Association learning: Dopamine and the formation of backward associations

Michael D. Kendig and Laura A. Bradfield*
School of Life Sciences, Faculty of Science, University of Technology Sydney, Sydney, NSW 2007, Australia
*Correspondence: laura.bradfield@uts.edu.au
https://doi.org/10.1016/j.cub.2022.06.007

The activity of dopamine neurons is critical for the ability to learn and update cue–reward associations. New work in rats shows that dopamine transients are also critical for the formation of backward associations in which the reward precedes the neutral stimulus.

A central research question asks which neural processes facilitate the learning of associations between cues in the surrounding environment and biologically significant outcomes. Cognitively, this process is thought to be driven by 'prediction error': the difference between the actual value of the reward and that expected in the presence of the cue(s). Although activity of midbrain dopaminergic neurons has long been considered a close neural correlate of prediction error, the seminal finding that generated this hypothesis1 and much subsequent work has relied on experimental preparations in which cue presentations precede reward delivery, whereas associations between neutral cues and rewards can form through a multitude of temporal arrangements. For instance, as illustrated in Figure 1, if you were to regularly eat ice cream on your way to the cinema (because you pass your favourite shop on the way), but only buy popcorn once you arrive, you would likely form both forward cue–reward associations between the cinema's cues and popcorn, as well as backward reward– cue associations between the cinema and ice cream. The cinema could even become a specific inhibitory cue for ice cream, as once you arrive you learn that it will be a certain amount of time before you will eat ice cream again. Dopamine's role in forming cue–reward associations is well established, but its role in forming reward– cue associations has been unclear. As they report in this issue of Current Biology, Seitz et al.2 have now demonstrated such a role by optogenetically inhibiting dopaminergic neurons in the ventral tegmental area (VTA) of rats during a Pavlovian backward conditioning procedure.

To enable the targeted inhibition of dopamine neurons with temporal precision, Seitz et al.2 injected an inhibitory halorhodopsin (NpHR) virus and implanted optic fibres into the VTA in transgenic rats expressing Cre-recombinase under control of the tyrosine hydroxylase promotor, used to ensure the specific transfection of dopamine neurons. Backward conditioning procedures began with rats receiving intermixed presentations of two palatable rewards — pellets and maltodextrin solution — each followed by unique auditory cues presented 10 seconds later. All animals received a green light delivered to the VTA for 2.5 seconds at the onset of each cue presentation, effectively inhibiting the activity of VTA dopamine neurons for the NpHR-injected group but not for the group injected with a control virus containing only enhanced yellow fluorescent protein (group eYFP).

Following backwards conditioning, rats were next trained to respond on one lever for the pellet reward and on a second lever for the maltodextrin reward in the absence of optogenetic inhibition. This procedure allowed Seitz et al.2 to probe the content of the learned associations. To do this, they gave rats a Pavlovian-to-instrumental transfer (PIT) test in which both levers were extended but no food rewards delivered. In control rats, the presentation of each cue drove responding on the lever that had been associated with the alternative outcome. Specifically, the backward cue that had followed pellet presentations increased responding on the maltodextrin lever, and the maltodextrin backward cue increased responding on the pellet lever, suggesting that each cue had become inhibitory of the outcome it preceded. Optogenetic inhibition of VTA dopamine transients during backwards conditioning abolished these effects, indicating that the ability of a backward-paired cue to guide responding toward the alternative reward depended on intact VTA dopamine neuron activity. Seitz et al.2

conducted additional tests to identify the inhibitory and/or excitatory associations that were disrupted by inhibiting VTA dopamine neurons. First, they paired two new visual cues with the same food rewards in a conventional forward association (where cue precedes reward). In control rats, responding to these visual food cues was suppressed when they were presented simultaneously with the backward-paired auditory cues.

Moreover, the backward-paired cues' inhibitory effects were reward-specific because responding during the pellet paired visual cue was more suppressed by the backward-paired pellet cue than the backward-paired maltodextrin cue, and vice versa. By contrast, in the NpHR-injected group, responding for the visual food-paired cues was unaffected by the presence and identity of the backward-paired auditory cues, suggesting that the inhibition of VTA dopamine at cue onset prevented the formation of specific inhibitory associations in backward conditioning. If we relate this finding to our example above, it suggests that dopamine signals are necessary to learn that the cinema is inhibitory of ice cream specifically, rather than of rewards more generally. A final experiment confirmed that inhibiting VTA dopaminergic neurons during cue-onset during forward conditioning did not prevent learning, though the initial rate of learning may have been delayed. The authors interpreted their findings to suggest that VTA dopaminergic neurons facilitate learning about contiguous events that may or may not include value. This is a bold claim, and represents a significant departure from the aforementioned theories of dopamine function that centre around the belief that it represents a teaching signal containing reward-prediction error information of the kind captured in Sutton and Barto's temporal difference reinforcement learning model3. As such, it deserves to be unpacked. The central facet of the contiguity claim is that, in contrast to the predictions of the temporal difference reinforcement learning model, in which value is non-directional and defined entirely by reward magnitude, VTA dopamine is clearly necessary for learning then reward value is both multifaceted (that is, defined by magnitude and identity) and can be altered depending on its directional relationship with a cue (forwards or backwards).

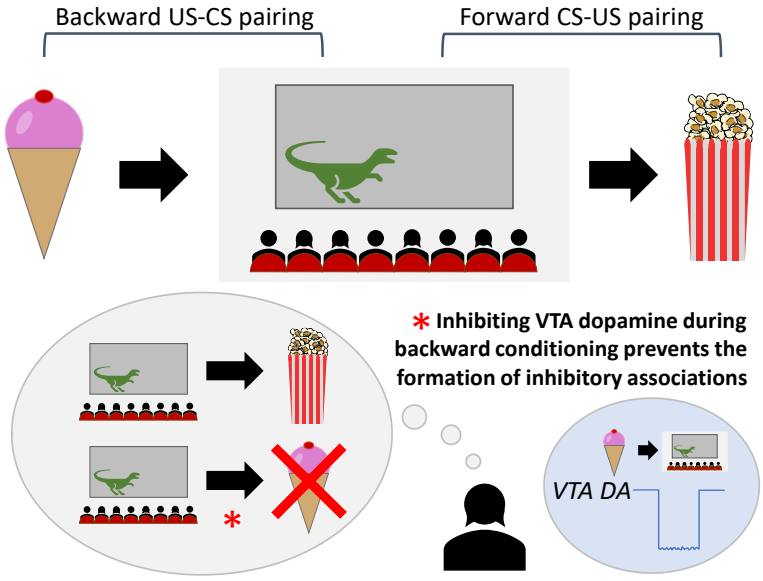

Figure 1. A depiction of backward and forward associations formed in relation to food consumed before and after attending the cinema. If your Sunday routine involves stopping by the ice cream parlour en route to the cinema, where you finish your cone and enter the movie to buy popcorn, the initially neutral cinema 'cues' are both backward-paired with ice cream and forward-paired with popcorn. Over repeated pairings, an excitatory association may form between the cinema and popcorn (the cinema predicts the forthcoming consumption of popcorn) while an inhibitory association might form between the cinema and ice cream (the cinema predicts the impending absence of ice cream). Forward cue–reward associations have long been understood to rely on dopamine transients in the brain. In their new paper, Seitz et al.2 show that inhibiting ventral tegmental area (VTA) dopamine transients during backward reward–cue pairings also prevents the formation of specific inhibitory associations.

Nevertheless, the claim that VTA dopamine drives learning through a contiguity-based mechanism is challenged by previous findings of Waelti et al.4, who demonstrated an absence of dopamine responding to a 'blocked' stimulus despite its contiguous pairings with both an excitatory cue and reward. Specifically, midbrain dopamine activity was measured in monkeys who had first learned to associate the visual cue A with juice, then received compound AX presentations also followed by juice. Because the juice was already well predicted by A, learning

about the association between X and juice was 'blocked', and dopamine neurons did not respond to later presentations of X alone. That is, dopamine neurons did not appear to 'learn' about cue X despite its contiguous pairings with both cue A and juice. More recently, Maes et al.5 demonstrated that inhibiting VTA dopamine signals during cue onset was likewise insufficient to prevent blocking in rats. Together, these findings suggest that the simple presentation of contiguous events is insufficient to engage dopamine. Rather, and as suggested by Seitz et al.2, it seems that dopamine transients are only necessary to associate contiguous events if there is some kind of prediction error between them, regardless of whether that error is driven by differences in predicted value, identity, or salience.

An alternative possibility is that both contiguity and scalar prediction error may be encoded by midbrain dopamine signals. Indeed, it is becoming increasingly recognised, including by Seitz et al.2, that dopamine does not necessarily do just 'one thing' in the brain, but that its role likely spans multiple functions including motivation, motor control (including action selection and inhibition), arousal, and reward encoding. Future studies will determine whether the functional heterogeneity of dopamine is defined by anatomical or projection specificity, as suggested by Lammel et al.6 and others7, by the density and distribution of different dopamine receptors throughout the brain, by interactions with other neurotransmitters and neuromodulators, and/or by some other, as-yet-undetermined factor. We believe that future studies that parse the roles of receptor subtypes in these effects might be particularly fruitful, given evidence that dopamine D1 and D2 receptors exert opposing effects on approach/avoidance behaviour8, conditioned place preference9 and nicotine ingestion and withdrawal10.

Finally, future studies that continue to elucidate the role of dopamine in backwards conditioning could have important clinical implications. While Seitz et al.2 used appetitive outcomes to show that backward-paired cues bias responding towards actions earning alternative rewards, studies using aversive outcomes show that backward conditioning may imbue cues with desirable associations. For example, studies of wheel running in rodents show that although forward flavour–running pairings generate a conditioned taste aversion, flavours consumed after running are relatively more preferred11, and a conditioned place preference develops for neutral contexts exposed after running12. It has also been suggested that foods consumed after chemotherapy sessions may be less likely to become aversive than those consumed before13. Thus, in addition to the potential implications of this work for pathologies such as schizophrenia noted by Seitz et al.2, if dopamine activity is necessary for the formation of aversive as well as appetitive backwards associations, it is also possible that these useful backwards associations could be facilitated through the use of pharmaceuticals that target dopamine function.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science 275, 1593–1599.

2. Seitz, B.M., Hoang, I.B., DiFazio, L.E., Blaisdell, A.P., and Sharpe, M.J. (2022). Dopamine errors drive excitatory and inhibitory components of backward conditioning in an outcome-specific manner. Curr. Biol. 32, 3210–3218.

3. Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (MIT Press).

4. Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. Nature 412, 43–48.

5. Maes, E.J.P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G., and Iordanova, M.D. (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. Nat. Neurosci. 23, 176–178.

6. Lammel, S., Lim, B.K., and Malenka, R.C. (2014). Reward and aversion in a heterogeneous midbrain dopamine system. Neuropharmacology 76, 351–359.

7. Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A., Crow, A.K., Malenka, R.C., Luo, L., Tomer, R., et al. (2015). Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. Cell 162, 635–647.

8. Nguyen, D., Alushaj, E., Erb, S., and Ito, R. (2019). Dissociative effects of dorsomedial striatum D1 and D2 receptor antagonism in the regulation of anxiety and learned approach-avoidance conflict decision-making. Neuropharmacology 146, 222–230.

9. White, N.M., Packard, M.G., and Hiroi, N. (1991). Place conditioning with dopamine D1 and D2 agonists injected peripherally or into nucleus accumbens. Psychopharmacology 103, 271–276.

10. Grieder, T.E., George, O., Tan, H., George, S.R., Le Foll, B., Laviolette, S.R., and van der Kooy, D. (2012). Phasic D1 and tonic D2 dopamine receptor signaling double dissociate the motivational effects of acute nicotine and chronic nicotine withdrawal. Proc. Natl. Acad. Sci. USA 109, 3101–3106.

11. Salvy, S.J., Pierce, W.D., Heth, D.C., and Russell, J.C. (2004). Taste avoidance induced by wheel running: effects of backward pairings and robustness of conditioned taste aversion. Physiol. Behav. 82, 303–308.

12. Lett, B.T., Grant, V.L., Koh, M.T., and Smith, J.F. (2001). Wheel running simultaneously produces conditioned taste aversion and conditioned place preference in rats. Learn. Motivat. 32, 129–136.

13. Jacobsen, P.B., Bovbjerg, D.H., Schwartz, M.D., Andrykowski, M.A., Futterman, A.D., Gilewski, T., Norton, L., and Redd, W.H. (1993). Formation of food aversions in cancer patients receiving repeated infusions of chemotherapy. Behav. Res. Therap. 31,739–748.