

## **Dopamine and the neural circuit underlying learning**

We are all prediction-making machines. Granted, our predictions are often wrong—as the old saying goes, “It’s tough to make predictions—especially about the future.” But even wrong predictions serve a purpose: they help us learn. Each time we make a choice, we predict the outcome of that choice. When the outcome matches our prediction, there is no need to learn. When the outcome is unexpected, however, we update our predictions, hoping to do better next time.

The idea that we learn by comparing predictions to reality has been a mainstay of animal learning theory since the 1950s (1–3), and is one of the foundations of machine learning (4). Remarkably, the brain has evolved a simple mechanism to make precisely these comparisons. In the 1990s, Wolfram Schultz and colleagues found that dopamine neurons in the midbrains of monkeys showed a curious response to reward (5). When the monkeys received an unexpected reward (in this case, a squirt of juice), dopamine neurons fired a burst of action potentials. When that same reward was expected, the neurons no longer fired. And if Schultz et al. played a trick on the monkeys, making them expect reward but ultimately withholding that reward, the dopamine neurons dipped below their normal firing rate (6). Together, these results demonstrated that dopamine neurons signal prediction error, or the difference between actual and predicted value. If an outcome is better than predicted, dopamine neurons fire; if an outcome is the same as predicted, there is no change in firing; and if an outcome is worse than predicted, dopamine neurons dip below baseline. The level of dopamine release then informs the rest of the brain when a prediction needs to be fixed, and in what direction.

This basic finding — that dopamine neurons signal errors in reward prediction— revolutionized the study of learning in the brain, supplying a powerful, mechanistic model for how reinforcement affects behavior (7). Despite extensive study, however, little is known about how dopamine neurons actually calculate prediction error. What inputs do dopamine neurons combine and how do they combine them? To answer these questions, we merged molecular biology, electrophysiology, and computational analysis.

We focused on the ventral tegmental area (VTA), a small brainstem nucleus that produces dopamine. Although a majority of neurons in this region are dopamine neurons, a substantial minority use the inhibitory neurotransmitter GABA instead. A recent study from our lab showed that these GABA neurons do not signal prediction error; rather, they encode reward expectation (8).

This finding raised a fascinating question: could dopamine neurons use the GABA expectation signal to calculate prediction error? To find out, we used a virus to introduce the light-sensitive protein channelrhodopsin (ChR2) selectively in VTA GABA neurons. This enabled us to control the activity of VTA GABA neurons with light, a technique called optogenetics. We then implanted a set of electrodes surrounding a fiber optic cable into the VTA. Once the mice recovered from surgery, we recorded from the VTA and manipulated VTA GABA neuron activity, all while the mice performed simple learning tasks.

Optogenetics offers formidable precision, but there are potential pitfalls. In particular, it is easy to manipulate neural activity in ways that never occur in real life, producing results that are difficult to interpret. Our system avoided this pitfall because we knew how VTA GABA neurons normally fire in our task. By recording during the manipulation, we made sure to mimic natural firing patterns.

When we stimulated VTA GABA neurons, dopamine neurons responded to unexpected rewards as if they were expected (9). Conversely, when we inhibited VTA GABA neurons, dopamine neurons responded to expected rewards as if they were unexpected. Finally, if we manipulated VTA GABA neurons simultaneously on both sides of the brain, we even changed the animals' behavior. After training mice to expect a certain size of reward, we artificially increased the expectation level by stimulating VTA GABA neurons during the anticipation period. The reward level, meanwhile, stayed the same. After several trials in which expectation exceeded reality, the disappointed mice stopped licking in anticipation of reward. When we turned off the laser, their behavior slowly returned to normal. We concluded that VTA GABA neurons convey to dopamine neurons how much reward to expect. In short, they put the 'prediction' in 'prediction error'.

The VTA GABA expectation signal is only part of the puzzle. Another vital question is how dopamine neurons actually use this input. What arithmetic do they perform? Again, we used molecular techniques to 'tag' neurons with ChR2, but this time, we tagged dopamine neurons instead of GABA neurons. In each recording session, we shined pulses of light and identified neurons as dopaminergic if they responded reliably to each pulse. This ensured that the recorded neurons were indeed dopamine neurons, eliminating the need for other, less accurate identification methods (10).

Using insights from the sensory literature (11), we designed a task to assess the input-output function of identified dopamine neurons and determine how expectation transforms this function. We found that dopamine neurons use simple subtraction (9). Although this arithmetic is assumed in computational models, it is remarkably rare in the brain; division is much more common, as exemplified by gain control. However, subtraction is an ideal calculation because it allows for consistent results over a wide range of rewards. Moreover, we found that individual dopamine neurons calculated prediction error in exactly the same way (12). Each neuron produced an identical signal, just scaled up or down. In fact, even on single trials, individual neurons fluctuated together around their mean activity. Such uniformity greatly simplifies information coding, allowing prediction errors to be broadcasted robustly and coherently throughout the brain—a prerequisite for any learning signal. Presumably, target neurons rely on this consistent prediction error signal to guide optimal behavior.

Our work begins to uncover both the arithmetic and the local circuitry underlying dopamine prediction errors. The method of evidence accumulation, the inputs that signal reward, and the biophysics underlying subtraction remain to be discovered—prime material for fresh predictions and unforeseen rewards.

## References

1. R. R. Bush, F. Mosteller, A mathematical model for simple learning. *Psychol Rev.* **58**, 313–323 (1951).
2. L. Kamin, in *Fundamental issues in associative learning* (1969), pp. 42–64.
3. R. A. Rescorla, A. R. Wagner, in *Classical conditioning II: current research and theory*, A. Black, W. Prokasy, Eds. (1972), pp. 64–99.
4. R. S. Sutton, A. G. Barto, *Reinforcement learning: An introduction* (Cambridge Univ Press, 1998), vol. 1.
5. W. Schultz, P. Dayan, P. R. Montague, A neural substrate of prediction and reward. *Science.* **275**, 1593–1599 (1997).
6. J. R. Hollerman, L. Tremblay, W. Schultz, Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* **80**, 947–963 (1998).
7. W. Schultz, Updating dopamine reward signals. *Curr. Opin. Neurobiol.* **23**, 229–238 (2013).
8. J. Y. Cohen, S. Haesler, L. Vong, B. B. Lowell, N. Uchida, Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* **482**, 85–88 (2012).
9. N. Eshel *et al.*, Arithmetic and local circuitry underlying dopamine prediction errors. *Nature.* **525**, 243–246 (2015).
10. E. B. Margolis, H. Lock, G. O. Hjelmstad, H. L. Fields, The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol. (Lond.).* **577**, 907–924 (2006).
11. S. R. Olsen, V. Bhandawat, R. I. Wilson, Divisive normalization in olfactory population codes. *Neuron.* **66**, 287–299 (2010).
12. N. Eshel, J. Tian, M. Bukwich, N. Uchida, Dopamine neurons share common response function for reward prediction error. *Nat. Neurosci.* **19**, 479–486 (2016).