

Changing how we explore:

Investigating the cognitive and neural basis of different exploration strategies

How humans explore can be shifted by the structure of the environment

The shift might be due to changes to attention, uncertainty detection, and control processes in the brain

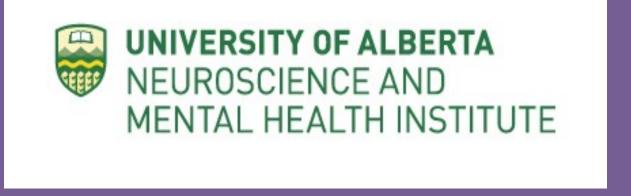
Poster

Website





Funding Thanks!







Background

Rationale

Experiment 1: - Humans use different exploration strategies - but why?

- RQ 1: Can shifting the reward distributions change how people explore within a task?

Experiment 2:

- If we can shift exploration strategies can neural-cognitive differences explain the mechanism of why?
- RQ 2: Are there EEG differences between exploration strategies?

Environment

Points: 45

Model

Models

Value eGreedy Random Probabilistic Softmax Value Value and uncertainty Directed Probabilistic Gradient Action Preference

Action

Selection

Exploration

Points: 60

Trial Structure - Learning Feedback Current Current Current Points: 45

Manipulations

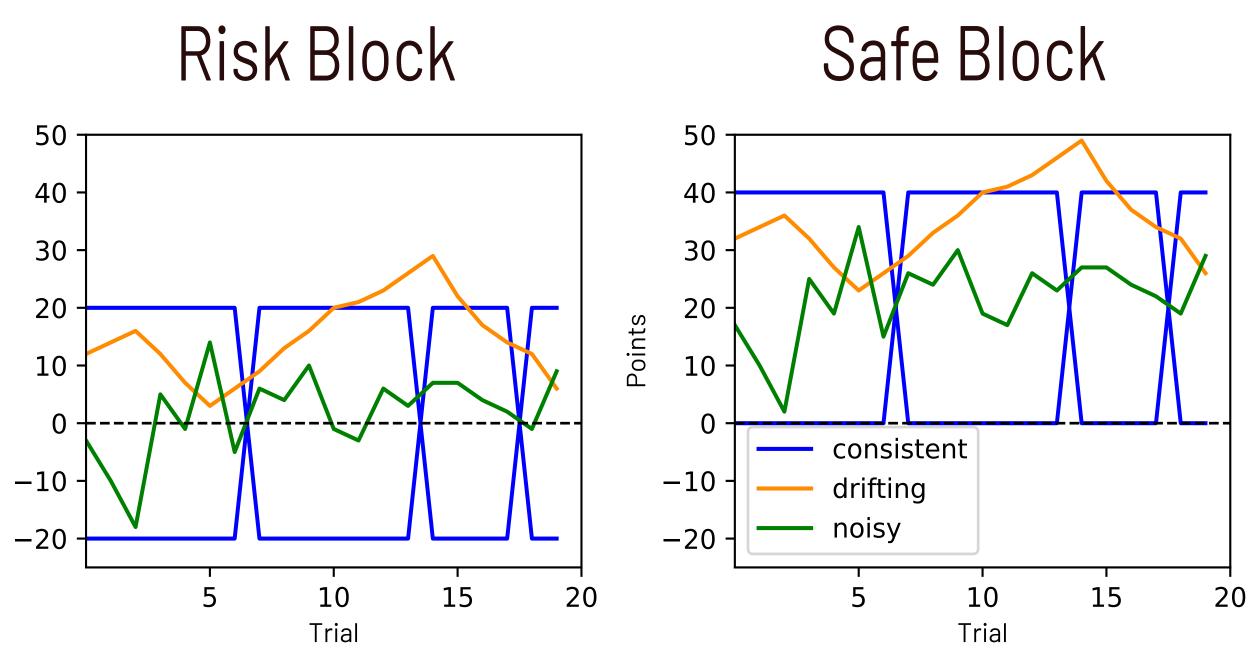
Experiment 1 (n = 57 people):

- 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)

Experiment 2 (n = 49 people):

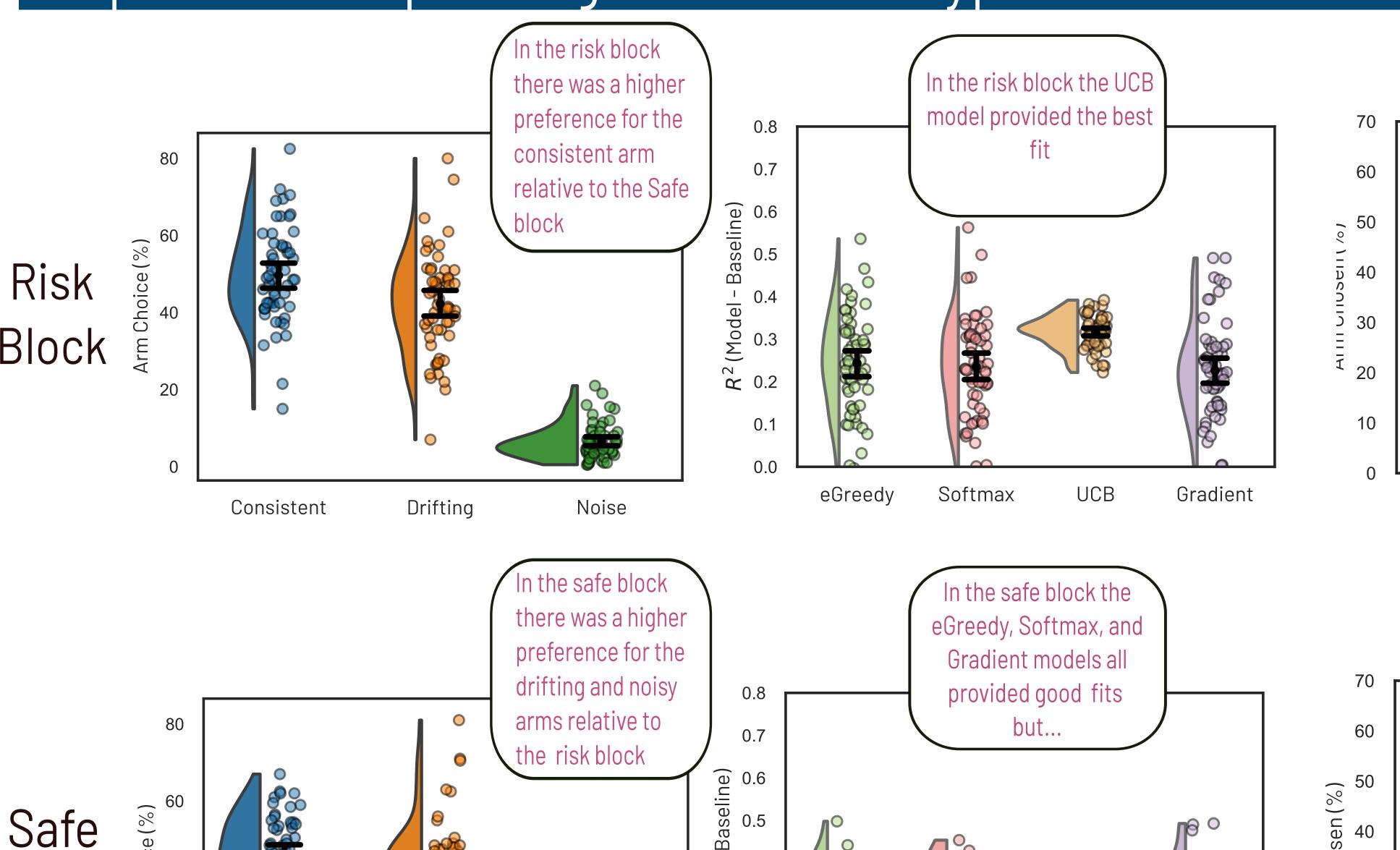
- Same as experiment 1 but we collected EEG to determine if there are differences in neural signals tied to: (1) attention, (2) uncertainty detection, and (3) cognitive control

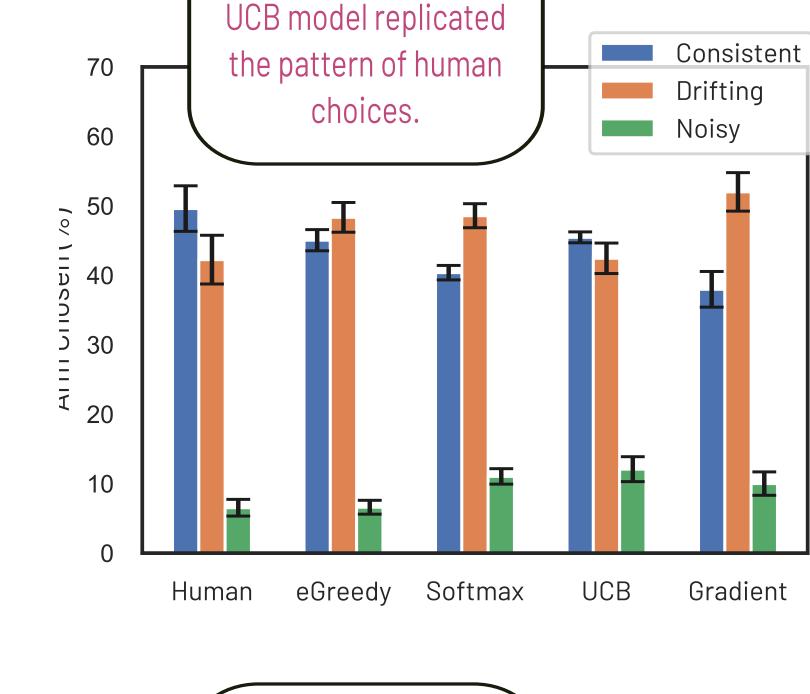
Reward Structure - Learning



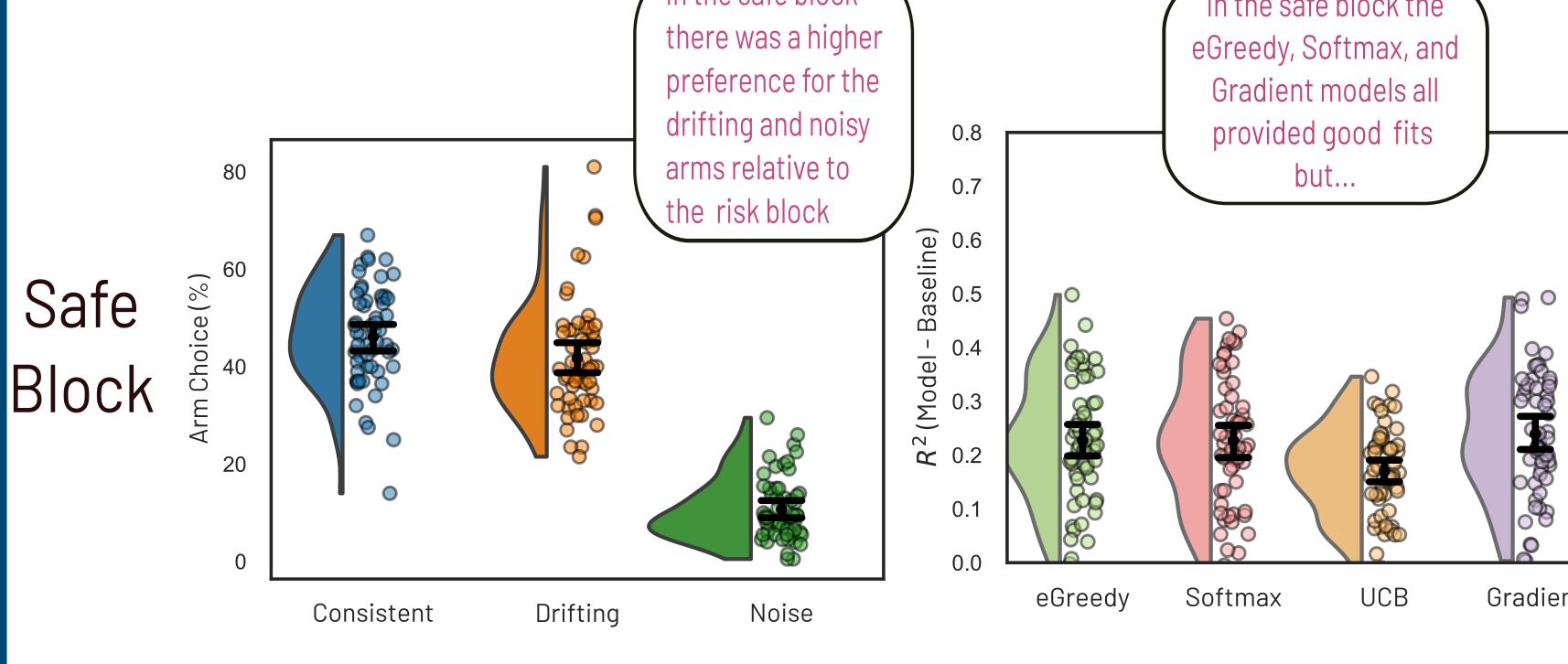
Results

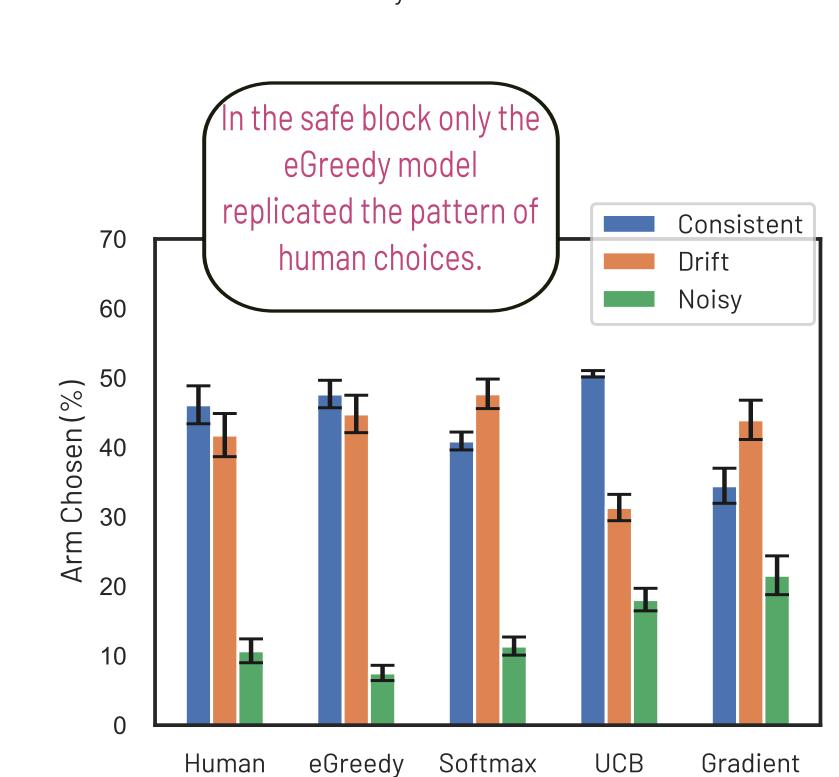
Exp 1 - People use directed (UCB) and random (eGreedy) exploration depending on the block type





In the risk block only the





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

Uncertainty Detection Control Signals Experiment 1 replicated Risk Block = Directed Exploration Risk Risk Safe Block = Random Exploration **Attention Signals** During forced exploration, attention signals were greater for directed exploration Time (ms) Time (ms) During learning, control signals During learning, uncertainty were greater for directed detection was greater for directed exploration exploration on explore trials Safe Safe Explore Exploit Time (ms)

Note: 0 ms is presentation of feedback, gray bar is where attention signal (the LPP) was measured, yellow star is the electrode used

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

Time (ms)

Time (ms) Note: 0 ms is presentation of feedback, models were used to classify trials as exploration, black box is where

control signal (theta; 4 to 7 Hz) was measured

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