Planning in the Hippocampus: Linking Actions and Outcomes to Guide Behavior

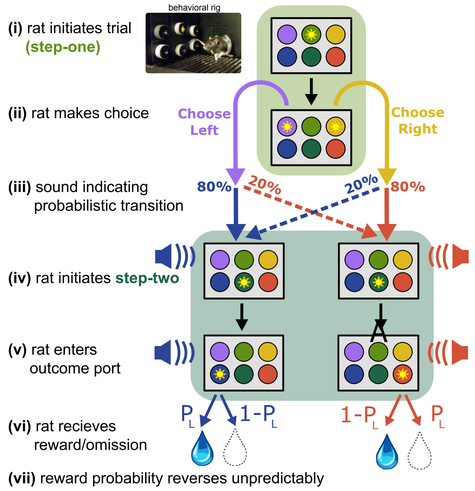
Sarah Jo Venditto[1], Kevin Miller[2,3*], Nathaniel Daw[1,2*], Carlos Brody[1,4*] * These authors contributed equally

1 Princeton Neuroscience Institute, Princeton University; 2 Google DeepMind; 3 University College London; 4 Howard Hughes Medical Institute

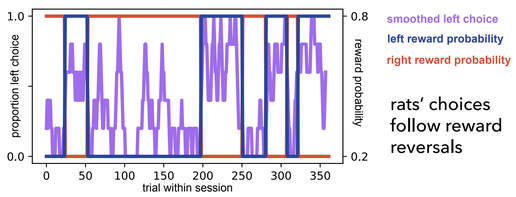
Summary

- Planning requires an **internal model** of the world that can be flexibly utilized to **link actions and subsequent consequences** across time and space
- The **hippocampus has been hypothesized to support this internal action-outcome model**, and has been causally linked to planning in both humans and rodents (Miller et al., 2017; Vikbladh et al 2019). However, *how hippocampal representations carry out this function is poorly understood*.
- We record from dorsal hippocampus while rats perform a multi-step reward-guided task that employs probabilistic transitions between actions and outcomes, the rat two-step task, which has been shown to reliably elicit planning.
- **Hippocampal activity provides distinct representations of state dependent on action**.
- In-between trials, we find state representations that oscillate between possible actions, where overall **action encoding is influenced by previous transition, reward, and upcoming action, suggestive of model-based credit assignment and planning**.

Rat two-step task



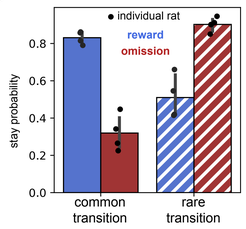
Example behavioral session



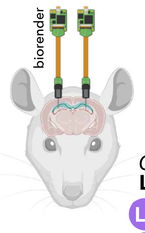
rats' choices follow reward reversals

Rats are primarily model-based

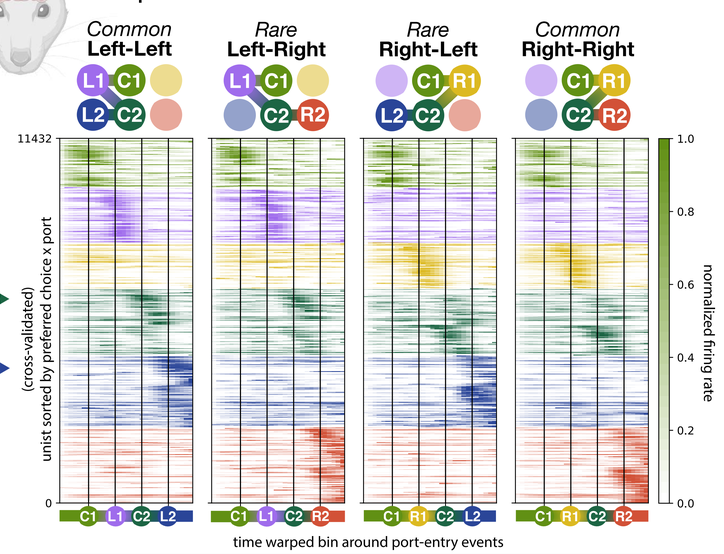
choices show a model-based influence of reward and transition from previous trial:
- repeat choices after rewarded common-transition trials and omission rare-transition trials
- switch choices after omission common-transition trials rewarded rare-transition trials

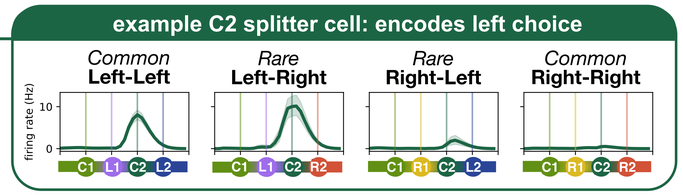


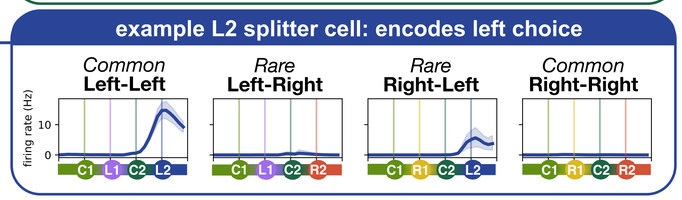
Dorsal hippocampus jointly encodes choice and task state



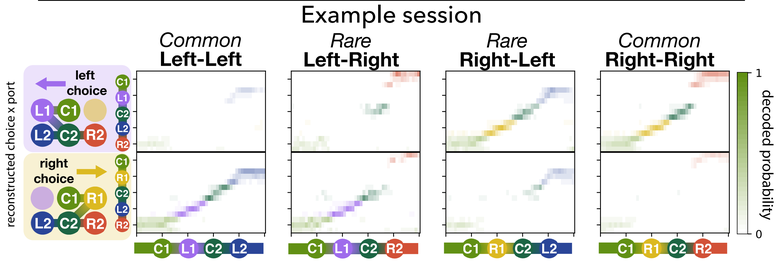
- Bilateral neuropixels 2.0 probes targeted at dorsal CA1 of hippocampus
- 4 rats, 131 sessions, 80-100+ filtered units per session



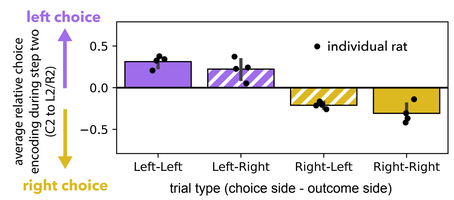




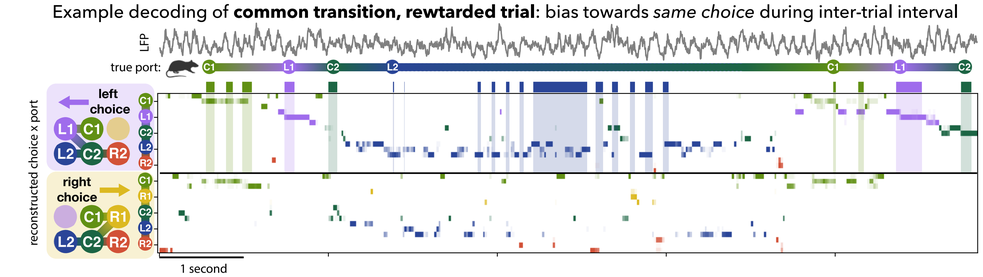
Bayesian decoding reconstructs choice x task state

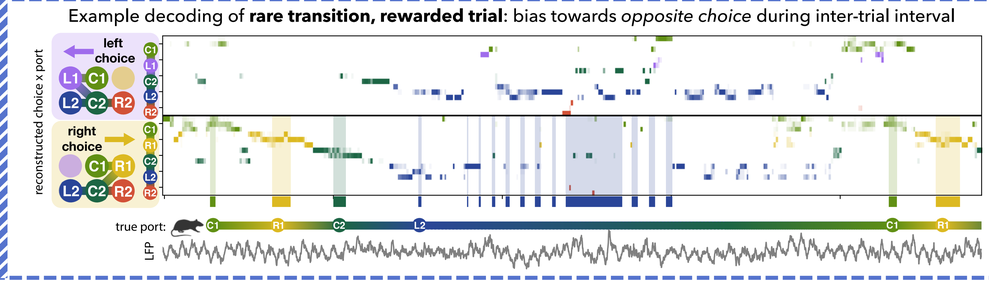


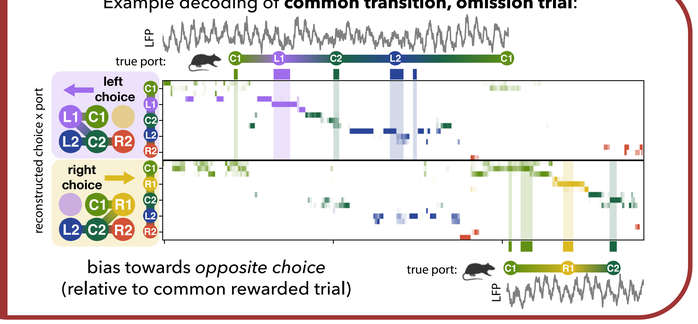
Choice can be decoded during second step task states



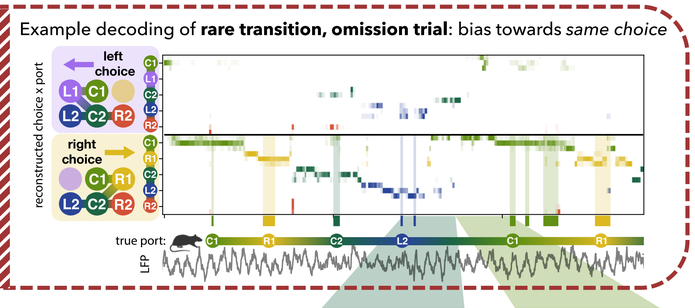
Choice encoding during the inter-trial interval supports model-based credit assignment and planning







- Representations of task state during the inter-trial interval (ITI) oscillate between posisble choices, coincident with theta oscillations in the LFP, consistent with *theta flickering* (Jezek et al. 2011) or *theta cycling* (Kay et al. 2020)



- Average choice encoding is influenced by previous reward, previous transition, and upoc-ming choice, reflective of both model-based credit assignment and choice planning.

Choice encoding biased by transition and reward