Three wrongs make a right: reward underestimation mitigates idiosyncrasies in human bandit behavior

Abstract:
Combining a multi-armed bandit task and Bayesian computational modeling, we find that humans systematically under-estimate reward availability in the environment. This apparent pessimism turns out to be an optimism bias in disguise, and one that compensates for other idiosyncrasies in human learning and decision-making under uncertainty, such as a default tendency to assume non-stationarity in environmental statistics as well as the adoption of a simplistic decision policy. In particular, reward rate underestimation discourages the decision-maker from switching away from a “good” option, thus achieving near-optimal behavior (which never switches away after a win). Furthermore, we demonstrate that the Bayesian model that best predicts human behavior is equivalent to a particular class of reinforcement learning models, thus giving statistical, normative grounding to phenomenological models of human behavior.

Host: Ruth Williams
Thursday, January 23, 2020
3:00 PM
AP&M 7218