Title: Reinforcement learning of abstract rules involves the dorsal striatum and prefrontal cortex

Authors: Carles Sindreu, Rafael Marín, Daniel Duque, Ainhoa Hermoso-Mendizabal, Manuel Molano-Mazón, Lejla Bektic, David Robbe, Alex Hyafil and Jaime de la Rocha

Reinforcement learning offers a conceptual framework of how agents can learn different tasks by promoting the actions that maximize expected reward. Tasks, however, are often based on abstract rules that classify stimuli or actions into categories inferred from the latent task structure. Although rule-guided choices can improve performance beyond that mediated by stimulus-response association, the neural substrates underlying rule learning are poorly understood. Here, we used a two-alternative auditory discrimination task that included trial blocks generated with two different statistical rules: a tendency to repeat or alternate the previous stimulus category. Rats developed a choice bias that gradually scaled with the recent history of rewarded repetitions versus alternations. To understand rat's behavior, we fitted three reinforcement learning (RL) models based on (1) action values (Left versus Right), (2) action values dependent on previous choice (e.g. coming from the left, Left-to-Left versus Left-to-Right sequences) and (3) rule values (Repetition versus Alternation). The rule RL accounted for the rat's behavior best. Moreover, we found that sequences starting on either side similarly contributed to the updating of the rule, providing a quantitative estimate of how different choices generalize to an abstract rule. To assess the role of different brain areas in rule encoding, we photo-inhibited neural activity during the inter-trial-interval and found that inactivation of the dorsal striatum or medial prefrontal cortex, but not parietal cortex, decreased the impact of the rule value on the impending choice while sparing the stimulus impact. Further, inactivation impaired the rule updating in illuminated trials affecting choices in several upcoming trials. Preliminary analysis of neural recordings in dorsomedial striatum suggests that a significant fraction of neurons encoded the rule value provided by the RL model. Our results identify an abstract response rule that can guide perceptual choices in volatile environments, and point to the frontostriatal system as a key site for reward-based updating of rule value.