Should we explore new rewards, or exploit an old ones? It might depend on happiness.

Real-life rarely hands us a 'menu', so decisions are serial

Classic economic theory assumes we can directly compare options on a 'menu', but this is not always straightforward:

- Go on a second date…or see if someone better comes along?
- Eat soup for the 5th day in a row…or go to a new restaurant?
- Watch one more episode…or try that meditation app?

Quick Thought Experiment

Let's imagine you are given 2 hours on a mystical island, where money grows on trees in abundance

Each time you shake a tree, gold coins fall out!

But if you keep shaking the same tree, the number of coins decreases…

Exploring a new tree will involve a time cost

The 'richness' of each tree is unknown before sampling, and once you leave a tree, it disappears forever

We call this a 'foraging' task

Exploration/Exploitation Dilemma

You have to choose between 'exploiting' the same tree, or 'exploring' a different one

As money in your tree depletes, the prospect of 'exploring' becomes more attractive

How do you know when to switch?

Disaster Strikes!

A tornado sweeps the island, and many trees' coins are lost in the ocean

You notice that the average reward rate on the island has decreased

How will you change your behaviour?

Marginal Value Theorem

The optimal choice rule is given by the Marginal Value Theorem (McNamara & Houston, 1985):

Leave the current tree when its reward rate drops below the estimated environmental average

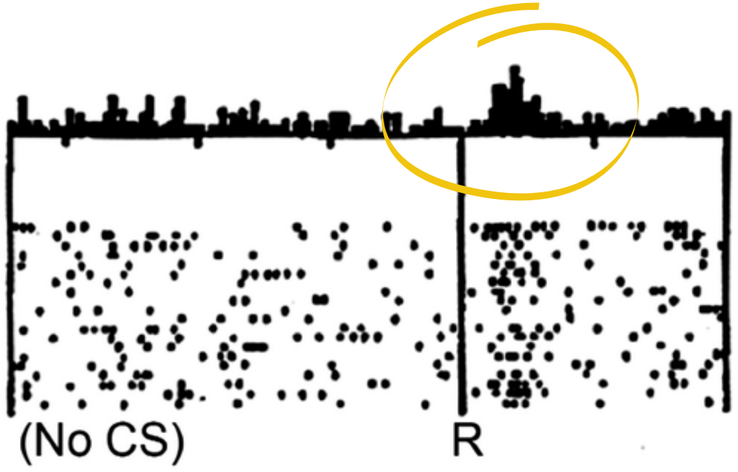
After the storm hits, we should explore less

Mood as a representation of reward momentum

How do we keep track of the average reward rate?

When a new tree is better than expected, we get a positive prediction error

Happiness has been hypothesised to represent a running average of recent reward prediction errors, it signals environmental richness (Eldar *et al.*, 2016; Villano *et al.*, 2018)

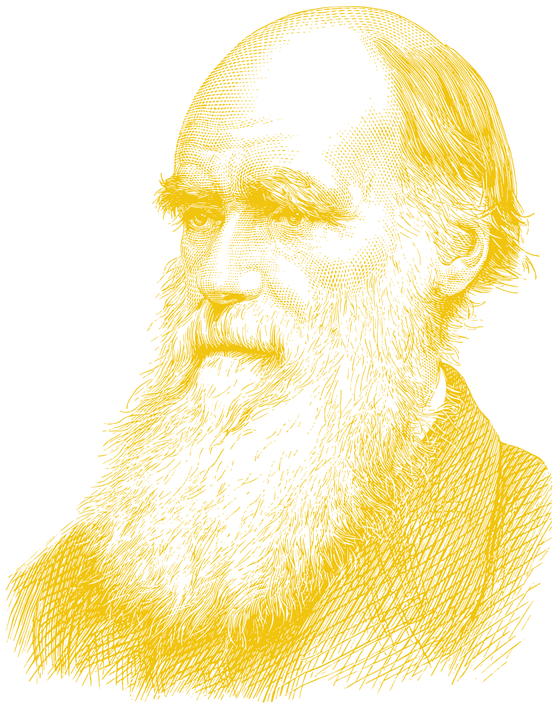


Evolutionary Perspectives

If happiness is correlated with reward abundance, then increased exploration is adaptive (Eldar et al., 2021)

For a forager, it is more efficient to update beliefs about the availability fruit in *all* trees when spring arrives, not just those sampled

In a lab-based foraging experiment, participants were less likely to settle for the current option when the overall reward environment was richer (Garrett & Daw, 2020)



Emotional Bias

If happiness is unrelated to environmental richness, happier agents will be biased towards over-exploration (Eldar et al., 2021)

Foragers entering into a new environment carry over behavioural bias from the previous environment

Effect is stronger when moving to a worse average reward rate (Garrett & Daw, 2020)

This Asymmetric belief updating or 'optimism bias' is widely observed across learning domains (Sharot *et al.*, 2011)

Depression may discourage exploration

Huys *et al.* (2015) have argued that many depressive symptoms link to a pessimistic estimate of the environmental reward rate

We can predict that depressed foragers will be biased towards exploitation

Evidence that the optimism bias is reduced in depressed patients, but this has not yet been tested in foraging tasks (Garrett et al, 2014; Korn et al, 2014)

Initial Research Questions

Depression may discourage exploration

Does experienced happiness arising from environmental richness increase exploratory foraging behaviour?

- The link between mood, environment quality and exploration has not been tested experimentally

Are depressive symptoms associated with an exploitation bias in foraging tasks?

- Novel application of established 'pessimistic avg. reward rate estimation' theory to foraging

Is there an interaction with happiness?

Experimental Design Sketch

Study 1

Participants

Online data collection, facilitating access to a wider subject pool with depressive symptoms

Participants will fill out several self-report questionnaires assessing anhedonic symptoms: Beck Depression Inventory-II (BDI-II; Beck et al., 1988), Snaith Hamilton Pleasure Scale (SHPS;Snaith et al., 1995), Positive and Negative Affect Schedule (PANAS; Watson et al., 1988), Mood and Anxiety Symptom Questionnaire (MASQ;Watson et al., 1995), and Perceived Stress Scale (PSS; Cohen et al., 1983).

Patch Leaving

Participants will perform a patch leaving task:

- Maximise reward in a limited amount of time when sources of reward clump together in patches and the value of a patch decreases with harvesting (Stephens and Krebs, 1986)

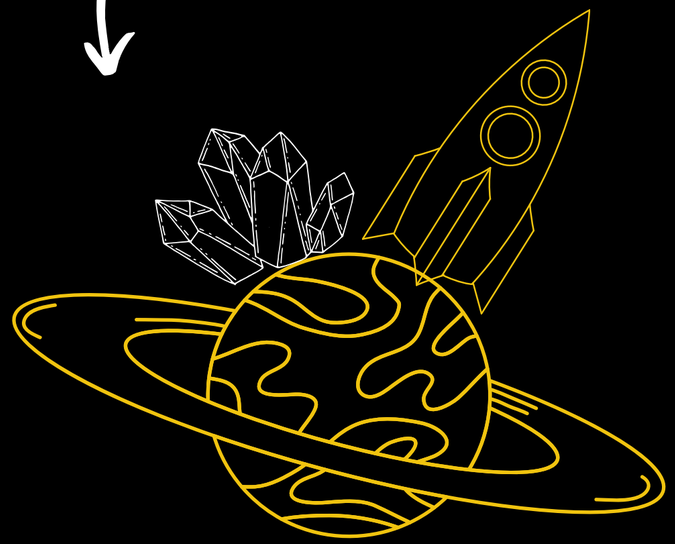
Decisions will be incentivised:

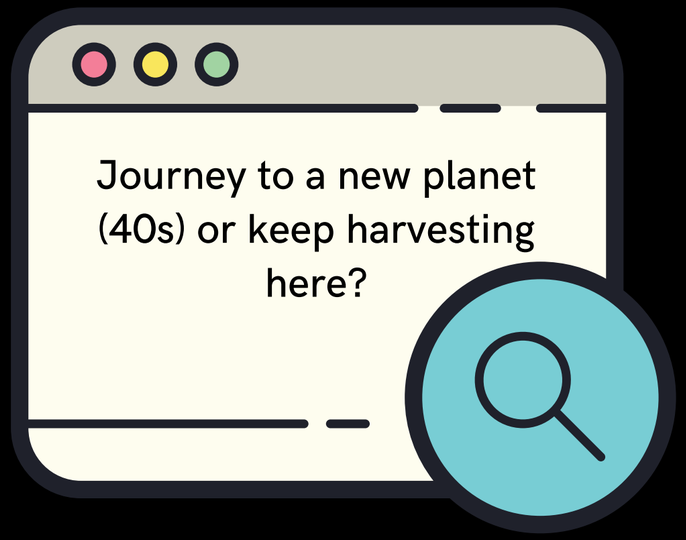
- Participants know from the outset that a subset of trials will be randomly selected and converted into real earnings

10:00

Total Gems: 8

Gems collected this round





Environmental Richness

Reward environment in each experimental block will either be Rich or Poor

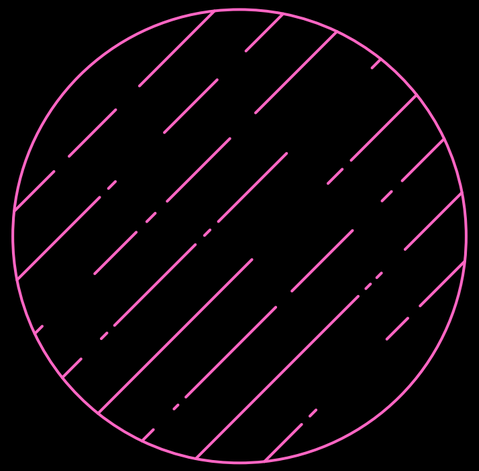
Maximum reward capacity of patches will be normally distributed in both environments, centred around a higher mean in the rich environment

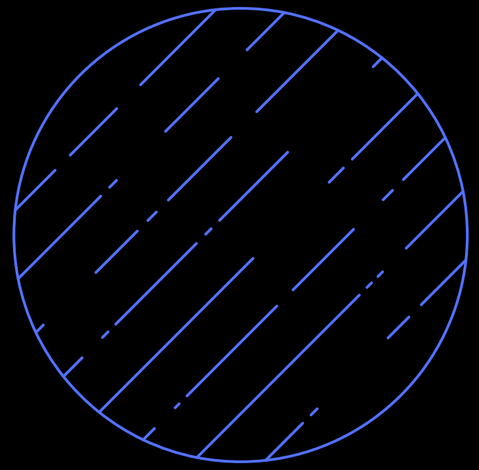
Participants will be exposed to two counterbalanced sets of Rich and Poor environments:

Allows for within-subject analysis

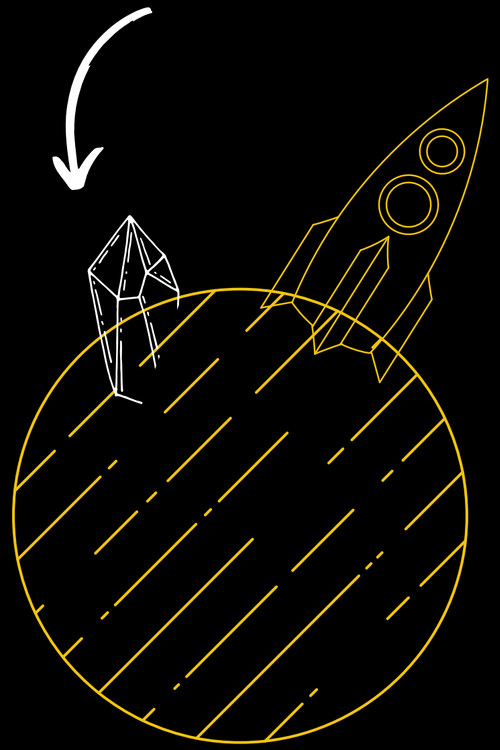
10:00

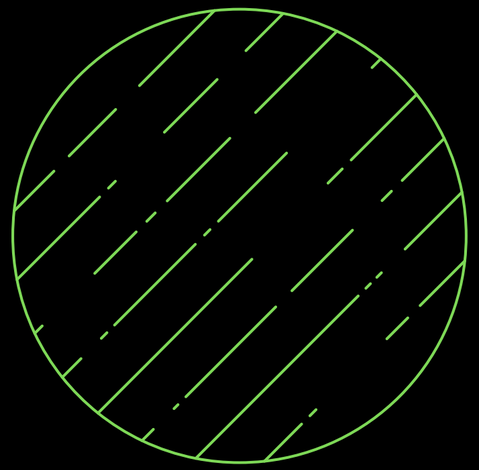
Total Gems: 2





Gems are less abundant





Mood Sampling

Subjective mood will be sampled at regular intervals using a visual analogue rating scale ranging from ‘very negative’ (0) to ‘very positive’ (100)

Hypotheses

Participants will be happier in the RICH environment (H1.1) and mood will be positively correlated with propensity to explore (H1.2)

Self-reported depressive symptoms will be associated with reduced effect of environment on exploration (H2.1) and mood modulation (H2.2)

We will also test whether the optimistic learning bias Garrett & Daw (2020) observe disappears, or even reverses, with high depressive symptoms

Analyses

We will use a hierarchical linear mixed-effects model (maximum likelihood estimation method), to account for between-and within-subject effects

P(explore) is the dependent variable, all fixed effects of interest and their interactions will be included:

- Patch quality
- Environment (avg. reward rate at time *t*)
- Mood
- Block order
- Depressive symptom level