nature reviews neuroscience

https://doi.org/10.1038/s41583-025-00925-2

Perspective

Check for updates

From avoidance to new action: the multifaceted role of the striatal indirect pathway

Jaeeon Lee[1,3,4] & Bernardo L. Sabatini[1,2,4]

Abstract

A hallmark of optimal reinforcement learning is that an agent learns to avoid actions that lead to negative outcomes while still exploring alternative actions that could lead to better outcomes. Although the basal ganglia have been hypothesized to contribute to this computation, the mechanisms by which they do so are still unclear. Here, we focus on the function of the striatal indirect pathway and propose that it is regulated by a synaptic plasticity rule that allows an animal to avoid actions that lead to suboptimal outcomes. We consider current theories of striatal indirect pathway function in light of recent experimental findings and discuss studies that suggest that indirect pathway activity is potentiated by the suppression of dopamine release in the striatum. Furthermore, we highlight recent studies showing that activation of the indirect pathway can trigger an action, allowing animals to explore new actions while suppressing suboptimal actions. We show how our framework can reconcile previously conflicting results regarding the indirect pathway and suggest experiments for future investigation.

Sections

[1]Howard Hughes Medical Institute, Department of Neurobiology, Harvard Medical School, Boston, MA, USA. [2]Kempner Institute for the Study of Natural and Artificial Intelligence, Harvard University, Boston, MA, USA. [3]Present address: Department of Molecular and Cellular Biology, Center for Brain Science, Harvard University, Cambridge, MA, USA. [4]These authors contributed equally: Jaeeon Lee, Bernardo L. Sabatini. e-mail: bsabatini@hms.harvard.edu

Introduction

Organisms adapt to their environment by learning to repeat sets of actions that lead to positive outcomes and to avoid those that lead to negative outcomes. This so-called 'law of effect', first proposed by Edward Thorndike, is a key tenet in behavioural psychology and has inspired neuroscientists to investigate its neural implementation[1–4]. Much work has pointed to the basal ganglia as the key set of brain regions implementing this computation[5–9]. Specifically, the two pathways through which striatal neurons modulate downstream behaviours, the direct pathway (consisting of direct striatal projection neurons (dSPNs)) and the indirect pathway (consisting of indirect striatal projection neurons (iSPNs)), are classically thought to mediate the reinforcement and punishment aspects of learning, respectively, by promoting or suppressing relevant actions[5,7,9]. Although studies generally agree that the activity of dSPNs is pro-kinetic and reinforcing, they disagree on the exact function of the striatal indirect pathway, which has been variously proposed to have roles in movement suppression[6,10–13], motor refinement[14–16], punishment[7,17,18], action switching[19,20], exploration[21], movement sequencing[19,22] and risk avoidance[23,24].

Given the diversity of these proposed functions, one fruitful approach to understand the function of iSPNs might be to look at the synaptic learning rule that governs when and how they undergo synaptic plasticity. Once the learning rule is established, one can start to infer whether it would allow iSPNs to subserve the various functions proposed by previous studies (for example, learning to avoid negative outcomes). Recent studies have begun to shed light on the conditions under which long-term potentiation (LTP) of the glutamatergic synapses that drive the activity of iSPNs occurs in rodents[25,26]. These experiments have validated a hypothesis that had long been assumed to be true, but not previously directly tested: that iSPNs are sensitive to below-baseline dips in the concentration of the neuromodulator dopamine in the striatum. Studies investigating dopamine have also found that negative outcomes might be represented via a dip in dopamine concentrations in certain parts of the striatum[25,27–30]. Together, these experiments suggest that a reduction in dopamine concentration that occurs when a particular behaviour leads to a negative outcome could open an important window for learning via plasticity in iSPNs. This window, in turn, might allow iSPNs to suppress the actions that led to the dip in dopamine concentration in the first place. By contrast, a few studies have shown that optogenetically activating iSPNs can suppress one behaviour while inducing a de novo action, distinct from the action that is being suppressed[19,21,31]. Although this might seem paradoxical at first, we propose that this might allow the animal to smoothly transition from a suboptimal action to an alternative action, suggesting a complex and multifaceted function for iSPNs.

Our goal in this Perspective is to highlight recent studies that have revealed details about the learning rules that govern the activity of iSPNs and the effects of activating iSPNs, and to propose how these two sets of findings provide a good starting point to understand iSPN function. We will first consider current models of iSPNs and describe the studies that either support or challenge each model. We will then highlight recent experiments investigating dopaminergic modulation of LTP of glutamatergic synapses onto iSPNs. Taking these findings into account, and inspired by previous models of the basal ganglia, we present a three-factor learning rule for iSPN synaptic plasticity that allows for correct suppression of suboptimal motor programmes. We present predictions of the three-factor learning rule, and consider whether they are supported by studies. Finally, we discuss the ability of iSPNs to induce alternative actions and the potential disinhibitory mechanisms that underlie these effects. In order to better explain the framework proposed, we have decided to focus primarily on studies related to sensorimotor learning in rodents. Nevertheless, we also consider how this framework may extend to non-motor learning in striatal regions such as the dorsomedial striatum (DMS) and ventral striatum (VS). We conclude by addressing the framework's limitations, highlighting studies that challenge its premises and proposing future experiments to help reconcile these conflicting findings.

Models of iSPN function

Several influential models of iSPN function in sensorimotor control have been proposed. Here, our goal is not to present a detailed historical account of all models of basal ganglia function but, rather, to highlight current working models of iSPNs that will provide useful context for understanding iSPN-dependent learning rules (we refer the reader to other excellent reviews[32–35] for much more comprehensive coverage of this topic).

One of the oldest models of iSPN function that is relevant for our discussion is the centre–surround model[36,37] (also called the complementary model[33]) (Fig. 1a). In this model, dSPNs activate a target motor programme, whereas iSPNs provide blanket inhibition of competing motor programmes. For example, dSPNs in a localized region of the striatum may activate the motor programme for locomotion, whereas iSPNs in the same region may suppress competing motor programmes that could interfere with locomotion (such as forelimb movement or licking). The centre–surround model is attractive in that it can explain why dSPNs and iSPNs seem to be coactive during movement onset[38–40]. However, the model assumes that iSPNs in a local region of the striatum can provide a relatively broad inhibition of downstream regions, something that has not been supported by anatomical and functional studies. For instance, anatomical studies have found that regions downstream of iSPNs such as the globus pallidus externus (GPe) and subthalamic nucleus have a topographical organization that is consistent with the existence of segregated parallel pathways[41–43]. In agreement with this, the direct and indirect pathways arising from the same local region of the striatum converge onto the same neurons within the substantia nigra reticulata (SNr)[43]. Functionally, a recent study from our laboratory has shown that activating iSPNs in different regions of the striatum is not equally efficacious in suppressing licking – only iSPNs in the ventrolateral striatum ((VLS) are able to suppress licking behaviour[21]. Overall, the centre–surround model does not seem to be supported by anatomical and functional data.

An updated model inspired by the centre–surround model is what we will here call the refinement model (Fig. 1b). In this model, iSPNs still provide inhibition to refine the target programme but do so by providing very focal inhibition in the movement space, an abstract space of all possible movement/motor programmes the animal could generate[34,44]. Thus, rather than providing blanket inhibition of all competing motor programmes, iSPNs inhibit motor programmes that are similar to but kinematically distinct from the selected action. For instance, in a mouse that is learning to reach for a pellet using its forelimb, iSPNs might suppress certain trajectories of the forelimb in order to make the resulting movement more precise and effective. Unlike the centre–surround model, the refinement model does not require iSPNs to project broadly to downstream areas. In fact, in this model, the direct and indirect pathways should converge anatomically onto the same area within the SNr (and thus same set of neurons within the SNr), so that the pathways can work together to sharpen the representation in the

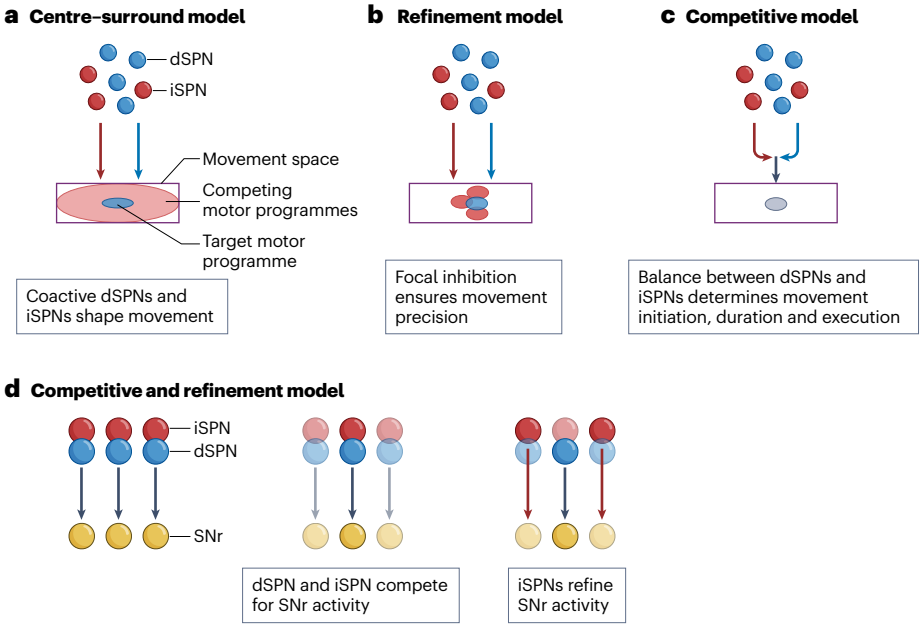


Fig. 1 | Models describing the function of direct and indirect striatal projection neurons. a, In the centre–surround model, direct striatal projection neurons (dSPNs) within a particular region of the striatum activate a target motor programme and indirect striatal projection neurons (iSPNs) in the same region provide blanket inhibition of all competing motor programmes, thus creating an inhibitory surround in the movement space[36]. **b**, The refinement model is an updated version of the centre–surround model in which the iSPNs refine the target movement by inhibiting only the subset of motor programmes that are closest to the target movement within the movement space. **c**, In the competitive model, dSPNs and iSPNs compete to control the same motor programme, with the balance between the two pathways determining key aspects of movement, such as its initiation, duration and execution[33]. **d**, Both the competitive and refinement models may be valid, and each model might be emphasizing distinct aspects of iSPN function that might partially depend on the postsynaptic targets in the substantia nigra reticulata (SNr): three triplets of dSPNs, iSPNs and the SNr target connected in parallel (left); dSPNs and iSPNs that project, directly or indirectly via the globus pallidus externus (GPe), to the same SNr neuron are thought to compete with each other, in alignment with the competitive model (middle); and iSPNs that do not share the postsynaptic SNr target with a dSPN are thought to refine the SNr activity by working cooperatively with dSPNs, in alignment with the refinement model (right).

relevant region of the SNr. Anatomical studies are indeed consistent with this view[41–43]. Functionally, a key prediction of the refinement model is that inactivation of iSPNs should lead to less precise and less effective movement generation. Indeed, iSPN inactivation has been shown to increase variability in lever press movements[15]. Various studies have also shown that inactivating iSPNs can slow motor learning or impair movement generation[16,31]. Furthermore, iSPN activation can refine movement velocity, suggesting a potential role of iSPNs in shaping movement kinematics[14]. Thus, although studies precisely measuring movement kinematics following iSPN ablation are limited, the existing evidence generally supports the notion that iSPNs play a crucial role in movement refinement.

The last, and perhaps most popular, model of iSPN function is the competitive model[33] (Fig. 1c). In this model, iSPNs oppose and compete against dSPNs for the control of the behavioural output. The balance between dSPN and iSPN activity is thought to determine the resulting motor programme, including its occurrence, duration and vigour. Most studies of iSPNs agree with this model. For instance, modulating the activity of iSPNs seems to result in phenotypes that are symmetrically opposite to those resulting from the modulation of dSPN activity[6,10,11,45,46]. A recent study that recorded dSPN and iSPN activity using dual-colour photometry showed that although dSPNs and iSPNs are generally coactive around movement onset, the balance between the two pathways during the coactive period can predict the magnitude of turning behaviour in a freely moving mouse, consistent with the competitive model[40].

We note that the refinement and competitive models are not necessarily mutually exclusive but might emphasize different aspects of iSPN function. Consider a hypothetical case in which three pairs of dSPNs and iSPNs project to three unique SNr neurons downstream (Fig. 1d). The iSPN and dSPN in a particular pair might compete against each other because they share the same postsynaptic SNr neuron target. However, the other two iSPNs do not share this postsynaptic target, and their activity will thus sharpen the representation in the downstream SNr without competing with the dSPN. Thus, the function of an iSPN might depend on its exact anatomical target in relation to that of a dSPN. Sets of iSPNs could thus both refine a motor programme and compete with dSPNs to determine the probability and duration of the motor programme.

Dopaminergic modulation of iSPN plasticity

In addition to anatomical and functional studies, another way to understand iSPN function is to consider the learning rule that dictates when synapses onto iSPNs undergo LTP. Although the manner in which iSPNs modulate behaviour might be multifaceted and striatal region-dependent, the learning rule that dictates when synapses onto iSPNs undergo LTP (and thus boosts their activity) is likely more or less uniform across the striatum. This is because iSPNs in different regions

of the striatum display similar expression profiles of neuromodulator receptors, are innervated by neurons in similar regions (such as the cortex and thalamus) and all receive dopaminergic input[47–50]. Therefore, understanding the conditions under which iSPN input synapses undergo LTP may provide insights into the underlying principles that unify the diverse contributions of iSPNs to behaviour.

Synaptic plasticity of striatal neurons has been extensively studied in the context of dopamine[26,47,51–54]. This is because the striatum is the major target of dopaminergic neurons located in the ventral tegmental area and substantia nigra pars compacta[49,55]. dSPNs and iSPNs are dynamically regulated by dopamine acting via dopamine receptors, a subclass of G-protein-coupled receptors (GPCRs). iSPNs predominantly express D2-like dopamine receptors, which – upon activation by dopamine – inhibit adenylyl cyclase (AC) and decrease the activity of protein kinase A (PKA). This influences many cellular properties[47] including the potential for synaptic plasticity[52,54,56,57]. Given the high affinity of D2 receptors for dopamine, it had long been assumed that PKA in iSPNs would be inhibited by the baseline tonic dopamine levels present in the striatum in vivo, thus preventing LTP induction in baseline conditions[47]. Thus, the brief decreases in dopamine concentration below baseline that have been observed in vivo during the occurrence of negative outcomes such as a foot shock or consumption of bitter water would hypothetically allow iSPNs to undergo LTP by reducing the activation of D2 receptors and upregulating PKA[58,59].

Two recent studies have indeed shown evidence for this hypothesis. Recent work from our laboratory used a newly developed lifetime-photometry system to measure the net activity of PKA in dSPNs and iSPNs in the nucleus accumbens in mice performing a food-motivated place conditioning task[25]. PKA in iSPNs was consistently activated when dopamine concentrations fell below baseline, a phenomenon observed when well-trained mice experienced reward omission. Optogenetic inactivation of dopamine neurons in naive mice confirmed that reductions in dopamine concentration are sufficient to activate PKA in a graded manner, with longer durations causing a higher increase in PKA level. Interestingly, optogenetic activation of dopamine neurons did not modulate PKA, suggesting that D2 receptors might be saturated at baseline levels of dopamine in vivo. In another study, Iino et al. used glutamate photolysis (glutamate uncaging) to induce LTP at postsynaptic terminals in the dendritic spines of individual iSPNs while optogenetically activating dopaminergic axons in brain slices[26]. The authors showed that tonic excitation of dopaminergic fibres prevented LTP (determined by measuring the enlargement of the spine head, a surrogate for LTP). LTP was rescued by a D2-receptor antagonist, suggesting that tonic dopamine excitation can block LTP. Furthermore, glutamate uncaging during a pause in dopamine firing induced LTP in a pause duration-dependent manner, with a longer pause enabling a higher change in spine volume. Although these two studies were conducted in the nucleus accumbens, it is likely that a similar learning rule would be observed in other regions of the striatum, given that iSPNs in different striatal regions all express D2-like receptors.

Three-factor rule for iSPN plasticity

Based on the studies described in the previous section, which highlighted the role of dopamine dips in gating LTP of glutamatergic synapses on iSPNs, we now propose a three-factor rule for iSPN LTP. This learning rule has many similarities to previously proposed models of synaptic plasticity in the basal ganglia[12,60–67]. However, by specifically considering iSPNs, we can outline the predictions of the learning rule with regards to the activity of iSPNs, the types of input impinging on individual iSPNs and the functional consequences of the heterogeneity of dopamine dips that have been observed in various behaviours and laboratory tasks.

We assume that an individual iSPN, located in the matrix compartment of a motor region of the striatum, receives three kinds of inputs: a dopaminergic input, an efference copy input and a state input (Fig. 2a). The dopaminergic input is a teaching signal that gates LTP. Consistent with previous findings, we assume that the dopaminergic signal needs to dip below baseline to deactivate D2 receptors and allow LTP to occur. Efference copies have been observed throughout the brain (see Supplementary Box 1) and we assume that each iSPN receives a dedicated efference copy input for a particular motor programme. More specifically, we assume that if an individual iSPN is capable of suppressing a motor programme, it will receive an efference copy input for that particular motor programme. Lastly, the state input provides information about the context in which the motor action is carried out, which includes sensory information, the internal state and latent representations of the world that the animal has inferred through experience. The information carried by the state input likely varies by striatal subregion, given the parallel organization of the cortical input onto the cortex (see Supplementary Box 2).

We propose that synapses carrying the state input are selectively potentiated when a dopamine dip occurs and there is a coincident activity in the state and efference copy inputs (Fig. 2a). To illustrate this process, consider a hypothetical example in which a mouse that has previously learned to lick a left spout in response to a cue for a water reward undergoes extinction training (Fig. 2b). In trial 1, the mouse licks the left spout after hearing the cue (tone A) to receive the water. However, extinction begins from trial 2 onwards, meaning that licking the left spout after hearing tone A no longer delivers a reward. After two trials of such reward omission, the mouse no longer licks the left spout after hearing tone A (trial 4), effectively suppressing a motor programme that is no longer optimal. How could this extinction occur via the learning rule described above? Consider an iSPN that is capable of suppressing the left lick and receives a dopaminergic input, an efference copy related to a left lick and a state input that signals the presence of tone A (Fig. 2c). We assume for the sake of illustration that dopamine in this example would increase during the receipt of a water reward and decrease during a reward omission. In trial 1, the mouse licks the left spout and receives a reward. This means that all three inputs (dopamine, efference copy and tone) are high, and LTP does not occur (Fig. 2c). In trials 2 and 3, the mouse does not receive a water reward after licking the left spout. This is signalled by a dip in dopamine concentration. Given that all three conditions are now met, this induces LTP at the synapse connecting tone A to the iSPN, allowing the iSPN to become more active after tone onset. In trial 4, the iSPN activity is high enough to suppress the licking of the left spout and the mouse no longer licks after tone A, effectively extinguishing the behaviour. Although the above example is highly simplistic, it illustrates how a dopamine-dependent learning rule could shape iSPN activity in order to suppress a non-optimal motor programme that has previously led to a dip in dopamine level.

This three-factor learning rule is consistent with both the refinement model and the competitive model. For the refinement model, one could imagine that iSPNs undergo LTP to refine the motor programme by inhibiting the non-optimal movement trajectories that previously lead to a dip in dopamine level. Alternatively, in the competitive model, iSPNs will undergo LTP in order to overcome the influence of the dSPNs. In the latter case, the magnitude and polarity of the experience-dependent synaptic plasticity in each striatal pathway

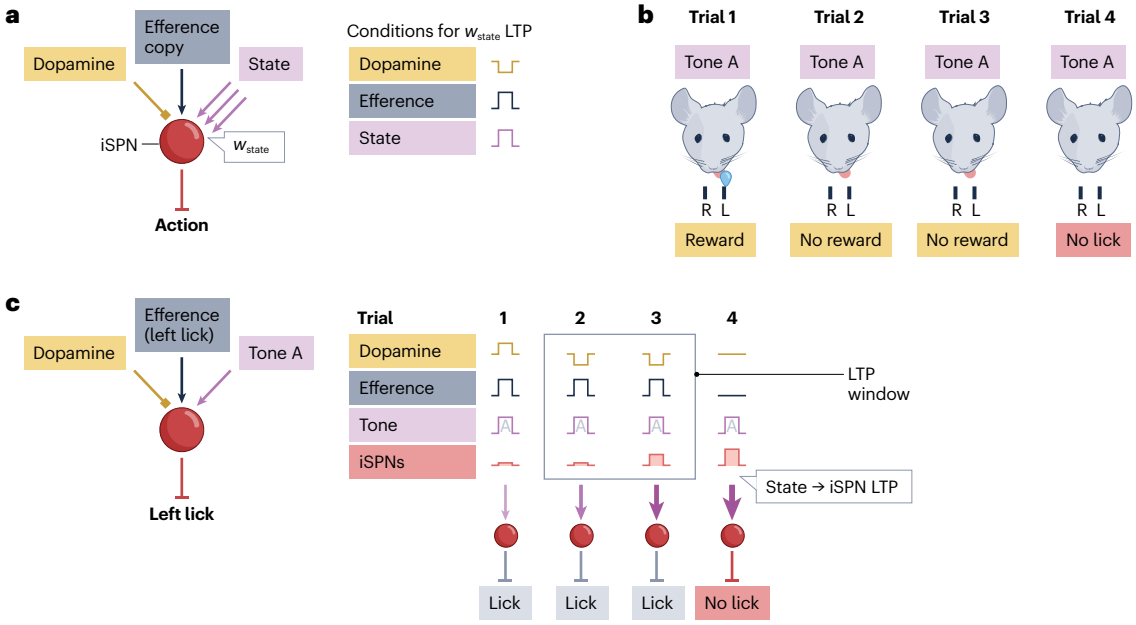


Fig. 2 | A framework for reinforcement learning in indirect striatal projection neurons. a, In the proposed three-factor learning rule, each indirect striatal projection neuron (iSPN) receives three distinct inputs: a dopamine signal, an efference copy signal and a state signal (left). Multiple arrows from the state represent the high-dimensional nature of this signal as compared with the other types of signals and show that iSPNs receive multiple inputs corresponding to different states. The function of iSPNs is to learn when to suppress a specific action using these three inputs. The conditions under which long-term potentiation (LTP) occurs for a state input–iSPN synapse (w_{state}) are a dip in dopamine that coincides with a state input and efference copy inputs (right). **b**, A mouse learning to suppress licking to an unrewarded spout after hearing a cue tone A. In trial 1, the mouse licks the left spout after tone onset to collect a water reward. We assume that the mouse has already learned the tone A–left lick association. In trials 2 and 3, the mouse licks after tone A onset, but no longer receives a water reward. In trial 4, the mouse no longer licks the left spout after hearing tone A. **c**, The inputs to a hypothetical iSPN whose activity can suppress a left lick (left) and how the three-factor learning rule could allow an iSPN to learn to suppress licking (right). A dip in dopamine during trials 2 and 3 allows the iSPN to undergo LTP. The LTP occurs at the synapse between the state (tone A) and the iSPN, allowing the tone to activate the iSPN much more efficiently in trial 4 and suppress left licks. Part **b** adapted from ref. 21, Springer Nature Limited.

will ultimately determine which pathway ends up winning, and thus controlling behaviour.

Predictions of the learning rule

We now discuss predictions of the three-factor learning rule with respect to the activity of iSPNs and the plasticity of distinct inputs. We compare the predictions of the learning rule and consider whether they are consistent with previous studies.

iSPN activity

One key prediction of the three-factor learning rule is that iSPNs receive an efference copy signal that indicates that the movement has been initiated. More specifically, if an iSPN is cable of suppressing motor programme A, then the iSPN will receive an efference copy whenever motor programme A is initiated. If this efference copy input is strong enough, it might paradoxically activate iSPNs at movement onset. Indeed, many studies have found that iSPNs are active around movement onset, consistent with this prediction[38–40,44,68]. However, this poses a challenge for the brain: although iSPN activity that reflects the efference copy signal may be necessary for learning about the environment, this activity could also interfere with ongoing movement. This is because iSPN activity is generally thought to be anti-kinetic. Thus, the efference copy might itself suppress the movement that has just been initiated. How could one solve this problem? A recent computational study demonstrated an elegant solution to this problem[64] in which the efference copy signal is provided to both dSPNs and iSPNs so that their activities perfectly cancel each other downstream. Any surplus activity in either pathway will still influence behaviour. In this way, action selection and learning can be multiplexed without interference. Alternatively, the efference copy signal might depolarize the iSPN, but not enough to cause the cell to spike and cause interference[44,68,69]. This idea was proposed by Fee, who hypothesized that the efference copy innervates the dendritic shaft of the neurons, gating plasticity but not causing spiking[70].

iSPNs receiving an efference copy of the ongoing action might also explain, to some degree, why activity in populations of iSPNs encodes information about the ongoing movement or behavioural syllable (an action motif that the animal naturally displays)[44,68,69,71–73]. Although controversies still exist as to what exactly striatal neurons encode as well as how this differs between dSPNs and iSPNs, previous studies generally support the idea that iSPNs encode both the onset and type of movement, potentially reflecting an efference copy[38,44,68,71]. Further work is necessary to examine if and when the activity in iSPNs represents an actual efference copy necessary for LTP.

Potentiation of state versus efference copy input

The three-factor learning rule also makes specific predictions about the types of inputs to iSPNs and their role in LTP. Specifically, it assumes

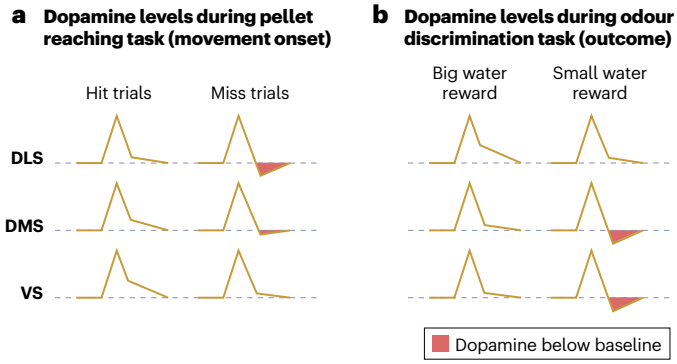


Fig. 3 | Task-dependent heterogeneity in dopamine suppression. a, Dopamine response profile around movement onset in the dorsolateral striatum (DLS), dorsomedial striatum (DMS) and ventral striatum (VS) measured via photometry during a pellet reaching task, aligned to hit trials (left) or miss trials (right)[86]. Red area indicates the period during which the dopamine concentration dipped below baseline (defined as the level during the period before cue onset). The dopamine dip in miss trials was most prominent in the DLS but less so in the DMS and VS. **b**, Dopamine response profile measured via photometry during an odour discrimination task[81]. Dopamine dips in response to the outcomes of trials in which the animals were rewarded with smaller amounts of water were most prominent in the VS and DMS but not present in the DLS.

that the efference copy synapse does not undergo LTP per se but that the state input does. To date, there is no study that has directly compared the ability of different inputs onto striatal neurons to undergo LTP. However, Fee proposed that thalamostriatal neurons preferentially projecting onto the dendritic shafts of striatal neurons could provide an efference copy signal whereas cortical input preferentially projecting onto the dendritic spines of striatal neurons could provide a state input[65]. Other anatomical studies have proposed that different deep cortical layer cell types, such as pyramidal tract neurons and intratelencephalic neurons, could provide distinct inputs to the striatum based on the type of synapses they make and their connectivity with different cell types[74–76]. Overall, more work is needed to tease apart which anatomical projections might be responsible for providing the state versus efference copy input onto iSPNs.

On the functional side, studies in rats have shown that motor memories are stored within thalamostriatal projections, and that corticostriatal projections are only necessary for learning a motor skill[77,78]. A closer look at the three-factor learning rule reveals that state–action memories are predicted to be stored within the state–iSPN synapses, which are plastic, but not within efference copy–iSPN synapses, with the latter being only necessary for learning. Taken together, this suggests that the motor cortex might provide the efference copy signal necessary for learning, with the thalamostriatal projection providing the state signal and ultimately storing the memories of a motor skill. It would be interesting to examine whether similar findings can be replicated in different regions of the striatum in a non-motor learning context in which corticostriatal and thalamostriatal projections are important for learning and storing task specific memories, respectively.

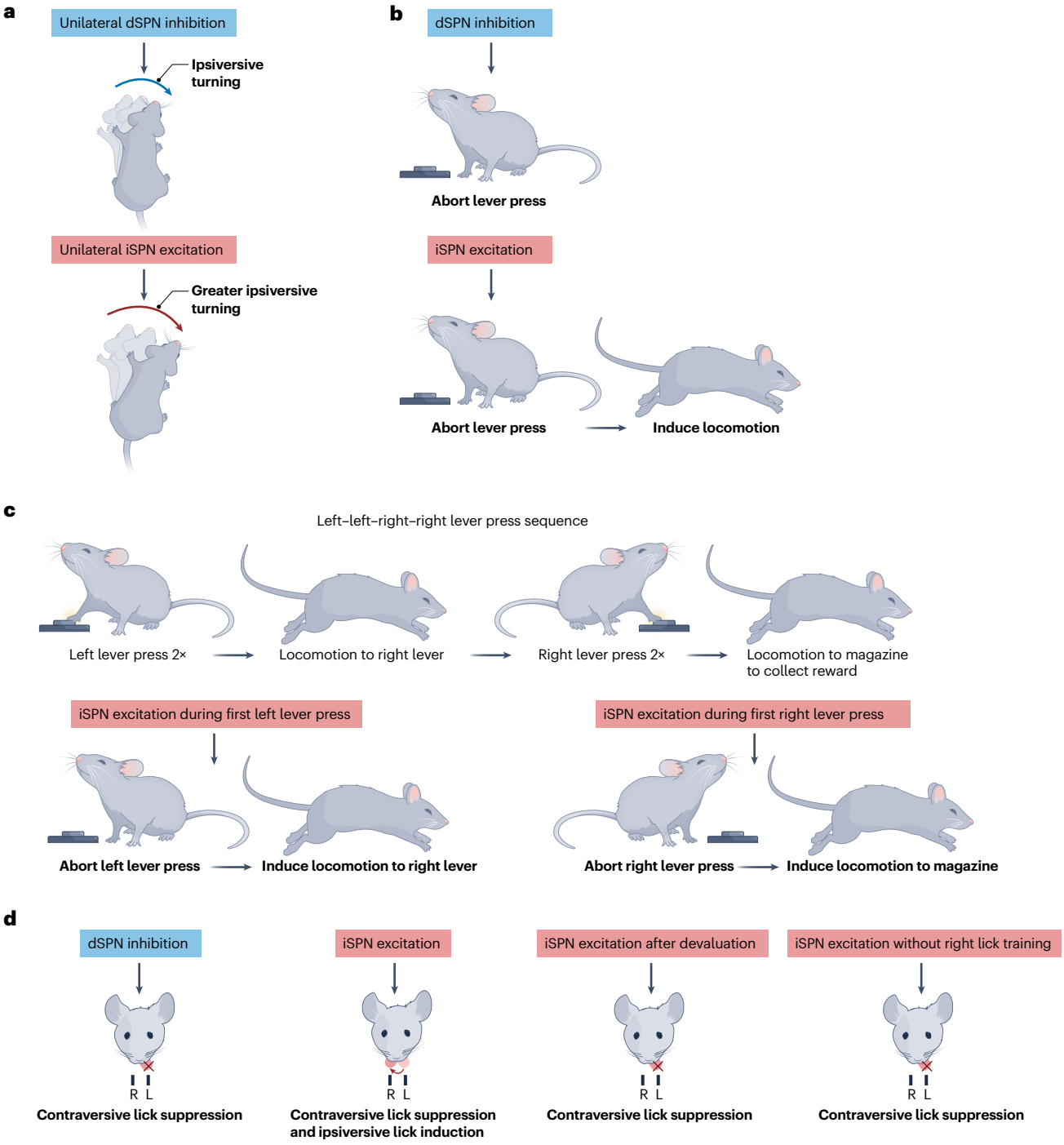
Regional heterogeneity in dopamine signals

In the three-factor learning rule, we have assumed that a negative outcome (such as reward omission) is represented via a dip in dopamine concentration in the striatum (Fig. 2c). Indeed, many studies have found that a negative outcome (or worse than expected outcome) is usually signalled by a dip in dopamine. These include delivery of a foot shock[79] or noxious stimuli[80], an airpuff[27–29], a smaller-sized water reward[81], reward omission[25,30,82–85], failure to grab a pellet[86] and bitter water[87]. However, there is considerable evidence that not all striatal regions encode negative outcomes via a dip in dopamine levels[79,81,86,88–92]. For instance, a study in primates has shown that the delivery of an aversive stimulus such as an airpuff can also excite dopaminergic neurons located in the dorsal lateral part of the substantia nigra pars compacta, potentially signalling the salience of the stimulus[29]. Negative outcome signalling is also spatially heterogeneous in mice, such that foot shocks induce a dip in dopamine only in the lateral but not in the medial nucleus accumbens shell[79]. Furthermore, dopamine in other striatal subregions, such as the tail of the striatum, appears to encode something akin to a threat prediction error[87,93].

The three-factor rule does not by itself assume that a negative outcome should be encoded via a dip in dopamine levels across the whole striatum. In fact, by restricting the striatal regions in which a particular negative outcome is represented via a dip in dopamine, the basal ganglia might spatially restrict the striatal region in which iSPNs can undergo LTP. This point can be illustrated by considering two studies that have simultaneously recorded dopamine levels in three different striatal regions. Phillips et al. trained mice to grab a pellet with their forelimb while recording dopamine levels via photometry in the dorsolateral striatum (DLS), DMS and VS[86] (Fig. 3a). Comparing trials in expert mice in which a pellet was successfully grabbed (hit trials) versus those in which it was missed revealed that dopamine dipped below baseline in the DLS but not in the DMS or VS. Thus, a negative outcome, which could be used to correct and refine the forelimb trajectory, is only represented in the DLS. Coincidentally, the DLS has been shown to be necessary for skilled forelimb movement[77,94,95]. Thus, an interesting possibility is that iSPNs in the DLS are preferentially recruited to learn the correct movement trajectory to perform the pellet reaching task.

Fig. 4 | Activation of indirect striatal projection neurons leads to alternative action. Findings of studies in which optogenetic activation of indirect striatal projection neurons (iSPNs) led to specific actions in various task settings **a**, Behavioural effect of unilateral optogenetic inhibition of direct striatal projection neurons (dSPNs) (top) versus unilateral optogenetic excitation of iSPNs (bottom) in the dorsolateral striatum (DLS)[39]. Both manipulations caused ipsiversive turning but the magnitude of turning was greater for iSPN excitation. **b**, Behavioural effect of bilaterally inactivating dSPNs (top) versus bilaterally activating iSPNs (bottom) in the DLS in mice trained to press a lever eight times for a reward[31]. dSPN inactivation suppressed lever pressing whereas iSPN activation both suppressed lever pressing and induced locomotion away from the lever. **c**, Mice were trained to perform a lever press sequence in order to receive a reward[19]. Bilaterally activating iSPNs in the DLS during the first left lever press caused mice to abort the sequence and induced locomotion towards the right lever. Activating the iSPNs during the first right lever press also caused mice to abort the sequence but induced locomotion towards the magazine. **d**, Behavioural effect of unilaterally manipulating dSPNs and iSPNs in the ventrolateral striatum (VLS) of mice trained to lick either the left or right spout after hearing a tone[21]. dSPN inactivation suppressed contraversive licking. iSPN activation suppressed contraversive licking and induced ipsiversive licking. The same iSPN manipulation did not lead to ipsiversive licking when the contraversive spout was devalued or when the mouse had never been trained to lick the contraversive spout.

In another study, Tsutsui-Kimura et al. recorded dopamine levels using photometry in the DLS, DMS and VS in mice trained to perform an odour discrimination task[81] (Fig. 3b). Although the regions sampled were similar to those in the study by Phillips et al., the authors found that a small water reward caused a dip in dopamine concentrations compared with a large water reward, but only in the VS and DMS. Coincidentally, the VS and DMS have been implicated in updating state value, as well as orienting behaviour in sensory-guided evidence accumulation tasks[96–98]. Thus, iSPNs in the VS and DMS might be specifically engaged to learn to discriminate the odours and guide the orienting behaviour in their task. Overall, these two studies highlight the fact that dips in dopamine concentration during specific behavioural events are not uniform across the striatum. This likely allows iSPNs in specific striatal regions to undergo LTP and suppress specific actions related to the outcomes encoded by the local dopamine dip. Thus, each striatal domain could represent an independent reinforcement learning



module, with a unique teaching signal, efference copy and state input, allowing much more targeted learning for a given behavioural task (see Supplementary Box 2).

Generating alternative actions

Being able to suppress behaviour that previously led to a negative outcome is necessary for optimal behaviour, and the three-factor learning rule provides a potential mechanism for achieving this. However, an animal still needs to perform an alternative action in place of the action that was suppressed, in order to keep searching for the best course of action. Interestingly, a few studies that have optogenetically activated iSPNs in various tasks have come to the seemingly surprising conclusion that activation of iSPNs can induce a new action in a context-dependent manner.

Many studies have shown that unilateral activation of either dSPNs or iSPNs in the dorsal striatum causes rotational behaviour[6,11,39,42,99,100]. In 2014, Tecuapetla et al. compared the effect of unilaterally inactivating dSPNs with that of unilaterally activating iSPNs[39] (Fig. 4a). Although both manipulations caused the mouse to turn ipsilaterally, the magnitude of turning was about three times higher in the iSPN activation group compared with the dSPN inactivation group. Although it is difficult to compare the results of the two manipulations head to head given the differences in the experimental protocols used, it is interesting to speculate why this would be the case. If both manipulations suppress contraversive turning, the resulting turning behaviour should be comparable. Thus, the fact that iSPN activation caused much more turning than dSPN inactivation suggests that iSPN activation might not be equivalent to dSPN inactivation, challenging the competitive model. One potential explanation for the more potent effect of activating iSPNs, compared with dSPN inactivation, would be that iSPNs simultaneously suppress contraversive turning and induce ipsiversive turning, whereas dSPN inactivation only suppresses contraversive turning. To test this idea, it would be interesting to repeat the experiments in a context in which mice are trained to turn ipsilaterally or contralaterally (for example, in a T-maze). If correct, it would be predicted that iSPN activation would cause mice that would usually decide to turn contralaterally to switch to turn ipsilaterally, whereas dSPN inactivation would simply abort the decision to turn contralaterally.

Perhaps a more convincing study on inducing alternative action was conducted by Tecuapetla et al. in 2016 (ref. 31). In this study, the authors trained mice to press a lever eight times in order to collect a reward. One interesting finding of this study was the effect of bilaterally inactivating dSPNs versus bilaterally activating iSPNs in the DLS. The authors found that inactivating dSPNs caused mice to abort pressing the lever, but activating iSPNs not only suppressed lever pressing but also induced locomotion away from the lever (Fig. 4b). This suggests that activation of iSPNs can induce an alternative action, something that cannot be induced by inactivating dSPNs alone.

Geddes et al. used a similar paradigm but instead of training on a single lever, mice were trained inside a chamber with two distinct levers. They learned to press the levers in a sequence consisting of left–left–right–right lever presses, after which a reward was delivered via a food magazine[19] (Fig. 4c). When iSPNs were activated bilaterally in the DLS during the first left lever press, mice immediately moved to the right lever without pressing the left lever the second time. The same manipulation applied to the first right lever press made the mice move to the magazine (Fig. 4c). Thus, similar to the results of the study by Tecuapetla et al., activating iSPNs induced locomotion but in a context-dependent manner. One possible explanation for these findings is that mice maintain an internal representation of upcoming motor programmes (such as locomotion towards the right lever) that is triggered when iSPNs are activated. This mechanism would enable mice to smoothly transition to a new action during a motor sequence. Supporting this idea, the authors found that mice with ablated iSPNs struggled to learn the switch from the left to the right lever.

Finally, a study from our laboratory in which mice were trained to lick left or right spouts in order to obtain a reward demonstrated similar results[21]. When iSPNs were activated unilaterally in the VLS, the stimulation suppressed contraversive licking but also induced ipsiversive licking (Fig. 4d). Interestingly, unilaterally inactivating dSPNs only suppressed contraversive licking without inducing any ipsiversive licking. This difference between iSPN activation and dSPN inactivation seems to parallel the more potent rotational behaviour observed by Tecuapetla et al. when activating iSPNs compared with inactivating dSPNs[39]. Furthermore, the induction of ipsiversive licking by iSPN activation went away if either the corresponding spout was devalued or if mice were never trained to lick the ipsilateral spout (relative to the stimulation hemisphere), indicating that the ability of iSPN activation to trigger the motor programme for ipsiversive licking was both learning-dependent and context-dependent (Fig. 4d). We also showed that when mice underwent training to drive extinction of the learnt licking of one spout by omitting rewards after a correct lick, they naturally explored the other spout, and that this extinction-driven exploration was driven by iSPN activity.

These four studies suggest that, in certain circumstances, iSPN activation is not the same as dSPN inactivation. Furthermore, iSPNs can induce a de novo action, distinct from the action that they suppress. This can be used to smoothly transition from one action to the next during a motor sequence or to guide exploration of an alternative action during extinction. How could this new action be generated via iSPN activation alone? We propose that this effect might be mediated by competitive inhibitory circuitry in regions downstream of iSPNs that track relevant motor programmes and mediate competition (Fig. 5a). The superior colliculus (SC) possesses local connectivity and long-range connections that could allow this disinhibition to occur for different kinds of motor programmes[101–104] (Fig. 5b). The SNr could also mediate a form of local competition, although a previous study has shown that axon collaterals within the SNr do not extend far beyond the soma, thus limiting their ability to implement competition between actions regulated in distinct regions of the SNr[105,106]. It is possible that other circuits downstream of the striatum or within the striatum (via axon collaterals)[107] might also be able to implement competition. Overall, further studies are needed to understand the implications of this aspect of iSPN function, and the potential mechanism that might underlie disinhibition mediating the generation of new actions via iSPN activation.

Generalization to other types of learning

In this Perspective, we have mainly focused on the role of iSPNs in the context of sensorimotor learning. Indeed, many computational models of the basal ganglia assume that what is being learned in the striatum is the state–action relationship[62–64,108]. However, there is also a large amount of work suggesting that the dorsal striatum can be divided into a goal-directed system in the DMS and a habitual system in the DLS[109–113]. For instance, lesions in the DMS make rats less sensitive to contingency degradation and outcome devaluation, two commonly used paradigms to quantify habitual behaviour in which the contingency between the action and the outcome is degraded (contingency

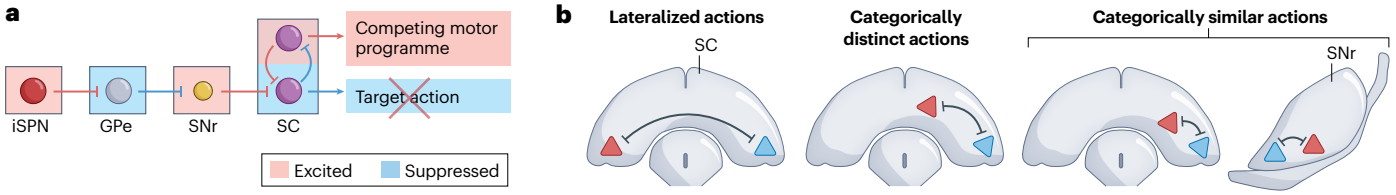


Fig. 5 | Circuit mechanisms mediating exploration via competition. a, The basal ganglia–collicular circuit that mediates exploration[21]. The suppression of a target action by indirect striatal projection neurons (iSPNs) disinhibits a competing motor programme that is represented in the superior colliculus (SC) to allow exploratory behaviour. **b**, Examples of putative competitive circuit motifs that could allow such disinhibition of an alternative programme. The selection of lateralized actions (such as leftward versus rightward turning) could be mediated by competition between the two SC hemispheres, whereby each hemisphere promotes the contraversive movement and suppresses the opposite hemisphere. Categorically distinct actions (such as a lever press versus licking) could involve competition within one SC hemisphere via mid-range axonal innervation. Categorically similar actions (such as a lever push versus a lever pull) which engage the same muscle groups could be encoded by distinct neurons within the same local region, with competition between those neurons occurring locally via short-range axonal innervations within the SC or substantia nigra reticulata (SNr). GPe, globus pallidus externus. Parts **a** and **b** adapted from ref. 21, Springer Nature Limited.

degradation) or the outcome is devalued (by providing unlimited access to food rewards before training)[112]. In contrast, rats with DLS lesions fail to develop habitual behaviour and quickly stop pressing the lever for a devalued outcome compared with control rats[111]. Inactivation of iSPNs in the DMS has also been shown to affect decision flexibility and goal-directed learning[114–116]. Despite these classic findings, it is still unclear what exactly is being learned in the DMS to allow for goal-directed behaviour. Are there abstract goals represented in the DMS that are reinforced, just as actions are reinforced in the DLS? If this is the case, then one could imagine a similar dopamine-dependent learning rule in the DMS operating not over actions but over abstract representations. Another possibility is that both the DMS and the DLS undergo reinforcement learning with a similar learning rule, but that the DLS tends to represent actions in 'chunks', making them less sensitive to outcome devaluation than those in the DMS[117–119]. For instance, it has been shown that the habitual system can be understood as forming and generating long action sequences as opposed to individual actions, a model that can recapitulate the formation of habits as well as sensitivity to contingency degradation as seen in real animals[118]. Another hierarchical view of the striatum has also been explored computationally in models in which the representations of state in the DMS and DLS are distinct, with the DMS having access to a much more abstract representation of the environment than the DLS, and the DMS and DLS work cooperatively to either promote or suppress behaviours[12].

On the other hand, there is also considerable evidence that the DMS can generate motor actions. For instance, stimulation of either the direct or indirect pathway in the DMS generates rotational behaviour[6,8,11,42,99,100], and inhibition of the DMS can impair orienting behaviour based on sensory evidence accumulation[96,98]. The DMS projects to medullary gigantocellular neurons which are required for turning gait programmes via the SNr[42,100], and DMS simulation activates the mesencephalic locomotor region[11]. The DMS also receives inputs from cingulate and secondary motor cortex areas[48] (also known as the frontal orienting field) that are involved in orienting behaviour in rats[120,121]. Thus, in contrast to the studies that point to an associative role for the DMS, many studies indicate that the DMS is involved in action control. A future challenge is to reconcile this motor function of the DMS with the goal-direct function described above.

Although we have omitted studies on the VS, it very likely that a similar three-factor learning rule operates in this region, given that the two studies that have investigated the effects of dopamine dips on iSPN LTP were conducted in the VS[25,26]. Previous studies suggest that the role of the VS is complex, with the range of functions in which it is involved including motivational control, action selection, goal-direct learning, state-value (critic) learning and feeding (see excellent reviews in refs. 122–125 for more information). This heterogeneity of function might partially reflect the cell-type heterogeneity of the VS, as demonstrated by single-cell sequencing analysis[126].

To conclude, understanding the exact functional role of non-motor regions, such as the DMS and VS, will help us better understand how iSPNs operate in a non-sensory context, and whether a similar three-factor learning rule could be generalized to non-sensorimotor learning.

Caveats and future directions

In this Perspective, we have presented a three-factor learning rule, inspired by recent experimental findings and previous computational models, and discussed how the learning rule fits with models of iSPN function. We have also discussed predictions of the learning rule and considered how well they are supported by previous studies on iSPNs. Lastly, we have discussed how iSPNs can generate alternative actions, adding a layer of complexity to their function. Below, we discuss potential caveats associated with these ideas, as well as several studies on iSPNs that might challenge the framework presented.

A few studies that have used an optogenetic self-stimulation paradigm in which iSPN activation is paired with an action or entry into an area of a testing chamber have shown that this can induce punishment[7,14,17,18] (although see ref. 127). Although these studies are interesting and generally align with the view that iSPN activity can suppresses movement/choice, the underling mechanism behind these learning effects is still unclear. At first glance, these studies seem to be at odds with the three-factor learning rule, according to which iSPN activation should not by itself be a teaching signal that can guide learning. One possibility is that self-stimulation itself can induce LTP in the stimulated iSPNs, which in turn will induce a change in behaviour such as avoidance. Previous studies have shown that optogenetic activation paired with a specific stimulus can induce new place fields or visual receptive fields in hippocampal or visual cortical neurons, respectively[128,129]. Another possibility is that regions downstream of iSPNs are remodelled to mediate learning. However, without knowing the locus of learning, it is difficult to determine whether these studies support or challenge the three-factor learning rule. Another

complication associated with self-stimulation studies is that the stimulation protocol itself seems to change the behavioural effect observed. A recent study reported that brief versus prolonged iSPN self-stimulation differentially induces reinforcement versus aversion, respectively, and that the effect of prolonged stimulation is mediated by δ-opioid receptors in the ventral pallidum[130]. This shows that self-stimulation can induce an array of effects, depending on the nature of the stimulation protocol, and suggests that researchers should try to understand exactly what is being changed or remodelled in their paradigm in order to better interpret the results of the studies using iSPN self-stimulation.

In the proposed framework, we have assumed that dopaminergic neurons encode a teaching signal to modulate synaptic plasticity in the iSPNs. Recent studies have suggested alternative explanations for what dopaminergic neurons encode, including value, salience, the learning rate, causal association and the impulse vector[58,84,85,131–133] (see refs. 134,135 for excellent reviews on this topic). Although it is beyond the scope of this Perspective to fully discuss the controversies regarding what dopaminergic neurons encode, we believe that these alternatives do not negate a function of dopamine as a teaching signal that induces synaptic plasticity. Indeed, the role of dopamine neurons in modulating synaptic plasticity in slice has been widely replicated[26,52,53,136,137]. What remains controversial is what exactly dopamine neurons encode. Future studies that consider the teaching component of dopamine neurons should yield an understanding of how their activity modulates iSPNs, and whether the three-factor learning rule is a good conceptual framework.

Directly testing the three-factor learning rule in vivo might require experiments in which dopamine neurons are artificially inactivated, in order to see whether LTP in iSPNs can be observed. In such experiments, the state signal could be provided via a sensory stimulus, and the inhibition of dopaminergic neurons could be paired with an action performed by the mouse. This would create all three of the conditions that the rule proposes are necessary for LTP in iSPNs. Directly observing LTP in vivo is technically challenging and might require two-photon structural imaging of spines or slice physiology in trained mice in order to determine whether synaptic plasticity has occurred. Despite these difficulties, artificially recreating the conditions necessary for LTP in vivo might yield important insights into whether iSPNs follow the three-factor learning rule.

Studies using careful analysis of behaviour might also be fruitful for refining our model of iSPN function. Recent machine learning tools for pose tracking have enabled researchers to understand the fine kinematics of subtle movements, including tongue kinematics during licking[138–142]. These studies have begun to reveal how complex seemingly simple movements are. Adopting a motor-centric approach to investigate how iSPNs contribute to sensorimotor control may provide deeper insights into the fundamental principles governing iSPN learning, which could then be generalized to other non-motor functions. For instance, one could ablate iSPNs either during or after a mouse has learned a skilled movement. By carefully tracking the movement using high-speed cameras and pose tracking tools, one might reveal interesting features of motor control that go awry when iSPNs are ablated. This could then be followed up by studies examining how dopamine activity contributes to the refinement of iSPN activity. Does a dip in dopamine coincide with a specific event during motor learning? Does disrupting this dip hinder motor learning by disrupting iSPN LTP? Overall, we believe studies investigating fine motor control might be fruitful for testing some of the predictions outlined in this Perspective.

On the other hand, testing how iSPNs can generate a new action might require the use of a task in which a mouse has to choose between categorically distinct actions. For instance, a paradigm in which a mouse has to perform either action A or action B, which are categorically distinct actions (for example, locomotion versus forelimb reaching), would allow us to test whether iSPN-mediated suppression of action A could induce action B. Thus, studies that allow mice to generate categorically distinct actions might start to reveal how iSPNs can flexibly use downstream circuity to generate new actions via disinhibition.

In conclusion, many challenges remain, and further experiments are needed to fully elucidate the exact function of iSPNs. Here, we have presented two conceptual frameworks that may help advance our understanding of iSPN function. The first framework outlines a simple learning rule that allows iSPNs to undergo dopamine-dependent LTP, allowing suppression of non-optimal motor programmes. The second framework outlines the potential role of iSPNs to generate a new alternative action, allowing exploration of new motor programmes. We hope that future research will refine, expand or challenge the ideas proposed in this Perspective, ultimately contributing to a deeper understanding of iSPN function during behaviour.

Published online: 07 May 2025

References

1. Thorndike, E. L. Animal intelligence: an experimental study of the associative processes in animals. *Psychol. Rev.* **5**, 551–553 (1898).
2. Kravitz, A. V. & Kreitzer, A. C. Striatal mechanisms underlying movement, reinforcement, and punishment. *Physiology (Bethesda)* https://doi.org/10.1152/physiol.00004.2012 (2012).
3. Tye, K. M. Neural circuit motifs in valence processing. *Neuron* **100**, 436–452 (2018).
4. Berridge, K. C. Affective valence in the brain: modules or modes? *Nat. Rev. Neurosci.* **20**, 225–234 (2019).
5. Hikida, T., Kimura, K., Wada, N., Funabiki, K. & Nakanishi, S. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* **66**, 896–907 (2010).
6. Kravitz, A. V. et al. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622–626 (2010).
7. Kravitz, A. V., Tye, L. D. & Kreitzer, A. C. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* **15**, 816–818 (2012).
8. Tai, L.-H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* **15**, 1281–1289 (2012).
9. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
10. Freeze, B. S., Kravitz, A. V., Hammack, N., Berke, J. D. & Kreitzer, A. C. Control of basal ganglia output by direct and indirect pathway projection neurons. *J. Neurosci.* **33**, 18531–18539 (2013).
11. Roseberry, T. K. et al. Cell-type-specific control of brainstem locomotor circuits by basal ganglia. *Cell* **164**, 526–537 (2016).
12. Cruz, B. F. et al. Action suppression reveals opponent parallel control via striatal circuits. *Nature* **607**, 521–526 (2022).
13. Oldenburg, I. A. & Sabatini, B. L. Antagonistic but not symmetric regulation of primary motor cortex by basal ganglia direct and indirect pathways. *Neuron* **86**, 1174–1181 (2015).
14. Yttri, E. A. & Dudman, J. T. Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* **533**, 402–406 (2016).
15. Sheng, M., Lu, D., Shen, Z. & Poo, M. Emergence of stable striatal D1R and D2R neuronal ensembles with distinct firing sequence during motor learning. *Proc. Natl Acad. Sci. USA* **116**, 11038–11047 (2019).
16. Durieux, P. F., Schiffmann, S. N. & de Kerchove d'Exaerde, A. Differential regulation of motor control and response to dopaminergic drugs by D1R and D2R neurons in distinct dorsal striatum subregions. *EMBO J.* **31**, 640–653 (2012).
17. Bonnavion, P. et al. Striatal projection neurons coexpressing dopamine D1 and D2 receptors modulate the motor function of D1- and D2-SPNs. *Nat. Neurosci.* **27**, 1783–1793 (2024).
18. Isett, B. R. et al. The indirect pathway of the basal ganglia promotes transient punishment but not motor suppression. *Neuron* **111**, 2218–2231.e4 (2023).
19. Geddes, C. E., Li, H. & Jin, X. Optogenetic editing reveals the hierarchical organization of learned action sequences. *Cell* **174**, 32–43.e15 (2018).
20. Nonomura, S. et al. Monitoring and updating of action selection for goal-directed behavior through the striatal direct and indirect pathways. *Neuron* **99**, 1302–1314.e5 (2018).

21. Lee, J. & Sabatini, B. L. Striatal indirect pathway mediates exploration via collicular competition. *Nature* **599**, 645–649 (2021).
22. Jin, X. & Costa, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
23. Zalocusky, K. A. et al. Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* **531**, 642–646 (2016).
24. LeBlanc, K. H. et al. Striatopallidal neurons control avoidance behavior in exploratory tasks. *Mol. Psychiatry* **25**, 491–505 (2020).
25. Lee, S. J. et al. Cell-type-specific asynchronous modulation of PKA by dopamine in learning. *Nature* **590**, 451–456 (2021).
26. Iino, Y. et al. Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature* **579**, 555–560 (2020).
27. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
28. Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *eLife* **5**, e17328 (2016).
29. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
30. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
31. Tecuapetla, F., Jin, X., Lima, S. Q. & Costa, R. M. Complementary contributions of striatal projection pathways to action initiation and execution. *Cell* **166**, 703–715 (2016).
32. Nelson, A. B. & Kreitzer, A. C. Reassessing models of basal ganglia function and dysfunction. *Annu. Rev. Neurosci.* **37**, 117–135 (2014).
33. Bariselli, S., Fobbs, W. C., Creed, M. C. & Kravitz, A. V. A competitive model for striatal action selection. *Brain Res.* **1713**, 70–79 (2019).
34. Klaus, A., Alves da Silva, J. & Costa, R. M. What, if, and when to move: basal ganglia circuits and self-paced action initiation. *Annu. Rev. Neurosci.* **42**, 459–483 (2019).
35. Park, J., Coddington, L. T. & Dudman, J. T. Basal ganglia circuits for action specification. *Annu. Rev. Neurosci.* **43**, 485–507 (2020).
36. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
37. Alexander, G. E. & Crutcher, M. D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* **13**, 266–271 (1990).
38. Cui, G. et al. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* **494**, 238–242 (2013).
39. Tecuapetla, F., Matias, S., Dugue, G. P., Mainen, Z. F. & Costa, R. M. Balanced activity in basal ganglia projection pathways is critical for contraversive movements. *Nat. Commun.* **5**, 4315 (2014).
40. Meng, C. et al. Spectrally resolved fiber photometry for multi-component analysis of brain circuits. *Neuron* **98**, 707–717.e4 (2018).
41. Jeon, H. et al. Topographic connectivity and cellular profiling reveal detailed input pathways and functionally distinct cell types in the subthalamic nucleus. *Cell Rep.* **38**, 110439 (2022).
42. Lee, J., Wang, W. & Sabatini, B. L. Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. *Nat. Neurosci.* **23**, 1388–1398 (2020).
43. Foster, N. N. et al. The mouse cortico-basal ganglia–thalamic network. *Nature* **598**, 188–194 (2021).
44. Klaus, A. et al. The spatiotemporal organization of the striatum encodes action space. *Neuron* **95**, 1171–1180.e7 (2017).
45. Bateup, H. S. et al. Distinct subclasses of medium spiny neurons differentially regulate striatal motor behaviors. *Proc. Natl Acad. Sci. USA* **107**, 14845–14850 (2010).
46. Carvalho Poyraz, F. et al. Decreasing striatopallidal pathway function enhances motivation by energizing the initiation of goal-directed action. *J. Neurosci.* **36**, 5988–6001 (2016).
47. Tritsch, N. X. & Sabatini, B. L. Dopaminergic modulation of synaptic transmission in cortex and striatum. *Neuron* **76**, 33–50 (2012).
48. Hunnicutt, B. J. et al. A comprehensive excitatory input map of the striatum reveals novel functional organization. *eLife* **5**, e19103 (2016).
49. Azcorra, M. et al. Unique functional responses differentially map onto genetic subtypes of dopamine neurons. *Nat. Neurosci.* **26**, 1762–1774 (2023).
50. Gokce, O. et al. Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
51. Plotkin, J. L., Day, M. & Surmeier, D. J. Synaptically driven state transitions in distal dendrites of striatal spiny neurons. *Nat. Neurosci.* **14**, 881–888 (2011).
52. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
53. Yagishita, S. et al. A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* **345**, 1616–1620 (2014).
54. Kreitzer, A. C. & Malenka, R. C. Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* **445**, 643–647 (2007).
55. Matsuda, W. et al. Single nigrostriatal dopaminergic neurons form widely spread and highly dense axonal arborizations in the neostriatum. *J. Neurosci.* **29**, 444–453 (2009).
56. Calabresi, P. et al. Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J. Neurosci.* **17**, 4536–4544 (1997).
57. Gerdeman, G. L., Partridge, J. G., Lupica, C. R. & Lovinger, D. M. It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* **26**, 184–192 (2003).
58. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* **68**, 815–834 (2010).
59. Hikida, T. et al. Pathway-specific modulation of nucleus accumbens in reward and aversive behavior via selective transmitter receptors. *Proc. Natl Acad. Sci. USA* **110**, 342–347 (2013).
60. González-Redondo, Á. et al. Reinforcement learning in a spiking neural model of striatum plasticity. *Neurocomputing* **548**, 126377 (2023).
61. Mikhael, J. G. & Bogacz, R. Learning reward uncertainty in the basal ganglia. *PLOS Comput. Biol.* **12**, e1005062 (2016).
62. Collins, A. G. E. & Frank, M. J. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol. Rev.* **121**, 337–366 (2014).
63. Blackwell, K. T. & Doya, K. Enhancing reinforcement learning models by including direct and indirect pathways improves performance on striatal dependent tasks. *PLOS Comput. Biol.* **19**, e1011385 (2023).
64. Lindsey, J., Markowitz, J. E., Gillis, W. F., Datta, S. R. & Litwin-Kumar, A. Dynamics of striatal action selection and reinforcement learning. *eLife* **13**, RP101747 (2024).
65. Fee, M. Oculomotor learning revisited: a model of reinforcement learning in the basal ganglia incorporating an efference copy of motor actions. *Front. Neural Circuits* **6**, 38 (2012).
66. Frank, M. J. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *J. Cogn. Neurosci.* **17**, 51–72 (2005).
67. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* **7**, 967–975 (2006).
68. Markowitz, J. E. et al. The striatum organizes 3D behavior via moment-to-moment action selection. *Cell* **174**, 44–58.e17 (2018).
69. Parker, J. G. et al. Diametric neural ensemble dynamics in parkinsonian and dyskinetic states. *Nature* **557**, 177–182 (2018).
70. Fee, M. S. The role of efference copy in striatal learning. *Curr. Opin. Neurobiol.* **25**, 194–200 (2014).
71. Varin, C., Cornil, A., Houtteman, D., Bonnavion, P. & de Kerchove d'Exaerde, A. The respective activation and silencing of striatal direct and indirect pathway neurons support behavior encoding. *Nat. Commun.* **14**, 4982 (2023).
72. Weglage, M. et al. Complete representation of action space and value in all dorsal striatal pathways. *Cell Rep.* **36**, 109437 (2021).
73. Barbera, G. et al. Spatially compact neural clusters in the dorsal striatum encode locomotion relevant information. *Neuron* **92**, 202–213 (2016).
74. Reiner, A., Hart, N. M., Lei, W. & Deng, Y. Corticostriatal projection neurons—dichotomous types and dichotomous functions. *Front. Neuroanat.* **4**, 142 (2010).
75. Deng, Y. et al. Differential organization of cortical inputs to striatal projection neurons of the matrix compartment in rats. *Front. Syst. Neurosci.* **9**, 51 (2015).
76. Lei, W., Jiao, Y., Mar, N. D. & Reiner, A. Evidence for differential cortical input to direct pathway versus indirect pathway striatal projection neurons in rats. *J. Neurosci.* **24**, 8289–8299 (2004).
77. Wolff, S. B. E., Ko, R. & Ölveczky, B. P. Distinct roles for motor cortical and thalamic inputs to striatum during motor skill learning and execution. *Sci. Adv.* **8**, eabk0231 (2022).
78. Kawai, R. et al. Motor cortex is required for learning but not for executing a motor skill. *Neuron* **86**, 800–812 (2015).
79. de Jong, J. W. et al. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101**, 133–151.e7 (2019).
80. Yang, H. et al. Pain modulates dopamine neurons via a spinal–parabrachial–mesencephalic circuit. *Nat. Neurosci.* **24**, 1402–1413 (2021).
81. Tsutsui-Kimura, I. et al. Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *eLife* **9**, e62390 (2020).
82. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
83. Rios, A. et al. Reward expectation enhances action-related activity of nigral dopaminergic and two striatal output pathways. *Commun. Biol.* **6**, 914 (2023).
84. Mohebi, A. et al. Dissociable dopamine dynamics for learning and motivation. *Nature* **570**, 65–70 (2019).
85. Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2016).
86. Phillips, C. D., Hodge, A. T., Myers, C. C., Leventhal, D. K. & Burgess, C. R. Striatal dopamine contributions to skilled motor learning. *J. Neurosci.* **44**, e0240242024 (2024).
87. Menegas, W., Akiti, K., Amo, R., Uchida, N. & Watabe-Uchida, M. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat. Neurosci.* **21**, 1421–1430 (2018).
88. Engel, L. et al. Dopamine neurons drive spatiotemporally heterogeneous striatal dopamine signals during learning. *Curr. Biol.* **34**, 3086–3101.e4 (2024).
89. Lerner, T. N. et al. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
90. Yuan, L., Dou, Y.-N. & Sun, Y.-G. Topography of reward and aversion encoding in the mesolimbic dopaminergic system. *J. Neurosci.* **39**, 6472–6481 (2019).
91. Hamid, A. A., Frank, M. J. & Moore, C. I. Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell* **184**, 2733–2749.e16 (2021).
92. Horvitz, J. C. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96**, 651–656 (2000).
93. Green, I., Amo, R. & Watabe-Uchida, M. Shifting attention to orient or avoid: a unifying account of the tail of the striatum and its dopaminergic inputs. *Curr. Opin. Behav. Sci.* **59**, 101441 (2024).
94. Dhawale, A. K., Wolff, S. B. E., Ko, R. & Ölveczky, B. P. The basal ganglia control the detailed kinematics of learned motor skills. *Nat. Neurosci.* **24**, 1256–1269 (2021).

95. Lemke, S. M., Ramanathan, D. S., Guo, L., Won, S. J. & Ganguly, K. Emergent modular neural control drives coordinated motor actions. *Nat. Neurosci.* **22**, 1122–1131 (2019).
96. Bolkan, S. S. et al. Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. *Nat. Neurosci.* **25**, 345–357 (2022).
97. Rothenhoefer, K. M. et al. Effects of ventral striatum lesions on stimulus-based versus action-based reinforcement learning. *J. Neurosci.* **37**, 6902–6914 (2017).
98. Yartsev, M. M., Hanks, T. D., Yoon, A. M. & Brody, C. D. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *eLife* **7**, e34929 (2018).
99. Guo, L., Walker, W. I., Ponvert, N. D., Penix, P. L. & Jaramillo, S. Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nat. Commun.* **9**, 1534 (2018).
100. Cregg, J. M., Sidhu, S. K., Leiras, R. & Kiehn, O. Basal ganglia–spinal cord pathway that commands locomotor gait asymmetries in mice. *Nat. Neurosci.* **27**, 716–727 (2024).
101. Takahashi, M., Sugiuchi, Y. & Shinoda, Y. Topographic organization of excitatory and inhibitory commissural connections in the superior colliculi and their functional roles in saccade generation. *J. Neurophysiol.* **104**, 3146–3167 (2010).
102. Takahashi, M., Sugiuchi, Y., Izawa, Y. & Shinoda, Y. Commissural excitation and inhibition by the superior colliculus in tectoreticular neurons projecting to omnipause neuron and inhibitory burst neuron regions. *J. Neurophysiol.* **94**, 1707–1726 (2005).
103. Doykos, T. K., Gilmer, J. I., Person, A. L. & Felsen, G. Monosynaptic inputs to specific cell types of the intermediate and deep layers of the superior colliculus. *J. Comp. Neurol.* **528**, 2254–2268 (2020).
104. Sooksawate, T., Isa, K., Behan, M., Yanagawa, Y. & Isa, T. Organization of GABAergic inhibition in the motor output layer of the superior colliculus. *Eur. J. Neurosci.* **33**, 421–432 (2011).
105. Mailly, P., Charpier, S., Menetrey, A. & Deniau, J.-M. Three-dimensional organization of the recurrent axon collateral network of the substantia nigra pars reticulata neurons in the rat. *J. Neurosci.* **23**, 5247–5257 (2003).
106. Brown, J., Pan, W.-X. & Dudman, J. T. The inhibitory microcircuit of the substantia nigra provides feedback gain control of the basal ganglia output. *eLife* **3**, e02397 (2014).
107. Dobbs, L. K. et al. Dopamine regulation of lateral inhibition between striatal neurons gates the stimulant actions of cocaine. *Neuron* **90**, 1100–1113 (2016).
108. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* **10**, 732–739 (2000).
109. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476 (2006).
110. Balleine, B. W., Delgado, M. R. & Hikosaka, O. The role of the dorsal striatum in reward and decision-making: fig. 1. *J. Neurosci.* **27**, 8161–8165 (2007).
111. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19**, 181–189 (2004).
112. Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* **22**, 513–523 (2005).
113. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behav. Brain Res.* **166**, 189–196 (2006).
114. Kwak, S. & Jung, M. W. Distinct roles of striatal direct and indirect pathways in value-based decision making. *eLife* **8**, e46050 (2019).
115. Peak, J., Chieng, B., Hart, G. & Balleine, B. W. Striatal direct and indirect pathway neurons differentially control the encoding and updating of goal-directed learning. *eLife* **9**, e58544 (2020).
116. Matamales, M. et al. Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. *Science* **367**, 549–555 (2020).
117. Dezfouli, A. & Balleine, B. W. Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* **35**, 1036–1051 (2012).
118. Dezfouli, A. & Balleine, B. W. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLOS Comput. Biol.* **9**, e1003364 (2013).
119. Dezfouli, A., Lingawi, N. W. & Balleine, B. W. Habits as action sequences: hierarchical action control and changes in outcome value. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130482 (2014).
120. Kopec, C. D., Erlich, J. C., Brunton, B. W., Deisseroth, K. & Brody, C. D. Cortical and subcortical contributions to short-term memory for orienting movements. *Neuron* **88**, 367–377 (2015).
121. Sinnamon, H. M. & Galer, B. S. Head movements elicited by electrical stimulation of the anteromedial cortex of the rat. *Physiol. Behav.* **33**, 185–190 (1984).
122. van der Meer, M. A. & Redish, A. D. Ventral striatum: a critical look at models of learning and evaluation. *Curr. Opin. Neurobiol.* **21**, 387–392 (2011).
123. Floresco, S. B. The nucleus accumbens: an interface between cognition, emotion, and action. *Annu. Rev. Psychol.* **66**, 25–52 (2015).
124. Kelley, A. E. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neurosci. Biobehav. Rev.* **27**, 765–776 (2004).
125. Mannella, F., Gurney, K. & Baldassarre, G. The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Front. Behav. Neurosci.* **7**, 135 (2013).
126. Chen, R. et al. Decoding molecular and cellular heterogeneity of mouse nucleus accumbens. *Nat. Neurosci.* **24**, 1757–1771 (2021).
127. Vicente, A. M., Galvão-Ferreira, P., Tecuapetla, F. & Costa, R. M. Direct and indirect dorsolateral striatum pathways reinforce different action strategies. *Curr. Biol.* **26**, R267–R269 (2016).
128. El-Boustani, S. et al. Locally coordinated synaptic plasticity of visual cortex neurons in vivo. *Science* **360**, 1349–1354 (2018).
129. Rolotti, S. V. et al. Local feedback inhibition tightly controls rapid formation of hippocampal place fields. *Neuron* **110**, 783–794.e6 (2022).
130. Soares-Cunha, C. et al. Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. *Mol. Psychiatry* **25**, 3241–3255 (2020).
131. Hughes, R. N. et al. Ventral tegmental dopamine neurons control the impulse vector during motivated behavior. *Curr. Biol.* **30**, 2681–2694.e5 (2020).
132. Coddington, L. T., Lindo, S. E. & Dudman, J. T. Mesolimbic dopamine adapts the rate of learning from action. *Nature* **614**, 294–302 (2023).
133. Jeong, H. et al. Mesolimbic dopamine release conveys causal associations. *Science* **378**, eabq6740 (2022).
134. Gershman, S. J. et al. Explaining dopamine through prediction errors and beyond. *Nat. Neurosci.* **27**, 1645–1655 (2024).
135. Coddington, L. T. & Dudman, J. T. Learning from action: reconsidering movement signaling in midbrain dopamine neuron activity. *Neuron* **104**, 63–77 (2019).
136. Reynolds, J. N. J., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).
137. Reynolds, J. N. J. & Wickens, J. R. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* **15**, 507–521 (2002).
138. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
139. Pereira, T. D. et al. SLEAP: a deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022).
140. Marshall, J. D. et al. Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire. *Neuron* **109**, 420–437.e8 (2021).
141. Bollu, T. et al. Cortex-dependent corrections as the tongue reaches for and misses targets. *Nature* **594**, 82–87 (2021).
142. Xu, D. et al. Cortical processing of flexible and context-dependent sensorimotor sequences. *Nature* **603**, 464–469 (2022).

Acknowledgements
The authors thank members of the Sabatini laboratory for helpful discussions.

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41583-025-00925-2.

Peer review information *Nature Reviews Neuroscience* thanks Minoru Kimura and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.