Mechanisms Of Working Memory Allocation In Reward Learning

Daniel Ehrlich
University of California Berkeley, Berkeley, California, United States

Anne Collins
UC Berkeley, Berkeley, California, United States

Abstract

Working memory (WM) is a core driver of cognition, supporting executive control, decision-making, and learning. In reward learning, WM works alongside slower reinforcement learning (RL) to establish associations between states, actions, and rewards. WM's capacity is highly limited, necessitating careful allocation of WM resources to optimize performance. How humans manage this WM constraint during reward learning, storing valuable information while discarding superseded data, remains an open question. In this study, we utilize a dynamic reward learning task to isolate rapid WM processes from slower RL mechanisms during reward learning. Through computational modeling we explore the operations humans use to allocate their limited WM resources efficiently. Our findings show strategies including (1) reward-dependent memory operations (write, forgetting, and over-write probabilities) and (2) strategic clearance policies (removing redundant and task-inconsistent data). This research clarifies WM's role in reward learning, highlighting the importance of WM operations in supporting complex behavior.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*