



Published in final edited form as:

Nat Neurosci. 2013 July ; 16(7): 777–779. doi:10.1038/nn.3448.

Dopamine signals mimic reward prediction errors

Geoffrey Schoenbaum,

Behavioral Neuroscience Branch, National Institute on Drug Abuse, Intramural Research Branch, US National Institutes of Health, Baltimore, Maryland, USA. Department of Anatomy and Neurobiology, University of Maryland, School of Medicine, Baltimore, Maryland, USA

Guillem R Esber, and

Behavioral Neuroscience Branch, National Institute on Drug Abuse, Intramural Research Branch, US National Institutes of Health, Baltimore, Maryland, USA

Mihaela D Iordanova

Department of Psychology, University of Maryland Baltimore County, Baltimore, Maryland, USA

Geoffrey Schoenbaum: geoffrey.schoenbaum@nih.gov

Abstract

Modern theories of associative learning center on a prediction error. A study finds that artificial activation of dopamine neurons can substitute for missing reward prediction errors to rescue blocked learning.

Dark clouds are present when it rains. So are umbrellas—but not always. Sometimes dark clouds come, rain falls and there are no umbrellas. As a result, we eventually learn that clouds predict rain and that we cannot use umbrellas as rain predictors. This concept, that associative learning is driven not by contiguity or simple co-occurrence, but by contingency and predictiveness, is at the heart of modern theories of associative learning^{1–3}. It is captured in these theoretical accounts by the dependence of learning on the presence of a prediction error: a discrepancy between what is expected and what is received. In this issue of *Nature Neuroscience*, Steinberg and Janak⁴ provide the most conclusive evidence to date that the associative learning circuits of the brain recognize brief bursts of activity in the midbrain dopamine neurons as a signal that a prediction error has occurred and that they learn accordingly.

A transformative idea in systems neuroscience has been the claim that rapid phasic changes in the firing of midbrain dopamine neurons signal these errors⁵. Starting with work by Schultz and colleagues⁶, a succession of studies has presented evidence showing progressively better correlations between the firing of dopamine neurons, across species and in both substantia nigra and ventral tegmental area, and errors in outcome prediction. These neurons increase firing to an unpredicted, but not a predicted, outcome and suppress firing when that predicted outcome is omitted (Fig. 1a). After training, they also fire to earlier

predictors—cues or events—when these occur unexpectedly. It would be hard to overstate the influence of these results, not just on theories surrounding dopamine function, but more broadly on how associated brain systems are thought to operate. To some extent, these findings have validated the entire approach of using principles derived from learning theory and computational science to explain how neural circuits function.

Yet despite the influence of the idea that dopamine neurons encode prediction errors, it is founded almost entirely on neural correlates. Evidence that the phasic dopamine signal actually serves as a prediction error has been much harder to come by. Until recently, causal experiments have had to rely on long-lasting pharmacological or genetic manipulations^{7–9}. These approaches have a fundamental flaw for addressing this specific hypothesis, which is that they are not temporally specific. Thus, although they can yield data consistent with the hypothesis, they are ill suited to invalidate it, as any failure to meet its predictions could be chalked up to the lack of temporal specificity. The recent advent of optogenetics, along with transgenic animals that allow one to insert these light-activated ion channels selectively into dopamine neurons, provided a potential solution to this problem. However, previous studies that applied this approach used behaviors, such as conditioned place preference and free, operant responding, that do not directly address the specific claim that dopamine neuron firing signals prediction errors as opposed to simply serving as a new rewarding outcome^{10,11}.

One way to get at this problem is to ask whether dopamine neurons can restore learning in a behavioral setting in which the cue and outcome remain contiguous and the only missing element is the prediction error. Blocking, first explored by Palladin¹² and fully developed by Kamin¹³, provides such a behavioral setting. In blocking (Fig. 1b), a target cue (umbrellas) is paired with an outcome (rain) whose occurrence is already predicted by another, previously conditioned cue (dark clouds). Even though the target cue is contiguous with the outcome, it fails to become a signal for the latter; subjects essentially ignore that cue if it is presented again later. This blockade of learning is thought to occur because the outcome is already fully predicted by the other cue, so there is no prediction error (and no phasic activation of dopamine neurons¹⁴) to drive learning.

Blocking provides a perfect vehicle to test whether restoring the prediction error, in the form of appropriately timed activation of the dopamine neuron, will rescue normal learning. Using the *tyrosine hydroxylase* promoter to drive expression of Cre recombinase in dopamine neurons in rats, which in turn triggered recombination and expression of a construct encoding channelrhodopsin, a light-sensitive sodium channel that causes neurons to fire when activated, Steinberg and Janak⁴ were able to use light to selectively trigger midbrain dopamine neurons in rats. In the critical experiment, the authors activated the dopamine neurons briefly when the rewarding outcome was delivered, mimicking the effect of a positive prediction error. They found that this pattern of activation had no effect on the rats' behavior at the time it was delivered; however, these rats showed conditioned responses to the normally blocked cue when it was presented later in a probe test. This rescue of learning, or unblocking, was not observed when the dopamine neurons were activated between the trials, unpaired from the reward, or in rats without Cre that received the same treatment but in whom channelrhodopsin was not expressed.

This experiment and the results are remarkable for several reasons. The first is simple: they are not correlative. Although single-unit recording, fast-scan cyclic voltammetry and functional magnetic resonance imaging are powerful techniques, the results are correlational. No matter how tight the correspondence between these measures and theoretical accounts of error signaling, they cannot demonstrate causality or show that the signal is actually used as an error. Here, Steinberg and Janak⁴ provide clear causal evidence that briefly activating dopamine neurons in a manner similar to the activity elicited by a prediction error is sufficient to drive learning.

These results are also remarkable for their temporal specificity. Although previous studies have shown that normal dopaminergic function is necessary for error-based learning (learning in procedures in which errors are isolated as the causal factor), these studies have used techniques, pharmacological or genetic, that lack temporal resolution⁷⁻⁹. Although the results in these reports are generally consistent with the idea that phasic dopamine signals can support error-based learning, alternative explanations involving more general effects on baseline or tonic levels remained plausible. The data from Steinberg and Janak⁴ rule out these possibilities.

Finally, and perhaps most importantly, the experiment by Steinberg and Janak⁴ addresses far more directly than previous attempts the question of whether the phasic activation of dopamine neurons serves as prediction error versus promoting learning for another reason, such as effects on salience or surprise or simply by serving as a new reward. The behavioral approach in the present experiment controls for many alternatives. This begins with the use of the blocking procedure. As noted above, in blocking, the target cue and the outcome are perfectly arranged to promote learning. The only thing missing is the prediction error. The prediction error is missing because there is a second cue, previously trained to predict the outcome. As a result, when the outcome is delivered, there is no difference between what is expected and what is received to drive learning to the target. The authors show that the lack of an error can be overcome by briefly activating the dopamine neurons when the outcome is delivered, which is the same time at which dopamine neurons would normally signal an error. The activation of the dopamine neurons did not alter the behavior of the rats to the compound cue, nor did it change their behavior at the time of stimulation or introduce a place preference to the food cup. Indeed, the form and shape of the conditioned response acquired to the target cue was the same as that exhibited to the other cue that was trained to predict the reward normally. Furthermore, pairing dopamine neuron stimulation with consumption of a flavored sucrose solution did not alter the rats' preference for that solution versus another flavored reward. Although absence of evidence is not evidence of absence, these various control data are not what one would expect if the transient activation of the dopamine neurons were perceived as a new reward. Instead, it is as if the brief activation of the dopamine neurons at the time of the reward caused the rats to discover that the normally blocked cue, which was otherwise contiguous with the reward, was predictive of the reward. This is consistent with the correlational data showing that dopamine neurons are activated not by reward *per se*, but by circumstances in which reward occurs unpredicted.

The authors also conducted several follow-up experiments to show that activation of the dopamine neurons could mitigate extinction of responding when the reward was reduced in

value or omitted. This shows in a different context that activating dopamine neurons can prevent or retard learning. It is particularly relevant given a recent report that some dopamine neurons increase firing to both unexpectedly good and bad events, possibly reflecting surprise in general or salience rather than the actual error signal reported more widely¹⁵. The effect on extinction demonstrated here suggests that, at least as a group, dopamine neurons are not signaling surprise or salience, as this would have been expected to enhance, rather than retard, learning.

As with any landmark study, the work of Steinberg and Janak⁴ opens the door to new questions as it closes the door on old ones. By showing that phasic activation of dopamine neurons is sufficient to replace a missing reward prediction error in such a well-controlled setting as blocking, this study allows us to now ask questions beyond this sticking point. For example it would be interesting to determine whether dopamine is also necessary for learning in a similar context. This could be readily done using unblocking. It would also be of interest to test whether learning unblocked by activation of the dopamine neurons reflects knowledge of the form and features of the specific reward that is present during the learning; this might be demonstrated by showing that responding to the unblocked cue is sensitive to devaluation of the reward or by showing that this cue can itself serve as a blocker. Such a demonstration would bolster the assertion that phasic dopamine acts as an error signal and not as a new reward.

Furthermore, it would be of interest to test more fully, using both causal and correlational approaches, whether the phasic dopamine signals have access to higher order information beyond what is contained in the present theoretical accounts mapped onto the firing of these neurons. Again, reliance on well-controlled learning theory tasks that force animals to rely on inference or model-based reasoning would provide a definitive answer. And finally, by combining these approaches with new molecular and genetic tools, it should be possible to also resolve questions about the specificity and possible heterogeneity of midbrain dopamine signals: whether the iconic, bidirectional reward prediction error is really all these neurons convey or whether the actual information represented in the release of dopamine (and perhaps co-release of other neurotransmitters in some areas) is more variable and perhaps targeted to different neural circuits. By resolving any real controversy over whether transient dopamine signals can replace missing reward prediction errors, the work by Steinberg and Janak⁴ provides a firm foundation from which such studies can proceed.

References

1. Rescorla, RA.; Wagner, AR. *Classical Conditioning II: Current Research and Theory*. Black, AH.; Prokasy, WF., editors. Appleton-Century-Crofts; New York: 1972. p. 64-99.
2. Pearce JM, Hall G. *Psychol Rev.* 1980; 87:532–552. [PubMed: 7443916]
3. Sutton RS, Barto AG. *Psychol Rev.* 1981; 88:135–170. [PubMed: 7291377]
4. Steinberg EE, et al. *Nat Neurosci.* 2013; 16:966–973. [PubMed: 23708143]
5. Schultz W, Dayan P, Montague PR. *Science.* 1997; 275:1593–1599. [PubMed: 9054347]
6. Mirenowicz J, Schultz W. *J Neurophysiol.* 1994; 72:1024–1027. [PubMed: 7983508]
7. Takahashi YK, et al. *Neuron.* 2009; 62:269–280. [PubMed: 19409271]
8. Iordanova MD, Westbrook RF, Killcross AS. *Eur J Neurosci.* 2006; 24:3265–3270. [PubMed: 17156387]

9. Darvas M, Palmiter RD. *Proc Natl Acad Sci USA*. 2009; 106:14664–14669. [PubMed: 19667174]
10. Tsai HC. *Science*. 2009; 324:1080–1084. [PubMed: 19389999]
11. Adamantidis AR, et al. *J Neurosci*. 2011; 31:10829–10835. [PubMed: 21795535]
12. Razran G. *Psychol Bull*. 1965; 63:42–64. [PubMed: 14274542]
13. Kamin, LJ. *Punishment and Aversive Behavior*. Campbell, BA.; Church, RM., editors. Appleton-Century-Crofts; New York: 1969. p. 242-259.
14. Waelti P, Dickinson A, Schultz W. *Nature*. 2001; 412:43–48. [PubMed: 11452299]
15. Matsumoto M, Hikosaka O. *Nature*. 2009; 459:837–841. [PubMed: 19448610]

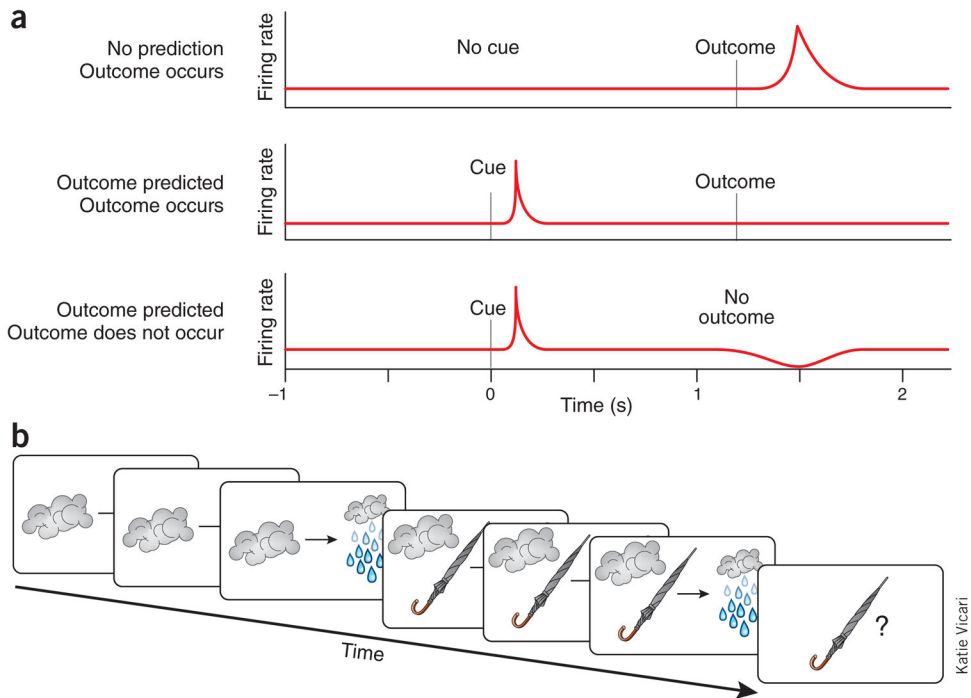


Figure 1. Prediction errors and blocking. **(a)** Illustration of the iconic dopaminergic prediction error, based on ref. 5. Dopamine neurons fire to unpredicted outcomes (top) and to cues that predict outcomes (middle); the same neurons will not fire to predicted outcomes (middle) and will suppress firing when predicted outcomes are omitted. This pattern of response is the fingerprint of a prediction error. **(b)** Illustration of blocking, using dark clouds, umbrellas and rain.