

Contrasting foraging and reinforcement learning computations in the prefrontal cortex of macaques



Zsombor Ungvárszki^{1,2,*}, C. Goussi-Denjean^{2,*}, A. Székely¹, M. di Volo^{2,°}, E. Procyk^{2,°} and Gergo Orbán^{1,°}

¹ Univ Lyon, Université Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, Bron, France
² Computational Systems Neuroscience Lab, Wigner Research Center for Physics, Budapest, Hungary
* Authors contributed equally as co-authors ° Authors jointly supervised this work



Overview

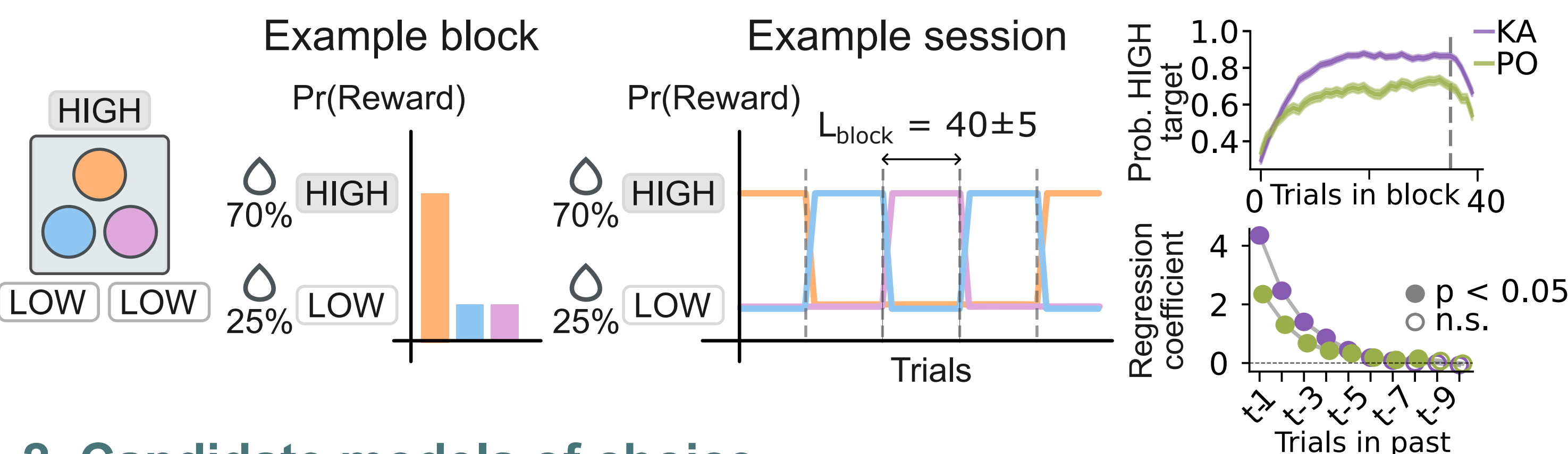
- Flexible decision-making is often formalized by reinforcement learning (RL), but foraging models provide an ecologically relevant alternative.
- We recorded neural population activity in frontal cortex while macaques performed a probabilistic bandit task, a standard paradigm used to study RL.
- Neural population analyses revealed a subspace in MCC that precisely tracked value computations.
- Dynamics in this value subspace showed signatures of foraging computation but conflicted with RL requirements.

Question

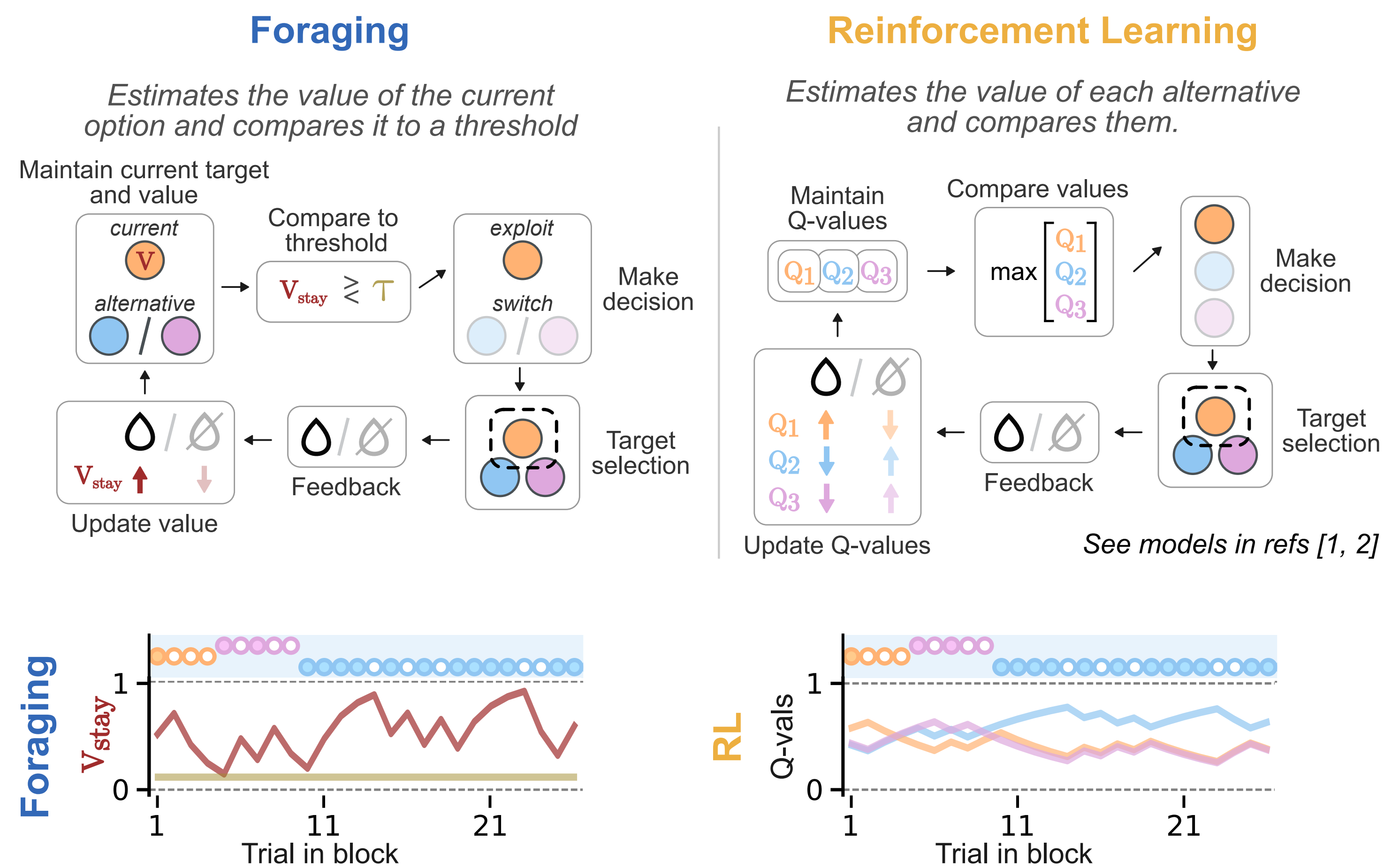
Does the brain compare all options' values, as in RL, or focus only on the current preferred choice, as in foraging models?

1. Experimental protocol

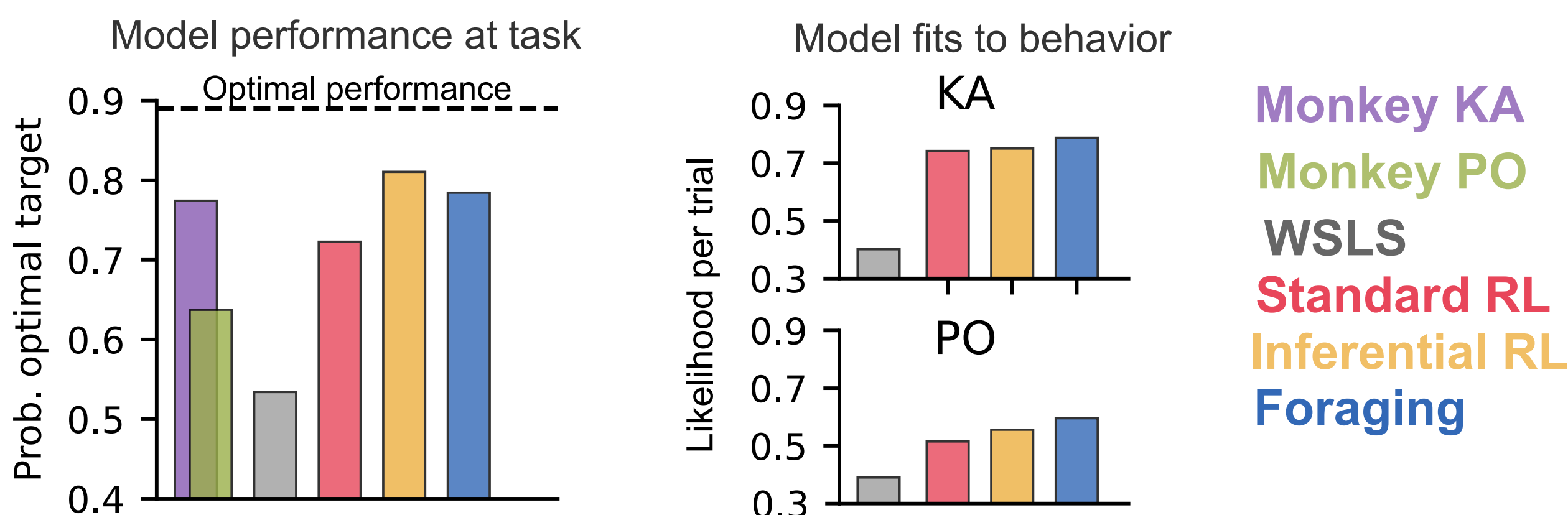
Two macaques performed a probabilistic three-armed bandit task, where one target had a high reward rate (the others low) that switched approximately every 40 trials.



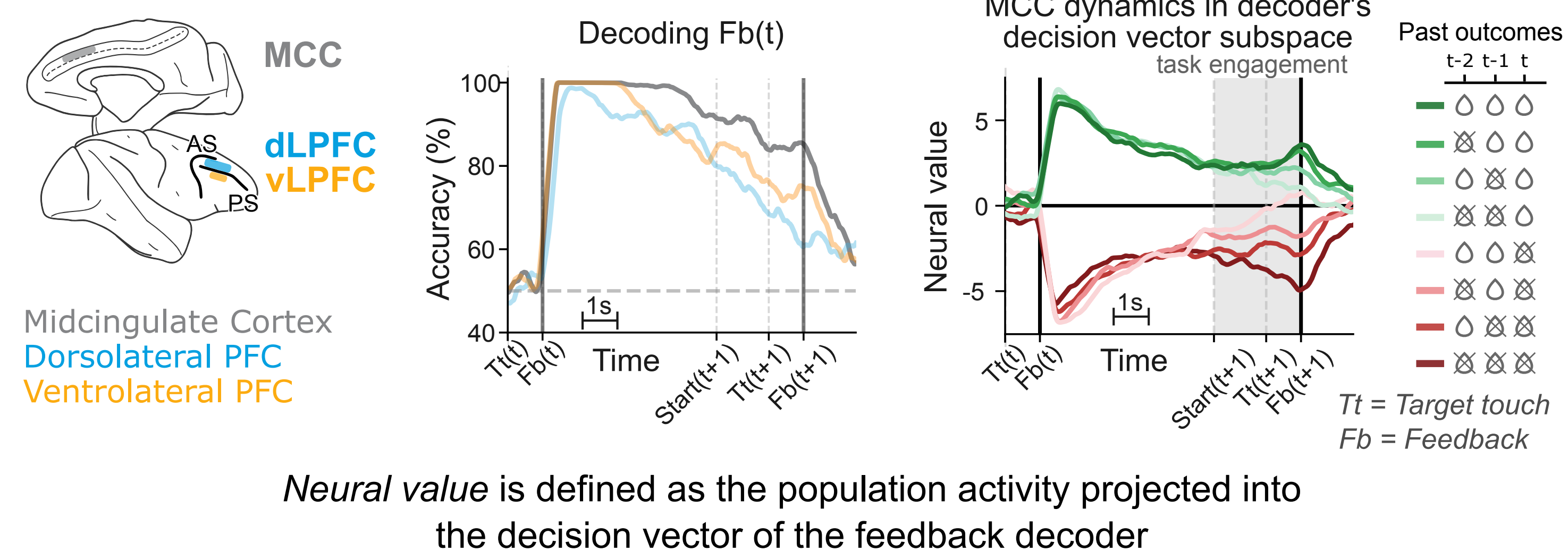
2. Candidate models of choice



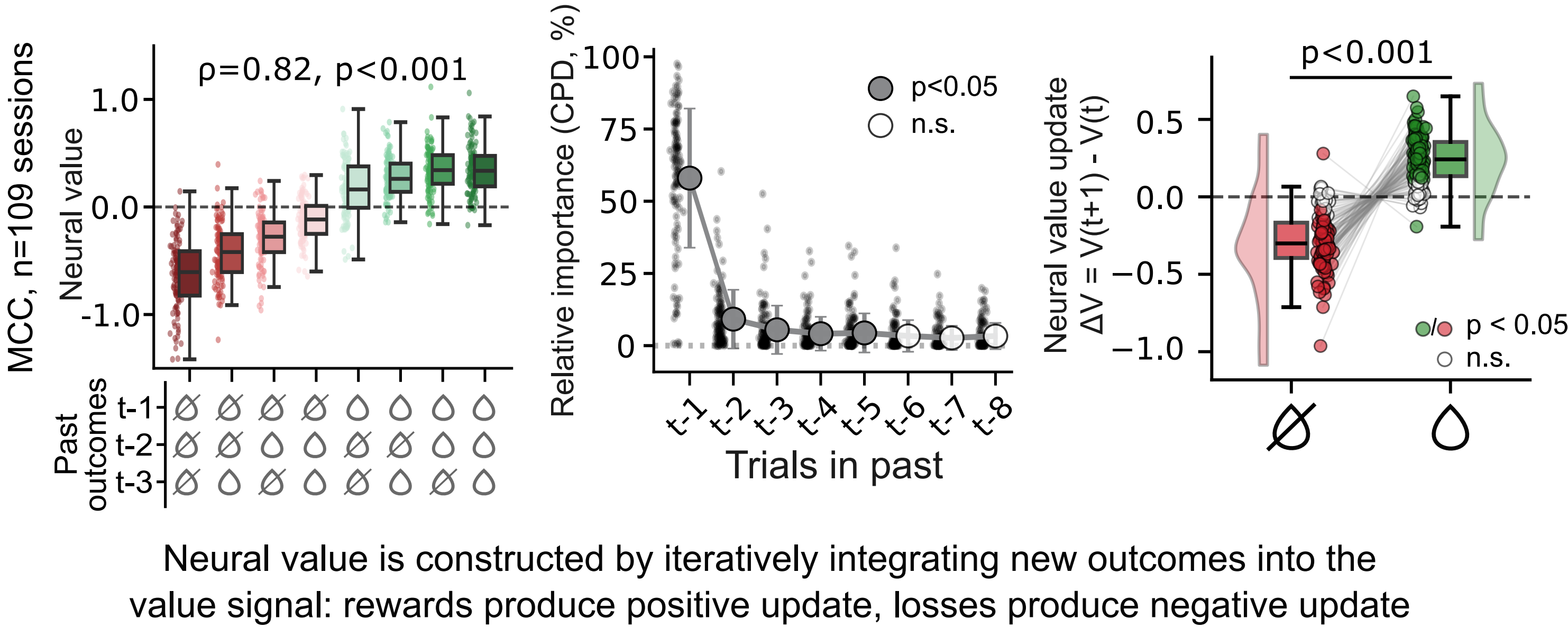
3. Both models perform close optimally, but the Foraging model explains behavior slightly better



4. Population activity in frontal cortex reveals a dedicated subspace for outcome integration

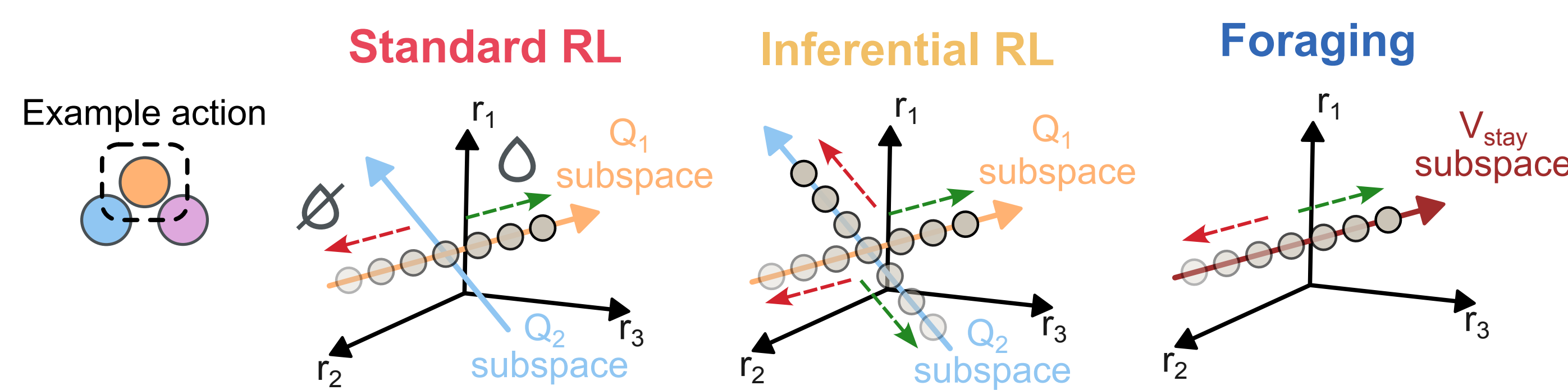


5. Neural populations update value estimates trial-by-trial using prediction-error-like learning

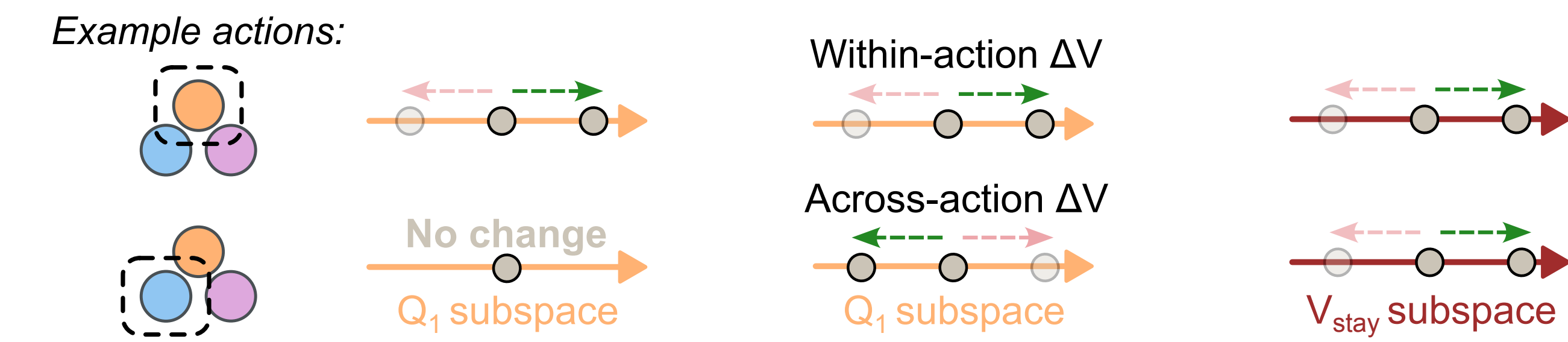


6. Neural markers of Foraging strategy

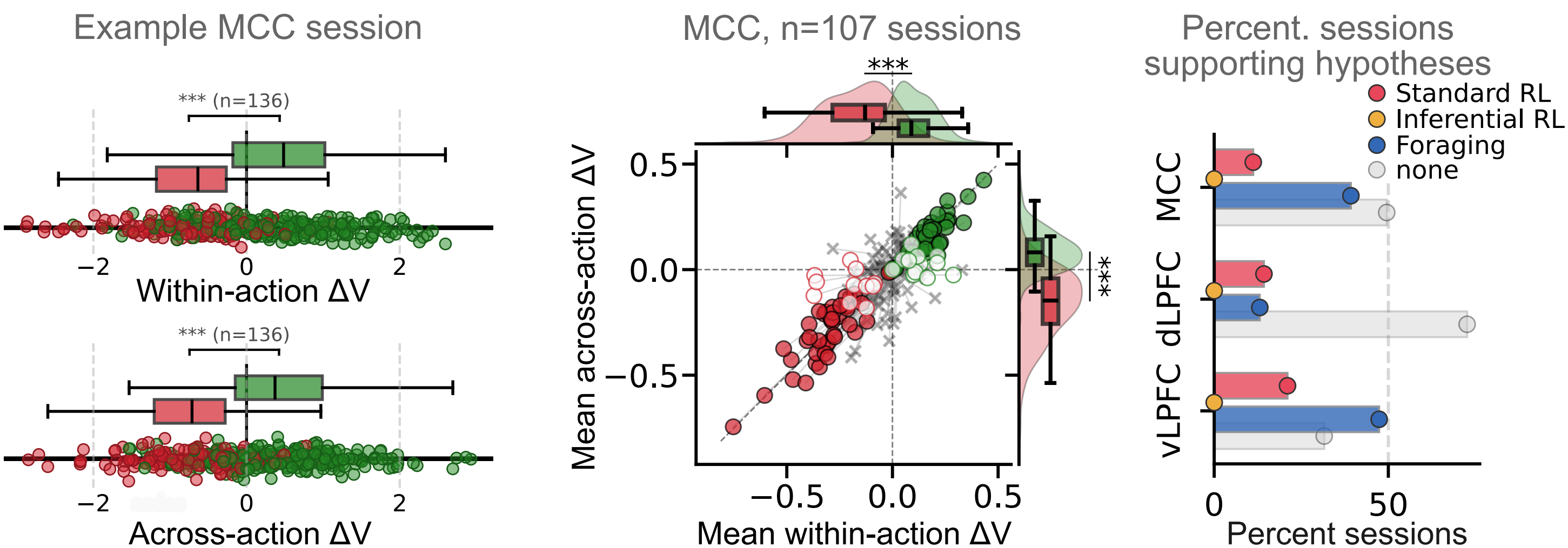
- H1 Standard RL:** Outcome updates only the action value of the chosen action
- H2 Inferential RL:** Outcome updates counterfactual action values with opposite sign
- H3 Foraging:** Outcome updates value independently of the chosen action



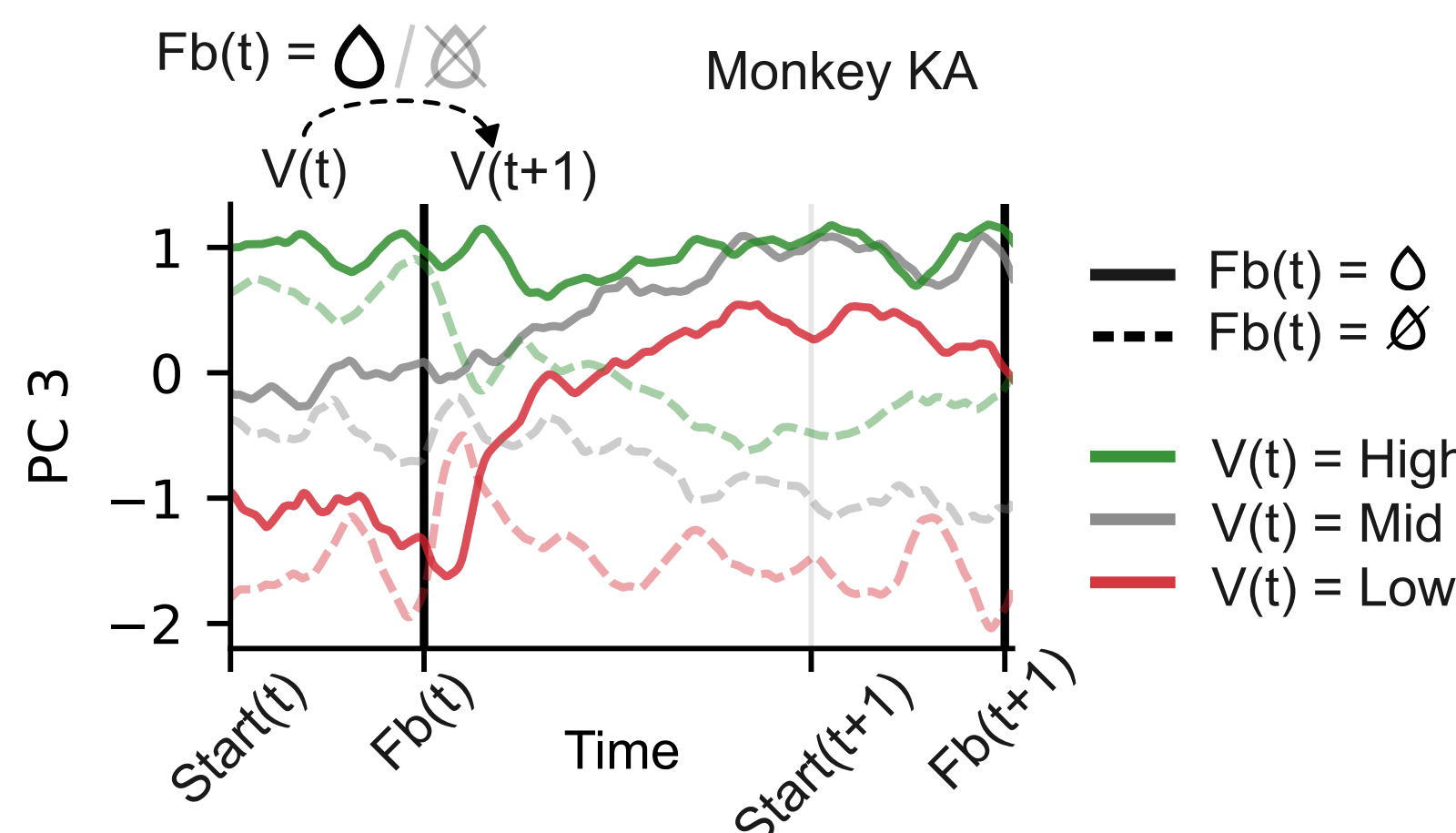
We test hypotheses by identifying the value subspace associated with one action and examining how values are updated within this subspace when a different action is taken



Across sessions, results consistently indicate an action-independent value subspace, supporting the Foraging hypothesis over RL accounts



7. Value updating unfolds smoothly over time, suggesting that MCC is involved in value computation.



Conclusion

- Monkeys use foraging (not RL) in a typical economic task
- Knowing the true strategy was key to interpreting MCC activity
- MCC encodes past outcomes in a dedicated subspace
- New outcomes are integrated into this subspace
- Strong evidence that value might actually be computed within this subspace, not just represented

Mechanistic role of MCC (proposed): MCC integrates outcomes within a dedicated neural subspace to compute a foraging value, and the activity in this subspace signals to the brain when to switch

1. Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.
2. Zid, Meriam, et al. "Humans forage for reward in reinforcement learning tasks." bioRxiv (2024).

Manuscript available at

