Mean of effective  $\bar{\delta}$  upon value convergence with TD RPE Daw et al. DA/5HT integrator, optimal policy, stationary environment

mean(effective  $\bar{\delta}$ , last 100 trials) -0.2-0.4-0.6-0.8-1.00.6 0.0 0.20.4 0.8 1.0 Reward Probability

Smoothed across 30 runs