EEG signals sensitive to cognitive control and uncertainty differ depending on the exploration strategy used

Main Point 1: How humans explore can be shifted by the structure of the environment

Main Point 2: Different exploration strategies are related to changes in uncertainty/surprise signals and control processes in the brain

Website

Poster

Funding Thanks!

CIFAR

NSERC CRSNG

UNIVERSITY OF ALBERTA NEUROSCIENCE AND MENTAL HEALTH INSTITUTE

Background

Rationale

- Humans use different exploration strategies to learn
- **Directed Exploration:** people explore to reduce uncertainty
- **Random exploration:** people simply explore randomly due to stochasticity
- **But... it's unclear why people change how they explore and what cognitive processes help explain different strategies**

- **Experiment 1 -** Can modifying the environment structure change how people explore?

- **Experiment 2 -** If we can shift exploration strategies - can neural-cognitive differences help us understand why?

Approach

- 4 arm bandit with **10 "Risk" blocks (points can be lost)** & **10 "Safe" blocks (no points lost)**
- Before each block humans were forced to explore each arm twice (forced explore)
- Paradigm was identical for experiment 1 and 2

Research Q's

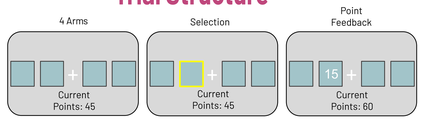
Experiment 1 (n = 55 people):
- RQ 1: Do people adopt different strategies in "risk" and "safe" blocks?

Experiment 2 (n = 49 people):
- RQ 2: Are there EEG differences between exploration strategies?

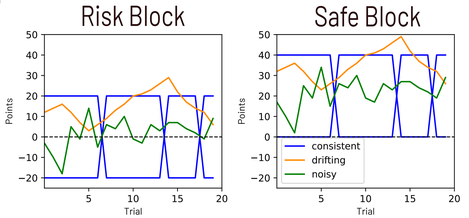
Method

Environment

Trial Structure



Reward Structure



Models

We used a series of **models adapted from reinforcement learning** to **classify humans' exploration strategies**

Model	Action Selection	Exploration
eGreedy	Value	Undirected Random
Softmax	Value	Probabilistic Random
Upper Confidence Bound (UCB)	Uncertainty, Value	Directed
Gradient	Action Preference	Probabilistic Random

EEG

In experiment 2, we collected EEG to determine if there are differences in neural signals tied to:
(1) uncertainty/surprise - P300
(2) cognitive control - Theta oscillations
...depending on the exploration strategy used

Results

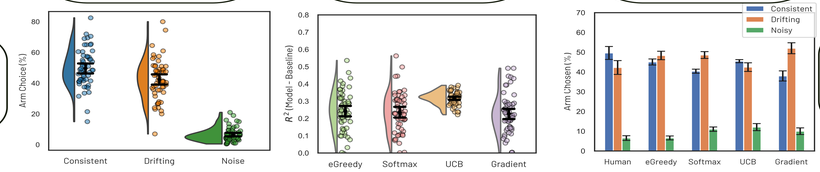
Exp 1 - People either use directed (UCB) or random (eGreedy) exploration depending on the block type

In the risk block...

Higher preference for the consistent arms (relative to safe)

The UCB model (directed exploration) provided the best fit

Only the UCB model replicated the human choices

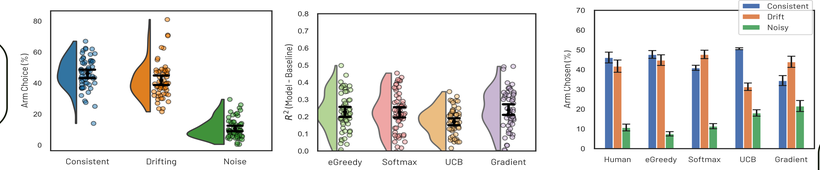


In the safe block...

Higher preference for the drifting/noisy arms (relative to risk)

The eGreedy, Softmax, Gradient models all provided good fits but...

Only the eGreedy model replicated the human choices



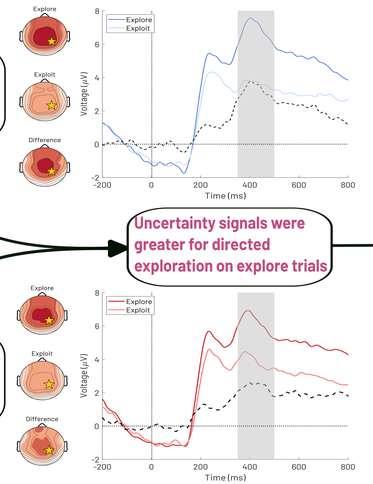
Exp 2 - The type of exploration strategy was related to changes in uncertainty and control signals

Experiment 1 was replicated
Risk Block = Directed Exploration
Safe Block = Random Exploration

In the risk block...

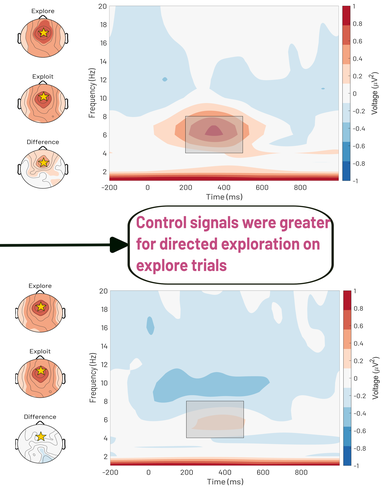
In the safe block...

Uncertainty Signals



Uncertainty signals were greater for directed exploration on explore trials

Control Signals



Control signals were greater for directed exploration on explore trials

Note: 0 ms is presentation of feedback, models were used to classify trials as explorations, gray bar is where uncertainty signal (the P300) was measured

Note: 0 ms is feedback, models were used to classify trials as explorations, black box is where control signal (theta; 4 to 8 Hz) was measured, data show difference wave

Thomas D. Ferguson, Alona Fyshe, Adam White (University of Alberta, Alberta Machine Intelligence Institute)
contact: tfergus2@ualberta.ca