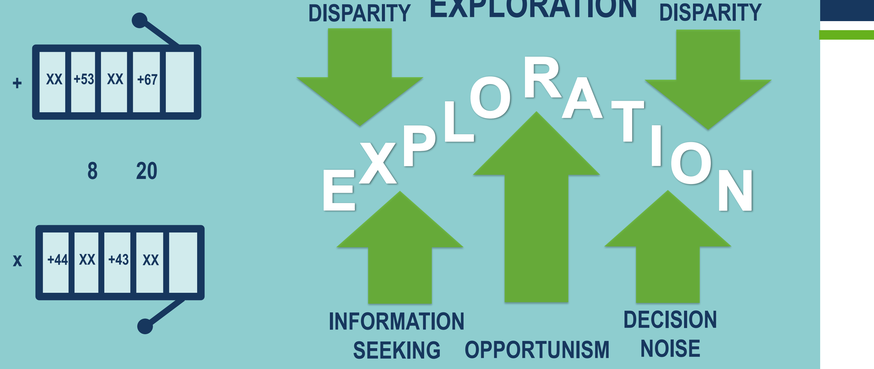
#1064

Opportunistic Exploration

Jack Dolgin[1], Bettina Bustos[2], Robert C. Wilson[3], & Wouter Kool[1]

[1] Department of Psychological & Brain Sciences, Washington University in St. Louis, [2] Department of Psychological & Brain Sciences, University of Iowa, [3] Department of Psychology, University of Arizona

GRAPHICAL ABSTRACT



INTRODUCTION

Exploration vs. Exploitation
Taking a chance on something new vs. sticking to what already works

How do humans solve this tradeoff?
Directed exploration: Deliberate prioritization
Random exploration: Stochastic sampling

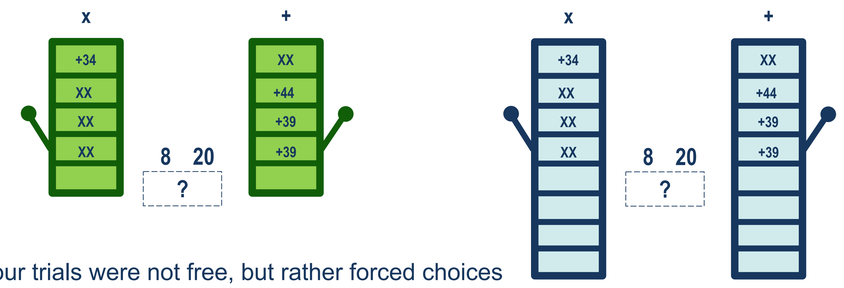
Is exploration planned at all?
Opportunities to explore abound, but with unequal costs
E.g., sports prospect call-ups during blowouts

What is the present goal?
(1) To evaluate if this tradeoff is about when and not just whether to explore and
(2) Replicate directed exploration in a dynamic setting

Support provided by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2139839.

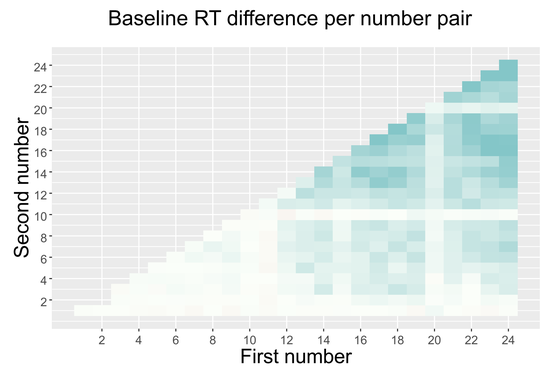
Contact info: jdolgin@wustl.edu

EXPERIMENTAL DESIGN



First four trials were not free, but rather forced choices
On remaining trials, selected a bandit by solving corresponding arithmetic problem
In above examples, typing “160” samples left bandit, typing 28 samples right bandit

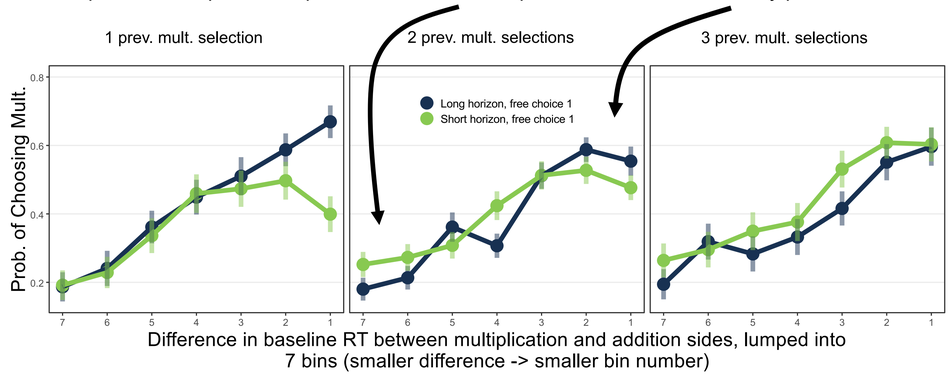
Indexing Difficulty



RESULTS

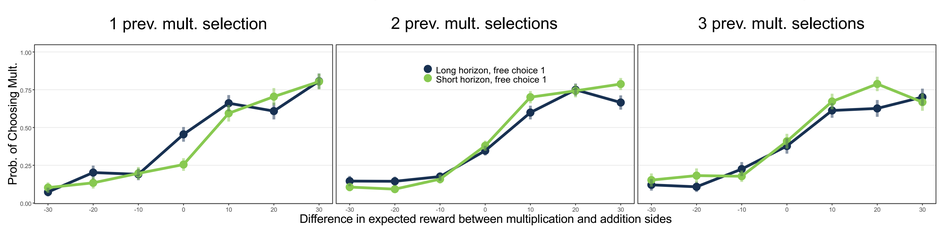
Exploiting when to Explore

People use a map of the experiment to avoid hard problems and seize on easy problems

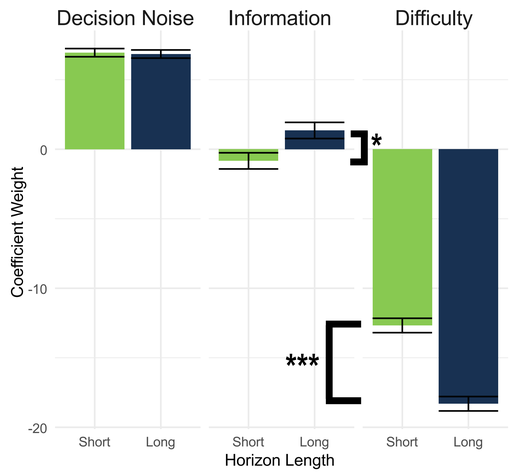


Directed Exploration

Preference for undersampled when more future trials remain, all else equal



Computational modeling reveals more sensitivity to difficulty and information when more future trials remain



$$p_{mult} = \frac{1}{1 + \exp(\frac{R_{mult} - R_{add} + \alpha(I_{mult} - I_{add}) + D(m_{add} - m_{mult}) + B}{\sqrt{2}\sigma_d})}$$

DISCUSSION

Expectations about future opportunities factor into current explore-exploit choice

Directed exploration in a dynamic bandit task

Echoes recent machine learning work on ‘cross-learning’

Future Questions

Is opportunistic exploration a form of directed exploration?

What are other considerations in an explore-exploit decision?