amii

Exploration Consistency

(Human) agent exploration in different bandit environments

People can **flexibly use different exploration strategies** depending on the environment

&

Exploration rate generalizes across tasks suggesting hierarchical control

Links

Poster

Website

Funding Thanks!

UNIVERSITY OF ALBERTA NEUROSCIENCE AND MENTAL HEALTH INSTITUTE

NSERC CRSNG

CIFAR

UNIVERSITY OF ALBERTA

Background

Rationale

- **Problem 1:** little work has examined if humans are consistent in how they explore across environments
- **Problem 2:** thus unclear if exploration generalizes across environments
- **Current experiment:** We had 30 humans complete two distinct bandits and we classified exploration strategies with RL models

Questions

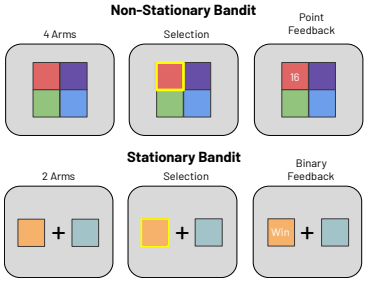
RQ 1: Do people flexibly use different exploration strategies depending on task demands?

RQ 2: Do any aspects of exploratory behaviour generalize across tasks?

Models

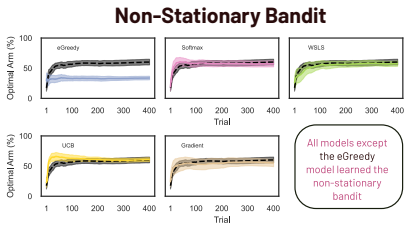
Model	Action Selection	Exploration
eGreedy	Value	Random
Softmax	Value	Probabilistic
UCB	Value and uncertainty	Directed
Win-Stay, Lose-Shift (WSLS)	Prior Trial	Heuristic
Gradient	Action Preference	Probabilistic

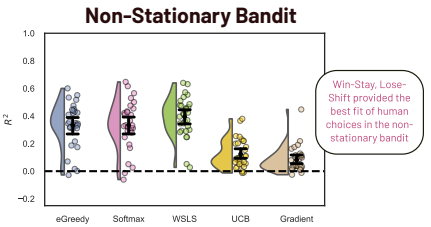
Environments

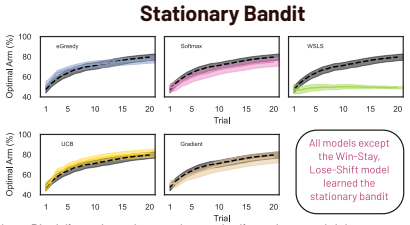


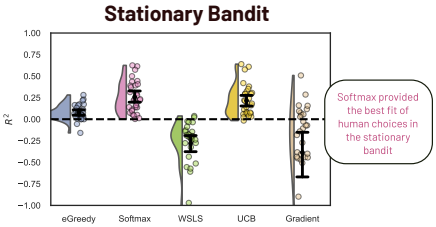
Results

RQ1: Yes! People use different exploration strategies per task demands



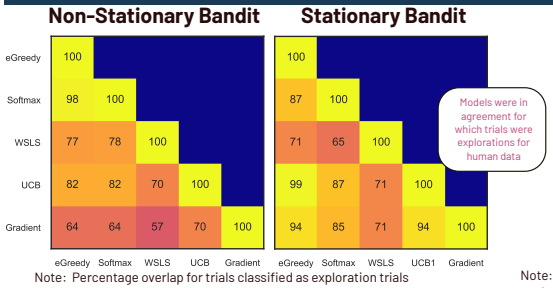




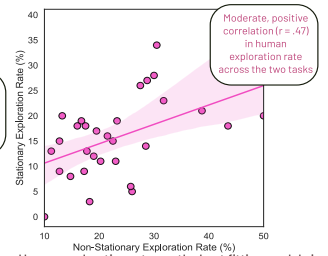


Note: Black lines shows human data, color lines show model data

RQ2: Yes! Exploration rate generalizes across bandits



Note: Percentage overlap for trials classified as exploration trials



Note: Human exploration rate per the best fitting models in each task

Thomas D. Ferguson, Alona Fyshe, Adam White (University of Alberta, AMII), Olave E. Krigolson (University of Victoria)