

ESSAY

CELL AND MOLECULAR BIOLOGY

Trial and error

Optogenetic techniques offer insight into the dopamine circuit underlying learning

By Neir Eshel

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 ma-

chine 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 a 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 [see the figure (A)]. 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 by 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 γ -aminobutyric acid (GABA) instead. A recent study from our laboratory showed that these GABA neurons do not signal prediction error; rather, they encode reward expectation (8).

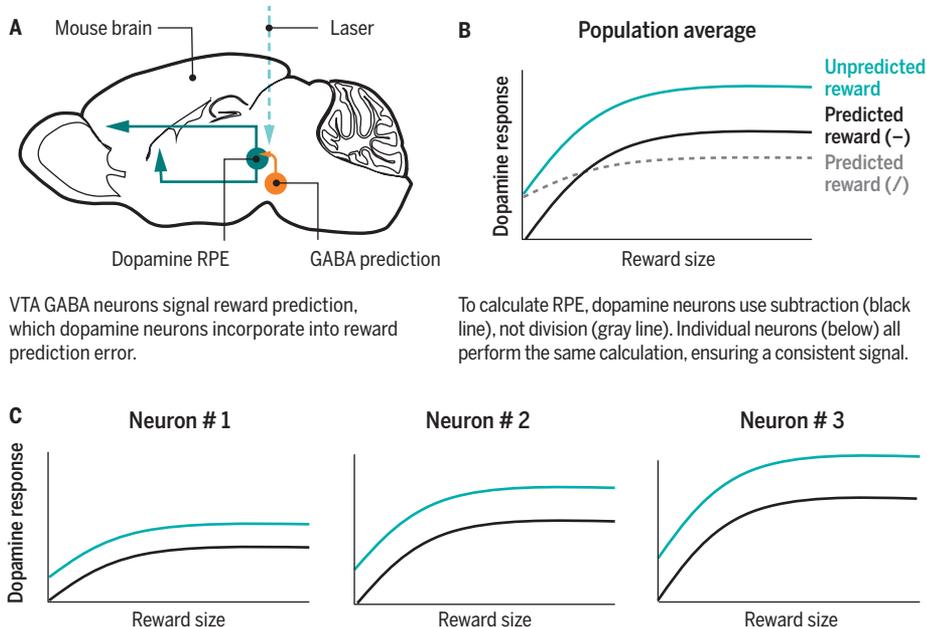
This finding raised a fascinating question: Could dopamine neurons use the GABA expectation signal to calculate prediction error? To find out, we used a virus to introduce the light-sensitive protein channelrhodopsin (ChR2) selectively in VTA GABA neurons. This enabled us to control the activity of VTA 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

Arithmetic and local circuitry of dopamine prediction errors

Dopamine neurons promote learning by conveying reward prediction error (RPE), the difference between actual and predicted reward. To probe how RPE is calculated, Eshel *et al.* recorded from dopamine neurons in mouse ventral tegmental area (VTA) while using optogenetics to manipulate nearby GABA neurons.



VTA GABA neurons signal reward prediction, which dopamine neurons incorporate into reward prediction error.

To calculate RPE, dopamine neurons use subtraction (black line), not division (gray line). Individual neurons (below) all perform the same calculation, ensuring a consistent signal.

Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94305, USA. E-mail: neshel@stanford.edu



GRAND PRIZE WINNER: CELL AND MOLECULAR BIOLOGY

Neir Eshel

Neir Eshel is a psychiatry resident at Stanford University, pursuing a career at the interface of research and clinical practice. He is interested in how we learn about rewards and punishments, how we make decisions based on this knowledge, and how these systems break down in neuropsychiatric disease. He has conducted research at the National Institutes of Health, Princeton University, the World Health Organization, University College London, and Harvard University. Outside the laboratory and clinic, Neir plays clarinet in chamber groups and orchestras and is a passionate advocate for lesbian, gay, bisexual, and transgender (LGBT) health equality. www.sciencemag.org/content/354/6316/1108



CATEGORY WINNER: TRANSLATIONAL MEDICINE

Canan Dagdeviren

Canan Dagdeviren was born in 1985 in Istanbul, Turkey. As a Fulbright Doctoral Fellow, she received her Ph.D. degree in Material Science and Engineering at the University of Illinois at Urbana-Champaign. Dagdeviren developed a conformable, piezoelectric, energy harvester that converts mechanical energy from internal organ movements into electric energy to power medical devices. It is soft and flexible and conforms to the heart as well as other soft tissues. This technology could extend the battery life of implanted electronics or eliminate the need of battery replacement, sparing patients from repeated operations and the risk of surgical complications. Beginning in January 2017, she will assume the role of assistant professor at the MIT Media Lab, where she will direct the Conformable Decoders research group. www.sciencemag.org/content/354/6316/1109.1



CATEGORY WINNER: GENOMICS AND PROTEOMICS

Sam Behjati

Originally from Germany, Sam Behjati read medicine at Oxford University and pursued postgraduate clinical training in London, UK. Funded by the Wellcome Trust, he joined the Cancer Genome Project of the Sanger Institute (UK) for doctoral research. Using a “mutational postcode” that cells acquire as they divide, Behjati has shown that it is feasible to reconstruct the early developmental process in the adult mouse. Building on ideas developed during his Ph.D. preparation, he now aims to define the embryonic origin and fate of childhood cancer cells. www.sciencemag.org/content/354/6316/1109.2



CATEGORY WINNER: ECOLOGY AND ENVIRONMENT

David Seekell

David Seekell is an environmental scientist based in Sweden. He holds a Bachelor of Science in Natural Resources from the University of Vermont, and a Ph.D. in Environmental Sciences from the University of Virginia. In his Ph.D. research, Seekell developed statistics to provide early warning that an ecosystem is passing a tipping point and is about to undergo a regime change. He is currently an assistant professor of Ecology in the Department of Ecology and Environmental Science at Umeå University. In 2015, he became a Wallenberg Academy Fellow, and in 2016, he received a Science for Solutions Award from the American Geophysical Union. www.sciencemag.org/content/354/6316/1109.3

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 do 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; this eliminated 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 to determine how expectation transforms this function. We found that dopamine neurons use simple subtraction (9) [see the figure (B)]. 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 in sensory systems. 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 [see the figure (C)]. 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 AND NOTES

1. R. R. Bush, F. Mosteller, *Psychol. Rev.* **58**, 313 (1951).
2. L. Kamin, in *Fundamental Issues in Associative Learning: Proceedings of a Symposium on Fundamental Issues in Associative Learning*, N. J. Macintosh and W. K. Honig, Eds., Dalhousie University, Halifax, Nova Scotia, June 1968 (Dalhousie Univ. Press, Halifax, 1969), pp. 42–64.
3. R. A. Rescorla, A. R. Wagner, in *Classical Conditioning II: Current Research and Theory*, A. Black, W. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64–99.
4. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction* (Cambridge Univ. Press, Cambridge, 1998), vol. 1.
5. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593 (1997).
6. J. R. Hollerman, L. Tremblay, W. Schultz, *J. Neurophysiol.* **80**, 947 (1998).
7. W. Schultz, *Curr. Opin. Neurobiol.* **23**, 229 (2013).
8. J. Y. Cohen, S. Haesler, L. Vong, B. B. Lowell, N. Uchida, *Nature* **482**, 85 (2012).
9. N. Eshel *et al.*, *Nature* **525**, 243 (2015).
10. E. B. Margolis, H. Lock, G. O. Hjelmstad, H. L. Fields, *J. Physiol.* **577**, 907 (2006).
11. S. R. Olsen, V. Bhandawat, R. I. Wilson, *Neuron* **66**, 287 (2010).
12. N. Eshel, J. Tian, M. Bukwich, N. Uchida, *Nat. Neurosci.* **19**, 479 (2016).

ACKNOWLEDGMENTS

I am grateful to the many colleagues and mentors who offered criticism and encouragement over the course of this research. Particular thanks to my Ph.D. adviser, N. Uchida, and to the members of the Uchida lab, especially M. Watabe-Uchida and J. Tian.

10.1126/science.aal2190

EXTENDED PDF FORMAT
SPONSORED BY



Trial and error

Neir Eshel (December 1, 2016)

Science **354** (6316), 1108-1109. [doi: 10.1126/science.aal2187]

Editor's Summary

This copy is for your personal, non-commercial use only.

- Article Tools** Visit the online version of this article to access the personalization and article tools:
<http://science.sciencemag.org/content/354/6316/1108>
- Permissions** Obtain information about reproducing this article:
<http://www.sciencemag.org/about/permissions.dtl>

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.