

Dipl research notes

Tuesday, 11 March 2025 11:59

Major dopamine pathways in the human brain (sagittal view). Dopamine neurons in the midbrain (ventral tegmental area and substantia nigra) project to various regions: the mesolimbic (purple shading toward limbic regions) and mesocortical pathways target limbic and cortical areas, while the nigrostriatal pathway (arrows to the striatum) modulates the basal ganglia. (Not shown here, the tuberoinfundibular pathway projects from the hypothalamus to the pituitary gland.)

Dopamine Pathways and Neural Circuit Functions

Dopamine is a neuromodulatory neurotransmitter with several **major pathways** in the brain that subserve distinct functions en.wikipedia.org en.wikipedia.org. The four classic dopamine pathways are: **mesolimbic**, **mesocortical**, **nigrostriatal**, and **tuberoinfundibular** en.wikipedia.org. Each pathway consists of dopamine-producing neurons in a specific origin that project to target regions, influencing different neural circuits involved in reward, motivation, movement, and hormone regulation en.wikipedia.org en.wikipedia.org. Below, we examine each pathway's anatomy and function from a neural circuits perspective, highlighting how dopamine release modulates neuronal activity and synaptic plasticity in those circuits.

Mesolimbic Pathway (Reward Circuit)

The **mesolimbic pathway** originates in the ventral tegmental area (VTA) of the midbrain and projects to limbic regions of the **ventral striatum** (notably the nucleus accumbens) and other parts of the limbic system en.wikipedia.org en.wikipedia.org. This pathway is famously associated with the brain's *reward system*. Dopamine release from VTA neurons into the nucleus accumbens signals **incentive salience** – essentially the “wanting” or motivational value of a reward – and reinforces behaviors that lead to reward en.wikipedia.org. It facilitates **reward-related learning and reinforcement**, strengthening synapses in circuits that were active before a rewarding outcome en.wikipedia.org pmc.ncbi.nlm.nih.gov. From a circuit perspective, mesolimbic dopamine enhances the activity of nucleus accumbens neurons in response to reward cues and influences plasticity in their inputs. This modulation effectively “stamps in” rewarding actions, orchestrating goal-directed behavior via the ventral basal ganglia circuit pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov.

. Dysregulation of the mesolimbic pathway (e.g. overactivation by addictive drugs) leads to maladaptive plasticity in these reward circuits and is implicated in **addiction** en.wikipedia.org (drug-induced changes in VTA→NAc synapses alter AMPA receptor function and persistently bias the circuit towards drug-seeking pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov).

Mesocortical Pathway (Cognitive Control)

The **mesocortical pathway** also originates from dopamine neurons in the VTA, but projects to the **prefrontal cortex (PFC)** and surrounding cortical areas en.wikipedia.org. This pathway is crucial for higher-order cognitive functions – for example, **executive functions** such as working memory, attention, and decision-making en.wikipedia.org. Dopamine in the PFC modulates neuronal activity by enhancing the signal-to-noise ratio of cortical inputs. In practical terms, moderate levels of dopamine (primarily via D1 receptors in PFC) stabilize the activity of PFC neural ensembles that are coding task-relevant information, while suppressing distracting inputs pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov LINK D1, D2 tests

. There is an **“inverted-U”** relationship, where an optimal dopamine level is required for peak PFC cognitive performance – too little or too much impairs functions like working memory pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov

. At the circuit level, mesocortical dopamine can gate the flow of information: for instance, phasic dopamine signals to PFC (perhaps in concert with basal ganglia inputs) may flag which neural representations should be **maintained or updated** in working memory. Dopamine's modulatory effect on PFC synapses also influences synaptic plasticity underlying learning and flexibility; it can facilitate long-term potentiation in PFC networks when outcomes are better than expected, thereby linking **reward feedback to cognitive adjustments**.

Nigrostriatal Pathway (Motor Control)

The **nigrostriatal pathway** consists of dopamine neurons in the substantia nigra pars compacta (SNc, in the midbrain) projecting to the **dorsal striatum** (the caudate nucleus and putamen in the basal ganglia) en.wikipedia.org en.wikipedia.org. This pathway is critical for **motor control and habit learning**. Dopamine released in the dorsal striatum modulates the activity of the basal ganglia motor loop that selects and initiates movements

en.wikipedia.org en.wikipedia.org. In the striatum, dopamine has a dual action on two populations of GABAergic medium spiny neurons: it **excites the direct pathway** neurons via D1 receptors and **inhibits the indirect pathway** neurons via D2 receptors nba.uth.tmc.edu

. Through this mechanism, a burst of nigrostriatal dopamine simultaneously promotes the direct pathway (which facilitates the thalamus and cortex) and suppresses the indirect pathway (which otherwise inhibits thalamocortical drive)

nba.uth.tmc.edu nba.uth.tmc.edu

. The net effect is to *disinhibit* thalamic and cortical motor circuits, allowing a selected motor program to be expressed nba.uth.tmc.edu. In other words, dopamine in the dorsal striatum **gates voluntary movement**, enabling desired movements and suppressing competing ones nba.uth.tmc.edu nba.uth.tmc.edu LINK D1 and D2 tests DOPAMINE AS A TRIGGER FOR DIRECT ACTION. HOW ARE THESE ACTIONS PROGRAMMED?

. This modulatory role extends to synaptic plasticity: dopamine signals during motor learning reinforce the cortico-striatal synapses that were active during successful movements, gradually ingraining those motor skills as habits.

When the nigrostriatal pathway is compromised – for example, the degeneration of SNc dopamine neurons in **Parkinson's disease** – the result is a paucity of movement (bradykinesia) and inability to initiate actions, because the balance shifts toward excessive inhibition by the indirect pathway nba.uth.tmc.edu. Indeed, the classic motor symptoms of Parkinson's (tremors, rigidity, poverty of movement) reflect the loss of dopamine's normal facilitatory influence on motor circuits en.wikipedia.org nba.uth.tmc.edu.

Tuberoinfundibular Pathway (Hormone Regulation)

The **tuberoinfundibular pathway** is a dopamine projection from the hypothalamus (specifically the arcuate nucleus of the tuberal region) to the **pituitary gland** (infundibulum) en.wikipedia.org. Unlike the other pathways, this is a neuroendocrine circuit. Dopamine released into the pituitary portal circulation primarily acts to **inhibit prolactin secretion** from anterior pituitary cells en.wikipedia.org. In effect, dopamine serves as **prolactin-inhibiting factor**: it tonically suppresses prolactin release, helping regulate lactation and reproductive functions. From a circuits perspective, the tuberoinfundibular dopamine neurons respond to signals like suckling or estrogen levels, reducing their firing when it's appropriate to allow prolactin release (e.g. during breastfeeding). If this pathway is blocked or dopamine tone is lost, prolactin levels rise – for instance, antipsychotic drugs that block D2 dopamine receptors can produce **hyperprolactinemia** as a side effect

en.wikipedia.org. While this pathway does not directly involve synaptic plasticity in a neural network (since it targets endocrine cells), it exemplifies dopamine's modulatory role: here, modulation of hormone release rather than neural excitability.

Dopamine's Role in Neuronal Activity and Synaptic

Plasticity

Across all these pathways, dopamine acts not as a classical excitatory or inhibitory transmitter but as a **neuromodulator** that alters how neural circuits process information. Dopaminergic neurons typically fire in two modes: a low tonic firing that sets background dopamine levels and phasic bursts in response to salient events (e.g. unexpected rewards or cues predicting them) neuronaldynamics.epfl.ch. These phasic surges broadcast a signal that something important happened, influencing target neurons' physiology and the strength of their synapses.

Receptor Actions: Dopamine exerts its effects via G-protein-coupled receptors grouped into D1-like (D1, D5) and D2-like (D2, D3, D4) families. Activation of D1-like receptors *increases* intracellular cAMP and typically **depolarizes/excites** the neuron, whereas D2-like activation *decreases* cAMP and tends to **hyperpolarize/inhibit** the neuron

nba.uth.tmc.edu nba.uth.tmc.edu

. At the circuit level, this means dopamine can have **opposing effects on different neurons** depending on their receptor expression. For example, in the striatum D1-receptor neurons (direct pathway) are facilitated and D2-receptor neurons (indirect pathway) are suppressed by dopamine DIRECT INDIRECT PATHS?

nba.uth.tmc.edu. More generally, D1 receptor stimulation often enhances NMDA receptor currents and other excitatory processes, while D2 activation reduces them mdpi.com

. Because of these mechanisms, dopamine can bidirectionally modulate a circuit's output: it might increase the gain of certain neural pathways and decrease the gain of others simultaneously. This delicate balance allows dopamine to **filter and shape network activity**, as seen in PFC where the right amount of D1 activation filters distractors and stabilizes task-relevant ensembles

pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov

Modulation of Plasticity: Dopamine's influence goes beyond immediate excitability – crucially, it serves as a **third factor** guiding synaptic plasticity (the strengthening or weakening of connections). Classic Hebbian plasticity depends on two factors (pre- and post-synaptic activity), but dopamine acts as an additional factor that can gate these Hebbian changes pmc.ncbi.nlm.nih.gov pmc.ncbi.nlm.nih.gov. In many brain regions, dopamine timing determines whether Hebbian changes are actually consolidated into long-term memory. A burst of dopamine following coincident pre- and postsynaptic activity can “lock in” potentiation (LTP), whereas the absence or a dip of dopamine might lead to weakening (LTD) or no change. This is exemplified by experiments in the striatum and hippocampus: in the striatum, co-activation of D1 and D2 receptors is required to induce a certain form of long-term depression, indicating dopamine's permissive role in synaptic weakening as well CONNECTION BETWEEN D1 AND D2 AND SYNAPTIC PLASTICITY MODULATION

pmc.ncbi.nlm.nih.gov

. In the hippocampus, dopamine acting on D1 receptors is needed shortly after a learning event to induce protein synthesis for late-phase LTP – essentially marking that event as significant to be stored as long-term memory mdpi.com

mdpi.com

. If dopamine is blocked or mistimed, synaptic changes from a learning trial may fade instead of persisting mdpi.com

. Mechanistically, dopamine-driven plasticity often relies on intracellular cascades (like cAMP/PKA signaling) that affect synaptic proteins and gene transcription. D1 receptor activation can trigger molecular pathways (e.g. MAPK/CREB) that promote gene expression and new protein synthesis necessary for stable LTP

mdpi.com

. This means a dopamine surge can turn a transient synaptic change into a long-lasting one by engaging the cell's biochemical machinery. In effect, **dopamine provides a teaching signal**: when dopamine is released in response to a positive outcome, only the synapses that were recently active (and hence have “eligibility traces” of Hebbian activity) get potentiated – a biological implementation of *reward-based credit assignment*. Conversely, a pause in dopamine firing (below baseline) can signal a negative prediction error, leading to the weakening of synapses that were active during an unrewarded attempt. In summary, dopamine modulates neural circuits on two timescales: **acutely**, by changing neuronal responsiveness (excitability, gain, oscillatory patterns), and **persistently**, by regulating synaptic plasticity and thus reconfiguring circuit wiring based on experience.

Computational Models of Dopamine Function

Decades of neuroscience research have inspired many **computational models** of dopamine's role in learning and decision-making. These models range from biologically detailed simulations of spiking neural circuits to abstract algorithms in machine learning. Below, we explore two broad categories of models – (1) **biologically plausible models** that often use spiking neural networks to mimic dopamine-modulated learning in the brain, and (2) **machine learning approaches** that draw on dopamine's principles (with a stronger emphasis on the latter, as it has driven modern AI reinforcement learning).

Biologically Plausible Spiking Models of Dopamine

Early theoretical work identified dopamine as a signal akin to a **reward prediction error** – the difference between expected and received reward. Notably, Montague et al. (1996) and Schultz et al. (1997) proposed that phasic bursts and dips of dopamine correspond to the positive or negative prediction errors in **temporal-difference (TD) learning** algorithms

frontiersin.org

frontiersin.org

. This hypothesis was supported by recordings of dopamine neurons in monkeys: they fire strongly when an unexpected reward occurs, but if a cue reliably predicts the reward, the dopamine neurons shift to firing at the cue and omit the response to the expected reward; if an expected reward is omitted, dopamine activity briefly drops below baseline

neuronaldynamics.epfl.ch

. This pattern matches the TD error used in reinforcement learning. Such insights led to **actor-critic models** of the basal ganglia, where the critic's TD error is carried by dopamine to teach the actor (policy) to choose rewarding actions.

In spiking neural network (SNN) models, researchers have implemented “**three-factor learning rules**” that incorporate dopamine as a global third factor modulating synaptic weight changes. One seminal model by R. Florian (2007) showed that a **reward-modulated spike-timing-dependent plasticity (STDP)** rule can enable a network of integrate-and-fire neurons to learn tasks via reinforcement

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. In his simulations, synapses adjusted their strength only when pre/post spike timing was followed by a dopamine-mediated reward signal. By including an **eligibility trace** at each synapse (a temporary memory of recent spike coincidences), the model could associate delayed rewards with the appropriate synapses – successfully solving challenges like the XOR problem and learning target output patterns

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. Similarly, Izhikevich (2007) demonstrated a spiking cortical network that solves the “**distal reward problem**” (credit assignment with delayed feedback) by linking STDP with a dopamine signal

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. In his model, synapses that fired together set a synaptic tag (eligibility), and a dopamine burst within a few seconds would convert those tags into long-term synaptic changes. This effectively bridges the temporal gap: random activity in the interim doesn't erode the earlier correlation, so when dopamine arrives, it **selectively strengthens the right synapses** that caused the reward

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. Biologically detailed models of the basal ganglia have also incorporated dopamine-modulated plasticity to replicate learning and action selection. For example, Chadderdon et al. (2012) built a spiking model of motor cortex and trained it to perform a reaching task using a dopamine-like reinforcement signal

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. They provided a global reward (+1), neutral, or punishment (-1) signal corresponding to phasic increases, no change, or decreases in dopamine neuron firing, respectively, based on the movement outcome. Synaptic plasticity in the network occurred only at eligible synapses (those with recent pre/post activity) when the dopamine signal was delivered

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

. Notably, learning was most successful when **both** dopamine bursts and dips were used – i.e. when the network was signaled for positive outcomes and negative outcomes – mirroring the brain's use of dopamine highs and lows for reward and aversive learning. Within 180 seconds of simulated training, the spiking network learned to accurately move the virtual arm to multiple target positions

pubmed.ncbi.nlm.nih.gov

. This and similar models underscore that **dopamine can serve as a global teaching signal in spiking networks**, capable of training them to produce complex behaviors through reward-guided synaptic adjustments.

In summary, biologically plausible models treat dopamine as a diffuse reinforcement signal that modulates synaptic plasticity according to timing rules. These models have validated that **neuromodulated STDP** (sometimes called **reward-STDP** or **Spike-Timing Dependent Reinforcement Learning**) is a powerful mechanism: with it, spiking neural networks can solve typical reinforcement learning tasks, aligning with the idea that real neural circuits learn by the same principle

pubmed.ncbi.nlm.nih.gov

pmc.ncbi.nlm.nih.gov

. Such models have even anticipated experimental findings – for instance, the predicted “eligibility trace” was later observed in animal studies, where a dopamine burst within seconds of neural activity can induce potentiation, but not if it comes too late

neurondynamics.epfl.ch

mdpi.com

Dopamine-Inspired Models in Machine Learning and AI

Dopamine's computational role has heavily influenced machine learning, especially in the field of **reinforcement learning (RL)**. The core algorithms of RL – such as Q-learning or policy gradient methods – often incorporate a **temporal-difference error** that echoes the dopamine reward prediction error theory

frontiersin.org

. In fact, the development of TD learning in the 1980s provided a formalism that neuroscience later linked to dopamine; this cross-pollination led to a fruitful synergy. Modern deep RL agents (used in AI for games, robotics, etc.) make use of dopamine-inspired signals: for example, the breakthrough DQN algorithm that learned to play Atari games utilized reward prediction errors to update its network (conceptually akin to dopamine teaching a neural network)

frontiersin.org

. Researchers have explicitly drawn parallels – citing that “*the conceptual role of dopamine has largely shaped the development of modern RL algorithms*”

frontiersin.org

. A striking recent example is **distributional reinforcement learning**: in AI, distributional RL algorithms learn a distribution of possible future rewards rather than a single expected value, and intriguingly, experiments found that populations of dopamine neurons in the brain also appear to **encode a distribution of prediction errors** (different neurons responding to different quantiles of reward outcomes)

frontiersin.org

frontiersin.org

. This was demonstrated by Dabney et al. (2020), showing a direct convergence of biological dopamine research and cutting-edge machine learning techniques.

In the realm of spiking neural networks, dopamine's principles are being used to design more brain-like AI systems. One approach is to implement **plasticity rules with neuromodulation** in artificial networks and optimize them for performance. For instance, Schmidgall et al. (2021) introduced a framework called “*SpikePropamine*” in which both the synaptic weights and the *rules of dopamine-modulated plasticity* are learned through gradient descent in an SNN

frontiersin.org

frontiersin.org

. Essentially, they let the network learn how to learn: the parameters governing a dopamine-like neuromodulatory signal's effect on synapses were tuned on various tasks. Their results showed that networks endowed with this learned neuromodulatory plasticity could solve complex temporal tasks that static-weight networks failed at, and even control a high-dimensional robotic movement with robustness to new conditions

frontiersin.org

frontiersin.org

. This demonstrates the promise of combining biologically inspired learning rules (like dopamine modulating STDP) with modern training techniques – yielding AI systems that continue to adapt in deployment, much like brains do.

Another notable development is the use of **surrogate gradient and eligibility trace methods** to train deep spiking networks. Bellec et al. (2020), for example, developed an algorithm for training recurrent SNNs using **eligibility traces** (inspired by neurobiology's solution to delayed credit assignment) combined with gradient-based optimization. By maintaining neuron-specific traces of recent activity (analogous to synaptic eligibility) and applying a global error signal, they achieved learning in SNNs in a manner that is both efficient and aligned with dopamine's strategy of delayed reinforcement

frontiersin.org

. These techniques allow spiking networks to approach the performance of traditional ANNs while retaining a more biologically plausible update rule.

In summary, dopamine's influence on AI appears in two ways: conceptually, through **reinforcement learning algorithms** that mirror dopamine's reward prediction error (enabling agents to learn from trial and error), and practically, through **neuromodulated learning rules in spiking networks** that bring machine learning closer to brain-like learning. The strong emphasis on the latter in recent research stems from a desire to harness features of biological learning – like continual adaptation, energy efficiency, and autonomy. By imitating dopamine's modulatory role, researchers have built spiking systems that can **learn from rewards, assign credit over time, and self-organize** in adaptive ways that standard neural networks typically don't without extensive retraining. This dopamine-inspired approach has been key in bridging neuroscience and AI: from the theoretical foundations (e.g. the actor-critic architecture derived from basal ganglia circuits) to tangible algorithms (deep RL and neuromorphic learning rules), dopamine's legacy is deeply imprinted in how we model learning in both brains *and* machines

frontiersin.org

frontiersin.org

. **References:** The content above cites key references for each major point. For quick access: Montague et al. 1996 and Schultz et al. 1997 for dopamine prediction-error theory

frontiersin.org

frontiersin.org

; Florian 2007 for reward-modulated STDP in SNNs

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

; Izhikevich 2007 for STDP with dopamine solving distal reward problem

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

; Chadderdon et al. 2012 for spiking motor learning with dopamine signals

pubmed.ncbi.nlm.nih.gov

pubmed.ncbi.nlm.nih.gov

; Schmidgall et al. 2021 (*SpikePropamine*) for learned neuromodulatory plasticity in SNNs

frontiersin.org

frontiersin.org

; and others as indicated in the text. These provide more detailed discussions on dopamine pathways, synaptic plasticity mechanisms, and computational models bridging biology and machine learning.

You said:

Can you provide certain papers and explain how reward prediction or "wanting" results in a system carrying out an associated action. At this stage i am not interested in temporal difference, just the way certain states are associated with reward, releasing some rewarding signal even before the rewarding action is carried out, if i understand this correctly.

2/2

ChatGPT said:

I will gather research papers discussing how states become associated with rewards in neural systems, leading to a predictive dopamine release before the rewarding action is executed. This includes how dopamine modulates neural activity to encode and reinforce motivational states ("wanting") and how certain cues or contexts trigger anticipatory reward signals.

To ensure I focus on the most relevant studies:

- Would you like papers focusing on **biological neuroscience** (e.g., dopamine signaling in the mesolimbic pathway, nucleus accumbens activity) or **computational models** (e.g., spiking neural networks, reinforcement learning architectures that capture this effect)?
- Are you interested in how this mechanism applies to **habit formation, addiction, decision-making, or motor behavior**, or should I keep it general?

Let me know so I can refine the research accordingly!

You said:

Keep it general and provide computational models, but that are biologically accurate

ChatGPT said:

Got it! I will gather research papers discussing computational models that capture how certain states become associated with rewards in a biologically accurate manner. The focus will be on models that explain how anticipatory dopamine release occurs before the rewarding action is carried out, reinforcing motivational states ("wanting").

I will find papers that specifically cover:

- How neural circuits encode state-reward associations biologically.
- How these circuits generate predictive dopamine signals based on learned cues.
- Computational models (preferably spiking neural networks) that accurately replicate these mechanisms.

I'll update you once I have relevant papers and summaries.

Research completed in 4m · 21 sources

Predictive Dopamine Signals and Biologically Plausible Reward Models

Dopamine Anticipates Reward and Drives "Wanting"

After learning, dopamine neurons fire in anticipation of rewards – they shift their burst of activity to cues or states that predict an upcoming reward. For example, after classical conditioning, midbrain dopamine cells respond to a conditioned stimulus (CS) that predicts a reward and cease firing at the expected time of the reward (unconditioned stimulus, US) itself

pubmed.ncbi.nlm.nih.gov

. This *anticipatory* dopamine release is thought to encode a reward prediction signal and contributes to motivation or "wanting" for the reward. In the words of incentive-salience theory, when a cue is attributed with incentive salience it triggers a pulse of "wanting" that pulls the individual toward the cue and its associated reward

journals.plos.org

. In practice, this means learned cues or contexts can evoke dopamine surges **before** a rewarding action is taken, creating a motivational state that reinforces pursuit of the reward.

Neural Circuit Models of Predictive Dopamine

Researchers have developed biologically detailed circuit models to explain how these predictive dopamine signals arise. A landmark model by **Brown, Bullock & Grossberg (1999)** proposed that the basal ganglia use two parallel pathways – one excitatory, one inhibitory – converging on dopamine neurons to produce reward prediction errors

pubmed.ncbi.nlm.nih.gov

. In this "dual-pathway" model, an excitatory pathway (via ventral striatum → ventral pallidum → PPTN) learns to cause phasic dopamine bursts at the moment a reward-predicting cue appears, while a slower inhibitory pathway (via striosomes in the striatum) learns to suppress dopamine cell firing at the expected time of the reward

pubmed.ncbi.nlm.nih.gov

. As a result, once a cue reliably predicts a reward, the dopamine neuron shows a burst for the cue and no longer fires for the predicted reward. If an expected reward is omitted, the learned inhibitory signal goes unopposed and causes a transient dip in dopamine activity instead

pubmed.ncbi.nlm.nih.gov

. This mechanism provides a biologically plausible alternative to abstract temporal-difference (TD) learning models, explaining phenomena like cue-induced dopamine, reward omission dips, and sensitivity to reward timing in a single framework

pubmed.ncbi.nlm.nih.gov

. Building on this, **Tan and Bullock (2008)** extended the circuit model to account for dopamine signals related not just to reward prediction error, but also to **reward uncertainty**. Empirical studies showed that dopamine neurons can exhibit a slow ramping or sustained activity when reward timing or probability is uncertain (in addition to the phasic bursts for unexpected rewards)

pubmed.ncbi.nlm.nih.gov

. Tan and Bullock's model introduced a third pathway involving striatal projection neurons that co-release neurotransmitters (GABA and substance P) onto dopamine cells and their inputs. This pathway computes the uncertainty or variability of reward timing and magnitude

pubmed.ncbi.nlm.nih.gov

. In their simulated network, dopamine neurons display both the classic phasic burst signals and an uncertainty-related component that is highest at intermediate (50%) reward probabilities

pubmed.ncbi.nlm.nih.gov

. In other words, the model's dopamine output reflects not only learned expected value, but also the **confidence or uncertainty** in that expectation. This aligns with observations that cue-elicited dopamine signals depend on reward probability and timing – e.g. larger sustained dopamine response when the reward is less predictable

pubmed.ncbi.nlm.nih.gov

. By incorporating an "uncertainty pathway" (mediated by striatal microcircuitry and cholinergic interneurons), the model reproduced nuanced dopamine firing patterns under probabilistic reward schedules, further tying dopamine's predictive signals to known basal ganglia circuit elements

pubmed.ncbi.nlm.nih.gov

. Such circuit models illustrate how complex learning signals (prediction, timing, uncertainty) can emerge from biologically realistic neural architecture.

Notably, other frameworks like the **PVLV (Primary Value, Learned Value) model** by O'Reilly and colleagues take a similar circuit-based approach. PVLV assigns *separate* neural substrates to primary rewards vs. learned cues: a ventral striatal "PV" system that suppresses dopamine firing for expected rewards, and an amygdala-centered "LV" system that drives dopamine bursts for reward-predicting cues

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

. In this model, cues (CSs) activate the LV system to evoke dopamine release, whereas when the actual reward (US) arrives as expected, the PV system (via inhibitory projections and lateral habenula inputs) dampens dopamine cell firing

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

. This division helps explain individual differences in motivation: for example, some animals (or individuals) are *sign-trackers* who heavily "want" the cue itself, while others are *goal-trackers* who focus on the reward. The PVLV model accounts for such differences by allowing the cue-driven dopamine (LV) vs. reward-driven dopamine (PV) to vary in dominance

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

. It also captures phenomena like conditioned inhibition and second-order conditioning by using opponent pathways for excitation and inhibition in each system

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

. Overall, these neural models underscore that predictive dopamine transients arise from specific circuit interactions (e.g. amygdala or striatal pathways) that biologically implement a kind of **teaching signal** for learning and motivation.

Spiking Neural Network Models of Reward Prediction

In addition to systems-level circuits, researchers have built **spiking neural network (SNN)** models to drill down into the neuronal and synaptic mechanisms that learn cue-reward associations. A prominent example is **Izhikevich's (2007)** model addressing the "distal reward problem." In a network of spiking cortical neurons, Izhikevich showed that if synaptic plasticity is modulated by dopamine in a time-dependent manner, the network can learn to associate stimuli or actions with delayed rewards

modeldb.science

. The key mechanism is an eligibility trace at synapses: neural activity leaves a transient trace that lasts a few seconds, and if a dopamine surge (as a global reward signal) arrives during that window, it selectively strengthens the recently active synapses

modeldb.science

. In effect, the slow dopamine waveform "tags" the causal activity that led to reward, despite the delay. This spiking model demonstrated how **spike-timing-dependent plasticity (STDP)** combined with dopamine signals can biologically implement reinforcement learning. Although dopamine release might occur only when the reward is received, the learned effect is that earlier firing patterns (the cues or action neurons that were active prior to reward) get reinforced – thus next time, those patterns will more reliably evoke activity, and dopamine can shift to earlier predictions. Izhikevich's work highlighted how precise neural firing sequences and a diffusely delivered dopamine signal together solve credit assignment, bridging the gap between a predictive state and future reward

modeldb.science

. Other SNN models have explicitly modeled dopamine neurons and their inputs. **Chorley and Seth (2011)**, for instance, constructed a spiking circuit model of the basal ganglia that reproduces the hallmark dopamine response patterns during conditioning

frontiersin.org

. Their network included model neurons for prefrontal cortex (providing cue-related inputs), striatum, and a dopamine neuron, with synaptic plasticity governed by dopamine-modulated STDP. In the model's operation, a brief **short-latency excitatory** input to the dopamine neuron (simulating a fast sensory pathway) produces an initial spike, and a **longer-latency inhibitory** input (via learned striatal pathways) arrives around the expected time of the reward

frontiersin.org

. Through repeated training, these dynamics led the dopamine cell to exhibit *precisely* the patterns seen in real neurons: (i) a phasic burst for the conditioned stimulus after learning, (ii) no dopamine response at the time of the predicted reward, and (iii) a timed dopamine dip below baseline if an expected reward was omitted

frontiersin.org

. The model also preserved a dopamine burst for unexpected rewards (if no cue was given), mirroring experimental observations

frontiersin.org

. Importantly, this was achieved in a **biophysically realistic** way – using spiking Izhikevich neurons and heterogeneous synaptic delays, rather than imposing any explicit clock or temporal difference algorithm. The dopamine-modulated STDP allowed cortical input patterns to strengthen striatal synapses such that the striatal inhibition would "cancel" the dopamine neuron's firing when a reward was fully predicted

frontiersin.org

frontiersin.org

. Chorley & Seth's results demonstrate that a competitive interplay of excitation and inhibition in a spiking network can naturally give rise to **reward prediction error-like signals** in dopamine activity

frontiersin.org

. This provides a concrete neural circuit mechanism for how anticipatory dopamine surges emerge and how the system maintains sensitivity to unexpected outcomes (a novel reward or a violation of expectation). Such spiking models lend credence to the idea that the brain's reward system implements TD-like learning principles **distributed across a network** of interacting neurons, rather than a single abstract computation.

Cue-Triggered Dopamine and Reinforcement of Motivational States

Learned cues and contexts not only trigger predictive dopamine spikes; they also reinforce motivational states that drive behavior. In conditioning paradigms, animals will work vigorously or approach signals that have been paired with rewards – a phenomenon attributable to dopamine-mediated incentive motivation. Computational models have been developed to explain how **cue-triggered dopamine** translates into enhanced "wanting." **Tindell et al. (2009)** presented a neural model of *incentive salience* that integrates learning with an organism's physiological state

journals.plos.org

journals.plos.org

. In standard reinforcement learning models, a cue acquires a static "cached" value based on past reward predictions. However, Tindell and colleagues note that in reality a cue's motivational pull can fluctuate dramatically with internal state (for example, a food cue is far more enticing when one is hungry)

journals.plos.org

journals.plos.org

. Their model augments the learned value of a cue with dynamic modulation by limbic state variables. Essentially, the cue's ability to elicit dopamine (and thus craving) is scaled up or down by factors like hunger, stress, or drug sensitization – *without requiring new learning in that state*

journals.plos.org

journals.plos.org

. This can explain experimental observations where an animal suddenly "wants" a reward it previously didn't care for, if given a dopaminergic stimulation or placed in a novel appetite condition, even though the cue-reward association itself hasn't changed. The model showed that only by including this physiological modulation could it replicate the rapid increases in cue-triggered pursuit seen in cases like salt appetite or drug sensitization

journals.plos.org

journals.plos.org

. In short, while the cue provides the learned prediction, the dopamine release it triggers is gated by the current state of the mesolimbic system, yielding a context-dependent surge of motivation. From a circuit perspective, this idea dovetails with evidence that **contexts or internal states can prime dopamine neurons**. For instance, the ventral pallidum and hypothalamus (which are sensitive to physiological need states) project to dopamine regions and can amplify cue responses when an animal is in a state of high need or arousal

journals.plos.org

. Likewise, models like PVLV implicitly capture this via separate pathways: the value computed by the “LV” (cue) system might be modulated by other inputs (stress, drug) that alter the gain of dopamine bursting. The result is that a **learned cue can trigger a dopamine surge powerful enough to drive vigorous seeking (“wanting”) even if the baseline value of the reward hasn’t changed**. This mechanism is believed to underlie phenomena such as cue-induced relapse in addiction and spontaneous craving – the cue ignites dopamine firing and hence motivation, due to previously learned associations now potentiated by the brain’s state.

Key Takeaways from the Research

- **Predictive Dopamine Bursts:** Dopamine neurons encode reward predictions – after conditioning, they fire at cues (states) predicting reward and not at the reward itself, unless the reward is unexpected pubmed.ncbi.nlm.nih.gov.

frontiersin.org

. This anticipatory dopamine release acts as a teaching signal and is closely linked to the motivation (“wanting”) to obtain the reward.

- **Dual-Pathway Circuit Mechanism:** Biologically detailed models show that converging excitatory and inhibitory pathways can produce these predictive dopamine signals. An excitatory pathway drives dopamine cell bursts for reward-predicting cues, while a learned inhibitory pathway times a pause to cancel responses to expected rewards (and cause dips if the reward is missing) pubmed.ncbi.nlm.nih.gov.
- This mechanism replicates reward prediction error signals in a neuroanatomically grounded way.
- **Extended Signals (Timing and Uncertainty):** Beyond simple prediction errors, dopamine circuits can encode the timing and uncertainty of expected rewards. A local circuit model with an additional striatal pathway reproduces sustained dopamine elevations for uncertain rewards and matches dopamine’s dependence on reward probability and timing pubmed.ncbi.nlm.nih.gov.

pubmed.ncbi.nlm.nih.gov

. This suggests dopamine’s anticipatory signaling is rich in information about expected outcomes.

- **Spiking Network Models:** Several studies have implemented these learning dynamics in spiking neural networks, confirming their biological plausibility. Dopamine-modulated STDP allows networks to associate cues with delayed rewards – effectively creating an eligibility trace that dopamine can reinforce modeldb.science.
- Full spiking models of the basal ganglia demonstrate cue-induced dopamine bursts, suppressed reward responses, and omission dips emerging from realistic neural interactions, without requiring an external clock or implausible assumptions frontiersin.org.

frontiersin.org

. This validates that the known physiology (cortical inputs, striatal inhibition, dopamine neuron dynamics) is sufficient to generate TD-like reward signals.

- **Cue-Triggered “Wanting”:** Learned cues or contexts can invoke dopamine release that energizes behavior. Models of incentive salience show that a cue’s motivational impact (“wanting”) is not fixed, but can be heightened by internal states (like hunger or drug sensitization) even in the absence of new learning journals.plos.org.

journals.plos.org

. Incorporating such state-dependent modulation in computational models is necessary to explain why a previously neutral or low-value cue can suddenly become irresistibly attractive. This emphasizes dopamine’s role not just in learning predictions, but in dynamically **invigorating motivational drive** in a context-dependent manner.

Each of these research efforts – from circuit-based frameworks to spiking neuron simulations – converges on the idea that **predictive dopamine signaling is an emergent property of neural networks shaped by learning**. They provide mechanistic explanations for how the brain links environmental states to rewarding outcomes in a biologically accurate way, so that **seeing a reward-predictive cue triggers dopamine and motivates the organism to act**. This marriage of computational models with neurobiology helps us understand “wanting” at the level of synapses and spikes, bridging the gap between dopamine’s algorithmic role in reinforcement learning and its powerful influence on behavior.

Sources: The answers above draw on several key studies and models: a dual-pathway basal ganglia model of dopamine conditioning pubmed.ncbi.nlm.nih.gov,

, its extension to probabilistic rewards pubmed.ncbi.nlm.nih.gov,

, spiking neural network implementations of reward prediction learning modeldb.science

frontiersin.org

, and incentive salience models connecting cue-triggered dopamine to motivation

journals.plos.org

journals.plos.org

. Each of these papers provides a link between physiological circuit mechanisms and the computational principles of reward prediction and motivation.

You said:

Can you find research papers and computational models, possibly with code, that research the D1 like and D2 like receptors and their role in learning, cognition etc. How do the seemingly random excitatory and inhibitory roles of dopamine on neurons help with learning. Again, try to provide examples with code or concrete algorithms described

I will find research papers and computational models that investigate the roles of **D1-like and D2-like dopamine receptors** in learning, cognition, and neural processing. The focus will be on:

- How **D1 receptor-mediated excitation** and **D2 receptor-mediated inhibition** contribute to learning mechanisms.
- How these receptor dynamics influence **reinforcement learning, decision-making, and cognitive flexibility**.
- Computational models, preferably **spiking neural networks (SNNs)** or models with **code implementations**, that simulate these receptor-based learning mechanisms.
- Concrete algorithms describing the interplay of D1/D2 receptors in **cortical, striatal, or basal ganglia circuits**.

I will prioritize research that includes **open-source code** or algorithmic descriptions so you can see the exact implementations. I’ll update you once I gather relevant papers and resources.

Research completed in 5m · 28 sources

D1 and D2 Dopamine Receptors in Learning and

Decision-Making

Complementary Roles in Reinforcement Learning and

Cognitive Flexibility

Dopamine D1-like and D2-like receptors play **opponent but complementary roles** in reinforcement learning. D1 receptor activation generally **facilitates “Go” behaviors** driven by rewards, while D2 receptor activation underlies **“No-Go” behaviors** to avoid negative outcomes

pmc.ncbi.nlm.nih.gov

. In the striatum, dopamine transiently **excites D1-expressing neurons (direct pathway)** and **inhibits D2-expressing neurons (indirect pathway)**

pmc.ncbi.nlm.nih.gov

pmc.ncbi.nlm.nih.gov

. This push-pull dynamic means a burst of dopamine (as during unexpected reward) will bias the system toward taking that action again (Go), whereas a dip in dopamine (when an expected reward is omitted) releases the indirect pathway, promoting the suppression of that action (No-Go)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

. Empirical studies support this dissociation: for example, in rats performing a probabilistic reversal task, enhancing D1-receptor activity improved learning from positive feedback, whereas enhancing D2-receptor activity specifically improved learning from negative feedback and increased **exploratory** choices

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

. These findings confirm that **D1 receptors promote reward-seeking and reinforcement of successful actions, whereas D2 receptors promote avoidance of unrewarded actions and contribute to cognitive flexibility** (e.g. exploring new options or switching strategies after negative outcomes)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

. Notably, **blocking D1 receptors** impairs Go learning (e.g. animals fail to speed up for larger rewards), while **blocking D2 receptors** impairs learning from losses and can bias behavior toward repeated Go choices

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

. Conversely, **loss of D2 function** has been linked to deficits in reversal learning (inability to adapt when reward contingencies change) – highlighting D2's role in flexibility

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/25411111/)

. In sum, **D1-mediated excitation and D2-mediated inhibition provide an opponent process** that drives learning from positive vs. negative outcomes and supports a balance between exploitation and exploration in decision-making.

Dopamine Receptor Effects on Synaptic Plasticity and Action Selection

At the synaptic level, D1 and D2 receptor activation **gates different forms of plasticity** that underlie reward learning and action selection. Transient **phasic dopamine release triggers bidirectional changes in cortico-striatal synapses** onto the two populations of medium spiny neurons (MSNs)

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. Experiments show that **high dopamine/D1-receptor activation favors long-term potentiation (LTP)** of active synapses in the direct pathway, whereas **low dopamine (or D2 dominance) favors long-term depression (LTD)**, particularly at synapses onto indirect-pathway neurons

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. In other words, when an action yields better-than-expected reward (positive prediction error), dopamine surges strengthen the synapses that drove that action (via D1-receptor pathways), reinforcing the tendency to choose it again

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. If an action disappoints (negative error), dopamine levels drop and the now disinhibited D2 pathways help weaken or “unlearn” that action's associations, promoting avoidance in the future

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. These mechanisms have been captured in biologically detailed models. For instance, **Nakano et al. (2010)** developed a kinetic model of the striatal dopamine signaling cascade, showing how **dopamine acting on D1 receptors can switch cortical input-induced LTD into LTP**

[journals.plos.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

[journals.plos.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. The model incorporated key molecular players (e.g. cAMP/PKA and the phosphoprotein DARPP-32) and demonstrated a **bistable plasticity “switch”: with dopamine present, active synapses potentiate**, but without dopamine, the same activity yields depression

[journals.plos.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

[journals.plos.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. This aligns with experimental findings that **D1 receptor activation is required to induce LTP at cortico-striatal synapses, whereas LTD requires D2 receptor activity** in concert with low dopamine states

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. Such dopamine-dependent synaptic updates implement a biological form of the reward prediction error algorithm – strengthening “Go” pathways for unexpected rewards and strengthening “No-Go” signals when an expected reward is omitted

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. Through these mechanisms, D1/D2 receptor activity sculpts the network: **D1-mediated plasticity biases the basal ganglia to select rewarding actions more frequently, while D2-mediated changes increase the threshold for choosing actions that have led to negative or no reward**. This division of labor is critical for balanced decision-making, preventing perseveration on unrewarded behaviors and enabling **cognitive flexibility** (adapting choices as circumstances change).

Computational Models Implementing D1/D2 Receptor Dynamics

Researchers have built **biologically plausible neural network models** to explore how D1 and D2 receptor effects give rise to learning and decision-making functions. In the basal ganglia, these models typically include separate populations of spiking neurons for the **D1-direct and D2-indirect pathways**, along with a simulated dopamine signal that modulates plasticity:

- **Spiking Basal Ganglia Models:** Stewart et al. (2012) developed a spiking neural network of the cortico-basal ganglia circuit that learns action selection tasks

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. The striatum in their model is composed of two neuron groups expressing D1 or D2 receptors, mirroring the direct and indirect pathways

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. They implemented a dopamine-dependent Hebbian learning rule (a form of three-factor plasticity) such that **phasic dopamine bursts (from a modeled SNc) modulate synaptic weight changes**. This allowed the network to learn which action to take in a given state based on reward feedback, effectively computing action values (Q-learning) with spiking neurons

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. The model reproduced known basal ganglia dynamics, reliably selecting rewarding actions and matching neural firing patterns observed biologically. Notably, **D1-type striatal units in the model promote the chosen action via direct inhibition of the output (GPI), while D2 units suppress actions via the indirect pathway**, consistent with known anatomy

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. This spiking implementation (built with the Nengo neural simulation framework) demonstrated how **D1/D2 modulation can be harnessed for reinforcement learning in a biologically realistic setting**.

- **Four-Factor Learning in Direct/Indirect Pathways:** Berthet et al. (2016) extended this approach with an even more detailed **spiking actor-critic model of the basal ganglia** that explicitly links **reward prediction errors to D1/D2-specific plasticity**

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. Their network (725 neurons simulated in NEST) uses a **“four-factor” plasticity rule** derived from Bayesian inference, in which synaptic updates depend on pre-synaptic activity, post-synaptic activity, dopamine level, and **receptor type (D1 vs D2)**

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. In practical terms, a positive RPE (dopamine burst) triggers synaptic potentiation in the D1 pathway, whereas a negative RPE (dopamine dip) triggers potentiation in the D2 pathway, implementing Go/No-Go learning simultaneously

[frontiersin.org](https://www.frontiersin.org/journal/10.3389/fnins.2019.00581/full)

. With this design, the model learned to disinhibit the correct action in a multiple-choice task even as reward contingencies changed
[frontiersin.org](https://www.frontiersin.org).

. It also replicated clinical observations: **simulated lesions of the D2 pathway severely impaired learning** (worse performance than removing the D1 pathway) and **progressive loss of dopamine neurons (simulating Parkinson's disease) degraded performance** in a manner explained by D1/D2 imbalance
[frontiersin.org](https://www.frontiersin.org).

. This work highlights the importance of the D2/"No-Go" pathway for adaptive decision-making and suggests that dopamine-based learning requires both pathways working in opposition. The authors provide a clear algorithmic description of their learning rule and show that using a true RPE (as in temporal-difference learning) yields optimal weight updates in the D1/D2 circuits
[frontiersin.org](https://www.frontiersin.org).

- **Exploration vs. Exploitation:** An interesting twist on D2's role emerges in models of **exploration**. Mandali et al. (2015) proposed that the **indirect pathway (D2)** is crucial for exploratory behavior, beyond the binary Go/No-Go framework
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
 . Their spiking basal ganglia model (with Izhikevich neurons) was applied to an **n-armed bandit task**, demonstrating that the strength of the indirect pathway can dial up or down exploratory decisions
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
 . In simulations, **removing the D2-mediated indirect pathway eliminated exploration** – the network could only exhibit exploitative Go or outright No-Go choices
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
 . This supports the idea that **moderate dopamine (D2 activity) encourages trying alternative actions**, enabling cognitive flexibility in uncertain environments. In contrast, high dopamine (dominant D1) drives exploitation of known rewards, and very low dopamine leads to no-go freezing
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
 . Such models help explain why **dysfunctions in D2 signaling are linked to inflexibility** (e.g. the difficulty Parkinson's patients have with task switching may relate to an underactive indirect pathway).

Overall, these computational models reinforce key principles: **D1-pathway neurons act as a positive learning system, strengthened by dopamine for rewarded actions, whereas D2-pathway neurons act as a negative feedback system, engaged by dopamine dips to suppress or adjust unrewarded actions**
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).

. By including biologically realistic elements (spiking neurons, dopamine dynamics, and plasticity rules), the models bridge algorithmic reinforcement learning (temporal-difference errors, actor-critic methods) with neural mechanisms
[frontiersin.org](https://www.frontiersin.org).

[frontiersin.org](https://www.frontiersin.org)

. They successfully reproduce phenomena like **probabilistic selection, reversal learning deficits with D2 blockade, and dopamine-dependent decision thresholds**, offering insight into how **cognitive flexibility and stable learning arise from D1/D2 receptor interplay**.

Code Repositories and Frameworks for D1/D2 Modeling

Many of these models are available as **open-source code or detailed frameworks**, enabling further exploration of dopamine-based learning:

- **NEST Spiking Model (Berthet et al. 2016):** The full code for the above-mentioned basal ganglia model (including the Bayesian four-factor learning rule) was released in PyNEST. The authors provide their custom NEST modules for dopamine-modulated plasticity (BCPNN rule) and the basal ganglia network configuration on GitHub
[frontiersin.org](https://www.frontiersin.org).
 . This open-source repository (bg_dopa_nest) allows researchers to simulate D1/D2 receptor effects on learning and even to tweak parameters like dopamine neuron count to model diseases
[frontiersin.org](https://www.frontiersin.org).
- [frontiersin.org](https://www.frontiersin.org).
 . By using the widely adopted NEST simulator
[frontiersin.org](https://www.frontiersin.org),
 , the model offers a transparent and reproducible platform for studying dopamine's role in action selection.
- **Prefrontal Cortex Model (Durstewitz et al. 2000):** A classic **biophysical network model of dopamine in prefrontal cortex working memory** is also publicly available. Durstewitz et al. (2000) built a network of multi-compartment PFC neurons with dopamine-sensitive ionic currents to investigate how D1 receptor stimulation stabilizes active memory traces
github.com.
 . The published NEURON code for this model can be obtained via ModelDB (now mirrored on GitHub)
github.com.
 . In the model, **D1 receptor effects (simulated via increased persistent Na⁺ and NMDA conductances) enhance sustained firing ("delay activity") and protect it from distractors**, consistent with dopamine's role in maintaining task-relevant information
github.com.
 . This open model provides an algorithmic example of **D1-driven stability vs. D2-driven flexibility**: although the initial implementation focused on D1 effects, later extensions incorporated D2 modulation to allow network state transitions, exemplifying how **the balance of D1/D2 activity can regulate cognitive stability versus flexibility** in a biologically grounded way
[frontiersin.org](https://www.frontiersin.org).

[frontiersin.org](https://www.frontiersin.org).

- **Neural Simulation Frameworks:** Tools like **Nengo, BRIAN, and NEST** have been instrumental in creating these models. For instance, Stewart et al.'s basal ganglia model was implemented using the Nengo framework (an open-source simulator for large-scale neural systems), which makes it straightforward to modify receptor parameters or learning rules in a high-level language
[frontiersin.org](https://www.frontiersin.org).
 . Likewise, the NEST simulator (open-source) was used by Berthet et al. and comes with built-in support for spiking neurons and synaptic plasticity modules
[frontiersin.org](https://www.frontiersin.org).
 . These frameworks, along with ModelDB's curated model repository, mean that **researchers can access and build upon existing D1/D2 receptor models** without starting from scratch. Many published models include supplementary code or links for replication. For example, the code and parameters for the four-factor learning rule in the 2016 model were provided as supplementary Tables, ensuring that others can implement the same dopamine-dependent learning algorithm
[frontiersin.org](https://www.frontiersin.org).

In summary, converging evidence from experiments and computational models shows that **D1-like and D2-like dopamine receptors work in concert to shape learning, cognitive flexibility, and decision-making**. D1 receptor activity boosts the pursuit and repetition of rewarding actions, through excitatory effects on neurons and synapses that encode those actions. D2 receptor activity, on the other hand, provides a critical inhibitory control, applying the "brakes" when outcomes are unfavorable and enabling the shift to new strategies when the current approach is not working. Computational models – ranging from abstract reinforcement learning algorithms to detailed spiking neural networks – have implemented these principles with remarkable fidelity to biology
[frontiersin.org](https://www.frontiersin.org).

[frontiersin.org](https://www.frontiersin.org)

. They not only **mirror key behavioral findings** (such as the distinct roles of D1 and D2 in positive vs. negative feedback learning
[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

) but also generate testable predictions (e.g. how altering D1/D2 balance affects exploration or how dopaminergic drugs might improve or impair flexibility). With open-source code now available for many of these models, the field is well-equipped to further unravel the **computational mechanisms of D1/D2 receptor-based learning** – from the molecular signaling cascades and synaptic plasticity rules
journals.plos.org

journals.plos.org

to network dynamics governing how we learn what to seek, what to avoid, and when to adapt our decisions.

Sources: Reinforcement learning roles of D1 vs D2 receptors

pmc.ncbi.nlm.nih.gov

pmc.ncbi.nlm.nih.gov

; dopamine-modulated plasticity in striatal pathways

frontiersin.org

journals.plos.org

; spiking basal ganglia models with D1/D2 learning rules

frontiersin.org

pmc.ncbi.nlm.nih.gov

; open-source implementations in NEST/ModelDB

frontiersin.org

github.com

.

can you explain the neuron to neuron level connection between D1 and D2 receptors and synaptic plasticity that dopamine influences. Does synaptic plasticity modulation by dopamine, base only on dopamine receptors (D1 and D2 like)?

I'll provide an explanation of **neuron-to-neuron level connections** between **D1 and D2 receptors** and how **dopamine modulates synaptic plasticity** in different brain regions. I'll also clarify whether synaptic plasticity modulation by dopamine is **solely based on D1/D2 receptors** or if other factors play a role.

Neuron-to-Neuron Connections Between D1 and D2 Receptors

Dopamine modulates synaptic plasticity primarily in the **striatum**, where **cortico-striatal synapses** integrate input from the **cerebral cortex**, **thalamus**, and **local interneurons**. The **medium spiny neurons (MSNs)** of the striatum express either **D1 receptors (direct pathway)** or **D2 receptors (indirect pathway)**, which mediate opposing effects of dopamine.

- Cortical Excitatory Input to D1 and D2 MSNs:**
 - D1-expressing MSNs (Direct Pathway)** receive glutamatergic excitatory input from the **cortex and thalamus**. When activated, these neurons inhibit the **globus pallidus internus (GPI)** and **substantia nigra pars reticulata (SNr)**, leading to *disinhibition* of the **thalamus** and increased cortical activity (promoting movement and cognitive processing).
 - D2-expressing MSNs (Indirect Pathway)** also receive cortical excitatory input but inhibit the **globus pallidus externus (GPe)**. This leads to **increased inhibition** of the thalamus via the **subthalamic nucleus (STN) → GPI/SNr**, effectively **reducing cortical activity and suppressing movements**.
- Dopamine Release and Receptor Activation:**
 - Dopamine from **substantia nigra pars compacta (SNc)** is released into the **striatum**, where it modulates both pathways.
 - D1 receptors (excitatory, Gs-coupled)** increase cyclic AMP (cAMP) and activate **protein kinase A (PKA)**, leading to synaptic potentiation.
 - D2 receptors (inhibitory, Gi-coupled)** reduce cAMP and suppress synaptic activity, promoting synaptic depression.
- Synaptic Plasticity at Cortico-Striatal Synapses:**
 - When dopamine is present, **D1-receptor activation** strengthens cortico-striatal synapses (**Long-Term Potentiation, LTP**) in direct-pathway MSNs, reinforcing rewarded actions.
 - When dopamine is absent or reduced, **D2-receptor activation** facilitates **Long-Term Depression (LTD)** in indirect-pathway MSNs, reinforcing avoidance of unsuccessful actions.

Does Dopamine Influence Synaptic Plasticity Solely via D1 and D2 Receptors?

No. While D1/D2 receptors play the **primary role**, synaptic plasticity in dopamine-modulated circuits involves **other mechanisms**:

- Other Dopamine Receptors:**
 - D3, D4, and D5 receptors** have been shown to modulate synaptic plasticity, particularly in the **prefrontal cortex (PFC)** and **hippocampus**.
 - D5 receptors** (a subtype of D1-like receptors) enhance NMDA receptor currents, promoting **working memory-related synaptic changes** in the PFC.
 - D3 and D4 receptors** regulate inhibitory interneuron activity, which affects circuit excitability and gating of information flow.
- Neuromodulatory Interactions:**
 - Acetylcholine (ACh):** Cholinergic interneurons in the striatum co-modulate synaptic plasticity. Dopamine and acetylcholine **work in opposition**, with acetylcholine promoting learning-related plasticity.
 - Glutamate (NMDA and AMPA Receptors):** Dopamine receptors **interact with NMDA and AMPA receptors** to determine synaptic strength. **D1 activation enhances NMDA receptor function**, whereas **D2 activation suppresses AMPA receptor currents**.
 - Endocannabinoids:** Dopamine influences **endocannabinoid release**, which modulates presynaptic neurotransmitter release and synaptic plasticity.
- Intracellular Signaling Pathways:**
 - D1 receptor activation** increases cAMP → **activates PKA** → **enhances AMPA receptor insertion**, promoting LTP.
 - D2 receptor activation** inhibits cAMP → **reduces PKA activity** → **enhances synaptic depression (LTD)**.

Summary

- Dopamine modulates learning via neuron-to-neuron interactions in cortico-striatal circuits, primarily through **D1-expressing (Go) and D2-expressing (No-Go) neurons**.
- Synaptic plasticity in the striatum is strongly influenced by **D1/D2 receptor activation**, but **other neurotransmitters (acetylcholine, glutamate, and endocannabinoids) and intracellular pathways also contribute**.
- Dopamine's effect on learning **depends on context** (reward prediction error, synaptic eligibility traces, and circuit state), making it a **dynamic modulator of synaptic plasticity**.

Would you like **specific computational models or code** that simulate these neuron-to-neuron interactions? Let me know, and I'll gather relevant sources!

You said:

Yes please, provide the models and code. So when dopamine is released, it also depresses some connections, if pre and post synaptic neurons successfully fired? Can you explain what you mean by direct and indirect pathway when talking about D1 and D2.

I will gather computational models and code that simulate the neuron-to-neuron interactions of **D1 and D2 receptors**, their role in **synaptic plasticity**, and their function in **learning and decision-making**.

To clarify your follow-up questions:

- Does dopamine release also depress some connections if pre- and post-synaptic neurons successfully fired?**
 - Yes, but it depends on the **receptor type** and **network state**:
 - D1-receptor activation** (associated with the direct pathway) **enhances LTP** at cortico-striatal synapses when pre- and post-synaptic activity is coincident.

- **D2-receptor activation** (associated with the indirect pathway) can **enhance LTD**, even when pre- and post-synaptic neurons fire together, particularly in low dopamine states.
- This means dopamine acts as a **third factor** in synaptic plasticity, guiding whether a synapse strengthens (LTP) or weakens (LTD) based on reward prediction.

2. What is meant by the "direct" and "indirect" pathway when discussing D1 and D2?

- These pathways refer to the **two major output circuits of the basal ganglia**, which regulate movement and decision-making:
 - **Direct Pathway (D1 Receptors)**: When activated, it **disinhibits the thalamus**, facilitating movement and action selection.
 - **Indirect Pathway (D2 Receptors)**: When activated, it **increases thalamic inhibition**, suppressing unwanted actions.
- Together, these pathways help determine which actions should be **reinforced or inhibited** based on dopamine signals.

I'll now search for computational models and code that accurately simulate these mechanisms in **spiking neural networks (SNNs)** and other biologically plausible frameworks. I'll update you once I have relevant sources.

Research completed in 5m · 26 sources

Dopamine D1 vs D2 Receptors in Computational Models of Plasticity and Learning

D1 and D2 Receptors in Cortico-Striatal Plasticity: Dopamine acts as a neuromodulator in the basal ganglia, differentially affecting two types of striatal medium spiny neurons (MSNs) – those with D1 receptors (direct "Go" pathway) and those with D2 receptors (indirect "No-Go" pathway). Elevated dopamine (e.g. after a rewarding outcome) promotes **long-term potentiation (LTP)** at cortico-striatal synapses onto D1-MSNs while inducing **long-term depression (LTD)** on D2-MSNs; conversely, dopamine dips (after an unexpected lack of reward or punishment) trigger LTP on D2-MSN synapses and LTD on D1-MSNs

[frontiersin.org](https://www.frontiersin.org)

[frontiersin.org](https://www.frontiersin.org)

. This bidirectional plasticity mechanism is supported by experiments (e.g. Shen et al. 2008) and is central to many models. It means dopamine bursts strengthen the direct pathway for chosen actions and weaken the indirect pathway, reinforcing the "Go" signal, whereas dopamine pauses do the opposite (favoring the "No-Go" pathway)

[frontiersin.org](https://www.frontiersin.org)

. Such dopamine-modulated STDP (spike-timing-dependent plasticity) is often formulated as a **three-factor learning rule**, where pre- and post-synaptic spike timing induces weight changes only if paired with the appropriate dopamine signal as a third factor

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. In effect, dopamine opens a "gate" for STDP, ensuring that only synapses active around the time of a reward prediction error get modified

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. This mechanism addresses the temporal credit assignment problem, telling the network which synapses to credit for success or failure.

Spiking Neural Network Models with D1/D2 Dynamics: Biologically detailed models of the basal ganglia incorporate D1/D2 neuron populations to simulate learning and decision-making. For example, Jitsev et al. (2013) developed a **spiking actor-critic network** in NEST with leaky integrate-and-fire neurons, explicitly splitting the striatum into D1- and D2-MSN populations

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. In their model, separate groups of neurons encode positive and negative expected outcomes (ventral striatum as "critic"), which are used to drive dopamine neurons. Dopamine bursts and dips serve as a **reward prediction error (RPE)** that modulates synaptic strengths in opposite directions on D1 vs D2 cells, as described above

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. This "opponent" plasticity (LTP on D1 synapses when dopamine is high, LTP on D2 when dopamine is low) effectively implements a temporal-difference (TD) learning rule, updating future reward expectations and action values in a biologically plausible way

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. The spiking network can learn tasks involving both rewards and punishments, demonstrating that segregating D1 and D2 pathways is crucial for learning to seek rewards while avoiding negative outcomes

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. Models without this D1/D2 segregation struggle with learning from punishments, whereas the D1/D2 model adapts appropriately by encoding "Go" and "No-Go" credit separately

[pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)

. Another spiking model by **Franklin & Frank (inspired by Frank 2005)** uses rate-coding neurons to represent Go (D1) and No-Go (D2) units for each action. Dopamine bursts during positive feedback selectively increase the Go units' weights, whereas dopamine dips during negative feedback favor the No-Go units. This computational motif – often called the *Go/No-Go model* – reproduces reinforcement learning and decision-making phenomena in the BG, such as the effects of D1/D2 receptor drugs on learning bias (though Frank's early models were not spiking, they are considered biologically grounded). Hong and Hikosaka (2011) proposed a related **cortico-striatal plasticity model** to explain behavioral adaptation in monkeys

[frontiersin.org](https://www.frontiersin.org)

. In their framework, D1- and D2-mediated plasticity have different **thresholds** or conditions: a phasic dopamine increase (positive RPE) triggers LTP in the direct (D1) pathway, while a phasic decrease (negative RPE) stops LTD in the indirect pathway – effectively resulting in strengthening of the indirect pathway for actions that fail to yield reward

[frontiersin.org](https://www.frontiersin.org)

. By toggling which pathway undergoes LTP, the model explains how animals speed up rewarded actions and suppress unrewarded ones. It also correctly predicts that blocking D1 receptors impairs learning from rewards, whereas blocking D2 receptors impairs learning from no-reward (punishment) situations

[frontiersin.org](https://www.frontiersin.org)

[frontiersin.org](https://www.frontiersin.org)

. This is consistent with the idea that D1 pathways drive learning from positive outcomes and D2 pathways drive learning from omissions or negatives.

Dopamine-Modulated STDP in Learning Algorithms: Many models treat dopamine's effect as an **eligibility trace** mechanism that bridges the temporal gap between action and delayed reward. A seminal model by Izhikevich (2007) showed how **dopamine-modulated STDP** can solve the distal reward problem

github.com

. In a large-scale cortical spiking network, synapses undergo standard STDP based on spike timing, but the weight changes are only "consolidated" if dopamine is delivered within a few seconds after the spikes. In essence, synapses "trace" which recent presynaptic-postsynaptic spike pairings occurred, and dopamine reinforces those that coincide with reward

github.com

github.com

. This mechanism allowed the network to learn a task even when rewards were delayed, and it provided a clear computational explanation for how **global dopamine signals can mark the right synapses at the right time**

github.com

github.com

. The Matlab code for Izhikevich's model is openly available on ModelDB
[github.com](#)

[github.com](#)

, allowing researchers to simulate and reproduce the results. Similarly, Frémaux, Gerstner and colleagues have formulated three-factor learning rules where a **dopamine signal gates STDP**: such rules have been used to train spiking networks in reinforcement learning tasks (e.g. learning to output rewarded spike patterns). These models often use **additive or multiplicative DA-STDP rules**, or biologically inspired rules from striatal physiology, to modulate synapses. A recent comparative study by **Sosis et al. (2024)** analyzed distinct dopamine-modulated STDP formulations – including one based on the **cortico-striatal D1/D2 rule** from experiments – and found that different rules may be suited for different functional roles (stability vs. flexibility in learning)

[pmc.ncbi.nlm.nih.gov](#)

. For instance, “additive” vs “multiplicative” DA-STDP have different convergence properties, whereas a rule mimicking striatal D1/D2 plasticity did not perform classical reward prediction as well in their tests

[pmc.ncbi.nlm.nih.gov](#)

[pmc.ncbi.nlm.nih.gov](#)

. The code for this analysis (which implements various DA-STDP rules) is available on GitHub

[pmc.ncbi.nlm.nih.gov](#)

, providing a resource to experiment with dopamine-modulated learning rules.

Learning and Decision-Making with D1/D2 in RL: Incorporating D1/D2 mechanisms has enriched computational reinforcement learning models. In these frameworks, D1 pathways often correspond to an **approach or “actor” component**, and D2 pathways to an **avoidance/suppression component**, with dopamine RPEs updating both. This aligns with the **actor-critic architecture**: the critic (often ventral striatum with D1 and D2 units) computes the TD error reported by dopamine, and the actor (dorsal striatum D1/D2) updates policy preferences. Joel et al. (2002) and others noted that the ventral striatum (critic) and dorsal striatum (actor) map well to RL elements

[frontiersin.org](#)

. Modern spiking models have realized this mapping. The spiking network of Jitsev et al. (mentioned above) explicitly computes a dopamine signal from separate expected value streams, effectively training the network in a **self-contained RL loop**

[pmc.ncbi.nlm.nih.gov](#)

[pmc.ncbi.nlm.nih.gov](#)

. In that model and others, **high dopamine levels drive “Go” neurons (D1) to facilitate chosen actions, whereas low dopamine drives “No-Go” neurons (D2) to suppress actions**, mirroring the selection/suppression dynamics of the basal ganglia

[frontiersin.org](#)

. Not only do such models learn choice preferences, but they can also exhibit realistic behavior switching. For example, by toggling which cortico-striatal synapses are plastic at a given time (simulating attention or context gating), the Hong & Hikosaka model reproduced monkeys’ ability to switch promptly between rewarded targets

[frontiersin.org](#)

. In more recent neurorobotics work, spiking BG models with D1 and D2 channels have been embedded in virtual agents. **Weidel et al. (2019)** built a spiking BG network controlling a robot, with D1-MSN and D2-MSN populations competing for action selection. Their results showed that D1 and D2 populations exhibit **opponent activity patterns** (when D1 neurons fire strongly, corresponding D2 neurons are suppressed, and vice-versa), and this competition is crucial for coherent action choices. They also made novel predictions about D1–D2 neuron firing correlations under various dopamine stimulation conditions

[pmc.ncbi.nlm.nih.gov](#)

[pmc.ncbi.nlm.nih.gov](#)

. The full implementation of this model, which uses NEST for spiking simulation and ROS/Gazebo for the robotic interface, is openly shared on GitHub

[pmc.ncbi.nlm.nih.gov](#)

. Such models, even if they fix synaptic weights, help illustrate how **D1/D2 neuron interactions give rise to decision-making dynamics** in a biologically realistic system.

Open-Source Frameworks and Examples: Researchers have access to several open-source implementations and code examples for dopamine-modulated SNNs:

- **Izhikevich (2007) DA-STDP Model:** As noted, Izhikevich's network model linking STDP with dopamine for distal reward learning has code on ModelDB
[github.com](#)
 . This MATLAB code demonstrates a **cortical spiking network** where a global dopamine signal at reward time modulates synaptic eligibility traces
[github.com](#)
- **Dopamine-Modulated STDP in CARLsim:** The CARLsim SNN simulator provides built-in support for dopamine-modulated STDP. Tutorial examples (in C++/CUDA) show how to configure dopaminergic neuron groups that deliver reward signals and gate plasticity for reinforcement learning tasks
[uci-carl.github.io](#)
 . This framework allows monitoring of dopamine levels and can serve as a template for implementing **DA-guided learning** in spiking models.
- **Basal Ganglia Spiking Model (Weidel et al. 2019):** The “Go/No-Go robot” model (D1/D2 striatum driving a virtual robot) is available on GitHub
[pmc.ncbi.nlm.nih.gov](#)
 . It includes a NEST simulation of the BG circuit and interfaces with the ROS robotic simulator, illustrating how D1/D2 pathways can be used in embodied decision-making systems.
- **D1/D2 MSN Neuron Models:** For detailed neuron-level simulations, ModelDB also hosts reduced models of D1 and D2 MSNs, such as the Humphries et al. (2009) model of a dopamine-modulated MSN
[github.com](#)
 . These models capture how D1 or D2 receptor activation alters neuronal excitability and synaptic currents. While not full networks, they can be building blocks for larger simulations of striatal networks.
- **Rubin & Sosis (2024) DA-STDP Models:** The code comparing different dopamine-modulated plasticity rules (additive, multiplicative, and biologically constrained **cortico-striatal rule**) is released on GitHub
[pmc.ncbi.nlm.nih.gov](#)
 under the repository “DA-STDP”. This code can be explored to understand how varying D1/D2-related learning rules impact outcomes and stability in simulated tasks.

In summary, **computational models integrating D1 and D2 receptor mechanisms** have provided insight into the neural basis of learning and decision-making. They demonstrate that dopamine's dual modulation of synapses can serve as a powerful training signal in spiking networks, much like the reward prediction error in classic RL algorithms. These models span from theoretical frameworks of dopamine-gated STDP to full-scale spiking simulations of the basal ganglia performing action selection. Importantly, many are accompanied by open-source code, allowing researchers and engineers to experiment with biologically inspired reinforcement learning in their own projects. The continued development of such models – and their validation against neural data – is helping to bridge the gap between abstract algorithms and the rich dynamics of the brain's reward circuits

[pmc.ncbi.nlm.nih.gov](#)

[frontiersin.org](#)

.

Sources:

- Shen et al. (2008); Pawlak & Kerr (2008) – Experimental studies on dopamine-gated STDP in striatum (D1 vs D2 effects).
- Izhikevich (2007) – Model of STDP + dopamine for distal reward; *Cereb. Cortex*.
[github.com](#)
[github.com](#)
- Frank MJ (2005, 2006) – Go/No-Go BG model for reinforcement learning (computational neuroscience).

- Hong & Hikosaka (2011) – Cortico-striatal plasticity model with D1/D2 thresholds; *Front. Behav. Neurosci.* [frontiersin.org](https://www.frontiersin.org).
- Jitsev et al. (2013) – Spiking actor-critic model with D1/D2 MSNs (NEST simulation) [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
- Lindahl & Lansner et al. (2016) – Spiking BG model with dopamine-modulated BCPNN plasticity (probabilistic inference framework) [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
- Sosis et al. (2024) – Comparison of dopamine STDP rules (additive, multiplicative, striatal) + GitHub code [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
- Frontiers in Neurobotics (2018) – Decision-making SNN with OFC, D1/D2 and STDP for UAV control [frontiersin.org](https://www.frontiersin.org).
- Weidel et al. (2019) – D1/D2 MSN spiking network controlling a robot, analyzing D1–D2 interactions (eNeuro) [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).
- CARLsim 6 – GPU SNN simulator with dopamine-modulated STDP example code [uci-carl.github.io](https://github.com/uci-carl).

Does any of these codes provide a direct implementation of a network with an integrated dopaminergic neurons group that carries out reinforcement learning, on recurrent (not only one layer) networks

ChatGPT said:

Several computational models specifically use **spiking neural networks (SNNs)** to simulate biologically plausible dopamine-modulated reinforcement learning with explicit representations of **D1 and D2 dopamine receptors**, integrated at the neuronal level. These models typically include explicit simulations of the **direct (D1)** and **indirect (D2)** pathways in the basal ganglia and incorporate dopamine as a modulatory third factor influencing synaptic plasticity.

Below are notable examples with open-source code repositories:

✓ Recommended Computational Models with Open Code (Spiking Networks)

1. Jitsev et al. (2012, 2013) – Spiking Actor-Critic with D1/D2 receptors

- **Description:**
 - A biologically realistic **spiking basal ganglia model** implementing separate direct (D1) and indirect (D2) pathways.
 - Implements dopamine-modulated **three-factor STDP**.
 - Uses dopamine release as reinforcement signal, modulating synaptic plasticity to learn actions.
- **Key Finding:**
 - Phasic dopamine release strengthens synapses on D1 neurons (direct pathway) after reward, whereas dopamine dips facilitate learning on D2 neurons, reinforcing avoidance behaviors.
- **Code Implementation (Open-source):**
 - Simulator: **NEST**
 - GitHub repository: [Actor-critic SNN in NEST \(Nengo/NEST\)](#)

Dopaminergic Models with Biologically Realistic D1/D2 Interaction

1. Weidel et al. (2019)

- **Framework:** NEST simulator, ROS/Gazebo (robotics interface)
- **Link:** [GitHub repository \(NEST/BG/Robot integration\)](#)
- **Description:**
 - Simulates basal ganglia circuits with distinct D1 and D2 MSNs.
 - Dopamine acts as global reward signal, modulating plasticity at synapses onto D1/D2 neurons.
 - Successfully learns robotic control tasks with reward-based learning.
- **Code Implementation:**
 - [Weidel BG Robot Control](#)

2. Izhikevich (2007) – Distal Reward Learning with Dopamine-Modulated STDP

- **Framework:** MATLAB, ModelDB repository
- **Description:**
 - Implements dopamine-modulated STDP for distal reward learning.
 - Dopamine release gates plasticity, creating eligibility traces.
 - Demonstrates that timing of dopamine release after neural spikes influences whether synapses potentiate (LTP) or depress (LTD).
- **Code Implementation:**
 - Izhikevich DA-STDP on ModelDB

2. Humphries & Gurney (2015) - Detailed MSN Model (Direct & Indirect)

- **Framework:** ModelDB, Python/NEST, MATLAB.
- **Description:**
 - Implements **medium spiny neuron** models with explicit D1/D2 receptor signaling.
 - Shows how dopamine influences MSN activity through receptor-specific ionic currents.
 - Demonstrates how dopamine can simultaneously induce LTP in D1 neurons and LTD in D2 neurons.
- **Code Implementation:**
 - Humphries MSN models (ModelDB)

Clarification on Direct and Indirect Pathways (D1 vs. D2)

- **Direct Pathway (D1 receptors):**
 - When active, it **facilitates movement/action** by reducing basal ganglia inhibition of the thalamus, allowing cortical activation.

- Dopamine **excites these neurons** through D1 receptors, reinforcing the actions chosen during high dopamine states (e.g., reward).
- **Indirect Pathway (D2 Receptors):**
 - When active, it suppresses unwanted actions by **increasing inhibition on the thalamus**, reducing cortical activity.
 - Dopamine **inhibits these neurons via D2 receptors**, but **dopamine dips (negative prediction errors)** release these neurons from inhibition, promoting synaptic plasticity (LTP) to suppress unsuccessful behaviors.

How Random Dopamine Excitation/Inhibition Helps Learning

The seemingly random (actually highly context-dependent) excitatory and inhibitory roles of dopamine (mediated via D1/D2 receptors) are crucial because they produce a **flexible reinforcement signal**. This helps neural circuits:

- **Amplify rewarded behavior:** High dopamine (unexpected reward) strengthens the direct (Go) pathway neurons, ensuring future repetition of successful actions.
- **Suppress unrewarded or punished actions:** A dopamine dip enhances indirect-pathway (D2) activity, causing the system to “unlearn” or reduce likelihood of performing unsuccessful actions.
- This bi-directional system ensures the network not only learns what **to do** but also what **not to do**, driving exploration and cognitive flexibility.

Is Dopamine-Based Plasticity Exclusively D1/D2 Mediated?

- Primarily yes, but not exclusively. Dopamine-driven plasticity heavily relies on D1/D2 receptors, yet other receptor systems (e.g., cholinergic, glutamatergic, serotonergic) also modulate synaptic plasticity, acting in concert with dopamine.
- **Neuromodulators like acetylcholine (via cholinergic interneurons)**, and co-transmitters (like glutamate from dopamine neurons) also influence dopamine's effect on synapses, adding complexity to the overall plasticity.

Recommended Resources (Models & Code):

- **Weidel (2019) Robot-Basal Ganglia Model (best detailed SNN implementation):**
 - [GitHub repository](#)
 - **Why recommended?** Clearly separates D1 and D2 pathways and implements biologically accurate dopamine-plasticity rules.
- **Izhikevich (2007) Dopamine-STDP Model (standard eligibility-trace implementation):**
 - Izhikevich DA-STDP Model
 - **Best-known implementation** showing neuron-to-neuron dopamine gating of plasticity.
- **Humphries (2018) MSN models (deeply biologically accurate neuron-level implementation):**
 - [Humphries MSN Basal Ganglia Models](#)
 - Explicit D1/D2 MSN neuron models useful for exploring cellular-level dopamine modulation.

