Dorocic I, Peck JR, Liu C, Kim CK, Tian L, Deisseroth K, and Lammel S (2019). A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. Neuron 101, 133–151.e7. [PubMed: 30503173]

73. Menegas W, Akiti K, Amo R, Uchida N, and Watabe-Uchida M (2018). Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. Nat. Neurosci 21, 1421–1430. [PubMed: 30177795]

74. Steinberg EE, Gore F, Heifets BD, Taylor MD, Norville ZC, Beier KT, Földy C, Lerner TN, Luo L, Deisseroth K, et al. (2020). Amygdala-Midbrain Connections Modulate Appetitive and Aversive Learning. Neuron 106, 1026–1043.e9. [PubMed: 32294466]

75. Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, and Uchida N (2012). Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. Neuron 74, 858–873. [PubMed: 22681690]

76. Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R, et al. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. Cell 162, 635–647. [PubMed: 26232229]

77. Menegas W, Bergan JF, Ogawa SK, Isogai Y, Umadevi Venkataraju K, Osten P, Uchida N, and Watabe-Uchida M (2015). Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. Elife 4, e10032. [PubMed: 26322384]

78. Tepper JM, and Lee CR (2007). GABAergic control of substantia nigra dopaminergic neurons. Prog. Brain Res 160, 189–208. [PubMed: 17499115]

79. Brazhnik E, Shah F, and Tepper JM (2008). GABAergic Afferents Activate Both GABAA and GABAB Receptors in Mouse Substantia Nigra Dopaminergic Neurons In Vivo. J. Neurosci 28, 10386–10398. [PubMed: 18842898]

80. Evans RC, Twedell EL, Zhu M, Ascencio J, Zhang R, and Khaliq ZM (2020). Functional Dissection of Basal Ganglia Inhibitory Inputs onto Substantia Nigra Dopaminergic Neurons. Cell Rep. 32, 108156. [PubMed: 32937133]

81. Samejima K, Ueda Y, Doya K, and Kimura M (2005). Representation of Action-Specific Reward Values in the Striatum. Science 310, 1337–1340. [PubMed: 16311337]

82. Lau B, and Glimcher PW (2008). Value Representations in the Primate Striatum during Matching Behavior. Neuron 58, 451–463. [PubMed: 18466754]

83. Hong S, and Hikosaka O (2008). The Globus Pallidus Sends Reward-Related Signals to the Lateral Habenula. Neuron 60, 720–729. [PubMed: 19038227]

84. Roesch MR, Singh T, Brown PL, Mullins SE, and Schoenbaum G (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. J. Neurosci 29, 13365–13376. [PubMed: 19846724]

85. Tachibana Y, and Hikosaka O (2012). The Primate Ventral Pallidum Encodes Expected Reward Value and Regulates Motor Action. Neuron 76, 826–837. [PubMed: 23177966]

86. Kim HF, Amita H, and Hikosaka O (2017). Indirect Pathway of Caudal Basal Ganglia for Rejection of Valueless Visual Objects. Neuron 94, 920–930.e3. [PubMed: 28521141]

87. Smith JB, Klug JR, Ross DL, Howard CD, Hollon NG, Ko VI, Hoffman H, Callaway EM, Gerfen CR, and Jin X (2016). Genetic-Based Dissection Unveils the Inputs and Outputs of Striatal Patch and Matrix Compartments. Neuron 91, 1069–1084. [PubMed: 27568516]

88. Crittenden JR, Tillberg PW, Riad MH, Shima Y, Gerfen CR, Curry J, Housman DE, Nelson SB, Boyden ES, and Graybiel AM (2016). Striosome–dendron bouquets highlight a unique striatonigral circuit targeting dopamine-containing neurons. Proc. Natl. Acad. Sci 113, 11318–11323. [PubMed: 27647894]

89. McGregor MM, McKinsey GL, Girasole AE, Bair-Marshall CJ, Rubenstein JLR, and Nelson AB (2019). Functionally Distinct Connectivity of Developmentally Targeted Striosome Neurons. Cell Rep. 29, 1419–1428.e5. [PubMed: 31693884]

90. Matsumoto M, and Hikosaka O (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447, 1111–5. [PubMed: 17522629]

91. Jhou TC, Fields HL, Baxter MG, Saper CB, and Holland PC (2009). The Rostromedial Tegmental Nucleus (RMTg), a GABAergic Afferent to Midbrain Dopamine Neurons, Encodes Aversive Stimuli and Inhibits Motor Responses. Neuron 61, 786–800. [PubMed: 19285474]

92. Hong S, Jhou TC, Smith M, Saleem KS, and Hikosaka O (2011). Negative Reward Signals from the Lateral Habenula to Dopamine Neurons Are Mediated by Rostromedial Tegmental Nucleus in Primates. J. Neurosci 31, 11457–71. [PubMed: 21832176]

93. Tian J, and Uchida N (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. Neuron 87, 1304–1316. [PubMed: 26365765]

94. Takahashi YK, Langdon AJ, Niv Y, and Schoenbaum G (2016). Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. Neuron 91, 182–93. [PubMed: 27292535]

95. Yang H, de Jong JW, Tak Y, Peck J, Bateup HS, and Lammel S (2018). Nucleus Accumbens Subnuclei Regulate Motivated Behavior via Direct Inhibition and Disinhibition of VTA Dopamine Subpopulations. Neuron 97, 434–449.e4. [PubMed: 29307710]

96. Tian J, Huang R, Cohen JY, Osakada F, Kobak D, Machens CK, Callaway EM, Uchida N, and Watabe-Uchida M (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. Neuron 91, 1374–1389. [PubMed: 27618675]

97. Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, and Schoenbaum G (2012). Reward Prediction Error Signaling in Posterior Dorsomedial Striatum Is Action Specific. J. Neurosci 32, 10296–305. [PubMed: 22836263]

98. Chen R, Puzerey PA, Roeser AC, Riccelli TE, Podury A, Maher K, Farhang AR, and Goldberg JH (2019). Songbird Ventral Pallidum Sends Diverse Performance Error Signals to Dopaminergic Midbrain. Neuron 103, 266–276.e4. [PubMed: 31153647]

99. Oemisch M, Westendorff S, Azimi M, Hassani SA, Ardid S, Tiesinga P, and Womelsdorf T (2019). Feature-specific prediction errors and surprise across macaque fronto-striatal circuits. Nat. Commun 10, 176. [PubMed: 30635579]

100. Ottenheimer DJ, Bari BA, Sutlief E, Fraser KM, Kim TH, Richard JM, Cohen JY, and Janak PH (2020). A quantitative reward prediction error signal in the ventral pallidum. Nat. Neurosci 23, 1267–1276. [PubMed: 32778791]

101. Watabe-Uchida M, Eshel N, and Uchida N (2017). Neural Circuitry of Reward Prediction Error. Annu. Rev. Neurosci 40, 373–394. [PubMed: 28441114]

102. Kramer PF, Twedell EL, Shin JH, Zhang R, and Khaliq ZM (2020). Axonal mechanisms mediating γ-aminobutyric acid receptor type A (GABA-A) inhibition of striatal dopamine release. Elife 9, e55729. [PubMed: 32870779]

103. Holly EN, Davatolhagh MF, España RA, and Fuccillo MV (2021). Striatal low-threshold spiking interneurons locally gate dopamine. Curr. Biol 31.

104. Collins AL, Aitken TJ, Greenfield VY, Ostlund SB, and Wassum KM (2016). Nucleus Accumbens Acetylcholine Receptors Modulate Dopamine and Motivation. Neuropsychopharmacology 41, 2830–2838. [PubMed: 27240658]

105. Sulzer D, Cragg SJ, and Rice ME (2016). Striatal dopamine neurotransmission: Regulation of release and uptake. Basal Ganglia 6, 123–148. [PubMed: 27141430]

106. Murakami M, Vicente MI, Costa GM, and Mainen ZF (2014). Neural antecedents of self-initiated actions in secondary motor cortex. Nat. Neurosci 17, 1574–82. [PubMed: 25262496]

107. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PEM, and Akil H (2011). A selective role for dopamine in stimulus–reward learning. Nature 469, 53–57. [PubMed: 21150898]

108. Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, and Janak PH (2013). A causal link between prediction errors, dopamine neurons and learning. Nat. Neurosci 16, 966–973. [PubMed: 23708143]

109. Chang CY, Esber GR, Marrero-Garcia Y, Yau H-J, Bonci A, and Schoenbaum G (2016). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. Nat. Neurosci 19, 111–116. [PubMed: 26642092]

110. Maes EJP, Sharpe MJ, Usypchuk AA, Lozzi M, Chang CY, Gardner MPH, Schoenbaum G, and Iordanova MD (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. Nat. Neurosci 23, 176–178. [PubMed: 31959935]

111. Knowlton BJ, Mangels JA, and Squire LR (1996). A Neostriatal Habit Learning System in Humans. Science 273, 1399–1402. [PubMed: 8703077]

112. Faure A, Haberland U, Conde F, and El Massioui N (2005). Lesion to the Nigrostriatal Dopamine System Disrupts Stimulus-Response Habit Formation. J. Neurosci 25, 2771–2780. [PubMed: 15772337]

113. Belin D, and Everitt BJ (2008). Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum. Neuron 57, 432–441. [PubMed: 18255035]

114. Wang LP, Li F, Wang D, Xie K, Wang D, Shen X, and Tsien JZ (2011). NMDA Receptors in Dopaminergic Neurons Are Crucial for Habit Learning. Neuron 72, 1055–1066. [PubMed: 22196339]

115. Hamid AA, Frank MJ, and Moore CI (2021). Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. Cell 184, 2733–2749.e16. [PubMed: 33861952]

116. Redish AD (2004). Addiction as a Computational Process Gone Awry. Science 306, 1944–7. [PubMed: 15591205]

117. Lüscher C, Robbins TW, and Everitt BJ (2020). The transition to compulsion in addiction. Nat. Rev. Neurosci 21, 247–263. [PubMed: 32231315]

118. Weintraub D, and Mamikonyan E (2019). Impulse Control Disorders in Parkinson's Disease. Am. J. Psychiatry 176, 5–11. [PubMed: 30848949]

119. Feinberg I (1978). Efference Copy and Corollary Discharge: Implications for Thinking and Its Disorders. Schizophr. Bull 4, 636–640. [PubMed: 734369]

120. Lindner A, Thier P, Kircher TTJ, Haarmeier T, and Leube DT (2005). Disorders of Agency in Schizophrenia Correlate with an Inability to Compensate for the Sensory Consequences of Actions. Curr. Biol 15, 1119–1124. [PubMed: 15964277]

121. Frith C (2012). Explaining delusions of control: The comparator model 20 years on. Conscious. Cogn 21, 52–54. [PubMed: 21802318]

122. Griffin JD, and Fletcher PC (2017). Predictive Processing, Source Monitoring, and Psychosis. Annu. Rev. Clin. Psychol 13, 265–289. [PubMed: 28375719]

123. Kort NS, Ford JM, Roach BJ, Gunduz-Bruce H, Krystal JH, Jaeger J, Reinhart RMG, and Mathalon DH (2017). Role of N-Methyl-D-Aspartate Receptors in Action-Based Predictive Coding Deficits in Schizophrenia. Biol. Psychiatry 81, 514–524. [PubMed: 27647218]

124. Cassidy CM, Balsam PD, Weinstein JJ, Rosengard RJ, Slifstein M, Daw ND, Abi-Dargham A, and Horga G (2018). A Perceptual Inference Mechanism for Hallucinations Linked to Striatal Dopamine. Curr. Biol 28, 503–514.e4. [PubMed: 29398218]

125. Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, Petrovic P, Uhlhaas P, Voss M, and Corlett PR (2018). The Predictive Coding Account of Psychosis. Biol. Psychiatry 84, 634–643. [PubMed: 30007575]

126. McCutcheon RA, Abi-Dargham A, and Howes OD (2019). Schizophrenia, Dopamine and the Striatum: From Biology to Symptoms. Trends Neurosci. 42, 205–220. [PubMed: 30621912]

127. Ford JM, and Mathalon DH (2019). Efference Copy, Corollary Discharge, Predictive Coding, and Psychosis. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 4, 764–767. [PubMed: 31495399]

128. Schmack K, Bosc M, Ott T, Sturgill JF, and Kepecs A (2021). Striatal dopamine mediates hallucination-like perception in mice. Science 372. [PubMed: 34437095]

129. Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, and Yin HH (2015). Beyond reward prediction errors: the role of dopamine in movement kinematics. Front. Integr. Neurosci 9, 39. [PubMed: 26074791]

130. Panigrahi B, Martin KA, Li Y, Graves AR, Vollmer A, Olson L, Mensh BD, Karpova AY, and Dudman JT (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. Cell 162, 1418–1430. [PubMed: 26359992]

131. Howard CD, Li H, Geddes CE, and Jin X (2017). Dynamic Nigrostriatal Dopamine Biases Action Selection. Neuron 93, 1436–1450.e8. [PubMed: 28285820]

132. Lopes G, Bonacchi N, Frazao J, Neto JP, Atallah BV, Soares S, Moreira L, Matias S, Itskov PM, Correia PA, et al. (2015). Bonsai: an event-based framework for processing and controlling data streams. Front. Neuroinform 9, 7. [PubMed: 25904861]

133. Keithley RB, Mark Wightman R, and Heien ML (2009). Multivariate concentration determination using principal component regression with residual analysis. TrAC Trends Anal. Chem 28, 1127–1136.

134. Keithley RB, and Wightman RM (2011). Assessing Principal Component Regression Prediction of Neurochemicals Detected with Fast-Scan Cyclic Voltammetry. ACS Chem. Neurosci 2, 514–525. [PubMed: 21966586]

135. Rodeberg NT, Sandberg SG, Johnson JA, Phillips PEM, and Wightman RM (2017). Hitchhiker's Guide to Voltammetry: Acute and Chronic Electrodes for in Vivo Fast-Scan Cyclic Voltammetry. ACS Chem. Neurosci 8, 221–234. [PubMed: 28127962]

136. Howe MW, Tierney PL, Sandberg SG, Phillips PEM, and Graybiel AM (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500, 575–579. [PubMed: 23913271]

137. Nichols TE, and Holmes AP (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Brain Mapp 15, 1–25. [PubMed: 11747097]

138. Maris E, and Oostenveld R (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190. [PubMed: 17517438]

HIGHLIGHTS

- Nigrostriatal dopamine is suppressed when outcomes result from goal-directed action

- Dopamine signals action-outcome prediction errors for food reward and opto-ICSS

- These action-outcome prediction errors are precisely timed and sequence-specific
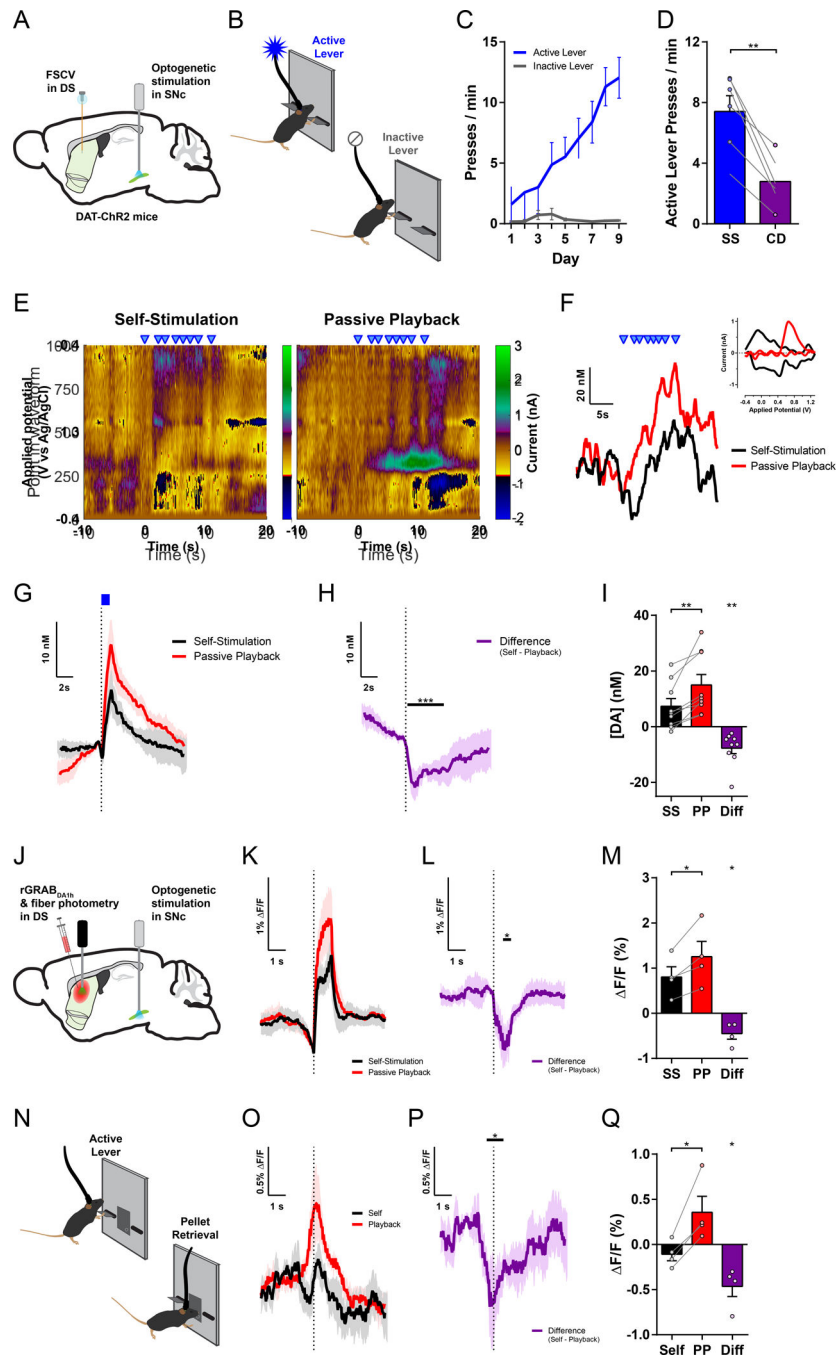
Figure 1. Goal-directed action suppresses nigrostriatal dopamine release

(**A**) Schematic of experimental preparation for optogenetic stimulation of SNc dopamine neurons and FSCV recording in dorsal striatum.

(**B**) Opto-ICSS behavioral task schematic.

(**C**) Acquisition of opto-ICSS ($n = 9$ mice; two-way repeated-measures ANOVA: main effect of Lever, $F_{1,8} = 16.84$, $P = 0.0034$; main effect of Day, $F_{8,64} = 9.314$, $P < 0.0001$; Lever by Day interaction, $F_{8,64} = 10.63$, $P < 0.0001$; Active Lever significantly greater than Inactive Lever for Days 4–9, Ps ≤ 0.008).

(**D**) Contingency degradation test: 30 min of opto-ICSS followed by 30 min contingency degradation test phase ($n = 6$ mice; paired t test, $t_5 = 5.441$, $P = 0.0028$).

(**E**) Representative voltammetric pseudocolor plots from a bout of stimulations (blue triangles) during the Self-Stimulation phase (left) and the matched stimulations from the Passive Playback phase (right).

(**F**) Dopamine responses to the series of stimulations in each session phase from the example in (E). Inset: cyclic voltammograms.

(**G**) Mean change in dopamine concentration evoked by Self-Stimulation and Passive Playback stimulations ($n = 9$ mice).

(**H**) Difference trace: Self-Stimulation minus Passive Playback, from traces in (G). Black bar indicates post-stimulation time points with significant difference vs. 0 (permutation test, $P = 0.0001$).

(**I**) Mean change in dopamine concentration (Self-Stimulation vs. Passive Playback, paired t test, $t_8 = 3.923$, $P = 0.0044$; equivalently, Difference vs. 0 one-sample t test: $t_8 = 3.923$, $P = 0.0044$).

(**J**) Schematic of experimental preparation for fiber photometry recordings of the red fluorescent dopamine sensor rGRAB$_{DA1h}$ in the dorsal striatum with optogenetic stimulation of SNc dopamine neurons.

(**K**) Mean   F/F rGRAB$_{DA1h}$ dopamine sensor response in dorsomedial striatum (DMS) evoked by Self-Stimulation and Passive Playback stimulations ($n = 4$ mice).

(**L**) Difference trace: Self-Stimulation minus Passive Playback, from traces in (K). Black bar, permutation test, $P = 0.0221$.

(**M**) Mean   F/F rGRAB$_{DA1h}$ dopamine sensor response (Self-Stimulation vs. Passive Playback paired t test, and Difference vs. 0 one-sample t test: $t_3 = 3.580$, $P = 0.0373$).

(**N**) Sucrose pellet CRF behavioral task schematic.

(**O**) Mean   F/F rGRAB$_{DA1h}$ dopamine sensor response aligned to sucrose pellet retrieval (1st head entry following reward delivery; $n = 4$ mice).

(**P**) Difference trace: Self minus Playback, from traces in (N). Black bar, permutation test, $P = 0.0196$.

(**Q**) Mean   F/F rGRAB$_{DA1h}$ dopamine sensor response to pellet retrieval (Self vs. Playback paired t test, and Difference vs. 0 one-sample t test: $t_3 = 4.127$, $P = 0.0258$).

FSCV, fast-scan cyclic voltammetry; DS, dorsal striatum; SNc, substantia nigra pars compacta; SS, Self-Stimulation; CD, Contingency Degradation; PP, Passive Playback; Diff, Difference (Self minus Playback). All error bars are SEM, same for below unless stated otherwise. See also Figure S1.
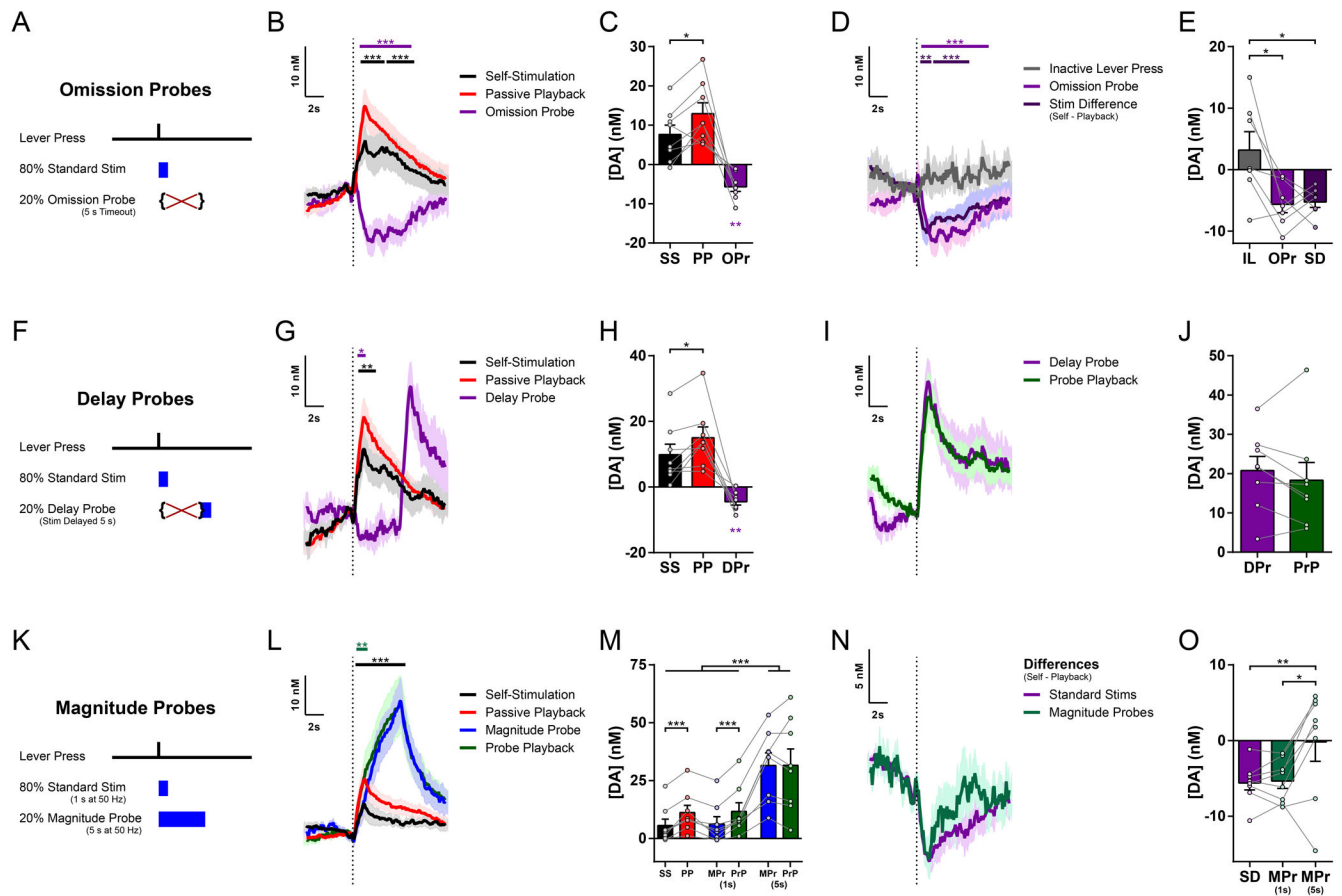
Figure 2. The inhibition of dopamine is action-specific and temporally precise

(**A**) Task schematic for Omission Probe FSCV sessions.

(**B**) Mean change in dopamine concentration for stimulations and probes in Omission Probe sessions ($n = 8$ mice; permutation tests: Omission Probe vs. 0, magenta bar, $P = 0.0001$; Self-Stimulation vs. Playback, black bar, $P = 0.0003$ for first time cluster and $P = 0.0001$ for second cluster).

(**C**) Mean change in dopamine concentration in Omission Probe sessions (Self-Stimulation vs. Passive Playback, paired t test, $t_7 = 3.114$, $P = 0.0170$; Omission Probe vs. 0, one-sample t test: $t_7 = 4.810$, $P = 0.0019$).

(**D**) Mean change in dopamine concentration for Inactive Lever presses, overlaid on Omission Probes from (B) and Difference traces (Self Stimulation minus Playback) from Fig. 1H ($n = 7$ mice; permutation tests: Inactive press vs. Omission Probe, magenta bar, $P = 0.0001$; Inactive press vs. Stim Difference, purple bar, $P = 0.007$ for first time cluster and $P = 0.0007$ for second cluster.

(**E**) Mean change in dopamine concentration for Inactive Lever presses versus Omission Probes from (C) and stimulation Differences from Fig. 1I (one-way repeated-measures ANOVA, $F_{2,12} = 7.419$, $P = 0.0080$; Tukey's multiple comparisons tests: Inactive Lever vs. Omission Probe, $P = 0.0133$; Inactive Lever vs. Stimulation Difference, $P = 0.0174$).

(**F**) Task schematic for Delay Probe FSCV sessions.

(**G**) Mean change in dopamine concentration for Standard Stimulations and probes in Delay Probe sessions ($n = 8$ mice; permutation tests: Delay Probe timeout period vs. 0, magenta bar, $P = 0.0103$; standard Self-Stimulation vs. Playback, black bar, $P = 0.0001$).

(**H**) Mean change in dopamine concentration in Delay Probe sessions (Self-Stimulation vs. Passive Playback, paired t test, $t_7 = 2.962$, $P = 0.0210$; Delay Probe vs. 0, one-sample t test: $t_7 = 4.341$, $P = 0.0034$).

(**I**) Mean change in dopamine concentration for Delay Probes and Probe Playback, each aligned to stimulation onset.

(**J**) Mean change in dopamine concentration for Delay Probe stimulation and Probe Playback.

(**K**) Task schematic for Magnitude Probe FSCV sessions.

(**L**) Mean change in dopamine concentration for standard stimulations and probes in Magnitude Probe sessions ($n = 8$ mice; permutation tests: Magnitude Probe vs. Probe Playback, teal bar, $P = 0.0052$; standard Self-Stimulation vs. Playback, black bar, $P = 0.0001$).

(**M**) Mean change in dopamine concentration in Magnitude Probe sessions (two-way repeated-measures ANOVA: main effect of Session Phase (Self vs. Playback), $F_{1,7} = 6.769$, $P = 0.0353$; main effect of Stimulation Type (Standard, early and late Magnitude Probe), $F_{2,14} = 31.32$, $P < 0.0001$; Session Phase by Stimulation Type interaction, $F_{2,14} = 7.724$, $P = 0.0055$; Sidak's multiple comparisons tests: late Magnitude Probes vs. Standard Stimulations and vs. early Magnitude Probes within both Self and Playback phases, all Ps < 0.0001; Self vs. Playback for Standard Stimulations, $P = 0.0006$; Self vs. Playback for early Magnitude Probes, $P = 0.0008$).

(**N**) Difference traces: Self-Stimulation minus Passive Playback for Standard Stimulations and Magnitude Probes, from traces in (L).

(**O**) Mean Differences comparing session phases (Self-Stimulation minus Passive Playback) for Standard Stimulations and Magnitude Probes at early (1 s) and late (5 s) time points from Difference traces in (N). (One-way repeated-measures ANOVA, $F_{2,14} = 7.653$, $P = 0.0057$; Tukey's multiple comparisons tests: Standard Stimulation Differences vs. late Magnitude Probe Differences, $P = 0.0099$; early vs. late Magnitude Probe Differences, $P = 0.0136$). SS, Self-Stimulation; PP, Passive Playback; OPr, Omission Probe; IL, Inactive Lever; SD, Standard Stimulation Difference (Self minus Playback); DPr, Delay Probe; PrP, Probe Playback; MPr, Magnitude Probe. See also Figure S2.
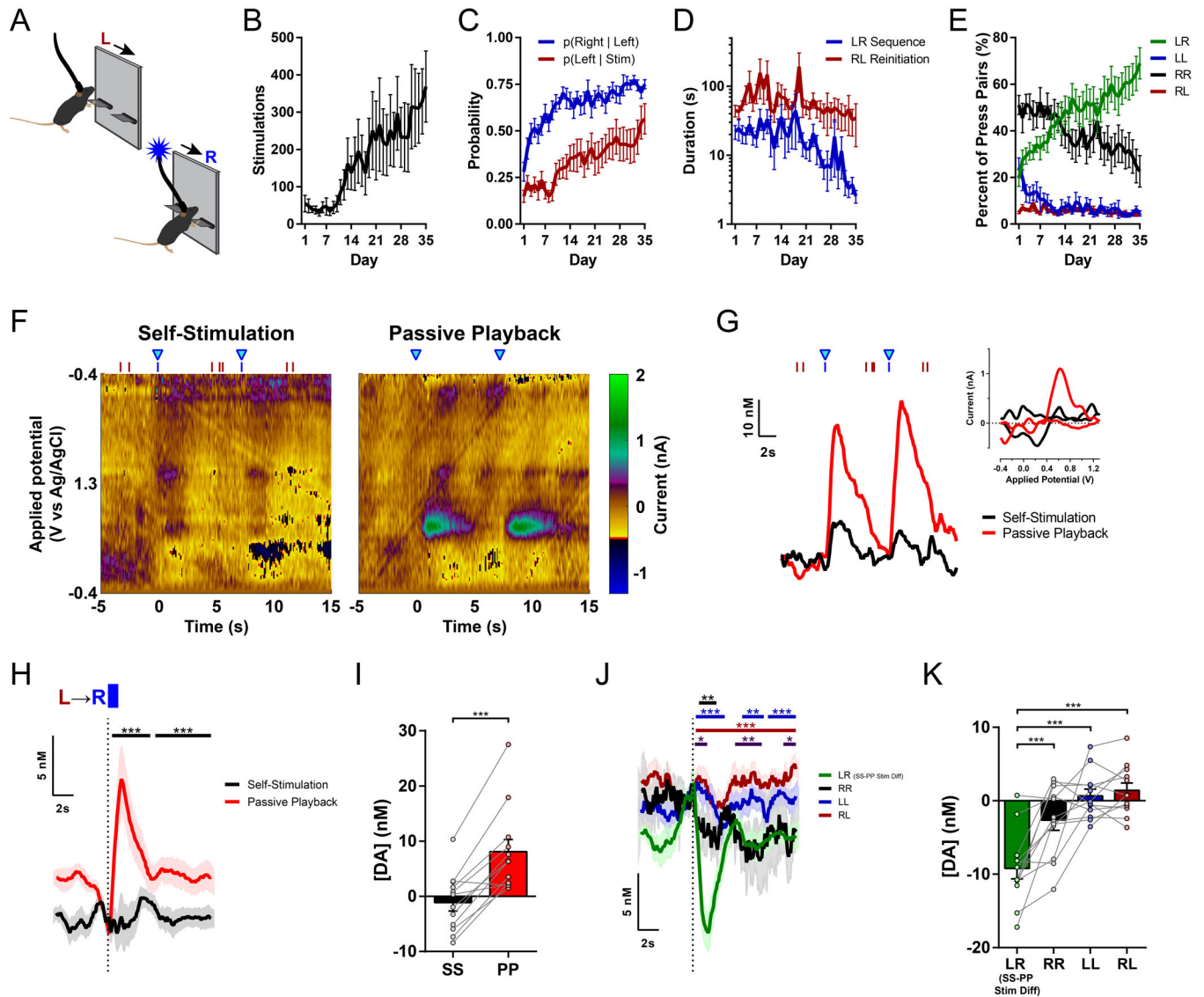
Figure 3. Sequence-specific inhibition of nigrostriatal dopamine during performance of learned LR sequence

(**A**) Left-Right Sequence Self-Stimulation task schematic.

(**B**) Stimulations earned across days of training ($n = 13$ mice; one-way repeated-measures ANOVA: $F_{34,408} = 3.550$, $P < 0.0001$).

(**C**) Transition probabilities: probability of pressing Right lever after each Left lever press, and probability of reinitiating Left press after stimulation (two-way repeated-measures ANOVA: main effect of Day, $F_{34,408} = 8.838$, $P < 0.0001$; main effect of Transition Type, $F_{1,12} = 47.18$, $P < 0.0001$; Dunnett's multiple comparisons tests vs. Day 1: p(Right | Left) first significantly differs on Day 3, $P = 0.0266$; p(Left | Stim) first significantly differs on Day 14, $P = 0.0265$).

(**D**) Median latencies to complete Left-Right sequences and to reinitiate a new sequence after previous stimulation (two-way repeated-measures ANOVA: main effect of Interval Type, $F_{1,12} = 18.79$, $P = 0.0010$).

(E) Relative frequency of each combination of lever press pairs (two-way repeated-measures ANOVA: main effect of Pair Type, $F_{3,36} = 32.58$, $P < 0.0001$; Pair Type by Day interaction, $F_{102,1224} = 4.315$, $P < 0.0001$).

(F) Representative voltammetric pseudocolor plots during two Left-Right sequences for optogenetic intracranial Self-Stimulation (left) and corresponding stimulations from the Passive Playback phase (right).

(G) Dopamine responses to the stimulations depicted in (F). Red and blue ticks denote left and right lever presses, respectively. Inset: cyclic voltammograms evoked by the first stimulation.

(H) Mean dopamine concentration changes evoked by stimulation in each phase ($n = 12$ mice; permutation test, Ps = 0.0001 for both time clusters).

(I) Mean change in dopamine concentration (paired t test, $t_{11} = 6.403$, $P < 0.0001$).

(J) Mean dopamine concentration changes evoked by non-stimulated pairs of lever presses (5-s maximum inter-press interval within pair). Traces are aligned to the second press in each pair type. The LR sequence stimulation Difference (Self-Stimulation minus Passive Playback) is overlaid for comparison (green). (Permutation tests: LR Stim Difference vs. RR, black bar, $P = 0.0046$; LR Stim Difference vs. LL, blue bars, Ps = 0.0002 for first time cluster, 0.0017 for second, and 0.0007 for third cluster; LR Stim Difference vs. RL, maroon bar, $P = 0.0001$; RR vs. RL, purple bars, Ps = 0.0185 for first time cluster, 0.0021 for second, and 0.0163 for third cluster).

(K) Mean change in dopamine concentration for each combination of non-reinforced press pairs during the Self-Stimulation phase, and the LR sequence stimulation Difference (Self minus Playback). (One-way repeated-measures ANOVA, $F_{3,33} = 20.34$, $P < 0.0001$; Tukey's multiple comparisons tests: LR Stim Difference vs. RR, $P = 0.0007$; LR Stim Difference vs. LL and vs. RL, Ps < 0.0001; one-sample t tests of each combination vs. 0: RR, $t_{11} = 1.919$, $P = 0.0813$; LL, $t_{11} = 0.6952$, $P = 0.5013$; RL, $t_{11} = 1.434$, $P = 0.1795$; LR Stim Difference, $t_{11} = 6.401$, $P < 0.0001$).

Stim, Stimulation; LR, Left-Right sequence; RR, Right-Right; LL, Left-Left; RL, Right-Left; SS, Self-Stimulation; PP, Passive Playback; Diff, Difference. See also Figure S3.
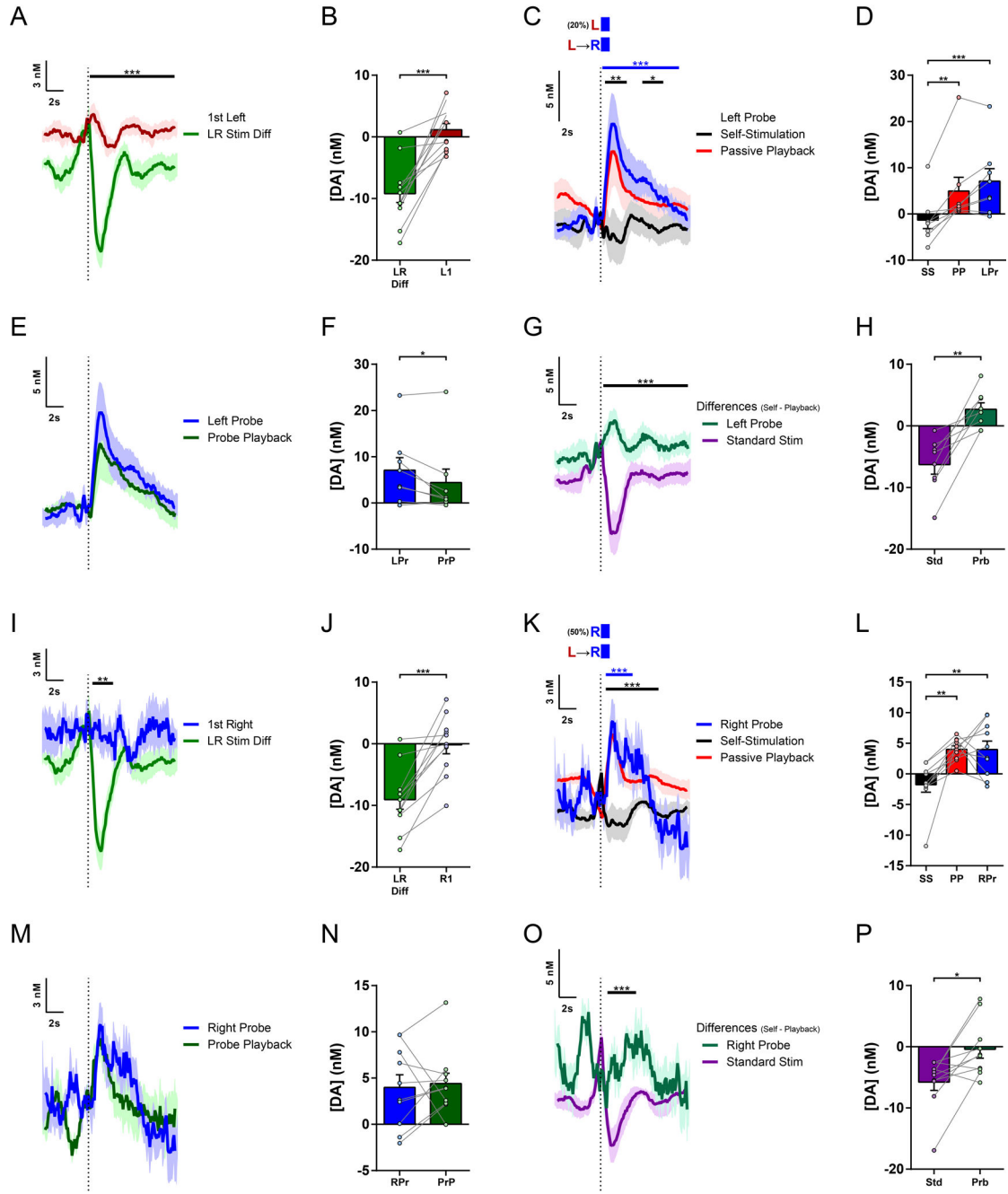
Figure 4. Different regulation of dopamine by individual actions of the LR sequence

(**A**) Mean dopamine concentration change to first Left lever press following previous stimulation during Left-Right sequence task performance, overlaid with the Difference Trace (Self minus Playback) for LR sequence stimulations from Fig. 3J for comparison ($n = 12$ mice; permutation test, $P = 0.0001$).

(**B**) Mean change in dopamine concentration for first Left lever presses and LR sequence stimulation Difference ($t_{11} = 6.325$, $P < 0.0001$).

(**C**) Mean dopamine concentration changes in Left-Right sequence sessions with Left Lever Probes ($n = 8$ mice; permutation tests: Left Probe vs. Standard LR Self-Stimulation, blue bar, $P = 0.0001$; Standard LR Self-Stimulation vs. Playback, black bars, $Ps = 0.0049$ for first and 0.011 for second time clusters, respectively).

(**D**) Mean change in dopamine concentration in Left Probe sessions (one-way repeated-measures ANOVA, $F_{2,14} = 17.19$, $P = 0.0002$; Tukey's multiple comparisons tests: Left Probe vs. LR Self-Stimulation, $P = 0.0002$; LR Self-Stimulation vs. Playback, $P = 0.0024$).

(**E**) Mean dopamine concentration changes in Left Probes and the Probe Playback.

(**F**) Mean change in dopamine concentration to Left Probes and Probe Playback (paired t test, $t_7 = 2.519$, $P = 0.0399$).

(**G**) Difference traces from Left Probe Sessions: Self-Stimulation minus Passive Playback for Standard LR Stimulations and Left Probes, from traces in (C) and (E); (permutation test, $P = 0.0001$).

(**H**) Mean Differences comparing session phases (Self-Stimulation minus Passive Playback) for Standard LR Stimulations and Left Probes (paired t test, $t_7 = 5.125$, $P = 0.0014$).

(**I**) Mean dopamine concentration change to first Right lever press following previous stimulations, overlaid with the Difference Trace (Self minus Playback) for LR sequence stimulations from Fig. 3J for comparison. First Right press was an additional press on the Right lever after a previous stimulation, without approaching the Left lever ($n = 11$ mice; permutation test, $P = 0.0012$).

(**J**) Mean change in dopamine concentration for first Right lever presses and LR sequence stimulation Difference ($t_{10} = 5.690$, $P = 0.0002$).

(**K**) Mean dopamine concentration changes in Left-Right sequence sessions with Right Lever Probes where animal did not approach the Left lever in the preceding inter-stimulation interval ($n = 10$ mice; permutation tests: Right Probe vs. Standard LR Self-Stimulation, blue bar, $P = 0.0001$; Standard LR Self-Stimulation vs. Playback, black bar, $P = 0.0001$).

(**L**) Mean change in dopamine concentration in Right Probe sessions (one-way repeated-measures ANOVA, $F_{2,18} = 10.47$, $P = 0.0010$; Tukey's multiple comparisons tests: Right Probe vs. LR Self-Stimulation, $P = 0.0026$; LR Self-Stimulation vs. Playback, $P = 0.0024$).

(**M**) Mean dopamine concentration changes in Right Probes and the Probe Playback.

(**N**) Mean change in dopamine concentration to Right Probes and Probe Playback.

(**O**) Difference traces from Right Probe Sessions: Self-Stimulation minus Passive Playback for Standard LR Stimulations and Right Probes, from traces in (K) and (M); (permutation test, $P = 0.0002$).

(**P**) Mean Differences comparing session phases (Self-Stimulation minus Passive Playback) for Standard LR Stimulations and Right Probes (paired t test, $t_9 = 3.080$, $P = 0.0131$).

LR Stim Diff, Left-Right Stimulation Difference (Self minus Playback); L1, 1st Left Press; R1, 1st Right Press; SS, Self-Stimulation; PP, Passive Playback; LPr, Left Probe; PrP, Probe Playback; Std, Standard LR Stimulation; Prb, Probe; RPr, Right Probe. See also Figure S4.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-tyrosine hydroxylase	Abcam	Cat#ab112; RRID:AB_297840
Chicken anti-GFP	Novus Biologicals	Cat#NB100-1614; RRID:AB_10001164
Donkey anti-rabbit Cy3	Jackson ImmunoResearch	Cat#711-165-152; RRID:AB_2307443
Donkey anti-chicken Alexa Fluor 488	Jackson ImmunoResearch	Cat#703-545-155; RRID:AB_2340375
Bacterial and virus strains		
AAV5-EF1a-DIO-ChR2(H134R)-mCherry	UNC vector core	RRID:SCR_002448
AAV1-hSyn-GRAB_rDA1h	[31] Addgene, Salk vector core	RRID:Addgene_140557
Deposited data		
Raw and analyzed data	This paper	Zenodo: https://doi.org/10.5281/zenodo.5501733
Experimental models: Organisms/strains		
Mouse: DAT-cre (*Slc6a3*)	Jackson Laboratory	RRID:IMSR_JAX:020080
Mouse: Ai32 (RCL-ChR2(H134R)/eYFP)	Jackson Laboratory	RRID:IMSR_JAX:024109
Software and algorithms		
Med-PC	Med Associates	Cat#SOF-735
Tarheel CV	UNC, via University of Washington	N/A
Bonsai	Bonsai	RRID:SCR_017218
MATLAB	Mathworks	Version 2017b
Prism	GraphPad Software	Version 6.07
OmniPlex	Plexon	Version 1.4.5
Offline Sorter	Plexon	Version 3.3.3