

Is human learning driven by Prediction Error?

Jiří Čevora

Cognition and Brain Sciences Unit, Medical Research Council

Máté Lengyel

Computational and Biological Learning Lab, Department of Engineering, University of Cambridge

Richard Henson

Cognition and Brain Sciences Unit, Medical Research Council

Abstract: Prediction Error [PE] is a core component of some of the most influential theories of how animals use experiences to update their knowledge (e.g, Rescorla & Wagner, 1972). The classic demonstration of PE is the single-cell recording done by Schultz and colleagues (1997). However, there is no evidence that this signal plays any role in learning.

Only two studies have related a neural correlate of PE to learning performance so far (Gläscher, Daw, Dayan, & O'Doherty, 2010; McGuire, Nassar, Gold, & Kable, 2014). We provide a formal analysis demonstrating that non-PE learning can also explain the results of these studies if the imaging signal they identify relates to the size of weight updates instead of PE.

We conclude that the case for PE driving many forms of animal learning is not yet sufficiently proven, and identify approaches which can potentially resolve this question in future.