Luis Luarte. DCCS. UDD. 2025-10-05.

Project title. Computational isomorphism in decision-making: a generalized mechanism for foraging in physical and abstract spaces.

Introduction.

Problem relevance. Decision-making under uncertainty is a fundamental cognitive process underlying a vast range of human and animal behaviors. A central component of this process is the exploration-exploitation dilemma. Exploitation involves leveraging current knowledge to maximize immediate returns, whereas exploration consists of gathering new information to improve future outcomes. Neural mechanisms that modulate spatial foraging (searching for resources in a physical environment) may have been evolutionary co-opted to govern search in abstract cognitive spaces, such as memory or problem spaces (Hills et al., 2008, 2012, 2015). This implies the existence of a generalized search algorithm, suggesting that the computational architecture managing the exploration-exploitation balance is isomorphic across different decision-making domains. However, direct empirical evidence demonstrating this isomorphism is limited. Establishing this connection would not only unify out understanding of cognition but could also have implications for understanding disorders characterized by atypical decision-making.

Research question. Is the computational architecture that governs the balance between exploration and exploitation in abstract foraging tasks isomorphic to the architecture that governs foraging in physical spaces?

General and specific objectives.

General Objective. To determine if the computational architecture regulating the exploration-exploitation balance is isomorphic between the domains of physical and abstract foraging.

Specific objectives. (1) to evaluate whether spontaneous behavioral variability (SBV), operationalized as the coefficient of variation (CV) of saccadic eye movements (SEM), is a function of environmental uncertainty, as measured by the reward prediction error (RPE). (2) To quantify the relationship between SBV and the exploration parameter (τ) in a decision model, independently for the physical and abstract domains. (3) To determine the existence of a significant within-subject correlation between the exploration parameters $(\tau_{physical}, \tau_{abstract})$ derived from both domains.

Contribution. This work seeks to provide direct empirical evidence for a generalized search mechanism in humans. By linking a neurophysiological variable (Feedback-related negativity (FRN), as a proxy for RPE), a motor behavior variable ($CV_{saccades}$), and computational parameters of decisions models (τ), this research offers a bridge between different levels of analysis. IF the isomorphism is confirmed, it would suggest that an individual's propensity for exploitation is a stable, domain-independent trait. A conclusion with profound implications for cognitive psychology, neuroeconomics, and computational psychiatry.

Theoretical framework.

The conceptual framework of this study is grounded in reinforcement learning (RL). RL models posit that an agent learns to estimate the value of different states or actions ($\mathrm{Qvalues}$) through experience, updating these estimates based on the RPE (δ), which is the difference between the obtained ($\mathrm{R_{actual}}$) and the predicted ($Q_{predicted}$) reward. The magnitude of this update is regulated by the learning rate (α), while the choice between actions is often modeled with a softmax function, where the temperature parameter (τ) controls the balance between exploitation (low τ) and exploration (high τ) (Sutton & Barto, 2020).

The central hypothesis is inspired by the proposal that goal-directed cognition is an evolutionary exaptation of spatial foraging (Hills et al., 2008). This idea is supported by the neurobiological evidence implicating dopaminergic circuits of the dorsolateral striatum (DLS) in both the sequencing of motor behaviors and the modulation of attention and decision-making. Dopamine fluctuations in the DLS structure spontaneous behavior even in the absence of explicit rewards, acting as a continuous teaching signal that reinforces the use of specific behavioral modules and increases the variability of behavioral sequences in the short term (Markowitz et al., 2023). We propose that this modulation of behavioral variability is the core of the generalized search mechanism.

Proposed mechanism. We postulate a causal mechanism: environmental uncertainty $H(states)$, that is, variability in environment state, is encoded by the RPE (δ). This RPE, associated with phasic dopamine activity and measurable through the FRN, modulates the activity of a central "variability controller". The output of this controller is manifested as variability is spontaneous behavior (measurable via $CV_{saccades}$), which, in turn, adjusts the exploration parameter τ in the RL models that describe the agent's decisions.

From this mechanism, three hierarchical hypotheses are derived:

1. *Behavioral variability is a function of uncertainty*. The activity of the variability controller is a function of environmental uncertainty. Specifically, a larger absolute RPE, indicative of greater uncertainty (as good predictions are

harder to come by), will cause an increase in spontaneous behavioral variability.

$$CV_{saccades} \sim f(|RPE|)$$

2. *Behavioral variability modulates exploration*. The levels of the variability controller, as measured by $CV_{saccades}$, modulate the exploration parameter τ in both domains. This proposes that the system's global "stochasticity setting" is the primary regulator of the local exploration-exploitation trade-off.

$$\tau_{abstract} \propto \tau_{physical} \propto CV_{saccades}$$

3. *Isomorphism of computational architecture*. Given that the same "stochasticity setting" parameter is applied to both tasks, the exploration parameters, $\tau_{physical}$ and $\tau_{abstract}$, will be significantly correlated withing subjects. This implies that a participant's propensity for exploration is a stable, domain-independent variable.

$$\rho(\tau_{physical}, \tau_{abstract}) > 0$$

Methods.

Experimental design description (see figures 2, 3). A 2x2 within-subjects design will be implemented, where each participant will perform tasks in both domains (factor 1: domain = {physical, abstract}) and under both uncertainty conditions (factor 2: RPE condition = {high, low}). Concurrent measurements of behavior (choices), eye-tracking (saccades), and electroencephalography (EEG) to record the FRN will be performed.

Physical foraging task: adapted from (Hills et al., 2008), participants will explore a "virtual forest" on a screen using their gaze. To "visit" a tree, they must fixate on it for 50 ms. The reward is binary (a red dot or nothing). The RPE condition will be manipulated through the spatial distribution of the rewards, controlled by a parameter β is a gaussian weighting kernel function. Therefore, two RPE conditions will be generated through the spatial distributions of rewards (this will determine how "easy" or "hard" is to generate a prediction), all estimation were with 60 trials (see figure 2 for details in what composes the trial).

- Low RPE (predictable environment): rewards are spatially clustered.
- Hight RPE (unpredictable environment): rewards are spatially diffuse.

Abstract foraging task. Adapted from (Stojić et al., 2020), this is a two-armed bandit task where participants choose between two abstract symbols. The RPE condition is manipulated through the rate of change of rewards probabilities, all estimation were with 60 trials(see figure 3 for details in what composes the trial).

- Low RPE (stable environment): reward probabilities change slowly.
- High RPE (volatile environment): reward probabilities change rapidly.

Protocol. Participants will receive general instructions for both games (see appendix 1). They will the complete the block for one task (either physical or abstract, latin-square balanced), followed by the blocks for the other. Within each task, the order of the RPE conditions will also be randomized. A 5-minutes break will be included between the two tasks, masked as equipment recalibration, to mitigate the spillover effects.

Task parametrization fit. Before the main experiment, a pilot study with a small group of participants will be conducted to ensure the chosen task parameters effectively induce the desired subjective experience of predictability (low RPE) and unpredictability (high RPE). After completing each task, pilot participants will undergo a structured questionnaire (see appendix 2). For the physical task, key parameters such as the hit box size for gaze detection and the time required to trigger a reward will be calibrated. Participants will be asked to rate how difficult it was to predict the location of rewards in both the clumpy (low RPE) and diffuse (high RPE) scenarios. For the abstract task, parameters like trial length will be adjusted. Participants will similarly be asked to rate the difficulty of predicting rewards in the stable (low RPE) and volatile (high RPE) conditions. The specific values for parameters such as β and ϕ (defined in the section below) are initial estimates and are subject to refinement based on the results of this pilot phase to maximize the effectiveness of the experimental manipulation.

Handling of biases and artifacts. For spillover effects, the randomized task order and the rest period are designed to minimize the carryover of specific strategies from one task to the next. Regarding social desirability and demand characteristics, instructions are framed in terms of performance maximization (“win as many points as possible”), without mentioning concepts like exploration or the hypothesized link between the tasks. The monetary incentive will be given a fictitious currency that is related to real reward, though is relation function is not specified, thus, mitigating effect of “having already too much reward” or socio-economic status incidence on reward evaluation.

Formulation of tasks and models.

1. Physical foraging task: the environment consists of Ntrees at positions $p_1, \dots, p_N$. The reward likelihood $L_i$for tree i is generated from a spatially correlated Gaussian process, conditioned on the rewards of all other trees L_{-i}:

$$L_i \mid L - i \sim \mathcal{N}\left(\beta \sum j \neq i\, w_{ij} L_j, 1\right)$$

Where the spatial correlation is determined by a Gaussian weighting kernel function:

$$w_{ij} = exp\left(-\frac{\left\|p_i - p_j\right\|^2}{2\lambda^2}\right)$$

The parameter β controls the strength of the spatial clustering, thereby manipulating the RPE condition.

The agent's choice is modeled based on a Q-value for moving from tree i to tree j:

$$Q_{i,j} = U_j - C_{i,j}$$

Where U_j is the expected utility of tree j and $C_{i,j} = |p_i - p_j|$ is the Euclidean distance cost. The probability of choosing tree j is given by the softmax function:

$$\pi_{j|i} = \frac{e^{Q_{i,j}/\tau_{physical}}}{\sum_{k \in A_i} e^{Q_{i,k}/\tau_{physical}}}$$

Where A_i is the set of available (unvisited) trees.

2. Abstract foraging task: the reward probability $p_{j,t}$, for each arm j at trial t follows a random walk in logit space. The logit of the probability is defined as $l_{j,t} = ln\left(p_{j,t}/\left(1 - p_{j,t}\right)\right)$. The process evolves according to:

$$l_{j,t+1} = l_{j,t} + \delta_{j,t+1}$$

$$\delta j, t + 1 = \phi \cdot \delta_{j,t} + \epsilon_j, where\ \epsilon_j \sim \mathcal{N}\left(0, \sigma_p^2\right)$$

The parameter ϕ controls the volatility of the environment, this manipulating the RPE condition. The reward $R_{j,t}$ is Bernoulli trial: $R_{j,t} \sim Bernoulli\left(p_{j,t}\right)$.

The agent learns using a standard Q-learning algorithm. The RPE at trial t for chosen action a_t is:

$$\delta_t = R_t - Q_t(a_t)$$

The Q-value is updated via:

$$Q_{t+1}(a_t) = Q_t(a_t) + \alpha \cdot \delta_t$$

The probability of choosing action a is given by the softmax function:

$$\pi_a = \frac{e^{Q_a/\tau_{abstract}}}{\sum_{k \in A} e^{Q_k/\tau_{abstract}}}$$

Operationalization of variables of interest.

- Independent variables:
  - Domain: categorical variable $\{Physical, Abstract\}$
  - RPE condition: categorical variable $\{High, Low\}$, manipulated via the parameters β and ϕ as formally defined above.
- Dependent variables:

- Exploration parameters $(\tau_{physical}, \tau_{abstract})$: the temperature parameters from the softmax functions for the physical and abstract tasks, respectively, as defined in the mathematical formulation section. These will be estimated by fitting the models to each participant's choice data.
- SBV $(CV_{saccades})$: measured during the ITI and calculated as the coefficient of variation of saccadic amplitudes $CV_{saccades} = \frac{\sigma_{saccades}}{\mu_{saccades}}$
- RPE:
  - Computational RPE: the trial-by-trial value of δ_t derived from the fitted Q-learning model for the abstract task.
  - Neural RPE: the amplitude of the feedback-related negativity (FRN) component of the ERP, measured from EEG data in a 200-350 ms. Window post-feedback.

Sample selection/recruitment and power analysis.

The power analysis for Hypothesis 3 was performed via Monte Carlo simulation with S iterations to find the minimum sample size N that achieves a target power $1 - \beta_{power}$ (set to 0.8 in this case) for a significance level α (see figure 1). We first fitted a linear mixed model on pilot data (Stojić et al., 2020), where

$$\widehat{\tau_{abstract}} = \beta_0 + \beta_1 \cdot CV_{saccades} + Covariates + (1 \mid participant)$$

With scaled variables, the target true correlation was estimated:

$$\rho_{true} \leftarrow \widehat{\beta_1}$$

And estimated measurement noise:

$$\sigma_{noise} \leftarrow \sigma(\epsilon)$$

For each sample N with 10.000 iteration each, N pairs of the latent exploration parameters were drawn from a multivariate normal distribution.

$$\left(\tau_{\text{physical}}, \tau_{\text{abstract}}\right)_i \sim \mathcal{N}\left(\mu = (00), \Sigma = \begin{pmatrix} 1 & \rho_{\text{true}} \\ \rho_{\text{true}} & 1 \end{pmatrix}\right) \quad \text{for } i = 1, \dots, N$$

Then, measurement noise is added to simulate an experimental observation.

$$\tau_{obs_physical,i} = \tau_{physical} + \epsilon_i, where\ \epsilon_i \sim \mathcal{N}(0, \sigma_{noise}^2)$$
$$\tau_{obs\ abstract,i} = \tau_{abstract} + \epsilon_i$$

Simulated observation is then used to compute the Pearson correlation r_s and its corresponding p-value p_s for the set of observed pairs $\{(\tau_{obsphysical,i}, \tau_{\{obsabstract,i)})\}_{i=1}^{N}$. Then, the statistical power is $Power(N) = \sum p_s < \alpha / S$, and the determined sample size is the minimum number of participants required for $Power \geq 0.8$. Based on this procedure a $N = 24$ was obtained, however, it should be taken into consideration that this was just a proxy for hypothesis 3 using only one domain of the task.

Participant recruitment. Participants will be recruited from the university's student and staff population. Recruitment will be conducted through two primary channels:

1. Flyers posted on campus notice boards and in common areas.
2. An online research participation sign-up system managed by the faculty.

All participants will provide written informed consent before the start of the experiment and will be compensated for their time, in addition to the performance-based monetary bonus.

Inclusion and exclusion criteria. To ensure data quality and sample homogeneity, inclusion criteria will be age between 18 and 35 years (assuming campus mean age, however, older participants are technically withing acceptance criteria), normal or corrected-to-normal vision, as verified by self-report. This is essential for the effective calibration and use of eye-tracking equipment. For exclusion, self-reported history of any neurological or psychiatric disorders, current use of any psychoactive medication or other medication known to affect the central nervous system or cognitive function.

Statistical models for hypothesis testing.

Linear mixed-effects models (LMMs) will be used to test the hypothesis, accounting for the within subject design. Let i denote the participant index. For hypothesis 1, this relationship will be tested using an LMM. The RPE can be operationalized either computationally $(\delta_{i,t})$ or neurally $(FRN_{i,t})$.

$$CV_{i,t} = \beta_0 + \beta_1 |RPE_{i,t}| + \beta_2 D_{i,t} + \beta_3 (|RPE_{i,t}| \times D_{i,t}) + (b_{01} + b_{1i} |RPE_{i,t}|) + \epsilon_{i,t}$$

Where $CV_{i,t}$ is the saccadic CV for participant i at trial t, D is a dummy variable for the task domain, b_{0i} and b_{1i} are random intercepts and slopes for each participant, and $\epsilon_{i,t}$ is the error term. A significant main effect of $|RPE|$ $(\beta_1 \neq 0)$ will support this hypothesis.

Hypothesis 2 will be tested by modeling the estimated τ parameters as a function of the average $CV_{saccades}$ for each participant in each domain. The data will be structured in a long format.

$$\tau_{i,d} = \gamma_0 + \gamma_1 \overline{CV}_{i,d} + \gamma_2 D_{i,d} + \gamma_3 (\overline{CV}_{i,d} \times D_{i,d}) + \left(g_{0i} + g_{1i\overline{CV}_{i,d}}\right) + \eta_{i,d}$$

Where $\tau_{i,d}$ is the estimated exploration parameter for participant i in domain d, and $\overline{CV}_{i,d}$ is their average saccadic CV in that domain. A significant main effect of $\overline{CV}_{i,d}$ will support this hypothesis.

Hypothesis 3 will be tested with a one-tailed Pearson correlation test on the paired, subject-level parameter estimates $\{\tau_{physical,i}, \tau_{abstract,i}\}_{i=1}^{N}$. A significant positive correlation will support the hypothesis of isomorphism.

Mediation analysis for RPE operationalization. To provide empirical grounds for selection the primary RPE measure (computational vs. neural) for the main analyses, a formal mediation analysis will be conducted. This analysis will test whether the effect of the experimental manipulation of uncertainty (X: RPE condition) on behavioral variability (Y: $CV_{saccades}$) is mediated be the RPE signal (M: either computational RPE or FRN amplitude). Two separate multilevel mediation models will be fitted, one for each proposed mediator:

1. Model 1: mediator $M = FRN_{amplitude}$
2. Model 2: mediator $M = |Computational\ RPE|$

For each model, the following paths will be estimated using mixed-effect regressions:

- Path a: the effect of the RPE condition on the mediator. $$M_{i,t} = a_0 + a_1 \cdot RPE_{Condition_{i,t}} + (u_{0i}) + \epsilon_{i,t}$$
- Path b and c': the effects of the mediator and RPE condition on the outcome, simultaneously.

$$\mathrm{CV}_{\mathrm{saccades},i,t} = c_0' + c_1' \cdot RPE_{Condition_{i,t}} + b_1 \cdot M_{i,t} + (v_{0i}) + f_{i,t}$$

The significance of the indirect effect $(a_1 \times b_1)$ will be assessed using bootstrapping (FRN or computational) that demonstrates a stronger and more significant mediation effect (i.e., account for a larger proportion of the total effect of the experimental condition on behavior) will be selected as the primary measure of RPE for the subsequent hypothesis tests.

Appendix I. Participant instructions.

“Welcome, and thank you for participating in this study. In this session, you will play two different short games. The goal in both games is to use the information you’re given to make choices and earn as many points as possible. We will give you detailed instructions for each right before it starts.

Game 1: the forest game. In this game, you will explore a virtual forest using just your eyes. Your goal will be to find hidden rewards by looking (and keeping your gaze within a tree for brief period) at different trees. When a tree is selected it will shake and show or not a point, after this the tree will remain faded to remind you that the tree was already shaken. Shaken trees will never provide points. You will have multiple instances of 15 seconds followed by a fixation cross.

Before starting the game. In this part of the experiment, you will play a simple game where your goal is to search for rewards in a virtual forest. The entire task is controlled by your eyes, so please try to remain still and comfortable. Your objective is to find as many rewards as possible in each round. Here’s how it works: (1) look to choose: you will see a forest of trees on the screen. To check a tree for a reward, simply look directly at it. (2) Shake the tree: if you hold your gaze on a tree for a moment, you will se the tree shake. This means you’ve successfully checked it. (3) Find the reward: after the tree shakes, it will either reveal a red dot (a point) or nothing. (4) Memory help: once you have checked a tree, it will fade slightly. This is a visual cue to help you remember where you've already looked. Please try to find as many rewards as you can. Let the experimenter know if you have any questions.

Game 2: *the symbols game*. In this game, you'll use a joystick to choose between two abstract symbols. Your goal is to learn which symbol is currently the best choice to win points.

Before starting the game. In this part of the experiment, you will play a choice game. Your goal is to figure out which of two symbols is the most rewarding and win as many points as you can. This task is controlled with a joystick. You win a point by choosing the symbol that is currently the 'best' option. Here's how it works: (1) See the choices: at the start of each round, you will see two different symbols appear on the screen. (2) Make your choice: use the joystick to move left or right to select the symbol you believe will give you a point. Please make your decision as quickly and accurately as you can while the symbols are visible. (3) Get feedback: after you make a choice, you will see feedback where the symbol was. A '+1' symbol means you won a point, and a '+0' means you did not win a point for that round. Please try to win as many points as possible. Let the experimenter know if you have any questions.

General in both tasks. For each point you earn in both games, you will earn 1 Zenny. Based on the total Zennys you win, you will be rewarded with a monetary bonus in CLP and the end of the session, in addition to $5.000 for transportation.

Appendix II. Questionnaires.

B.1 Questionnaire for The Forest Game (Physical Foraging Task)

Instructions: Please answer the following questions about your experience playing the Forest Game. Use the scale from 1 to 5 to indicate your rating.

1. Please rate how difficult it was to trigger a reward from a tree.

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

2. Please rate the difficulty of focusing your gaze on a tree to check it.

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

3. Please rate the likelihood of triggering a tree reward by accident (without intentionally looking at it).

(1) Very Unlikely

(2) Unlikely

(3) Neutral

(4) Likely

(5) Very Likely

4. Please rate the perceived duration of the reward animation.

(1) Very Short

(2) Short

(3) Neutral

(4) Long

(5) Very Long

5. In the scenarios (clustered/sparse, unknown to participant marked as either A or B), how difficult was it to predict where other rewards would be?

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

6. In the scenarios (the other one), how difficult was it to predict where other rewards would be?

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

B.2 Questionnaire for The Symbols Game (Abstract Foraging Task)

Instructions: Please answer the following questions about your experience playing the Symbols Game. Use the scale from 1 to 5 to indicate your rating.

1. Please rate the perceived length of each round (trial).

(1) Very Short

(2) Short

(3) Neutral

(4) Long

(5) Very Long

2. In the scenarios (either A or B scenario for different RPE), how difficult was it to predict which symbol would give a point?

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

3. In the scenarios (the other one) how difficult was it to predict which symbol would give a point?

(1) Very Easy

(2) Easy

(3) Neutral

(4) Difficult

(5) Very Difficult

Figures.

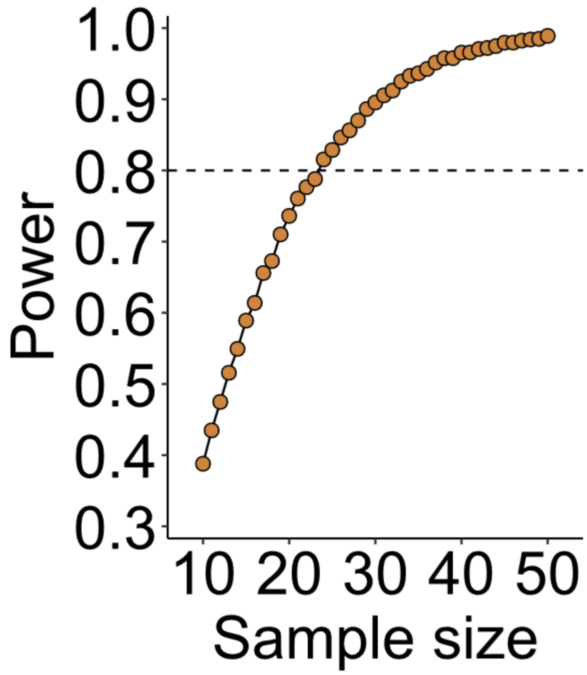

Figure 1Power calculation for different sample sizes

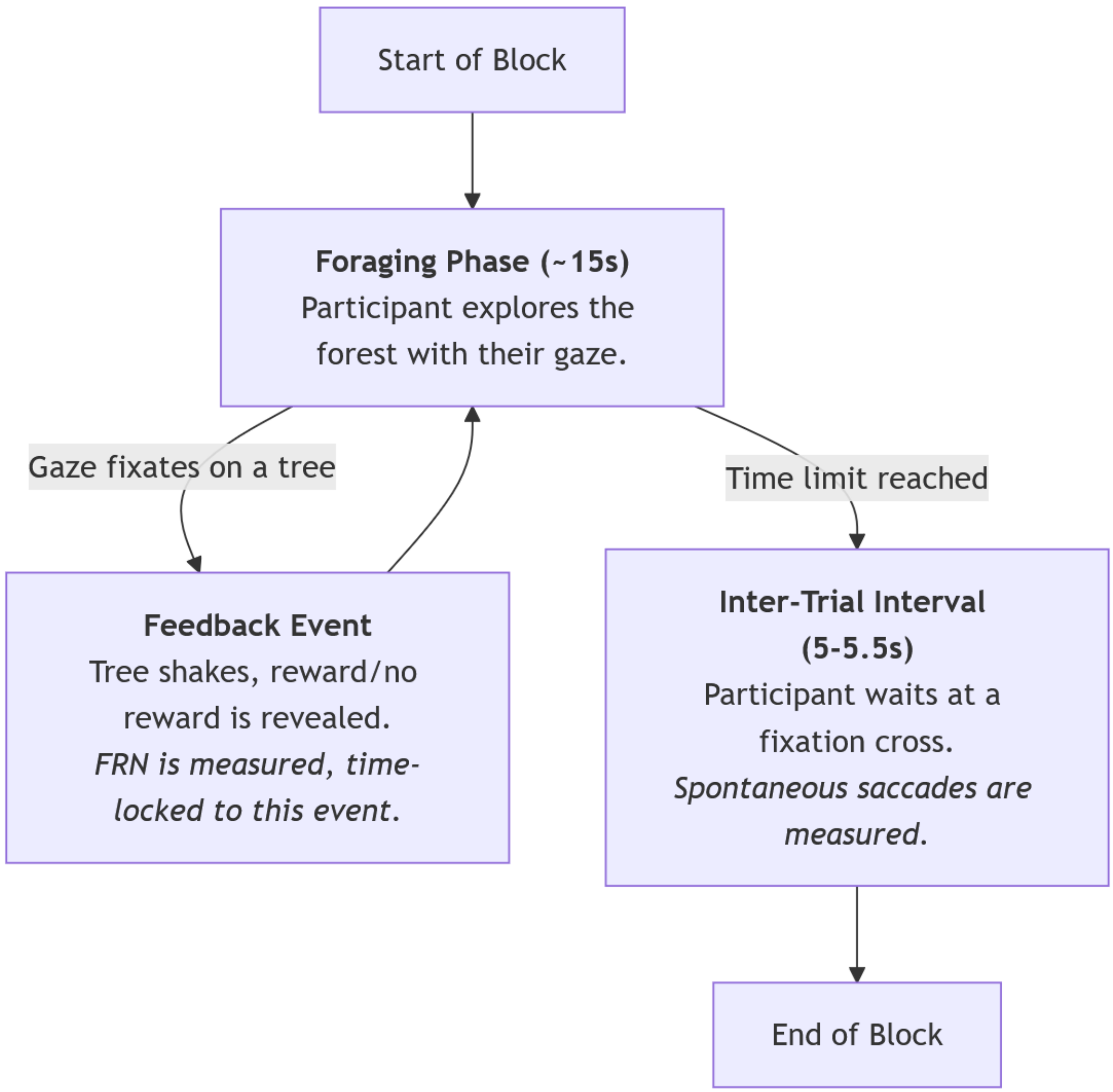

Figure 2Single trial of the physical task

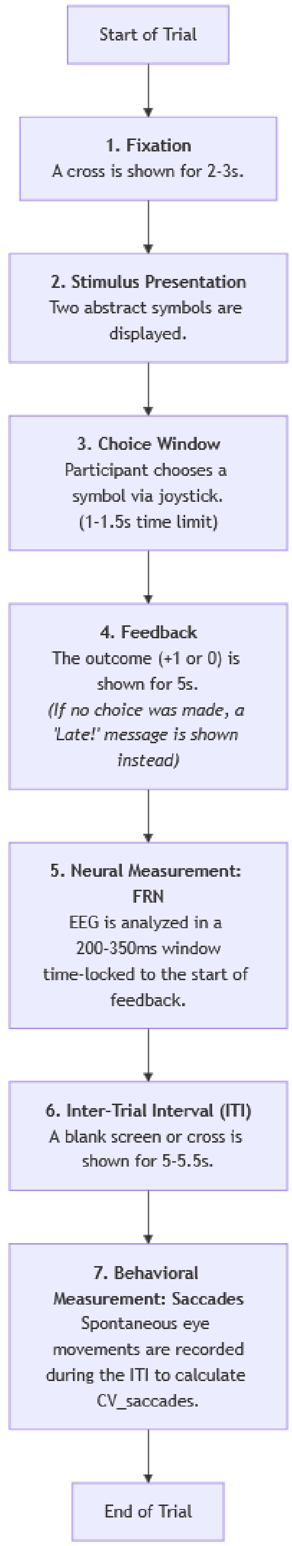

Figure 3 Abstract Foraging trial

References.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. https://doi.org/10.1037/a0027373

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in External and Internal Spaces: Evidence for Generalized Cognitive Search Processes. *Psychological Science*, *19*(8), 802–808. https://doi.org/10.1111/j.1467-9280.2008.02160.x

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54. https://doi.org/10.1016/j.tics.2014.10.004

Markowitz, J. E., Gillis, W. F., Jay, M., Wood, J., Harris, R. W., Cieszkowski, R., Scott, R., Brann, D., Koveal, D., Kula, T., Weinreb, C., Osman, M. A. M., Pinto, S. R., Uchida, N., Linderman, S. W., Sabatini, B. L., & Datta, S. R. (2023). Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature*, *614*(7946), 108–117. https://doi.org/10.1038/s41586-022-05611-2

Stojić, H., Orquin, J. L., Dayan, P., Dolan, R. J., & Speekenbrink, M. (2020). Uncertainty in learning, choice, and visual fixation. *Proceedings of the National Academy of Sciences*, *117*(6), 3291–3300. https://doi.org/10.1073/pnas.1911348117

Sutton, R. S., & Barto, A. (2020). *Reinforcement learning: An introduction* (Second edition). The MIT Press.