

An inductive bias for slowly changing features in human reinforcement learning

Noa Hedrich

Max Planck Institute for Human Development, Berlin, Germany

Eric Schulz

Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Sam Hall-McMaster

Harvard University, Cambridge, Massachusetts, United States

Nicolas Schuck

Universität Hamburg, Hamburg, Germany

Abstract

Distinguishing relevant features from noise is a central challenge for efficient behaviour. We asked whether humans address this challenge by leveraging the insight that behaviourally relevant processes change on a slower timescale than noise. To test this idea, participants were asked to learn the rewards of two-dimensional bandits when either a slowly or quickly changing feature of the bandit predicted reward. Participants accrued more reward and achieved better generalisation to unseen bandits when the reward-predictive feature changed slowly, rather than quickly. These effects were stronger when participants experienced the feature speed before learning about rewards. Computational modelling revealed that participants adjusted their learning rates based on feature speed. Those who learned better from slow features also had higher learning rates for it from the onset. These results provide evidence that human reinforcement learning favours slower features, suggesting a bias in how humans approach reward learning.