Assessing model-based and model-free Pavlovian-instrumental transfer using a novel two-stage paradigm

Laura Alexandra Wirth
Universität Potsdam, Potsdam, Germany

Daniel Schad
University of Potsdam, Potsdam, Germany

Abstract

Computational reinforcement learning models suggest that learning involves both model-free (MF) reward prediction errors and model-based (MB) state prediction errors, observed in instrumental and Pavlovian learning (Daw et al., 2011; Schad et al., 2020). Pavlovian-instrumental transfer (PIT) demonstrates Pavlovian values impacting instrumental responses. Single-lever PIT paradigms, often considered as MF, show correlations with reduced MB instrumental control (Garbusow et al., 2014; Review Cartoni et al., 2016; Sebold et al., 2016). To explore whether single-lever PIT effects are exclusively MF or also MB, we created a novel two-stage paradigm assessing MF and MB control trial by trial. Computational dual-control model simulations revealed a two-way interaction for MF and a three-way interaction for MB PIT. Thus far, Bayesian sequential analysis using Savage-Dickey density ratios (N=10) suggests the existence of MF (BF=3.93) and MB (BF=1.26) influences on PIT, aligning with Pavlovian learning and emphasizing the role of MB computations in single-lever PIT tasks.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*