

# Reinforcement Learning in Brain Systems: A Computational Bridge Between Neuroscience and Behavior

Mark T. Caldwell\*

Department of Computational Neuroscience Institute of Neural Science and Artificial Intelligence  
United States of America

## Commentary

**Received:** 01-Dec-2025, Manuscript No. RRNS-25-189145; **Editor assigned:** 03-Dec-2025, Pre-QC No. RRNS-25-189145 (PQ); **Reviewed:** 17-Dec-2025, QC No. RRNS-25-189145; **Revised:** 22-Dec-2025, Manuscript No. RRNS-25-189145 (R); **Published:** 29-Dec-2025, DOI: 10.4172/rrns.9.019

### \*For Correspondence

Mark T. Caldwell, Department of Computational Neuroscience Institute of Neural Science and Artificial Intelligence United States of America

**E-mail:** mark.caldwell@insa-neuro.edu

**Citation:** Mark T. Caldwell, Reinforcement Learning in Brain Systems: A Computational Bridge Between Neuroscience and Behavior. RRJ Hosp Clin Pharm. 2025.9.019.

**Copyright:** © 2025 Mark T. Caldwell, this is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Reinforcement learning (RL) has emerged as a unifying computational framework for understanding learning and decision-making in biological systems. In neuroscience, RL provides a formal description of how organisms adapt behavior through interactions with their environment by maximizing cumulative reward. This commentary explores how RL principles are implemented in brain systems, particularly within dopaminergic pathways, cortical-basal ganglia loops, and limbic structures. The article discusses reward prediction error signaling, neural representations of value, and the role of model-free and model-based learning systems in behavior. Furthermore, it highlights recent advances linking RL with neurophysiology, including evidence from electrophysiology, neuroimaging, and optogenetic studies. The integration of RL theory with brain function has significantly advanced understanding of psychiatric disorders, neurodegenerative diseases, and adaptive cognition. This article also discusses limitations of current RL-based neural models and suggests future directions involving hierarchical RL, meta-learning, and biologically plausible deep reinforcement learning systems.

## Keywords

Reinforcement learning, dopamine, reward prediction error, basal ganglia, decision-making, neurocomputational models, model-free learning, model-based learning, cognitive neuroscience, synaptic plasticity

## INTRODUCTION

Understanding how the brain learns from experience remains one of the central challenges in neuroscience. Organisms continuously interact with complex environments, making decisions that maximize rewards and minimize punishment. Reinforcement learning (RL), originally developed in machine

learning and behavioral psychology, provides a mathematical framework for describing this adaptive process.

In neuroscience, RL has become a dominant paradigm for explaining how neural circuits encode reward, evaluate actions, and update future behavior. Classical behavioral experiments in animals showed that learning depends not only on immediate outcomes but also on expectations of future reward. These findings laid the foundation for computational models of reward-based learning that now align closely with observed neural activity patterns.

A key breakthrough in neurobiology was the discovery that dopamine neurons encode a signal consistent with reward prediction error (RPE), a core component of RL theory. This discovery established a direct link between computational learning models and biological substrates.

### Theoretical Foundations of Reinforcement Learning

RL is based on the idea that an agent learns to take actions in an environment to maximize cumulative reward. The learning process is typically formalized through:

States (S): Representation of environment

Actions (A): Possible behavioral choices

Rewards (R): Feedback signals

Policy ( $\pi$ ): Strategy for action selection

Value function (V): Expected future reward

A central concept is the reward prediction error, which measures the difference between expected and received reward. This error drives learning by updating value estimates.

### **Mathematically:**

Positive RPE  $\rightarrow$  strengthens action-value association

Negative RPE  $\rightarrow$  weakens expectation

This computational structure closely mirrors neural learning mechanisms in the brain.

### **Dopamine and Reward Prediction Error Signaling**

One of the most influential discoveries in neuroscience is that midbrain dopamine neurons behave according to RL-like principles. These neurons, located primarily in the ventral tegmental area (VTA) and substantia nigra, respond not simply to reward, but to unexpected reward changes.

When reward is better than expected, dopamine activity increases. When it is worse, activity decreases. When reward is fully predicted, dopamine response diminishes.

This pattern corresponds closely to temporal-difference (TD) learning algorithms used in RL models.

Dopamine therefore acts as a teaching signal, guiding synaptic plasticity in downstream brain regions such as the striatum and prefrontal cortex. This mechanism forms the biological basis of reward-based learning in the brain.

### **Cortico-Basal Ganglia Circuits and Action Selection**

The basal ganglia play a central role in RL-inspired decision-making. These structures integrate cortical inputs and dopaminergic signals to determine which actions should be reinforced or suppressed.

#### **Key components include:**

Striatum: encodes action values

Globus pallidus: regulates output pathways

Substantia nigra: dopamine modulation

Thalamus: feedback to cortex

Within RL frameworks, the basal ganglia function as a policy selection system, where competing actions are evaluated based on expected reward.

Model-free RL is strongly associated with habitual behavior mediated by these circuits. Repeated reinforcement strengthens specific neural pathways, leading to automatic behavioral responses.

### **Model-Free vs Model-Based Learning in the Brain**

Neuroscience distinguishes two major RL systems:

#### **1. Model-Free Learning**

Relies on cached values from past experience

Fast but inflexible

Associated with dorsal striatum

Supports habitual behavior

#### **2. Model-Based Learning**

Uses internal simulation of environment

Flexible but computationally expensive

Associated with prefrontal cortex and hippocampus

Supports planning and goal-directed behavior

The brain dynamically balances these systems depending on task complexity and cognitive load.

Recent evidence suggests that prefrontal cortex may implement meta-RL functions, adjusting learning strategies based on context.

### **Role of the Prefrontal Cortex in Reinforcement Learning**

The prefrontal cortex (PFC) is essential for executive control and decision-making. It integrates information about goals, context, and past outcomes.

Computational models propose that PFC functions as a meta-reinforcement learning system, meaning it learns how to learn. Instead of storing fixed values, it dynamically adjusts learning rules based on experience.

This allows flexible adaptation in uncertain environments, enabling humans to solve complex tasks such as abstract reasoning and multi-step planning.

### **Limbic System and Emotional Reinforcement**

The limbic system, particularly the amygdala and hippocampus, contributes to RL by encoding emotional and contextual aspects of learning.

Amygdala: assigns emotional salience to stimuli

Hippocampus: encodes episodic memory of reward contexts

These structures influence how strongly rewards are remembered and how quickly behaviors are reinforced. Emotional learning often enhances RL efficiency, especially in survival-related contexts.

### **Synaptic Plasticity and Biological Implementation**

At the cellular level, RL corresponds to synaptic plasticity mechanisms such as:

Long-term potentiation (LTP)

Long-term depression (LTD)

Dopamine modulates these processes by influencing the strength of synaptic connections. When RPE is positive, synaptic strengthening occurs; when negative, weakening occurs.

This biological implementation ensures that learning is both adaptive and energy-efficient.

### **Reinforcement Learning in Neurological and Psychiatric Disorders**

Dysfunction in RL circuits is implicated in several disorders:

Parkinson's disease: dopamine depletion impairs reward learning

Addiction: exaggerated reward signaling leads to compulsive behavior

Depression: reduced reward sensitivity affects motivation

Schizophrenia: altered prediction error processing

Understanding these mechanisms has led to new therapeutic approaches targeting dopamine pathways and decision-making circuits.

### **Experimental Evidence Supporting RL in the Brain**

#### **Multiple experimental methods support RL theories:**

Single-neuron recordings show dopamine RPE signals

fMRI reveals reward-related activation in striatum and PFC

Optogenetics demonstrates causal roles of dopamine circuits

Behavioral experiments align with computational RL predictions

These findings strongly suggest that RL is not just a metaphor but a biologically grounded framework.

### **Limitations of Current Models**

#### **Despite success, RL models of brain systems have limitations:**

Oversimplified representation of neural states

Limited biological realism in artificial RL models

Difficulty modeling long-term planning and abstraction

Incomplete understanding of subjective reward

Additionally, real brain systems likely use multiple interacting RL algorithms rather than a single unified mechanism.

### **Future Directions**

#### **Future research is likely to focus on:**

Hierarchical reinforcement learning in cortical systems

Integration of RL with memory systems

Deep biologically plausible RL models

Multi-agent neural decision systems

Personalized computational psychiatry

Emerging frameworks suggest that brain learning is distributed, hierarchical, and continuously adaptive.

## **CONCLUSION**

Reinforcement learning provides a powerful framework for understanding how brain systems learn, adapt, and make decisions. From dopamine-driven reward prediction errors to complex cortical-basal ganglia interactions, RL principles are deeply embedded in neural computation.

While significant progress has been made, the brain remains far more complex than current computational models suggest. Continued integration of neuroscience, psychology, and artificial intelligence is essential for fully understanding reinforcement learning in biological systems.

## **REFERENCES**

1. Dalmau J, Armangue T, Planagumà J, Radosevic M, Mannara F, Leypoldt F, et al. An update on anti-NMDA receptor encephalitis for neurologists and psychiatrists: Mechanisms and models. *Lancet Neurol.* 2023;22(2):123-138.
2. Titulaer MJ, McCracken L, Gabilondo I, Armangue T, Glaser C, Iizuka T, et al. Treatment and prognostic factors for long-term outcome in patients with anti-NMDA receptor encephalitis: An updated cohort study. *Lancet Neurol.* 2021;20(7):573-584.
3. Graus F, Titulaer MJ, Balu R, Benseler S, Bien CG, Cellucci T, et al. Updated diagnostic criteria for autoimmune encephalitis. *Lancet Neurol.* 2020;19(9):748-763.
4. Abboud H, Probasco JC, Irani S, Ances B, Benavides DR, Bradshaw M, et al. Autoimmune encephalitis: Proposed best practice recommendations for diagnosis and acute management. *J Neurol Neurosurg Psychiatry.* 2021;92(7):757-768.
5. Nosadini M, Thomas T, Eyre M, Anlar B, Armangue T, Benseler SM, et al. International consensus recommendations for the treatment of pediatric anti-NMDA receptor encephalitis. *Neurology.* 2022;98(9):e1059-e1072.