

---

# Uncertainty computations explain value-based foraging

Sunreeta Bhattacharya\*<sup>1</sup> and Ross Williamson<sup>2</sup>

<sup>1</sup>Carnegie Mellon University – États-Unis

<sup>2</sup>University of Pittsburgh – États-Unis

## Résumé

Efficient foraging in unfamiliar environments under time and energy constraints requires animals to rapidly learn relevant parameters such as the distribution of resources and travel times between resource patches. Animals often rely on value-based strategies informed by recent foraging history, but these strategies may be suboptimal in terms of maximizing reward. In a hidden-state foraging task where mice alternated between two distinct reward locations, Vertechi et al (1) observed that mice acquire an optimal inference-based foraging strategy early in training. We analyzed their open-source dataset (2) to find that reward accumulation rates continued to improve steadily even after this optimal strategy was established. This discrepancy between knowledge and performance suggests that task-relevant information might be suppressed, possibly in areas like the orbitofrontal cortex (3). We hypothesized that uncertainty in the value-based representation of the task could be responsible for keeping the task-representations latent and under-utilized initially, with gradual exploitation of the optimal strategy unfolding over time. To investigate this, we employed a Bayesian Q-learning framework to model state-action value functions as distributions that guide action selection in the Vertechi et al dataset (2). We examined two action-selection policies – a "greedy exploitation" policy and an "uncertainty-minimizing" exploration policy, which could explain the observed foraging behavior. Unlike the traditional explore-exploit trade-off, this model emphasizes exploration towards safer, less uncertain options, than riskier, unknown choices. Our findings indicate that in the early stages of learning, animal behavior aligns with an uncertainty-minimizing policy. As learning progresses and the uncertainty around option values decreases, the animals increasingly shift to maximizing reward by consistently selecting the option with the highest value. By tracking the Q-values and associated variances for each state-action pair, we find that the evolution of uncertainty in these value-representations closely tracks foraging performance over learning.

Vertechi, P. et al. Inference-Based Decisions in a Hidden State Foraging Task: Differential Contributions of Prefrontal Cortical Areas. *Neuron* 106, 166-176.e6 (2020).

<https://zenodo.org/record/3607558>

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron* 81, 267-279 (2014).

---

\*Intervenant