

Russ Poldrack
UCLA Psychology

Modeling prediction error during classification learning

A prominent theory in neuroscience suggests reward learning is driven by the discrepancy between a subject's expectation of an outcome and the actual outcome itself. Furthermore, it is postulated that midbrain dopamine neurons relay this mismatch to target regions including the ventral striatum. Using functional Magnetic Resonance Imaging (fMRI), we tested striatal responses to prediction errors for probabilistic classification learning with purely cognitive feedback. We used a version of the Rescorla-Wagner model to generate prediction errors for each subject and then entered these in a parametric analysis of fMRI activity. Activation in ventral striatum/nucleus-accumbens (Nacc) increased parametrically with prediction error for negative feedback. This result extends recent neuroimaging findings in reward learning by showing that learning with cognitive feedback also depends on the same circuitry and dopaminergic signaling mechanisms.