COMMENTARY

Dopamine: Context and counterfactuals

Michael L. Platt[a,b,c,1] and John M. Pearson[d]

Dopamine and Reward Prediction Errors

How does the brain learn what feels good? How does it adapt to new foods, bad bets, and played-out songs? How is it tricked into wanting ever-increasing doses of heroin or cocaine? Is there a simple rule that links neural signaling to the fundamental approach and avoidance responses on which more complex behavior is built?

It is now almost 20 y since Schultz, Montague, and Dayan (1) gave a provocative answer to these questions, an answer that is now recognized as one of the cornerstones of reinforcement learning and neuroeconomics. In a series of prior papers (reviewed in ref. 2), Schultz and collaborators had demonstrated that dopamine-releasing cells in the ventral midbrain respond to the delivery of rewards with a burst of action potentials, that these cells also respond to cues predicting rewards, and that over the course of learning, dopamine responses shift from the delivery of the reward to the predictive cue itself. Schultz, Montague, and Dayan proposed a simple, but powerful, computational account of these findings: Dopamine encodes a key variable posited by theories of reinforcement learning. These theories posit that animals select behaviors on the basis of which ones they expect to result in reward, updating their beliefs on the difference between expectations and observed outcomes, good or bad (3). This reward prediction error is large when rewards are unexpected and small when rewards are fully predicted, and its magnitude and sign drive the speed and direction of learning, respectively.

The findings of Kishida et al. directly implicate dopamine in counterfactual learning, regret, and disappointment.

This hypothesis proved seminal in several ways. First, it linked a particular neural signal to a computational model, thus permitting the development of quantitative predictions about the physiology of reward and learning. Second, because that computational model was a model of both learning and choice, it inspired subsequent studies probing the neurobiology of decision making. Finally, because many drugs of abuse directly affect dopamine function, the model provided a mathematical explanation for addiction. If dopamine encodes a prediction error, and if more dopamine signals that a reward is better than predicted, then every dose of cocaine or methamphetamine is more rewarding than expected, and the cues associated with these chemicals become powerful motivators for drug-seeking behavior. In other words, addiction is a normal learning process gone awry (4).

Human Dopamine Recordings Challenge the Reward Prediction Error Hypothesis

The dopamine prediction error finding has been endorsed by dozens of neurobiological studies in animals, and indirectly supported by brain imaging studies in humans (5, 6). In PNAS, Kishida et al. (7) build upon this work by directly measuring dopamine release in the human striatum. Their remarkable findings directly challenge the now classic view that dopamine release simply encodes errors in reward prediction and instead suggest that dopamine neurons compute these errors within the context of alternative outcomes (7). The findings of Kishida et al. directly implicate dopamine in counterfactual learning, regret, and disappointment (8–11).

The current gold standard for measuring dopamine in vivo with high temporal resolution is fast-scan cyclic voltammetry, a method now widely used in rodents (12) but only rarely usable in humans due to its invasive nature. For this study, Kishida et al. (7) recruited patients undergoing surgery to implant deep brain stimulators (DBS) to treat the symptoms of Parkinson's disease. Because DBS surgery requires direct access to the brain, investigators were able to record dopamine levels inside the striatum (caudate and putamen) using a voltammetric probe they previously developed and validated (13).

To probe the computational role of dopamine in humans, the investigators used a strategic investment

[a]Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104; [b]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104; [c]Marketing Department, University of Pennsylvania, Philadelphia, PA 19104; and [d]Duke Institute for Brain Sciences, Duke University, Durham, NC 27708

Author contributions: M.L.P. and J.M.P. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 200.

[1]To whom correspondence should be addressed. Email: mplatt@mail.med.upenn.edu.

www.pnas.org/cgi/doi/10.1073/pnas.1522315113

game they had previously developed and used to image human brain activity during decision making (14). In the game, participants make a series of bets on a fictitious stock that rises and falls in value over time. On each turn, players can bet between 0% and 100% of the current value of their portfolios, with returns based on the percent change in the stock's value in the next period. Investigators then calculate a prediction error based on the difference of the observed outcome of each choice and a predicted value estimated from the mean and the volatility of the stock.

Surprisingly, the investigators found that dopamine levels in the striatum were uncorrelated with standard reward prediction errors. Although dopamine responses to large bets did follow the predicted pattern, responses to small bets showed the reverse pattern, with negative prediction errors resulting in larger dopamine responses. To account for these findings, the authors developed a computational model in which dopamine encodes a mixture of actual and counterfactual prediction errors. According to this model, dopamine levels reflect not only experienced rewards and losses but also the rewards and losses that might have been experienced if a different decision had been made. Counterfactual reward signals have been reported previously in cortical brain areas in humans and monkeys (10, 11) but have never been directly linked to dopamine. The current study thus directly links dopamine to the intensely human feelings of regret and relief, disappointment and satisfaction.

This remarkable study is significant in several important ways. First, it demonstrates the feasibility and importance of studying dopamine signaling directly in humans and the value of invasive recordings in humans where possible and practical. Second, it raises an important question about the translatability of key findings derived from studies in rodents to humans. Although the current report does not directly contradict previous findings, it does add layers of complexity not yet probed in rodent studies. Finally, this study demonstrates once again the importance of computational models in providing an interpretive grid for neural data and suggesting new computational hypotheses for brain function.

Although Kishida et al. (7) raise important issues for the field, they also provoke a number of unanswered questions to be addressed in future work. Most prominently, dopamine measurements were made in patients suffering from Parkinson's disease, a disorder characterized by impaired dopamine function. In addition, DBS implantation requires that patients forego levodopa therapy before surgery, so they were tested in an OFF (untreated) as opposed to an ON (treated) state. Also, although this is one of the few, rare circumstances in which such experiments are possible, care must be taken when generalizing from patients to the general population.

Moreover, one of the key strengths of the computational approach is also a weakness. By comparing physiological signals to variables in the model, one implicitly assumes that the model itself provides at least an approximate account of behavior. In this case, Kishida et al. (7) found that a combination of prediction errors and counterfactual reward signals could account for the pattern of dopamine responses, but there are, no doubt, alternate models to consider, and these might yield diverging interpretations. Because no such model yet developed constitutes a "ground truth" against which to compare, there is a profound need for competitive testing of models to ensure that behavior provides independent evidence for the putative computational correlates of physiological signals.

Finally, as Kishida et al. (7) note, there are several methodological differences that make direct comparison of human and animal studies difficult. Because of the limitations of the surgical plan, experimenters were only able to record from the caudate and, in a few subjects, the putamen. It is possible that other locations, in particular the ventral striatum, would have yielded more-classic dopamine prediction error signals. Moreover, animal studies nearly always involve recordings performed only after extensive training, and typically make use of much simpler tasks. Differences in human dopamine levels may well reflect these discrepancies in both training and task complexity.

The present study makes plain the continued value of experiments in multiple model systems, even under the difficult circumstances of the operating room. Given the central role of dopamine in contemporary theories of both learning and decision making, the work by Kishida et al. (7) offers an important message to those working with other methods and in other systems: The computational role of dopamine is perhaps not quite so clear, or so limited, as we might have thought. Until scientists challenge the brain with the types of computational problems people routinely face in the real world, we may continue to underestimate its real complexity, not least the key capacities that make us human.

1 Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275(5306):1593–1599.
2 Schultz W (2010) Review: Dopamine signals for reward value and risk: Basic and recent data. *Behav Brain Funct* 6(1):24.
3 Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
4 Redish AD (2004) Addiction as a computational process gone awry. *Science* 306(5703):1944–1947.
5 Glimcher PW (2011) Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proc Natl Acad Sci USA* 108(Suppl 3):15647–15654.
6 Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442(7106):1042–1045.
7 Kishida KT, et al. (2016) Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proc Natl Acad Sci USA* 113:200–205.
8 Camille N, et al. (2004) The involvement of the orbitofrontal cortex in the experience of regret. *Science* 304(5674):1167–1170.
9 Coricelli G, et al. (2005) Regret and its avoidance: A neuroimaging study of choice behavior. *Nat Neurosci* 8(9):1255–1262.
10 Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the anterior cingulate cortex. *Science* 324(5929):948–950.
11 Abe H, Lee D (2011) Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* 70(4):731–741.
12 Phillips PEM, Robinson DL, Stuber GD, Carelli RM, Wightman RM (2003) Real-time measurements of phasic changes in extracellular dopamine concentration in freely moving rats by fast-scan cyclic voltammetry. *Methods Mol Med* 79:443–464.
13 Kishida KT, et al. (2011) Sub-second dopamine detection in human striatum. *PLoS One* 6(8):e23291.
14 Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci USA* 104(22):9493–9498.