What is the role of orbitofrontal cortex in dopamine-dependent reinforcement learning?

Robert C. Wilson[1], Yuji K. Takahashi[2], Geoffrey Schoenbaum[2,3] and Yael Niv[1]
[1]Department of Psychology and Neuroscience Institute, Princeton University, [2]Department of Anatomy & Neurobiology, [3]Department of Psychiatry, University of Maryland School of Medicine

Orbitofrontal cortex (OFC) has been implicated in signalling reward expectancies, but its exact role , and how this differs from that the role of ventral striatum (VS), is an open question. One idea is that VS is the seat of value learning in model-free, dopamine-dependent reinforcement learning, while OFC represents values in dopamine-independent model-based learning. However, recent results [1] show that unilateral OFC lesions significantly alter dopaminergic prediction error signalling, as recorded from single units in the VTA of rats performing a learning task (described in [2]).

To understand the precise contribution of OFC to prediction error signalling, we compared predictions from different actor/critic architectures (embodying alternative hypotheses about OFC) to the experimental results. The electrophysiological results could not be explained by positing that OFC learns reward expectancies in a model-free system (i.e., OFC as the “critic”), either instead of, or alongside the VS. Similarly, positing that OFC learns expectancies in a model-based manner, and conveys these to dopamine, predicted results at odds with the data.

The only model that accounted for the full pattern of experimental results, was one in which the OFC enhances the representation of the task so that it includes not only tangible external inputs (e.g., presence of an odor) but also abstract knowledge (e.g., previous decisions). The latter, which we hypothesize is the OFC’s unique contribution, can be used to distinguish between states that are externally similar, but predict different rewards. We suggest that this enhanced representation inputs into VS and serves as a basis for model-free reinforcement learning and prediction error computation. It is also an ideal basis for model-based learning. Lesioning OFC thus impacts prediction error signalling by forcing learning to rely on simple, stimulus-bound, representations. Beyond these data, our hypothesis accounts for a variety of seemingly divergent findings in the OFC literature.

[1] Takahashi et al. Orbitofrontal cortex is required for expectancy-related changes in phasic firing of midbrain dopamine neurons. 404.2. 2010 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2010. Online.

[2] Roesch et al. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nature Neuroscience 10(12) 1615-1624 (2007)