The Effects of Time Horizon and Guided Choices on Explore-Exploit Decisions in Rodents

Siyu Wang[1], Blake Gerken[2], Julia R. Wieland[2], Robert C. Wilson[1,3], and Jean-Marc Fellous[1,4,5]

1. Department of Psychology, University of Arizona **2**. Neuroscience and Cognitive Science Program, University of Arizona **3**. Cognitive Science Program, University of Arizona
4. Program in Applied Mathematics, University of Arizona **5**. Department of Biomedical Engineering, University of Arizona

Introduction

Humans and animals have to balance the need for exploring new options and exploiting known options that yield good outcomes. This tradeoff is known as the **explore-exploit dilemma**. To better understand the neural mechanisms underlying how humans and animals address the explore-exploit dilemma, a good animal behavioral model is critical. Most previous rodent explore-exploit studies used ethologically unrealistic operant boxes and reversal learning paradigms in which the decision to abandon a bad option is confounded by the need for exploring a novel option for information collection, making it difficult to separate different drives and heuristics for exploration. In this study, we investigated how rodents make explore-exploit decisions using a spatial navigation **Horizon Task** (Wilson, Geana, White, Ludvig, & Cohen, 2014) adapted to rats to address the above limitations. We compared the rats' performance to that of humans using identical measures. We showed that r**ats use prior information to effectively guide exploration**. In addition, **rats use information-driven directed exploration like humans,** but the extent to which they explore has the **opposite dependence** on time horizon than humans. Moreover, **we found that free choices and guided choices have fundamentally different influences on exploration in rodents**, a finding that has not yet been tested in humans.

Behavioral tasks

Horizon adaptive exploration in rats vs humans

Behavior over the course of a game

Accuracy by guided reward

"win-stay, lose-shift"

model-free exploratory behavior

Quantify directed vs random exploration using hierarchical Bayesian modeling

$$\Delta Q = R_{guided} - \theta - b * s_{guided} + \alpha_{LG}\Delta R_{LG} + \alpha_{LS}\Delta R_{LS} \quad (1)$$

$$p(unguided) = \frac{1}{1+e^{-\frac{\Delta Q}{\sigma}}} \quad (2)$$

As horizon increases

-Humans increase directed exploration

-Rats decrease directed exploration

-Random exploration is not modulated by horizon in this task.

Rat and Human experiment (between-subject)

Rat experiment (within-subject)

Control experiment

Shorter horizon vs higher volatility

Volatility alone does not account for the horizon adaptive exploration in rats.

Self-guided vs Cue-guided

Self-guided explo ration is treated intrinsically differently than cue-guided choices.

Rats are more accurate when self-guided.

Rats explore less-when self-guided.

Acknowledgements

We thank Blaine Harper for giving insightful feedback on the task design, Maddie Souder for giving extensive help with rat training, and Kristine Gradisher and Blaine Harper for help with the rat experiments.

Funding sources:
Undergraduate Biology Research Program at the University of Arizona
NIH R01AG061888 (Robert C. Wilson)
NSF IIS-1703340 (Jean-Marc Fellous)