

## **Enhancing reward learning in the absence of an effect on reward**

**This scientific commentary refers to ‘Impulse control disorder in Parkinson’s disease is associated with abnormal frontal value signalling’ by Tichelaar *et al.***

**(<https://doi.org/10.1093/brain/awad162>).**

The central dopaminergic system is involved in a range of processes including motor control and vigour, reward learning and hedonic experience. It is of little surprise therefore that the dopaminergic medications used to treat Parkinson’s disease have a range of effects over and above their intended impact on action initiation. Dopamine’s role in reward-based learning and decision making has been of particular interest, with reasonable arguments being made that medication effects on this system may be responsible both for the often helpful additional action of dopaminergic medication in reducing symptoms of depression, (1) and the less helpful action of inducing impulse control difficulties. (2,3) While there has been a reasonable consensus on the impact of these medications on reward learning, a degree of heterogeneity in findings remains, in part related to the often modest sample sizes used in prior work. In this context, the article by Tichelaar and co-workers in this issue of *Brain*, (4) which reports the results of a substantial case-control study of reward learning in patients with Parkinson’s disease, is to be welcomed.

The authors use a simple instrumental reward learning task, with both behavioural and neuroimaging outcomes, to assess reward guided learning and decision making in groups of patients with Parkinson’s disease on and off medication and in non-clinical control participants. The patient group is further subdivided into those with and without significant symptoms of impulse control difficulties and depression. The authors report that the

previously described effect of dopaminergic medication on reward learning—an increase in the selection of more highly rewarded outcomes—is not apparent across all patients, but is rather restricted to those patients with higher levels of impulse control difficulties. While the associational nature of the study precludes strong conclusions about the causal effect of medication on reward driven choice, the effect of dopaminergic medication on rewarded choice in patients with impulse control difficulties has been previously reported in smaller patient samples using an experimental design (5) and so seems convincing.

An interesting question arising from this work is what mechanism might underlie the selective effect of medication in impulsive patients, and whether this might tell us something about the role of the medication in the production of impulsive symptoms, or perhaps the action of dopamine on reward learning and choice more generally. The functional neuroimaging results provide a hint at what may be going on—the difference in activity in the prefrontal cortex (PFC) of patients on vs. off medication was greater in those patients with, than without, impulse control difficulties. Importantly, no difference in outcome-locked response was observed, when reward prediction errors (RPE) would be expected. The authors interpret this as evidence that medication had influenced the degree to which learned values were used to drive choice, rather than the changes in the reward learning systems themselves (which would be expected to alter RPEs). In one sense this finding is not straightforward: previous studies have relatively consistently found a difference in RPEs between patients off medication and control participants, but this effect was not apparent in the current study, raising concerns over the sensitivity of the measure used. Further, other smaller studies, using both experimental (5) and associational (6) designs, have described an effect of dopaminergic medication on RPEs in similar populations with raised impulsivity. Balancing this however, the current study was able to detect a coherent association with medication at the time of

choice, and recruited a substantially larger sample than previous studies, so the lack of an RPE effect of dopaminergic medication cannot be simply written off.

What sort of process might lead to an increase in reward guided choice in the absence of an effect on RPEs? An interesting possibility is suggested by recent work on the role of the dopaminergic system in habit learning (Figure 1; 7). Actions that result in rewarding outcomes are generally repeated, with reinforcement learning providing a useful framework to understand how this preference develops. Reinforcement learning describes RPEs, the difference between the received and expected reward, as the signal which drives this learning. A great deal of previous work has demonstrated that RPEs are broadcast by dopaminergic neurons projecting from the ventral tegmental area to the ventral striatum. Over time however, actions that are frequently taken may become habitual, that is they persist even in the absence of further reward, and it is not clear how to incorporate such reward insensitive choice within existing reinforcement learning accounts. Recent theoretical work has suggested that such habitual actions may be driven by a subset of dopaminergic neurons that code 'action prediction errors' (APE), the difference between the action actually taken and that usually taken in a given state (7), rather than RPEs.

In support of this, early preclinical work has recently reported the presence of APEs in dopaminergic activity in the dorsal striatum of rodents (8). While APEs are conceptually similar to RPEs (they are both calculated as the difference between actual and expected outcomes), the key feature of APEs is that they drive the repetition of familiar actions regardless of whether reward is received. In effect APEs drive a learning system that attempts to mimic the choices previously made by the reward sensitive system, but do so in a more efficient manner. In the task used in the current study participants tend to select the more

highly rewarded outcome. An increase in dopaminergically mediated habit learning would be expected to exaggerate this tendency, but would do so independently of received rewards, and thus RPEs, as suggested by the neuroimaging findings reported. In other words, dopaminergic medication may increase the choice of rewarded outcomes, by influencing a system which itself is insensitive to rewards. A habit driven effect of dopaminergic treatment may also be important in understanding the development of impulse control problems during which patients repeat actions in the face of frankly negative consequences.

Regardless of the specific mechanism behind the effects reported in the paper by Tichelaar and colleagues, it provides a fascinating addition to the literature on the role of dopamine in rewarded and impulsive behaviour in Parkinson's disease and demonstrates once again the complex role of the dopaminergic system in human cognition.

*Michael Browning<sup>1,2\*</sup>*

*<sup>1</sup>Department of Psychiatry, University of Oxford, UK*

*<sup>2</sup>Oxford Health NHS Trust, Oxford, UK*

Email: michael.browning@psych.ox.ac.uk

## **Competing interests**

MB has acted as a consultant for Jansen, CHDR and Boehringer, has previously held shares in Plival Ltd and has received travel support for attending a conference from Lundbeck.

## **Funding**

MB is supported by the NIHR Oxford Health Biomedical Research Centre. The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health.

## References

1. Barone P, Poewe W, Albrecht S, Debieuvre C, Massey D, Rascol O, *et al.* (2010): Pramipexole for the treatment of depressive symptoms in patients with Parkinson's disease: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol* 9: 573–580.
2. Weiss HD, Marsh L (2012): Impulse control disorders and compulsive behaviors associated with dopaminergic therapies in Parkinson disease. *Neurol Clin Pract* 2: 267–274.
3. Halahakoon DC, Kaltenboeck A, Martens M, Geddes JG, Harmer CJ, Cowen P, Browning M (2023): Pramipexole Enhances Reward Learning by Preserving Value Estimates. *Biol Psychiatry* S0006-3223(23)01356–2.
4. Tichelaar JG, Sayalı C, Helmich RC, Cools R (2023): Impulse control disorder in Parkinson's disease is associated with abnormal frontal value signalling. *Brain* awad162.
5. Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010): Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron* 65: 135–142.
6. Piray P, Zeighami Y, Bahrami F, Eissa AM, Hewedi DH, Moustafa AA (2014): Impulse control disorders in Parkinson's disease are associated with dysfunction in stimulus valuation but not action valuation. *J Neurosci* 34: 7814–7824.

7. Bogacz R (2020): Dopamine role in learning and action inference ((T. Kahnt & K. M. Wassum, editors)). *eLife* 9: e53262.
8. Greenstreet F, Vergara HM, Pati S, Schwarz L, Wisdom M, Marbach F, *et al.* (2022, September 14): Action prediction error: a value-free dopaminergic teaching signal that drives stable learning. bioRxiv, p 2022.09.12.507572.

## Figure legend

### Figure 1 Parallel goal-directed and habitual learning systems (adapted from (7)).

Parallel learning systems are instantiated in the striatum. A goal directed learning system (red) and a habit learning system (blue). Both systems receive information about the current state, or context ( $s$ ) from cortical inputs. The goal directed learning system receives inputs about received rewards ( $r$ ) via the thalamus which it uses to update its estimate of the value of taking an action in a given state ( $P(R|s, a)$ ) by calculating a reward prediction error ( $\delta_g$ ). The value of the action can be used to influence choice ( $a$ ; green) in reward learning tasks such as that reported by Tichelaar and co-workers. The reward prediction error is encoded in the activity of dopaminergic neurons projecting from the ventral tegmental area to the ventral striatum. The habit learning system has an anatomically similar arrangement, but is learning a different quantity—the probability of taking a given action in a particular state ( $P(a|s)$ ). Here dopamine encodes an action prediction error, which is independent of received rewards. The habit learning system is effectively attempting to mimic the choices of the goal directed system and so will increase the choice of rewarded options in tasks such as that reported here. Dopaminergic drugs would be predicted to increase both reward sensitive, value learning and reward insensitive habit learning (for more details see (7)).