Previous research has shown that humans approximate Bayes-optimal solutions in various domains. What mechanisms underlie these complex computations? This paper aims to test key predictions of one mechanistic model, which employs hierarchically-structured reinforcement learning (RL). The model proposes that RL acts at two levels: a lower-level RL loop learns values for actions in response to stimuli, creating stimulus-response mappings; a higher-level loop learns values for entire clusters of stimulus-response mappings, called task sets (TS), and determines which TS to select to guide current behavior. Crucially, this model approximates Bayes-optimal inference. We report rich evidence for this model in humans: in our novel paradigm, TS values affected learning and performance, TS reactivation, error types, context preference, and TS creation. These findings support the existence of a second, higher-level RL loop, and its importance in structuring our interactions with the world.

Flexibly adapting behavior to different contexts is a critical component of human intelligence. It requires knowledge to be structured as coherent, context-dependent action rules, or task-sets (TS). Nevertheless, inferring optimal TS is computationally complex. This paper tests the key predictions of a neurally-inspired model that employs hierarchically-structured reinforcement learning (RL) to approximate optimal inference. The model proposes that RL acts at two levels of abstraction: a higher-level RL process learns context-TS values, which guide TS selection based on the context; a lower-level process learns stimulus-actions values within TS, which guide action selection depending on stimuli. In our novel task paradigm, we found evidence that values were indeed learned at both levels in parallel. Context-TS and stimulus-action values affected learning, whereas TS values affected TS reactivation and generalization. This supports the claim of a second, higher-level RL process, and its importance in structuring our interactions with the world.

Flexibly adapting behavior to different contexts is a critical component of human intelligence. It requires knowledge to be structured as coherent, context-dependent action rules, or task-sets (TS). Nevertheless, inferring optimal TS is computationally complex. This paper tests the key predictions of a neurally-inspired model that employs hierarchically-structured reinforcement learning (RL) to approximate optimal inference. The model proposes that RL acts at two levels of abstraction: a higher-level RL process learns context-TS values, which guide TS selection based on context; a lower-level process learns stimulus-actions values within TS, which guide action selection in response to stimuli. In our novel task paradigm, we found evidence that participants indeed learned values at both levels. Not only stimulus-action values, but also context-TS values affected learning and TS reactivation, and TS values alone determined TS generalization. This supports the claim of two RL processes, and their importance in structuring our interactions with the world.