Published in final edited form as:

Trends Cogn Sci. 2013 September ; 17(9): . doi:10.1016/j.tics.2013.06.010.

Opening the black box: dopamine, predictions, and learning

Neir Eshel, Ju Tian, and Naoshige Uchida

Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA

Abstract

Dopamine neurons are thought to promote learning by signaling prediction errors, that is, the difference between actual and expected outcomes. Whether these signals are sufficient for associative learning, however, remains untested. A recent study used optogenetics in a classic behavioral paradigm to confirm the role of dopamine prediction errors in learning.

To adapt to their environment, animals must learn which stimuli to approach and which to avoid. This type of learning has a rich pedigree in psychology, from the slobbering of Pavlov's dogs to the pecking of Skinner's pigeons. Associative learning is more complicated than simply pairing a cue with an outcome, however. In the 1960s, Kamin [1] reported that if stimulus A already predicts an outcome, then combining A with a new stimulus X fails to trigger an association between X and the outcome (Figure 1). This phenomenon, called blocking, revealed a previously overlooked component of learning: surprise. For associations to form, the outcome must be different than expected. Indeed, this difference between actual and predicted outcome – known as prediction error – has since appeared in a bewildering array of studies, ranging from animal learning to computer science.

How are prediction errors encoded in the brain? In the 1990s, Schultz and colleagues recorded from midbrain dopamine (DA) neurons in monkeys and found that these neurons are excited by unexpected reward, unaltered by expected reward, and inhibited when expected reward is omitted [2]. In other words, they encode a bidirectional reward prediction error, exactly as predicted by theory. Furthermore, this error signal varies with reward magnitude, probability, and timing, and develops in lock-step with behavior, all pointing to a possible contribution of dopamine to conditioning. Correlation does not imply causation, however; does this prediction error signal actually trigger learning? More specifically, is blocking (the lack of association with X) caused by the reduced DA response when stimulus A already predicts reward? With cell-type specific, temporally precise methods to manipulate neuronal activity, this prediction can now be tested. Over the past five years, several groups have used channelrhodopsin-2 (ChR2), a light-gated proton channel, to demonstrate that activating DA neurons can reinforce behavior [3]. The importance of the prediction error response, however, remained unclear. In a recent paper, Steinberg *et al.* [4] combined optogenetics with the blocking paradigm to fill this gap.

Steinberg *et al.* [4] established the blocking effect in a three-phase task. In the first phase, rats were trained to associate auditory cue A with sucrose delivery in a reward port. In the

© 2013 Elsevier Ltd. All rights reserved.

Corresponding author: Uchida, N. (uchida@mcb.harvard.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

second phase, visual cue X was presented together with auditory cue A, again followed by sucrose. Finally, in the third phase (test phase), cue X was presented alone, in the absence of sucrose. The amount of time the rats spent in the reward port during the cue was used as the conditioned response. As expected, rats showed almost no response to cue X in the test phase – cue A had blocked learning to X. How did this happen? Steinberg *et al.* hypothesized that blocking occurred because DA neurons failed to respond to the fully-predicted outcome. If this is true, then stimulating DA neurons during the reward should unblock learning, causing rats to respond to cue X. This is exactly what they found.

Taking advantage of transgenic rats that expressed Cre recombinase under the control of the tyrosine hydroxylase promoter, Steinberg *et al.* [4] expressed ChR2 specifically in DA neurons in the ventral tegmental area. They then used a laser to stimulate DA neurons whenever the rats received sucrose. As expected, these rats – but not their wildtype littermates – showed a conditioned response to cue X. This result indicates that blocking – the diminished power of a predicted reward to cause associations – is due to a reduced DA response, providing strong evidence that DA regulates prediction error-based learning.

In a complementary experiment, Steinberg *et al.* [4] also tested whether the dip in DA neuron activity that occurs when the reward is smaller than expected is necessary to weaken the association between stimulus and outcome. After training the rats to associate a cue with sucrose reward, sucrose was either replaced with water (a less rewarding outcome) or omitted altogether. In this case, DA neurons usually reduce their firing and the animal stops responding to the cue. But when Steinberg *et al.* [4] activated DA neurons during the presentation of the new, worse outcome, rats maintained their response to the cue for a significantly longer period of time. Together with previous studies that inhibited DA neurons to cause avoidance learning [5], these results support the conclusion that the DA dip is involved in learning from worse-than-predicted outcomes. It remains to be established whether the brief dip (<500 ms) that DA neurons generally exhibit is sufficient for learning, as previous studies used longer periods of DA inhibition.

Steinberg *et al.*'s [4] results may also speak to a longstanding debate in the field about the origin of blocking. In particular, does blocking occur because an unsurprising outcome cannot support conditioning [1,6] or because animals' attention is drawn to already learned cues [7,8]? In other words, is the crucial component the cue or the outcome? The fact that DA stimulation at the time of reward is sufficient to cause unblocking suggests that it may be the outcome, rather than the cue, that explains blocking. Indeed, because DA neurons respond similarly to the single cue A and the compound AX [9], it is unlikely that DA cue responses can control attention selectively to one cue vs the other. For a stronger test of this hypothesis, future studies may wish to temporally separate the cue from the outcome and test whether DA stimulation is more effective at one time than the other (note that in Steinberg *et al.* [4], the cue remained available as the rats received their rewards). Experimenters could also measure the animals' attention to stimulus X (e.g., their orienting response) to ensure that attention is not altered by stimulation.

Knowing that DA prediction-error signaling can cause learning raises a new set of questions. How do DA neurons calculate prediction error [10]? Which of their projection targets are crucial for learning and how does DA facilitate this learning? What other pathways might the brain use in parallel? We now have the tools to answer all of these questions, and finally connect venerable psychological theories with the 'black box' inside the brain.

Acknowledgments

NE was supported by NIGMS T32GM007753 and the Sackler Scholar Programme in Psychobiology. Supported in part by NIH grant 5R01MH095953.

References

1. Kamin, L. Selective association and conditioning. In: Mackintosh, NJ.; Honig, FWK., editors. *Fundamental Issues in Associative Learning*. Dalhousie; 1969. p. 42-64.
2. Schultz W, et al. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
3. Steinberg EE, Janak PH. Establishing causality for dopamine in neural function and behavior with optogenetics. *Brain Res*. 2013; 1511:46–64. [PubMed: 23031636]
4. Steinberg, EE., et al. A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*. 2013. <http://dx.doi.org/10.1038/nn.3413>
5. Tan KR, et al. GABA neurons of the VTA drive conditioned place aversion. *Neuron*. 2012; 73:1173–1183. [PubMed: 22445344]
6. Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.; Prokasy, W., editors. *Classical Conditioning II: Current Research and Theory*. 1972. p. 64-99.
7. Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol Rev*. 1975; 82:276–298.
8. Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev*. 1980; 87:532–552. [PubMed: 7443916]
9. Waelti P, et al. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 2001; 412:43–48. [PubMed: 11452299]
10. Cohen JY, et al. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*. 2012; 482:85–88. [PubMed: 22258508]

Figure 1.

Dopamine prediction errors in associative learning. In the classic ‘blocking’ paradigm, an animal learns to associate a cue (here the sound A) with an outcome (a chunk of meat for a hungry dog). After many pairings of cue and outcome, the animal mounts a conditioned response (salivation) to the cue. Next, the original cue (A) is combined with another cue (the light X), and this compound cue (AX) is paired with a reward. Finally, the animal is presented with the new cue (X) alone. If simple pairing of cue and reward is sufficient for learning, the animal should respond to X much like it responded to A. But it does not: learning about X has been ‘blocked’ by A. This effect demonstrates the importance of predictions in associative learning – when a reward is fully predicted (in this case by cue A), then further learning does not occur. What is the neural mechanism for blocking? Dopamine neurons are known to signal prediction errors: they respond to unexpected rewards more than to expected rewards [2]. In blocking tasks, their response to reward is greatly reduced when a compound cue (AX) is presented. Furthermore, they respond to the original reward-predicting cue A, but not to the blocked cue X [9]. Is blocking a result of reduced dopamine responses to expected rewards? To answer this question, Steinberg *et al.* [4] used optogenetics to stimulate dopamine neurons at the time of expected reward in compound trials (in their study, sucrose was the reward in thirsty rats). As predicted, adding dopamine activity ‘unblocked’ stimulus X, causing animals to mount a conditioned response.