Report

Current Biology

Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors

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

Authors

Nick G. Hollon, Elora W. Williams, Christopher D. Howard, Hao Li, Tavish I. Traut, Xin Jin

Correspondence

xjin@bio.ecnu.edu.cn

In brief

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.

Hollon et al., 2021, Current Biology *31*, 1–14
December 6, 2021
https://doi.org/10.1016/j.cub.2021.09.040

CellPress

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

Current Biology

CellPress

Report

Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors

Nick G. Hollon,[1] Elora W. Williams,[1] Christopher D. Howard,[1,4] Hao Li,[1] Tavish I. Traut,[1] and Xin Jin[1,2,3,5,*]
[1]Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[2]Center for Motor Control and Disease, Key Laboratory of Brain Functional Genomics, East China Normal University, Shanghai 200062, China
[3]NYU–ECNU Institute of Brain and Cognitive Science, New York University Shanghai, Shanghai 200062, China
[4]Present address: Neuroscience Department, Oberlin College, Oberlin, OH 44074, USA
[5]Lead contact
*Correspondence: xjin@bio.ecnu.edu.cn
https://doi.org/10.1016/j.cub.2021.09.040

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. 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 and 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. 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.

RESULTS

Suppression of optogenetically stimulated nigrostriatal dopamine by goal-directed action

Mice expressing channelrhodopsin-2 selectively in their dopamine neurons[16,17] (see Method details) 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, S1A, and S1B). 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 (Method details). The mice significantly reduced their performance rate (Figures 1D, S1C, and S1D), 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

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

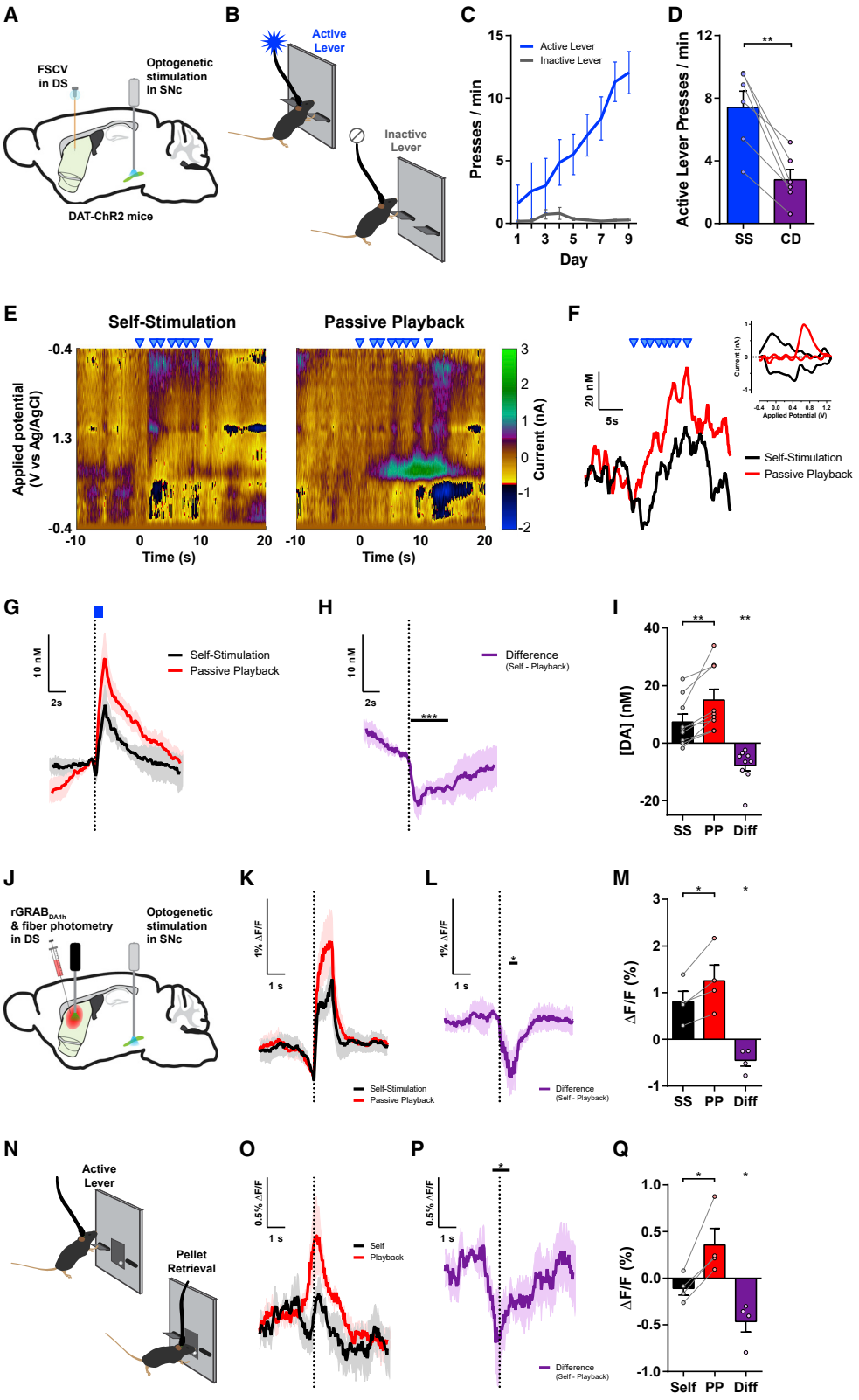

Figure 1. Goal-directed action suppresses nigrostriatal dopamine release

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

(B) Opto-ICSS behavioral task schematic.

(legend continued on next page)

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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, 1E, and 1F).

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 (Figures 1F–1I). 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 ($p < 0.0001$) and was a trend in the same direction for the remaining 2 mice ($p = 0.0623$ 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 differ significantly (Figure S1F), consistent with the transient suppression observed in the self-versus-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 may 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 (Figures S1G–S1I). Self-stimulations following shorter ISIs evoked less dopamine release than did those with long ISIs (Figures S1J and S1K), 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-versus-playback stimulation difference was not significantly related to the ISI (Figures S1J–S1M), highlighting the importance of temporally matched stimulations in the current experimental design. The self-versus-playback difference also was stable across early and late stimulations within this recorded session (Figures S1N–S1Q), 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–S1U), we recapitulated the suppression of self-stimulated dopamine release in the dorsomedial striatum (DMS; Figures 1K–1M), 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 (Figures 1N–1Q). We also found similar suppression of the optogenetically evoked $rGRAB_{DA1h}$ response in the dorsolateral striatum (DLS; Figures S1V–S1Y), 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 (Figures 2B and 2C), consistent with a neurochemical instantiation of a negative prediction error.[15] The time course for this omission probe dip was

(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, $p \leq 0.008$).
(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 voltametric 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 versus 0 (permutation test, $p = 0.0001$).
(I) Mean change in dopamine concentration (self-stimulation versus passive playback, paired t test, $t_8 = 3.923$, $p = 0.0044$; equivalently, difference versus 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 $\Delta F/F$ $rGRAB_{DA1h}$ 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 $\Delta F/F$ $rGRAB_{DA1h}$ dopamine sensor response (self-stimulation versus passive playback paired t test, and difference versus 0 one-sample t test: $t_3 = 3.580$, $p = 0.0373$).
(N) Sucrose pellet CRF behavioral task schematic.
(O) Mean $\Delta F/F$ $rGRAB_{DA1h}$ 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 (O). Black bar, permutation test, $p = 0.0196$.
(Q) Mean $\Delta F/F$ $rGRAB_{DA1h}$ dopamine sensor response to pellet retrieval (self versus playback paired t test, and difference versus 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 SEMs; same for below unless stated otherwise.
See also Figure S1.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

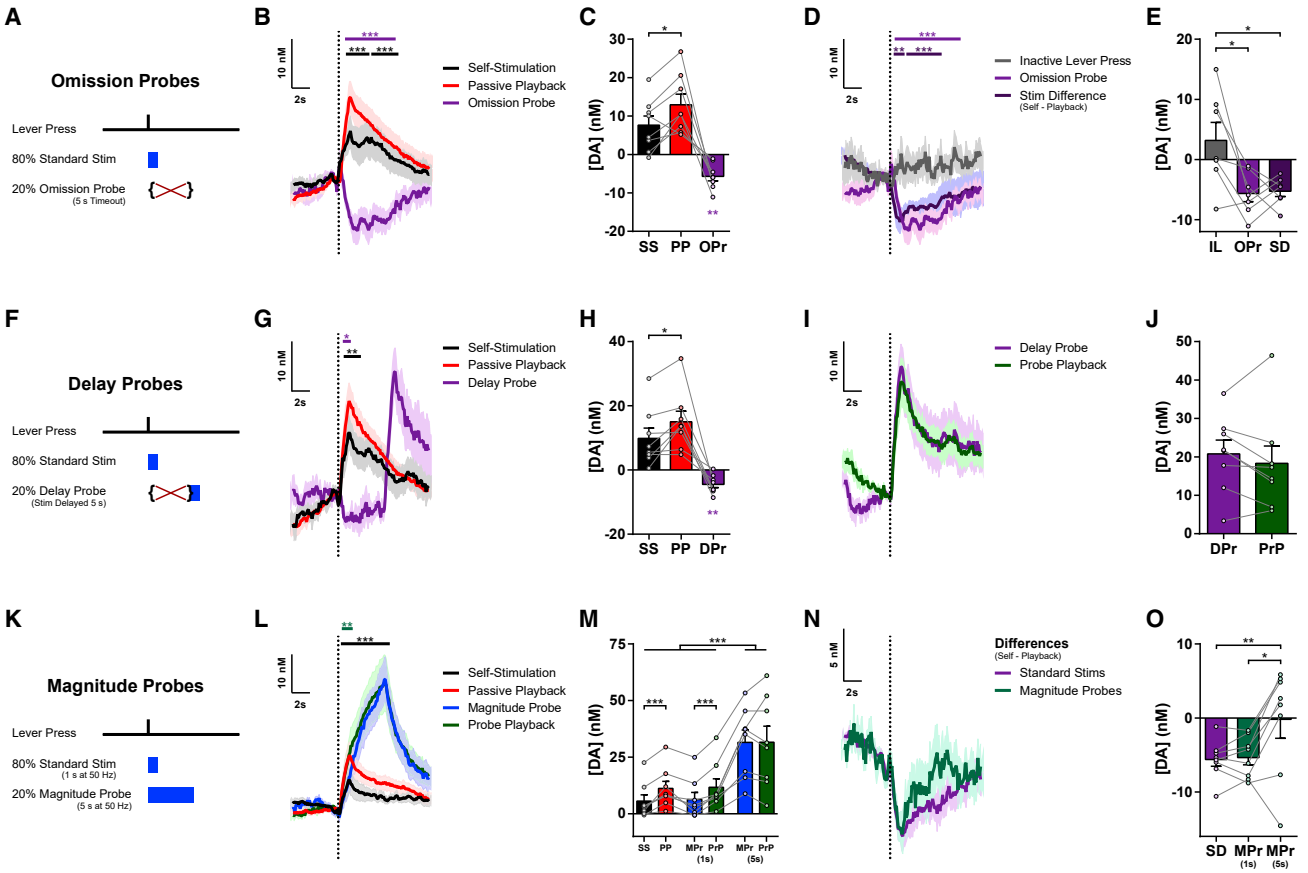

Figure 2. The inhibition of dopamine is action-specific and temporally precise

(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 versus 0, magenta bar, p = 0.0001; self-stimulation versus 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 versus passive playback, paired t test, $t_7 = 3.114$, p = 0.0170; omission probe versus 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 Figure 1H (n = 7 mice; permutation tests: inactive press versus omission probe, magenta bar, p = 0.0001; inactive press versus 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 Figure 1I (one-way repeated-measures ANOVA, $F_{2,12} = 7.419$, p = 0.0080; Tukey's multiple comparisons tests: inactive lever versus omission probe, p = 0.0133; inactive lever versus 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 versus 0, magenta bar, p = 0.0103; standard self-stimulation versus playback, black bar, p = 0.0016).

(H) Mean change in dopamine concentration in delay probe sessions (self-stimulation versus passive playback, paired t test, $t_7 = 2.962$, p = 0.0210; delay probe versus 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 versus probe playback, teal bar, p = 0.0052; standard self-stimulation versus 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 versus 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 versus standard stimulations and versus early magnitude probes within both self and playback phases, all p < 0.0001; self versus playback for standard stimulations, p = 0.0006; self versus 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).

(legend continued on next page)

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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-versus-playback difference (Figure S2A). This omission probe dip was not merely an artifact of FSCV background subtraction,[35] in which reuptake during the stimulation-free timeout period may 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 (Figures S2B and S2C). 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 (Figures S2D–S2G). *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 (Figures S2H–S2M). Collectively, these results demonstrate that the action 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 in which the animal pressed the inactive lever, which had never been reinforced throughout training (Figures 2D and 2E), 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 (Figures 2G and 2H). 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 (Figures 2I and 2J). 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, in which 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 evoked much greater dopamine release, as expected for longer-duration stimulation (Figures 2L and 2M). 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 (Figures 2N and 2O). 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. 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 lever and then the right lever (LR) to earn optogenetic nigrostriatal dopamine neuron self-stimulation (Method details; Figures 3A, S3A, and S3B). 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 (Figures S3C–S3F). Therefore,

(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 versus late magnitude probe differences, p = 0.0099; early versus 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.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

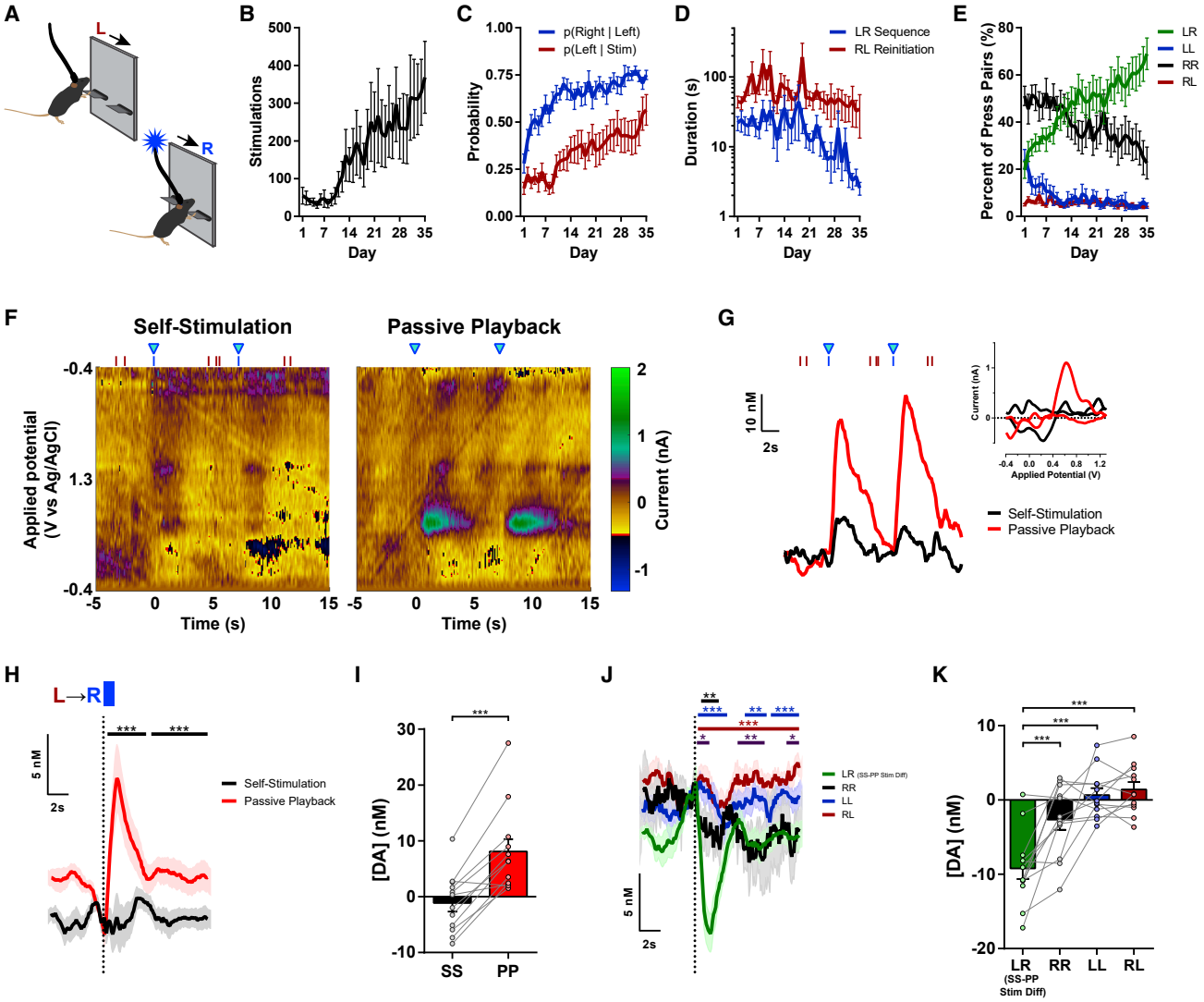

Figure 3. Sequence-specific inhibition of nigrostriatal dopamine during performance of learned LR sequence

(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 versus 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 voltametric 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, $p = 0.0001$ 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 versus RR, black bar, $p = 0.0046$; LR stim difference versus LL, blue bars, $p = 0.0002$ for first time cluster, 0.0017 for second, and 0.0007 for third cluster; LR stim difference versus RL, maroon bar, $p = 0.0001$; RR versus RL, purple bars, $p = 0.0185$ 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

(legend continued on next page)

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

rather than simply associating the reinforcing outcome with the most proximal action at the right lever, the animals' behavior suggested that they 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–3I, S3H, and S3I), recapitulating the main result from the single-lever CRF cohort (Figures 1E–1I). We also examined dopamine transmission aligned to the completion of other combinations of non-reinforced press pairs (Figures 3J and 3K), essentially the multi-press analogs of the non-reinforced inactive lever presses from the single-lever CRF cohort (Figures 2D and 2E). 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 (Figures 3J and 3K). 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 may 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, then we may expect similar inhibition for each individual left and right lever press, and the summation of each to the full inhibition at outcome delivery. Alternatively, since animals chunked these action elements into fully concatenated action sequences, we may expect the action-induced inhibition of dopamine to begin at sequence initiation and persist throughout the 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, 4B, 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 (Figures 4C–4H). Instead, these left probes actually evoked significantly greater dopamine release than their playback (Figure 4F). At the individual animal level, 5 of 8 mice (62.5%) showed a significant dopamine increase, and none showed a significant suppression. Outside of LR sequences, single right lever presses (Method details) did not result in the inhibition of dopamine, in stark contrast to the full inhibition of self-stimulated versus playback-evoked dopamine for correct LR sequences (Figures 4I and 4J). The dopamine response to probe stimulations for these isolated right presses also did not differ from their playback, exhibiting no significant inhibition (Figures 4K–4P). 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 (Figures 4K and 4L). 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] 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

difference versus RR, $p = 0.0007$; LR stim difference versus LL and versus RL, $p < 0.0001$; one-sample t tests of each combination versus 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.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

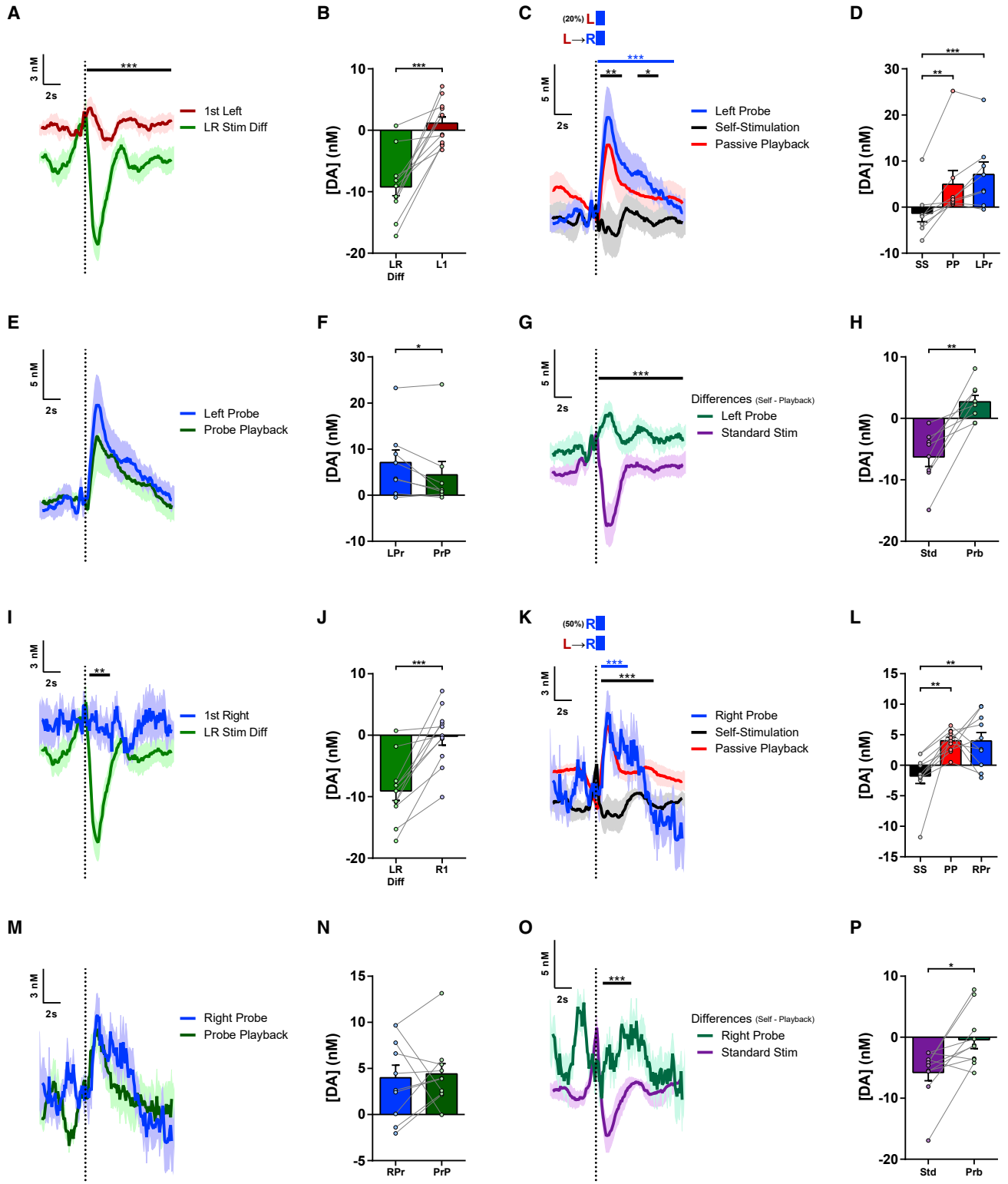

Figure 4. Different regulation of dopamine by individual actions of the LR sequence

(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 Figure 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).

(legend continued on next page)

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

regulation provides critical new insights 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 present 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 naive 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 may 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 the 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 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, 2E, 3J, and 3K). 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. The distinct regulation of dopamine to the same action depending on sequence membership (Figures 4I–4P) 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] 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

(C) Mean dopamine concentration changes in left-right sequence sessions with left lever probes (n = 8 mice; permutation tests: left probe versus standard LR self-stimulation, blue bar, p = 0.0001; standard LR self-stimulation versus playback, black bars, p = 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 versus LR self-stimulation, p = 0.0002; LR self-stimulation versus 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 Figure 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 in which animal did not approach the left lever in the preceding inter-stimulation interval (n = 10 mice; permutation tests: right probe versus standard LR self-stimulation, blue bar, p = 0.0001; standard LR self-stimulation versus 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 versus LR self-stimulation, p = 0.0026; LR self-stimulation versus 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.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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 FSCV recordings of the present study 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 the regulation of dopamine dynamics differs across distinct striatal subregions, how these dynamics evolve throughout learning, and how each may 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] 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 naive animals. Furthermore, 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] 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 a 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 present study would be largely similar in the accumbens core[30] and DLS, although we also may 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, the 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 present 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 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] although 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 may 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- versus passive optogenetic stimulation (Figures S2H–S2M) 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 present 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 disruption of 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 the 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 may regulate ongoing actions[36,39,129,130] and bias online action selection.[131] The present 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

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
- METHOD DETAILS
  - Surgical procedures
  - Behavioral training
  - Continuous reinforcement cohort
  - Left-Right sequence cohort
  - Contingency degradation
  - FSCV
  - Fiber photometry
  - *In vivo* electrophysiology
  - Histology
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.09.040.

ACKNOWLEDGMENTS

We thank Jared Smith, Jason Klug, Sho Aoki, Zhongmin Lu, Roy Kim, Kanchi Mehta, Anthony Balolong-Reyes, Sage Aronson, and Scott Ng-Evans for helpful discussions and technical assistance. The research reported in this article was supported by NIH grants K99MH119312 (to N.G.H.) and R01NS083815 (to X.J.), the Salk Institute Pioneer Postdoctoral Endowment Fund and the Jonas Salk Fellowship (to N.G.H.), and the McKnight Memory and Cognitive Disorders Award (to X.J.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS

X.J., N.G.H., and C.D.H. conceived the project. N.G.H. and X.J. designed the experiments. N.G.H., E.W.W., C.D.H., H.L., and T.I.T. conducted the experiments. N.G.H. and H.L. analyzed the data. N.G.H. and X.J. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 11, 2021
Revised: August 31, 2021
Accepted: September 15, 2021
Published: October 11, 2021

REFERENCES

1. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.
2. Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. Science *299*, 1898–1902.
3. Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85–88.
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.
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.
6. Hollon, N.G., Arnold, M.M., Gan, J.O., Walton, M.E., and Phillips, P.E.M. (2014). Dopamine-associated cached values are not sufficient as the basis for action selection. Proc. Natl. Acad. Sci. USA *111*, 18357–18362.
7. Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., and Witten, I.B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat. Neurosci. *19*, 845–854.
8. Coddington, L.T., and Dudman, J.T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nat. Neurosci. *21*, 1563–1573.
9. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., and Witten, I.B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature *570*, 509–513.
10. Kremer, Y., Flakowski, J., Rohner, C., and Lüscher, C. (2020). Context-dependent multiplexing by individual VTA dopamine neurons. J. Neurosci. *40*, 7489–7509.
11. Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature *459*, 837–841.
12. Kim, H.F., Ghazizadeh, A., and Hikosaka, O. (2015). Dopamine neurons encoding long-term memory of object value for habitual behavior. Cell *163*, 1165–1175.
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.
14. Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci. *10*, 1615–1624.
15. Hart, A.S., Rutledge, R.B., Glimcher, P.W., and Phillips, P.E.M. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J. Neurosci. *34*, 698–704.
16. Zhuang, X., Masson, J., Gingrich, J.A., Rayport, S., and Hen, R. (2005). Targeted gene expression in dopamine and serotonin neurons of the mouse brain. J. Neurosci. Methods *143*, 27–32.

17. Madisen, L., Mao, T., Koch, H., Zhuo, J.M., Berenyi, A., Fujisawa, S., Hsu, Y.-W.A., Garcia, A.J., 3rd, 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.

18. Sparta, D.R., Stamatakis, A.M., Phillips, J.L., Hovelsø, N., van Zessen, R., and Stuber, G.D. (2011). Construction of implantable optical fibers for long-term optogenetic manipulation of neural circuits. Nat. Protoc. *7*, 12–23.

19. Clark, J.J., Sandberg, S.G., Wanat, M.J., Gan, J.O., Horne, E.A., Hart, A.S., Akers, C.A., Parker, J.G., Willuhn, I., Martinez, V., et al. (2010). Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nat. Methods *7*, 126–129.

20. Saunders, B.T., Richard, J.M., Margolis, E.B., and Janak, P.H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat. Neurosci. *21*, 1072–1083.

21. Rossi, M.A., Sukharnikova, T., Hayrapetyan, V.Y., Yang, L., and Yin, H.H. (2013). Operant self-stimulation of dopamine neurons in the substantia nigra. PLoS ONE *8*, e65799.

22. Ilango, A., Kesner, A.J., Keller, K.L., Stuber, G.D., 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.

23. Keiflin, R., Pribut, H.J., Shah, N.B., and Janak, P.H. (2019). Ventral tegmental dopamine neurons participate in reward identity predictions. Curr. Biol. *29*, 93–103.e3.

24. Witten, I.B., Steinberg, E.E., Lee, S.Y., Davidson, T.J., Zalocusky, K.A., Brodsky, M., Yizhar, O., Cho, S.L., Gong, S., Ramakrishnan, C., et al. (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron *72*, 721–733.

25. Koralek, A.C., Jin, X., Long, J.D., 2nd, Costa, R.M., and Carmena, J.M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature *483*, 331–335.

26. Clancy, K.B., Koralek, A.C., Costa, R.M., Feldman, D.E., and Carmena, J.M. (2014). Volitional modulation of optically recorded calcium signals during neuroprosthetic learning. Nat. Neurosci. *17*, 807–809.

27. Neely, R.M., Koralek, A.C., Athalye, V.R., Costa, R.M., and Carmena, J.M. (2018). Volitional modulation of primary visual cortex activity requires the basal ganglia. Neuron *97*, 1356–1368.e4.

28. Yin, H.H., Ostlund, S.B., and Balleine, B.W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur. J. Neurosci. *28*, 1437–1448.

29. Balleine, B.W. (2019). The meaning of behavior: discriminating reflex and volition in the brain. Neuron *104*, 47–62.

30. Covey, D.P., and Cheer, J.F. (2019). Accumbal dopamine release tracks the expectation of dopamine neuron-mediated reinforcement. Cell Rep. *27*, 481–490.e3.

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.

32. Houk, J.C., Adams, J.L., and Barto, A.G. (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, , J.C. Houk, J.L. Davis, and D.G. Beiser, eds. (MIT Press), pp. 249–270.

33. Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. *16*, 1936–1947.

34. Sutton, R.S., and Barto, A.G. (2018). Reinforcement Learning: An Introduction, Second Edition (MIT Press).

35. Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J., and Berke, J.D. (2016). Mesolimbic dopamine signals the value of work. Nat. Neurosci. *19*, 117–126.

36. da Silva, J.A., Tecuapetla, F., Paixão, V., and Costa, R.M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. Nature *554*, 244–248.

37. Dodson, P.D., Dreyer, J.K., Jennings, K.A., Syed, E.C.J., Wade-Martins, R., Cragg, S.J., Bolam, J.P., and Magill, P.J. (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. Proc. Natl. Acad. Sci. USA *113*, E2180–E2188.

38. Gallistel, C.R. (1980). The Organization of Action (Lawrence Erlbaum Associates).

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

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

41. Geddes, C.E., Li, H., and Jin, X. (2018). Optogenetic editing reveals the hierarchical organization of learned action sequences. Cell *174*, 32–43.e15.

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

43. Jin, X., and Costa, R.M. (2015). Shaping action sequences in basal ganglia circuits. Curr. Opin. Neurobiol. *33*, 188–196.

44. Rao, R.P.N., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87.

45. Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive processing: a canonical cortical computation. Neuron *100*, 424–435.

46. Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). An internal model for sensorimotor integration. Science *269*, 1880–1882.

47. Crapse, T.B., and Sommer, M.A. (2008). Corollary discharge across the animal kingdom. Nat. Rev. Neurosci. *9*, 587–600.

48. Schneider, D.M., Sundararajan, J., and Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. Nature *561*, 391–395.

49. Wurtz, R.H. (2018). Corollary discharge contributions to perceptual continuity across saccades. Annu. Rev. Vis. Sci. *4*, 215–237.

50. Graybiel, A.M. (1998). The basal ganglia and chunking of action repertoires. Neurobiol. Learn. Mem. *70*, 119–136.

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.

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

53. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

54. Wassum, K.M., Ostlund, S.B., and Maidment, N.T. (2012). Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. Biol. Psychiatry *71*, 846–854.

55. Collins, A.L., Greenfield, V.Y., Bye, J.K., Linker, K.E., Wang, A.S., and Wassum, K.M. (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. Sci. Rep. *6*, 20231.

56. Garris, P.A., Kilpatrick, M., Bunin, M.A., Michael, D., Walker, Q.D., and Wightman, R.M. (1999). Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature *398*, 67–69.

57. Kilpatrick, M.R., Rooney, M.B., Michael, D.J., and Wightman, R.M. (2000). Extracellular dopamine dynamics in rat caudate-putamen during experimenter-delivered and intracranial self-stimulation. Neuroscience *96*, 697–706.

58. Owesson-White, C.A., Cheer, J.F., Beyene, M., Carelli, R.M., and Wightman, R.M. (2008). Dynamic changes in accumbens dopamine correlate with learning during intracranial self-stimulation. Proc. Natl. Acad. Sci. USA *105*, 11957–11962.

59. Rodeberg, N.T., Johnson, J.A., Bucher, E.S., and Wightman, R.M. (2016). Dopamine dynamics during continuous intracranial self-stimulation: effect of waveform on fast-scan cyclic voltammetry data. ACS Chem. Neurosci. *7*, 1508–1518.

60. Kruse, J.M., Overmier, J.B., Konz, W.A., and Rokke, E. (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. Learn. Motiv. *14*, 165–181.

61. Corbit, L.H., and Janak, P.H. (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.

62. Takahashi, Y.K., Batchelor, H.M., 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.

63. Wise, R.A. (2002). Brain reward circuitry: insights from unsensed incentives. Neuron *36*, 229–240.

64. Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. Science *354*, 1278–1282.

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

66. Fee, M.S. (2014). The role of efference copy in striatal learning. Curr. Opin. Neurobiol. *25*, 194–200.

67. Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci. *22*, 513–523.

68. Gremel, C.M., and Costa, R.M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. *4*, 2264.

69. Matamales, M., McGovern, A.E., Mi, J.D., Mazzone, S.B., Balleine, B.W., and Bertran-Gonzalez, J. (2020). Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. Science *367*, 549–555.

70. Yin, H.H. (2010). The sensorimotor striatum is necessary for serial order learning. J. Neurosci. *30*, 14719–14723.

71. Tsutsui-Kimura, I., Matsumoto, H., Akiti, K., Yamada, M.M., 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.

72. de Jong, J.W., Afjei, S.A., Pollak Dorocic, I., Peck, J.R., Liu, C., Kim, C.K., 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.

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.

74. Steinberg, E.E., Gore, F., Heifets, B.D., Taylor, M.D., Norville, Z.C., Beier, K.T., Földy, C., Lerner, T.N., Luo, L., Deisseroth, K., and Malenka, R.C. (2020). Amygdala-midbrain connections modulate appetitive and aversive learning. Neuron *106*, 1026–1043.e9.

75. Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A., and Uchida, N. (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. Neuron *74*, 858–873.

76. Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A., Crow, A.K., Malenka, R.C., Luo, L., Tomer, R., and Deisseroth, K. (2015). Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. Cell *162*, 635–647.

77. Menegas, W., Bergan, J.F., Ogawa, S.K., 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.

78. Tepper, J.M., and Lee, C.R. (2007). GABAergic control of substantia nigra dopaminergic neurons. Prog. Brain Res. *160*, 189–208.

79. Brazhnik, E., Shah, F., and Tepper, J.M. (2008). GABAergic afferents activate both GABAA and GABAB receptors in mouse substantia nigra dopaminergic neurons in vivo. J. Neurosci. *28*, 10386–10398.

80. Evans, R.C., Twedell, E.L., Zhu, M., Ascencio, J., Zhang, R., and Khaliq, Z.M. (2020). Functional dissection of basal ganglia inhibitory inputs onto substantia nigra dopaminergic neurons. Cell Rep. *32*, 108156.

81. Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337–1340.

82. Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. Neuron *58*, 451–463.

83. Hong, S., and Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral habenula. Neuron *60*, 720–729.

84. Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E., 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.

85. Tachibana, Y., and Hikosaka, O. (2012). The primate ventral pallidum encodes expected reward value and regulates motor action. Neuron *76*, 826–837.

86. Kim, H.F., Amita, H., and Hikosaka, O. (2017). Indirect pathway of caudal basal ganglia for rejection of valueless visual objects. Neuron *94*, 920–930.e3.

87. Smith, J.B., Klug, J.R., Ross, D.L., Howard, C.D., Hollon, N.G., Ko, V.I., Hoffman, H., Callaway, E.M., Gerfen, C.R., and Jin, X. (2016). Genetic-based dissection unveils the inputs and outputs of striatal patch and matrix compartments. Neuron *91*, 1069–1084.

88. Crittenden, J.R., Tillberg, P.W., Riad, M.H., Shima, Y., Gerfen, C.R., Curry, J., Housman, D.E., Nelson, S.B., Boyden, E.S., and Graybiel, A.M. (2016). Striosome-dendron bouquets highlight a unique striatonigral circuit targeting dopamine-containing neurons. Proc. Natl. Acad. Sci. USA *113*, 11318–11323.

89. McGregor, M.M., McKinsey, G.L., Girasole, A.E., Bair-Marshall, C.J., Rubenstein, J.L.R., and Nelson, A.B. (2019). Functionally distinct connectivity of developmentally targeted striosome neurons. Cell Rep. *29*, 1419–1428.e5.

90. Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. Nature *447*, 1111–1115.

91. Jhou, T.C., Fields, H.L., Baxter, M.G., Saper, C.B., and Holland, P.C. (2009). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. Neuron *61*, 786–800.

92. Hong, S., Jhou, T.C., Smith, M., Saleem, K.S., 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*, 11457–11471.

93. Tian, J., and Uchida, N. (2015). Habenula lesions reveal that multiple mechanisms underlie dopamine prediction errors. Neuron *87*, 1304–1316.

94. Takahashi, Y.K., Langdon, A.J., 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*, 182–193.

95. Yang, H., de Jong, J.W., Tak, Y., Peck, J., Bateup, H.S., and Lammel, S. (2018). Nucleus accumbens subnuclei regulate motivated behavior via direct inhibition and disinhibition of VTA dopamine subpopulations. Neuron *97*, 434–449.e4.

96. Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida, N., and Watabe-Uchida, M. (2016). Distributed and mixed information in monosynaptic inputs to dopamine neurons. Neuron *91*, 1374–1389.

97. Stalnaker, T.A., Calhoon, G.G., Ogawa, M., Roesch, M.R., and Schoenbaum, G. (2012). Reward prediction error signaling in posterior dorsomedial striatum is action specific. J. Neurosci. *32*, 10296–10305.

98. Chen, R., Puzerey, P.A., Roeser, A.C., Riccelli, T.E., Podury, A., Maher, K., Farhang, A.R., and Goldberg, J.H. (2019). Songbird ventral pallidum sends diverse performance error signals to dopaminergic midbrain. Neuron *103*, 266–276.e4.

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

100. Ottenheimer, D.J., Bari, B.A., Sutlief, E., Fraser, K.M., Kim, T.H., Richard, J.M., Cohen, J.Y., and Janak, P.H. (2020). A quantitative reward prediction error signal in the ventral pallidum. Nat. Neurosci. *23*, 1267–1276.

101. Watabe-Uchida, M., Eshel, N., and Uchida, N. (2017). Neural circuitry of reward prediction error. Annu. Rev. Neurosci. *40*, 373–394.

102. Kramer, P.F., Twedell, E.L., Shin, J.H., Zhang, R., and Khaliq, Z.M. (2020). Axonal mechanisms mediating γ-aminobutyric acid receptor type A (GABA-A) inhibition of striatal dopamine release. eLife *9*, e55729.

103. Holly, E.N., Davatolhagh, M.F., España, R.A., and Fuccillo, M.V. (2021). Striatal low-threshold spiking interneurons locally gate dopamine. Curr. Biol. *31*, 4139–4147.e6.

104. Collins, A.L., Aitken, T.J., Greenfield, V.Y., Ostlund, S.B., and Wassum, K.M. (2016). Nucleus accumbens acetylcholine receptors modulate dopamine and motivation. Neuropsychopharmacology *41*, 2830–2838.

105. Sulzer, D., Cragg, S.J., and Rice, M.E. (2016). Striatal dopamine neurotransmission: regulation of release and uptake. Basal Ganglia *6*, 123–148.

106. Murakami, M., Vicente, M.I., Costa, G.M., and Mainen, Z.F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. Nat. Neurosci. *17*, 1574–1582.

107. Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E.M., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. Nature *469*, 53–57.

108. Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., and Janak, P.H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nat. Neurosci. *16*, 966–973.

109. Chang, C.Y., Esber, G.R., 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.

110. Maes, E.J.P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G., and Iordanova, M.D. (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. Nat. Neurosci. *23*, 176–178.

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

112. Faure, A., Haberland, U., Condé, F., and El Massioui, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. J. Neurosci. *25*, 2771–2780.

113. Belin, D., and Everitt, B.J. (2008). Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. Neuron *57*, 432–441.

114. Wang, L.P., Li, F., Wang, D., Xie, K., Wang, D., Shen, X., and Tsien, J.Z. (2011). NMDA receptors in dopaminergic neurons are crucial for habit learning. Neuron *72*, 1055–1066.

115. Hamid, A.A., Frank, M.J., and Moore, C.I. (2021). Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. Cell *184*, 2733–2749.e16.

116. Redish, A.D. (2004). Addiction as a computational process gone awry. Science *306*, 1944–1947.

117. Lüscher, C., Robbins, T.W., and Everitt, B.J. (2020). The transition to compulsion in addiction. Nat. Rev. Neurosci. *21*, 247–263.

118. Weintraub, D., and Mamikonyan, E. (2019). Impulse control disorders in Parkinson's disease. Am. J. Psychiatry *176*, 5–11.

119. Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking and its disorders. Schizophr. Bull. *4*, 636–640.

120. Lindner, A., Thier, P., Kircher, T.T.J., Haarmeier, T., and Leube, D.T. (2005). Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. Curr. Biol. *15*, 1119–1124.

121. Frith, C. (2012). Explaining delusions of control: the comparator model 20 years on. Conscious. Cogn. *21*, 52–54.

122. Griffin, J.D., and Fletcher, P.C. (2017). Predictive processing, source monitoring, and psychosis. Annu. Rev. Clin. Psychol. *13*, 265–289.

123. Kort, N.S., Ford, J.M., Roach, B.J., Gunduz-Bruce, H., Krystal, J.H., Jaeger, J., Reinhart, R.M.G., and Mathalon, D.H. (2017). Role of N-methyl-D-aspartate receptors in action-based predictive coding deficits in schizophrenia. Biol. Psychiatry *81*, 514–524.

124. Cassidy, C.M., Balsam, P.D., Weinstein, J.J., Rosengard, R.J., Slifstein, M., Daw, N.D., Abi-Dargham, A., and Horga, G. (2018). A perceptual inference mechanism for hallucinations linked to striatal dopamine. Curr. Biol. *28*, 503–514.e4.

125. Sterzer, P., Adams, R.A., Fletcher, P., Frith, C., Lawrie, S.M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., and Corlett, P.R. (2018). The predictive coding account of psychosis. Biol. Psychiatry *84*, 634–643.

126. McCutcheon, R.A., Abi-Dargham, A., and Howes, O.D. (2019). Schizophrenia, dopamine and the striatum: from biology to symptoms. Trends Neurosci. *42*, 205–220.

127. Ford, J.M., and Mathalon, D.H. (2019). Efference copy, corollary discharge, predictive coding, and psychosis. Biol. Psychiatry Cogn. Neurosci. Neuroimaging *4*, 764–767.

128. Schmack, K., Bosc, M., Ott, T., Sturgill, J.F., and Kepecs, A. (2021). Striatal dopamine mediates hallucination-like perception in mice. Science *372*, eabf4740.

129. Barter, J.W., Li, S., Lu, D., Bartholomew, R.A., Rossi, M.A., Shoemaker, C.T., Salas-Meza, D., Gaidis, E., and Yin, H.H. (2015). Beyond reward prediction errors: the role of dopamine in movement kinematics. Front. Integr. Neurosci. *9*, 39.

130. Panigrahi, B., Martin, K.A., Li, Y., Graves, A.R., Vollmer, A., Olson, L., Mensh, B.D., Karpova, A.Y., and Dudman, J.T. (2015). Dopamine is required for the neural representation and control of movement vigor. Cell *162*, 1418–1430.

131. Howard, C.D., Li, H., Geddes, C.E., and Jin, X. (2017). Dynamic nigrostriatal dopamine biases action selection. Neuron *93*, 1436–1450.e8.

132. Lopes, G., Bonacchi, N., Frazão, J., Neto, J.P., Atallah, B.V., Soares, S., Moreira, L., Matias, S., Itskov, P.M., Correia, P.A., et al. (2015). Bonsai: an event-based framework for processing and controlling data streams. Front. Neuroinform. *9*, 7.

133. Keithley, R.B., Heien, M.L., and Wightman, R.M. (2009). Multivariate concentration determination using principal component regression with residual analysis. Trends Analyt. Chem. *28*, 1127–1136.

134. Keithley, R.B., and Wightman, R.M. (2011). Assessing principal component regression prediction of neurochemicals detected with fast-scan cyclic voltammetry. ACS Chem. Neurosci. *2*, 514–525.

135. Rodeberg, N.T., Sandberg, S.G., Johnson, J.A., Phillips, P.E.M., and Wightman, R.M. (2017). Hitchhiker's guide to voltammetry: acute and chronic electrodes for in vivo fast-scan cyclic voltammetry. ACS Chem. Neurosci. *8*, 221–234.

136. Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579.

137. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. *15*, 1–25.

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

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

CellPress

STAR★METHODS

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

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.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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 boxes, 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.

Continuous reinforcement cohort

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.

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

Left-Right sequence cohort

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.

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 −0.4 V 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 Quantification and statistical analysis). 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 (Figures S3H and S3I).

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

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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 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 (ΔF/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

Please cite this article in press as: Hollon et al., Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.09.040

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).

Current Biology, Volume *31*

Supplemental Information

Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors

Nick G. Hollon, Elora W. Williams, Christopher D. Howard, Hao Li, Tavish I. Traut, and Xin Jin

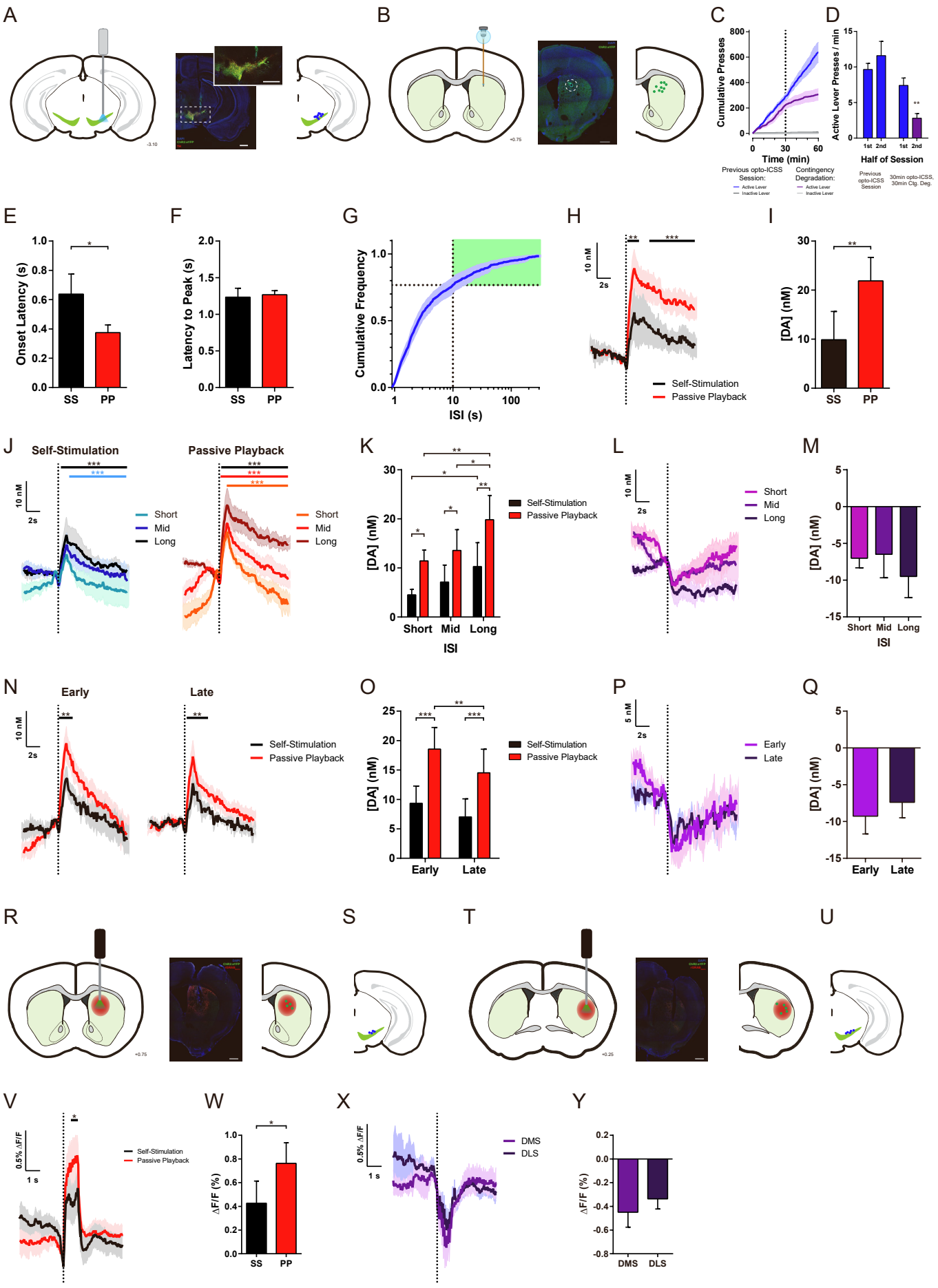
A
B
C
D
Cumulative Presses
Time (min)
Previous opto-ICSS Session:
Contingency Degradation:
Active Lever
Inactive Lever
Active Lever Presses / min
Half of Session
1st 2nd 1st 2nd
Previous opto-ICSS Session
30min opto-ICSS, 30min Ctg. Deg.
E
Onset Latency (s)
SS PP
F
Latency to Peak (s)
SS PP
G
Cumulative Frequency
ISI (s)
H
10 nM
2s
Self-Stimulation
Passive Playback
I
[DA] (nM)
SS PP
J
Self-Stimulation
Passive Playback
Short
Mid
Long
K
[DA] (nM)
Short Mid Long
ISI
L
M
N
Early
Late
O
Early Late
P
5 nM
Q
R
S
T
U
V
0.5% ΔF/F
1 s
W
ΔF/F (%)
X
DMS
DLS
Y

Figure S1. Opto-ICSS CRF cohort histology, contingency degradation, bout initiations, ISI, within-session comparisons, and DLS fiber photometry. Related to Figure 1.

(**A**) Left: Coronal schematic of fiber optic placement targeting the SNc (-3.10 mm posterior from bregma). Center: Representative image of fiber optic placement over the SNc in animal selectively expressing ChR2-eYFP in TH-positive dopamine neurons (scale bars = 500 μm). Right: Fiber optic placement for mice in the CRF cohort. (**B**) Left: Coronal schematic of FSCV carbon-fiber microelectrode placement in the dorsal striatum (+0.75 mm anterior to bregma). Center: Representative image of FSCV carbon-fiber microelectrode placement in the dorsal striatum (scale bar = 500 μm). The lesion was made by passing a current through the electrode just before perfusion after the conclusion of all experiments (see Methods). Right: FSCV electrode placement for mice in the CRF cohort. (**C**) Cumulative presses over time within the contingency degradation test session (30 min opto-ICSS followed by the 30 min contingency degradation test phase), overlaid with performance throughout the previous day's standard opto-ICSS session for comparison (n = 6 mice). (**D**) Summary of mean Active lever press rate during each phase of the contingency degradation test session (as in Fig. 1D), compared to the preceding day's standard opto-ICSS session (two-way repeated-measures ANOVA: main effect of Day, $F_{1,5}$ = 8.157, P = 0.0356; Day by Half of Session interaction, $F_{1,5}$ = 25.30, P = 0.0040; Sidak's multiple comparisons tests: Contingency Degradation 1st vs. 2nd Half, P = 0.0082; 2nd Half of Previous Day vs. Contingency Degradation test phase, P = 0.0004). (**E**) Latency for the onset of the stimulation-evoked dopamine response to exceed baseline: Time from stimulation onset until the dopamine response exceeded 2 standard deviations above the mean of baseline period (1 s preceding stimulation). Note that 1 animal's Self-Stimulation response did not cross this threshold within 2 s of stimulation onset, and was excluded from this analysis (n = 8 mice; paired t test: t_7 = 2.582, P = 0.0364). (**F**) Latency to the peak dopamine response, within 2 s following stimulation onset (n = 9 mice; paired t test: t_8 = 0.2722, P = 0.7924). (**G**) Cumulative frequency distribution of inter-stimulation intervals (ISIs) from the opto-ICSS FSCV recording session (n = 9 mice). Green shading indicates ISIs > 10 s, used to define bout initiation for the subset of stimulations analyzed in (H-I). (**H**) Mean dopamine concentration change to bout-initiating Self-Stimulation (ISI > 10s since previous stimulation) and corresponding Passive Playback stimulations. Black bars indicate time points where the Self-Stimulation response significantly differs from Passive Playback (permutation test, Ps = 0.007 and 0.0001 for first and second time clusters, respectively). (**I**) Mean change in dopamine concentration for the bout-initiating subset of stimulations in (H). (t_8 = 3.600, P = 0.0070). (**J**) Mean dopamine concentration changes to stimulations sorted by preceding ISI into short, medium (mid), and long ISI tertiles (permutation tests: Self-Stimulation Short vs. Long ISI, black bar, P = 0.0005; Short vs. Mid, blue bar, P = 0.0003; Playback Short vs. Long ISI, black bar, P = 0.0001; Mid vs. Long, red bar, P = 0.0001; Short vs. Mid, orange bar, P = 0.0009). (**K**) Mean change in dopamine concentration to stimulations sorted into ISI tertiles as in (J); (two-way repeated-measures ANOVA: main effect of Session Phase, $F_{1,8}$ = 16.07, P = 0.0039; main effect trend of ISI, $F_{2,16}$ = 3.295, P = 0.0633; Sidak's multiple comparisons tests: Self-Stimulation vs. Playback with Short ISIs, P = 0.0142; Mid, P = 0.0235; Long, P = 0.0011; Self-Stimulation Short vs. Long, P = 0.0439; Playback Short vs. Long, P = 0.0034; Playback Mid vs. Long, P = 0.0278). (**L**) Difference traces: Self minus Playback sorted by ISI from (J). (**M**) Mean Differences from (L); (one-way repeated-measures ANOVA is not significant, $F_{2,16}$ = 0.5891, P = 0.5664). (**N**) Mean dopamine concentration changes to the first 20 (Early) and final 20 (Late) stimulations within each phase of the recorded FSCV session (permutation tests: Early, P = 0.0038; Late, P = 0.0045). (**O**) Mean change in dopamine concentration to Early and Late stimulations as in (N); (two-way repeated-measures ANOVA: main effect of Session Phase, $F_{1,8}$ = 14.94, P = 0.0048; main effect of Early vs. Late, $F_{1,8}$ = 8.846, P = 0.0178; Sidak's multiple comparisons tests: Early Self vs. Playback, P < 0.0001; Late Self vs. Playback, P < 0.0001; Self-Stimulation Early vs. Late, P = 0.0788; Playback Early vs. Late, P = 0.0055). (**P**) Difference traces: Self minus Playback for Early and Late stimulations from (N). (**Q**) Mean Differences from (P) do not significantly differ (t_8 = 1.449, P = 0.1854). (**R**) Schematic (left), representative image (center), and fiber placement (right) for fiber photometry recordings of the red dopamine sensor rGRAB$_{DA1h}$ in the dorsomedial striatum (DMS) for data in Figure 1K-P. (**S**) Fiber placement for optogenetic stimulation of SNc dopamine neurons in this DMS rGRAB$_{DA1h}$ group. (**T-U**) Histology for dorsolateral striatum (DLS) rGRAB$_{DA1h}$ group, as in (R-S). (**V**) Mean ΔF/F rGRAB$_{DA1h}$ dopamine sensor response in DLS evoked by Self-Stimulation and Passive Playback stimulations (n = 5 mice; black bar, permutation test, P = 0.0162). (**W**) Mean ΔF/F rGRAB$_{DA1h}$ dopamine sensor response in DLS (paired t test, t_4 = 3.944, P = 0.0169). (**X**) Difference traces: Self minus Playback for DMS and DLS rGRAB$_{DA1h}$ groups. (**Y**) Mean Differences from (X) do not significantly differ (unpaired t test, t_7 = 0.7691, P = 0.467). Ctg. Deg., Contingency Degradation; ISI, Inter-stimulation interval; SS, Self-Stimulation; PP, Passive

Playback; DMS, dorsomedial striatum; DLS, dorsolateral striatum. Error bars are SEM here and for below figures.

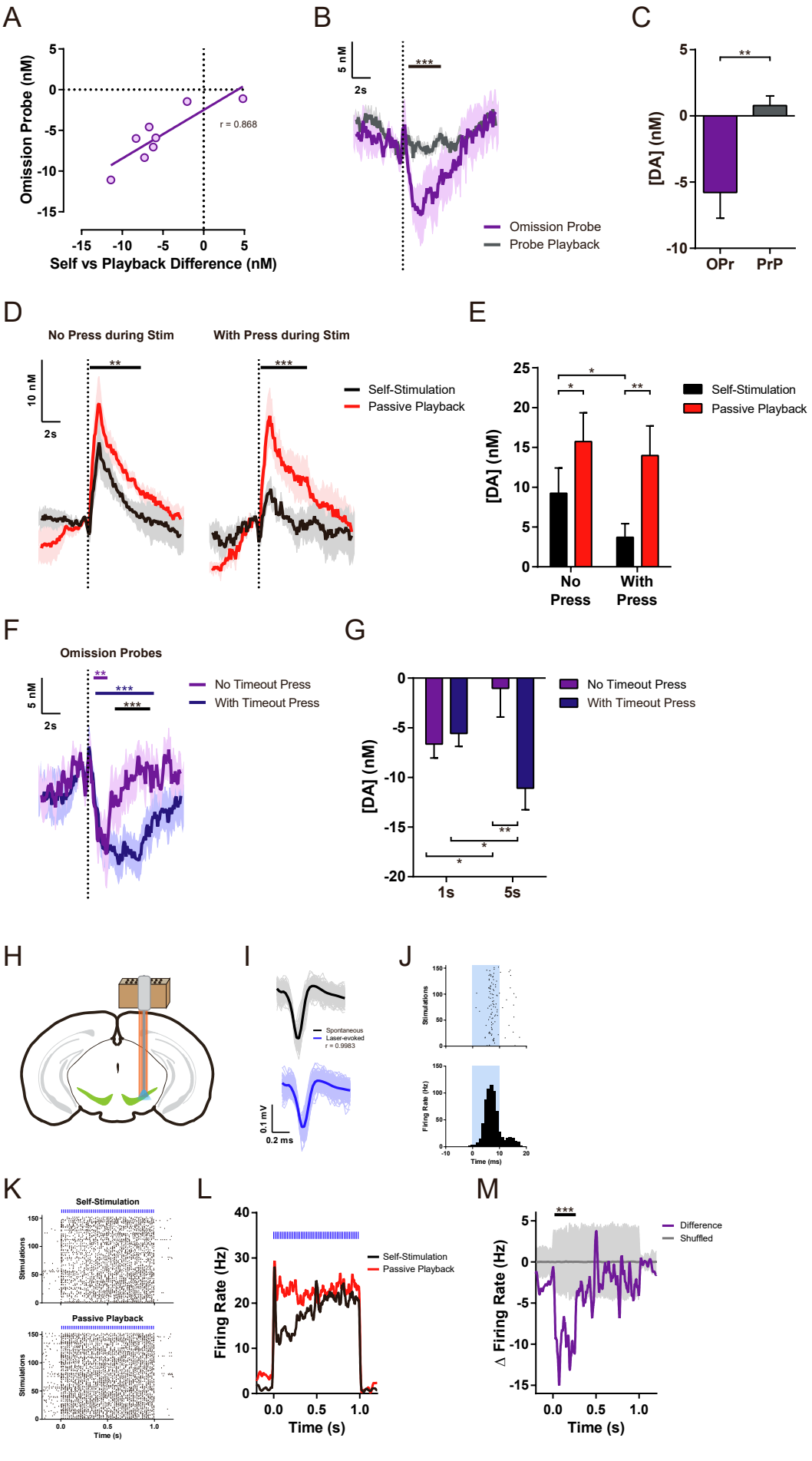
A
Omission Probe (nM)
Self vs Playback Difference (nM)
r = 0.868
B
5 nM
2s
***
Omission Probe
Probe Playback
C
[DA] (nM)
**
OPr
PrP
D
No Press during Stim
With Press during Stim
10 nM
2s
Self-Stimulation
Passive Playback
E
[DA] (nM)
No Press
With Press
Self-Stimulation
Passive Playback
F
Omission Probes
5 nM
2s
No Timeout Press
With Timeout Press
G
[DA] (nM)
1s
5s
No Timeout Press
With Timeout Press
H
I
Spontaneous
Laser-evoked
r = 0.9983
0.1 mV
0.2 ms
J
Stimulations
Firing Rate (Hz)
Time (ms)
K
Self-Stimulation
Passive Playback
Stimulations
Time (s)
L
Firing Rate (Hz)
Time (s)
Self-Stimulation
Passive Playback
M
Δ Firing Rate (Hz)
Time (s)
Difference
Shuffled

Figure S2. Inhibition of dopamine to isolated omissions, augmented suppression by additional presses, and *in vivo* electrophysiology. Related to Figure 2.

(**A**) Scatter plot depicting significant correlation between the amplitude of Omission Probe dopamine response and the stimulation response Difference (Self minus Playback; n = 8 mice, Pearson correlation r = 0.868, P = 0.0052). (**B**) Mean dopamine concentration change following temporally isolated Omission Probes (latency > 5 s since previous stimulation) and corresponding time points from the Passive Playback phase (n = 8 mice; permutation test, P = 0.0005). (**C**) Mean change in dopamine concentration following temporally isolated Omission Probes and equivalent time points from Playback phase (paired t test, t_7 = 3.511, P = 0.0098). (**D**) Mean dopamine concentration changes evoked by Self-Stimulations without (left) or with (right) an additional lever press during the ongoing stimulation, and the corresponding Passive Playback stimulations (n = 9 mice; permutation tests: Self-Stimulation with no press during stim vs. its Playback, P = 0.0011; Self-Stimulation with press during stim vs. its Playback, P = 0.0003; Self-Stimulation with vs. no press during stim, P = 0.0028). (**E**) Mean change in dopamine concentration for Self-Stimulations with or without additional presses during the stimulation, and their Playback (two-way repeated-measures ANOVA: main effect of Press, $F_{1,8}$ = 8.144, P = 0.0214; main effect of Session Phase, $F_{1,8}$ = 16.62, P = 0.0035; Sidak's multiple comparisons tests: Self-Stimulation with vs. no additional press, P = 0.0350; Self-Stimulation without additional press vs. Playback, P = 0.0163; Self-Stimulation with additional press vs. Playback, P = 0.0011). (**F**) Mean dopamine concentration changes during Omission Probes with or without additional presses during the probe timeout period (n = 8 mice; permutation tests: Omission Probe with no timeout press vs. 0, magenta bar, P = 0.002; Omission Probe with timeout press vs. 0, blue bar, P = 0.0001; Omission Probe with vs. without press, black bar, P = 0.0003). (**G**) Mean change in dopamine concentration at early (1 s) vs. late (5 s) time points during Omission Probes with or without additional presses during the timeout period (two-way repeated-measures ANOVA: main effect of Press, $F_{1,7}$ = 8.181, P = 0.0243; Press by Time interaction, $F_{1,7}$ = 16.70, P = 0.0047; Sidak's multiple comparisons tests: with vs. no press, late, P = 0.0025; no press, early vs. late, P = 0.0444; with press, early vs. late, P = 0.0479). (**H**) Schematic of experimental preparation for *in vivo* extracellular electrophysiology recordings with optogenetic identification of SNc dopamine neurons. (**I**) Waveforms of an optogenetically identified dopamine neuron for spontaneous (top) and laser-evoked (bottom) spikes (Pearson correlation, r = 0.9983, P < 0.0001). (**J**) Raster plot (top) and peri-event time histogram (bottom) of this dopamine neuron response to 10-ms optogenetic stimulation pulse. Each row in the raster represents one stimulation, and black ticks are spikes. (**K**) Raster plot of the same dopamine neuron responses aligned to Self-Stimulation (top) and Passive Playback stimulations (bottom). (**L**) Firing rate of the optogenetically identified dopamine neuron in (K) in response to Self-Stimulation versus Passive Playback stimulations. (**M**) Difference trace for the dopamine neuron in (K-L) depicting Self-Stimulation minus Playback difference in stimulation-evoked firing rate between session phases. Gray shaded region depicts the 95% confidence interval of the null distribution generated from shuffled data (permutation test, P = 0.0001). OPr, Omission Probe; PrP, Probe Playback.

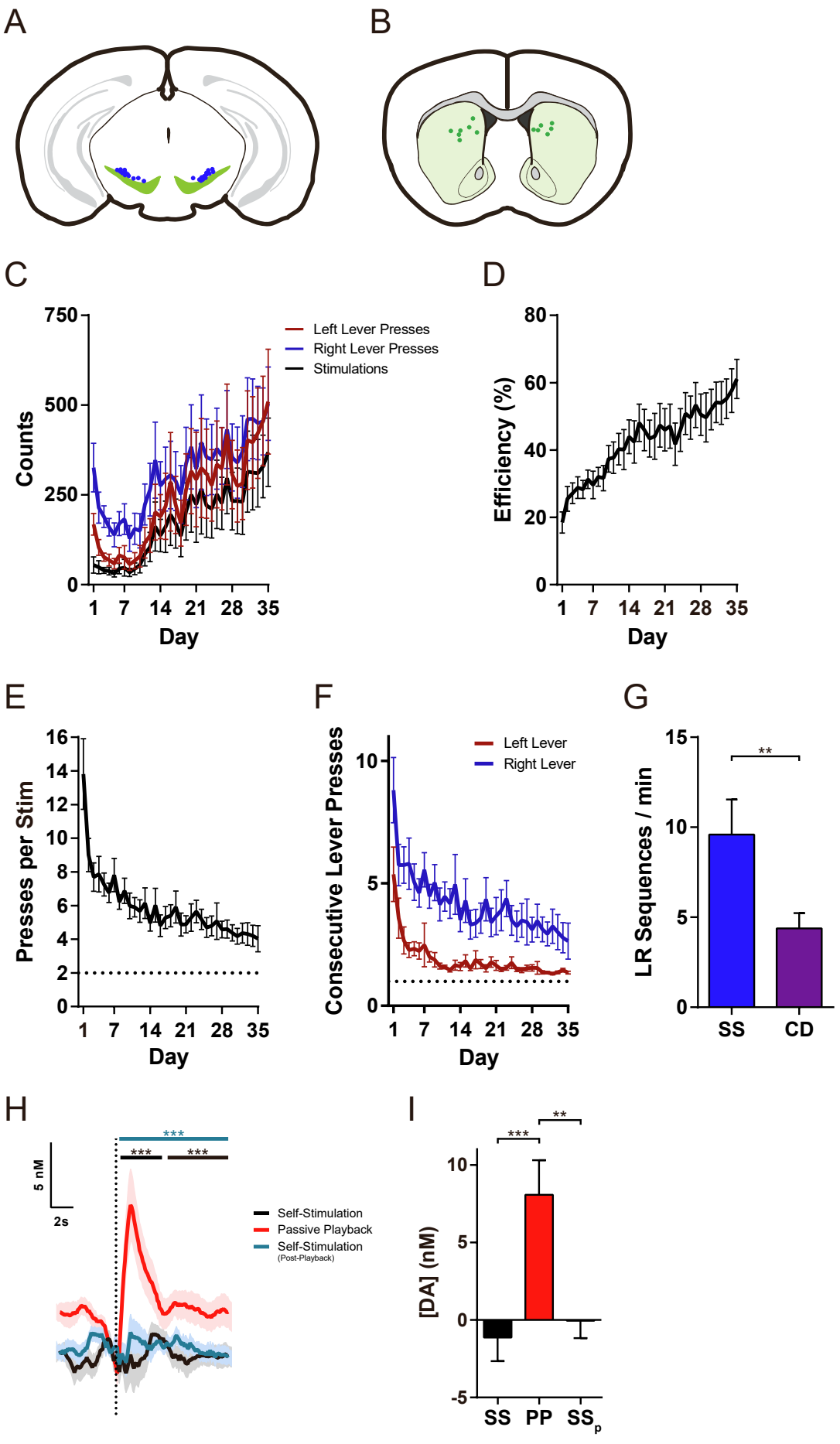
A
B
C
750
500
250
0
Counts
1 7 14 21 28 35
Day
Left Lever Presses
Right Lever Presses
Stimulations
D
80
60
40
20
0
Efficiency (%)
1 7 14 21 28 35
Day
E
16
14
12
10
8
6
4
2
0
Presses per Stim
1 7 14 21 28 35
Day
F
10
5
0
Consecutive Lever Presses
1 7 14 21 28 35
Day
Left Lever
Right Lever
G
15
10
5
0
LR Sequences / min
**
SS
CD
H
***
***
***
5 nM
2s
Self-Stimulation
Passive Playback
Self-Stimulation
(Post-Playback)
I
***
**
10
5
0
-5
[DA] (nM)
SS
PP
SS$_p$

Figure S3. Left-Right sequence cohort histology, additional behavior, and session phase order control. Related to Figure 3.

(**A**) Fiber optic placement for mice in the LR sequence cohort. (**B**) FSCV electrode placement for mice in the LR sequence cohort. (**C**) Presses on each lever and stimulations earned across days of training (n = 13 mice; statistics for stimulations presented in Fig. 3B; two-way repeated-measures ANOVA for Lever by Day: main effect of Day, $F_{34,408}$ = 3.189, P < 0.0001; main effect of Lever, $F_{1,12}$ = 16.67, P = 0.0015; Lever by Day interaction, $F_{34,408}$ = 1.535, P = 0.0307). (**D**) Efficiency across days of training, calculated as the number of stimulations per pair of lever presses (one-way repeated-measures ANOVA, $F_{34,408}$ = 6.936, P < 0.0001). (**E**) Total presses per stimulation (either lever) across days of training (one-way repeated-measures ANOVA, $F_{34,408}$ = 8.122, P < 0.0001). (**F**) Consecutive presses on each lever across days of training (two-way repeated-measures ANOVA: main effect of Day, $F_{34,408}$ = 8.430, P < 0.0001; main effect of Lever, $F_{1,12}$ = 23.07, P = 0.0004). (**G**) Contingency degradation test: 30 min of LR sequence opto-ICSS followed by 30 min contingency degradation test phase (n = 10 mice; paired t test, t_9 = 3.458, P = 0.0072). (**H**) Mean dopamine concentration change evoked by LR Self-Stimulation before or after the Passive Playback phase. The pre-playback Self-Stimulation (black) and Passive Playback responses are the same data as Fig. 3H, and the post-playback Self-Stimulation (teal) is an additional 30 min phase of LR sequence opto-ICSS following the Playback phase to control for potential order effects (n = 12 mice; permutation tests: Ps = 0.0001 for all time clusters, black bars for pre-playback Self-Stimulation vs. Playback, teal bar for post-playback Self-Stimulation vs. Playback). (**I**) Mean change in dopamine concentration for Self-Stimulation before or after Passive Playback (one-way repeated-measures ANOVA, $F_{2,22}$ = 26.39, P < 0.0001; Tukey's multiple comparisons tests: Self-Stimulation (pre) vs. Playback, P < 0.0001; Self-Stimulation (post) vs. Playback, P < 0.0001). SS, Self-Stimulation; CD, Contingency Degradation; PP, Passive Playback; SS_P, Self-Stimulation (post-Playback).

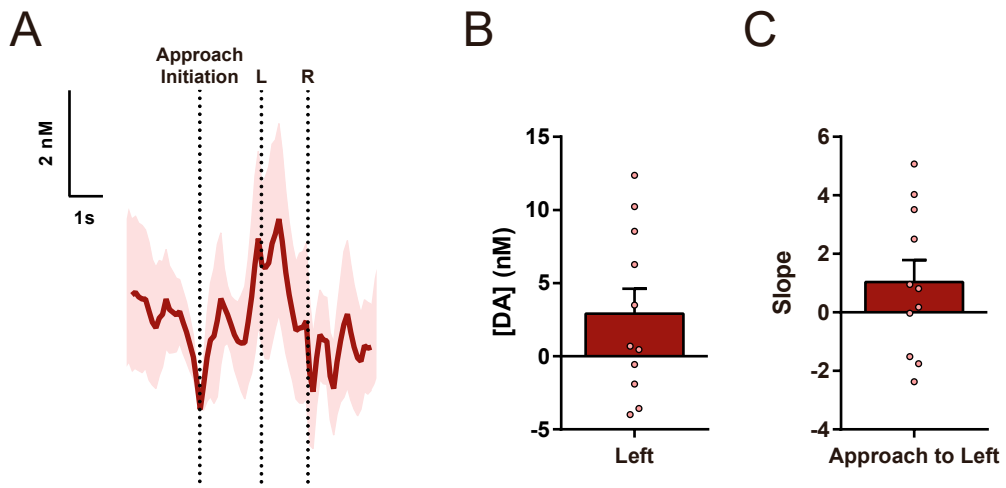

Figure S4. Left-Right sequences aligned to approach initiation. Related to Figure 4.
(**A**) Mean dopamine concentration change aligned to approach initiation (n = 11 mice). Intervals from approach initiation to Left lever press and from Left to Right press are scaled to each interval's median duration. (**B**) Mean change in dopamine concentration in the 0.5 s following the Left lever press, relative to the pre-approach baseline (one-sample t test vs 0, t_{10} = 1.700, P = 0.120). (**C**) Slope of linear regression fit to dopamine trace from approach initiation to Left press (one-sample t test vs 0, t_{10} = 1.388, P = 0.1952).