



UNIVERSITEIT VAN AMSTERDAM

MSc Biological Science  
Evolution of Behaviour and Mind

Research Project

---

Agent-based Model Reveals Flexible Risk Processing can  
Improve Navigation Efficiency in Primate Foraging

---

by

Muhan Li  
15387275

*July 11, 2025  
48 ECTS  
November 2024 - July 2025*

Supervisor: Dr. Emiel van Loon, Associate Professor Statistical Ecology

Assessor: Dr. Emiel van Loon, Associate Professor Statistical Ecology

Examiner: Dr. Karline Janmaat, Professor of Cognitive Behavioral Ecology



IBED - Institute for Biodiversity and Ecosystem Dynamics

# Contents

<b>Abstract</b> .....	1
<b>1. Introduction</b> .....	2
<b>2. Methods</b> .....	6
2.1 General Information of Experiments .....	6
2.2 General Description of the Agent-based Model .....	8
2.3 Learning and Strategy Updating Rule of the Agent .....	9
2.3.1 General logic of each strategy and their update .....	9
2.3.2 Cognitive mapping with Feature Association Matrix .....	10
2.3.3 Path learning rule and the selection rule of landmarks .....	12
2.3.4 The movement rule of avatar-based move and random move ....	13
2.3.5 Strategy updating rule and SoftMax probability calculation .....	16
2.4 Exploratory Data Analysis of Emerging Patterns .....	19
<b>3. Results</b> .....	21
3.1 Follow/Explore Dynamics .....	21
3.2 Foraging Performance .....	23
3.3 Spatial Learning .....	27
<b>4. Discussion</b> .....	28
4.1 Dynamic Risk Perception Enhances Adaptive Foraging in Agents ....	28
4.2 Social Implications of Dynamic Risk Perception in Public Information Use .....	32
4.3 Limitations, Applications, and Conclusions .....	33
<b>References</b> .....	36
<b>Ethics Approval</b> .....	47
<b>Acknowledgements</b> .....	47
<b>Appendices</b> .....	48

## **Abstract**

Understanding how animals make decisions under uncertainty is a fundamental question in behavioural ecology, particularly in fission–fusion societies where group composition fluctuate unpredictably. In these contexts, individuals must negotiate the trade-off between relying on private spatial memory and utilizing external public information. To investigate the cognitive mechanisms underlying this balance, we combined agent-based modelling (ABM) with virtual reality (VR) foraging experiments in humans and chimpanzees. Participants navigated a simulated environment in which the reliability of public information declined over time, requiring dynamic adjustments between individual exploration and public information use. Our ABM implemented two risk evaluation mode: a static risk-aversion model and a prediction error induce risk-seeking model. Each of these was further divided into two variants in which agents dynamically updated expected rewards based on experience, and another in which expected rewards remained fixed. The results show that individuals improve their foraging efficiency in unfamiliar environments by adjusting reliance on public information based on the predicted reliability of the informer. Models that allowed agents to update expected reward based on experience accordingly best reproduced observed foraging behaviours in both species. These findings indicate that risk-sensitive decision-making involves not merely the avoidance of uncertainty but the active regulation of risk to optimise information use. Our study highlights the flexible risk evaluation is an important cognitive adaptation for optimising navigation and foraging efficiency.

## 1. Introduction

The social dynamics of group-living animals arise from the interaction between individual decision-making and collective benefits. This tension becomes especially pronounced in fission-fusion societies. Hans Kummer (1971) first introduced this concept to describe species like chimpanzees (*Pan troglodytes*), whose group compositions change fluidly. Individuals may separate into smaller groups (fission) or merge into larger ones (fusion) based on ecological pressures, resource availability, and social needs. Although fission-fusion societies are rare among mammals, they are common in primates (Altmann, 1988; Madsen and de Silva, 2024). Living in fission-fusion system requires individuals to constantly adjust their foraging strategies as group sizes shift and social bonds form or dissolve (Couzin, 2006). In this context, their foraging decision-making must balance the uncertainty of their own environmental knowledge with the potential risks of relying on socially acquired information (Filippo Aureli et al., 2008). Social foraging theory, particularly the producer-scrounger framework (Giraldeau and Caraco, 2000; Koops and Giraldeau, 1996), posits that individuals adopt divergent roles: "producers" invest in costly resource discovery, while "scroungers" exploit others' efforts. However, studies testing the producer-scrounger model have primarily examined foraging patch accessibility (Beauchamp, 2008) and social cue transmission (Garg et al., 2022), often overlooking the integration of individual memory and public information use. In fission-fusion societies, subgroup dynamics and temporary associations create significant uncertainty in interactions between group members (Ramos-Fernandez et al., 2018). This uncertainty in social interaction can compromise the reliability of socially transmitted cues (Wright, 2006), including the potential for misinformation, such as deceptive signals of predation risk (Fahimipour et al., 2023) or food-related signal (Clay et al., 2012; McLinn and Stephens, 2006). Foraging decision-making thus becomes one of the key tensions within these social dynamics, requiring animals to weigh the credibility of both private and public information sources. This challenge is further amplified in novel environments, where prior knowledge may be limited or outdated. As such, these contexts raise critical questions about the cognitive and behavioural strategies that animals employ to adaptively navigate the shifting balance between individual exploration, social reliance, and how these strategies are shaped by the ecological and social structure of their groups.

Uncertainty presents a fundamental challenge for biological organisms, particularly in the foraging context, where animals must predict environmental conditions to maximize rewards while minimizing risks (Houston and Rosenström, 2024a). Effective decision-making demands the evaluation of available options, anticipation of conspecific behaviours, and continuous integration of personal information derived from environmental cues such as landmarks (Wise et al., 2023), olfactory signals (Murthy, 2024), and acoustic stimulus (Penar et al., 2020). Yet animals often forage under conditions of limited or noisy information (Lichtenberg et al., 2020), where even simple decisions can carry profound survival consequences. To mitigate such

uncertainty, individuals must invest time and energy into information acquisition, updating internal representations to adapt to dynamic ecological conditions. Although personal information remains crucial, the energetic and opportunity costs associated with its collection have favoured the widespread use of public information gleaned from observing conspecifics and heterospecifics (Kane and Kendall, 2017). Public information, particularly in foraging contexts, allows individuals to infer resource quality more rapidly, facilitating faster adaptation to novel environments (Dall et al., 2005). However, the reliance on public information can also trigger informational cascades (Danchin et al., 2004), resulting in suboptimal decisions when public information is incomplete, misleading, or deceptive—an ever-present risk in dynamic fission–fusion societies (Wright, 2006), where group composition and information reliability fluctuate. In such settings, for instance, in chimpanzees, fission-fusion dynamics necessitate sophisticated spatio-temporal memory to track dispersed food resources while mitigating the risk of being misled by unreliable group members. Despite recognition of these complexities, the cognitive mechanisms underpinning the assessment of public information reliability remain poorly understood. This gap is particularly salient given the evolutionary significance of fission–fusion dynamics in both non-human primates and early hominins, suggesting that the ability to navigate informational uncertainty may have been a cornerstone in the evolution of complex social cognition (Grove et al., 2012; Semple and Higham, 2013). Investigating how primates balance personal and public information under uncertainty thus offers critical insights into the selective pressures shaping the evolution of foraging cognition, social learning, and the emergence of complex societies.

However, studying the foraging cognition of primates in large-scale environments presents significant challenges, especially in balancing ecological validity with experimental control of confounding variables, as available information is difficult to quantify. While some studies have attempted to uncover the mechanisms underlying group decision-making by employing GPS collaring of entire baboon troops to capture all private information (Bracken et al., 2022; Strandburg-Peshkin et al., 2015), such approaches require long-term, full-group monitoring, which may pose ethical concerns and potentially interfere with the natural behaviour and survival (Portugal and White, 2018; Wilson et al., 2015). Recent advancements in virtual reality (VR) technology with captive chimpanzees have opened new avenues for investigating chimpanzee navigation in controlled, large-scale environments, as demonstrated by Allritz et al., 2022. This novel methodology offers a relatively comprehensive approach to investigate real-world navigation in captive primates by addressing the limitations in spatial scale inherent (Dolins et al., 2014) to previous methods, such as three-dimensional (3D) maze tasks (Dm et al., 2009) and barrier-and-detour tasks (Menzel Jr. and Menzel, 2007), thereby enabling the presentation of more ecologically valid problems to subjects. Moreover, VR technology offers a powerful tool for investigating foraging cognition across different taxa by providing a standardized and controlled environment. Comparative studies are essential for understanding the evolution of foraging cognition within a phylogenetic framework (Nunn and Barton,

2001), yet meaningful comparisons across primate species are often hindered by ecological differences of their home ranges. Moreover, the features of private and public information vary across ecological contexts, thereby complicating cross-species comparisons. VR technology creates a shared, ecologically valid space where primates can engage in foraging tasks under identical conditions, allowing for more direct comparisons of cognitive mechanisms. Additionally, to avoid the WEIRD (Western, Educated, Industrialized, Rich, and Democratic) bias that often limits cognitive research (Apicella et al., 2020), insights from hunter-gatherer societies are crucial for a comprehensive understanding of foraging cognition. By integrating VR-based comparative approaches into cognitive ecology, researchers can more effectively elucidate the cognitive trade-offs underlying information use in support of foraging behaviour across species.

Based on above background, foraging cognition offers a powerful window into the neural and evolutionary origins of risk sensitivity. While VR techniques allow researchers to present multiple primates from different taxa with identical foraging conditions, agent-based modeling (ABM) enables exploration of the cognitive and neurocomputational processes that underlie decision-making in such environments (Gribkova et al., 2024; Zedadra et al., 2017). Although animals are traditionally assumed to be risk-averse—a view consistent with risk-sensitive foraging theory (McNamara and Houston, 1992), which posits that individuals tend to avoid uncertainty to maximize expected returns—emerging evidence reveals that animals also exhibit subjective evaluations and systematic decision biases (De Petrillo and Rosati, 2021). One such example is the hot hand effect (Tversky and Kahneman, 1971), wherein individuals are more likely to engage in risk-seeking behaviour following a streak of rewards. This phenomenon has been linked to dopaminergic responses to positive prediction errors, leading to increased risk-taking (Moeller et al., 2021). Risk perception tendency, including loss aversion and the hot hand effect, has been observed in non-human primates (Pelé et al., 2014). Findings about similar risk cognition across primate taxa imply that the determinants of risk-related behaviour may be shaped by selective pressures encountered throughout evolutionary history, particularly those associated with foraging under uncertainty or within unfamiliar environments. However, little is known about how primates cognitively represent and evaluate risk associated with socially acquired information during foraging. Most studies on risk cognition tend to treat foraging-related risk and information use as separate domains, often overlooking the inherent risks associated with the information itself (Garber and Dolins, 2014; Garcia et al., 2021). Behavioural data alone are often insufficient to infer the underlying computational models or cognitive pathways involved in such evaluations (Janson and Byrne, 2007). ABM provides a powerful framework for simulating candidate neural encoding mechanisms, allowing researchers to identify plausible cognitive pathways used in risk evaluation (Herd et al., 2021; Hesslow, 2012). By integrating neurocomputational modeling with principles from behavioural ecology, ABM can elucidate how decision-making processes emerge from specific neural substrates and how these substrates are shaped

by ecological constraints.

ABMs have become increasingly powerful tools for testing foraging theory by comparing model-generated artificial data with empirical observations, thereby advancing our understanding of primate foraging cognition. However, the cognitive mechanisms underlying primate foraging decisions remain comparatively underexplored within this computational framework, as existing research has primarily focused on ecological constraints (Green et al., 2020; Sellers et al., 2007; Zappala and Logan, 2010). Prior ABM research on chimpanzees has demonstrated their ability to optimise travel routes using spatial memory, revealing sophisticated landscape knowledge and route planning abilities (Green et al., 2020; Robira et al., 2021). However, these models rarely examine how chimpanzees construct cognitive maps to learn foraging paths in new environments (Janmaat, 2019). (A sentence about problems in ABM research for chimpanzees). Similarly, ABM studies on human foragers have explored the interplay between individual knowledge and social cues in shaping food selection and exploration strategies (Barceló et al., 2014; Janssen and Hill, 2014; Wren et al., 2020). These studies highlight how the use of public information and social relationships may have enabled early humans to mitigate foraging risks and adapt to environmental change. However, they often overlook the inherent uncertainty associated with such information. The cognitive mechanisms governing the balance between independent exploration and social information reliance also remain poorly understood, particularly in comparative contexts. Although traditional observational approaches offer invaluable insights for understanding social learning in both species (Gilby and Machanda, 2022; Pretelli et al., 2022), they struggle to disentangle the specific cognitive costs associated with public information use, such as the reliability of learned cues or the limits of episodic memory. This limitation complicates direct comparisons of human and chimpanzee foraging strategies, despite their shared ecological challenges—navigating large territories, recalling patchy resource distributions, and weighing the trade-off between exploiting known food sources and exploring novel ones (Hills et al., 2015; Janmaat et al., 2014; Normand and Boesch, 2009). Neurocomputational ABMs can effectively test whether the evaluation of public information during foraging in humans and non-human primates is governed by shared cognitive mechanisms.

This study employs an ABM framework to investigate the cognitive adaptations underlying foraging behaviour in humans and chimpanzees when navigating unfamiliar environments. Specifically, the research aims to elucidate how primates perceive and regulate risk associated with public information use during foraging to optimise their decision-making. To examine how agents balance the use of public information and spatial memory, participants completed a virtual foraging task in which they navigated a simulated forest containing seven fruit-bearing trees, with the goal of collecting as much fruit as possible. At the start of each session, a knowledgeable avatar guided participants to fruit, but as trees were depleted, its reliability declined, increasing the risk of following. In the ABM, agents' perception

of this dynamic risk was modelled using two alternative mechanisms: a risk-aversion model, which assumes a fixed tendency to avoid uncertainty, and the Prediction Errors Induce Risk Seeking (PEIRS) model, which captures dynamic shifts in risk preference modulated by dopaminergic prediction error signals (Moeller et al., 2021). To further explore the role of reward expectation in risk learning, the model tested two assumptions: one in which the expected reward updated with experience, and another in which it remained fixed. All agents, regardless of model condition, shared strategy updating, spatial learning, and movement mechanisms to ensure comparability. Using four combinations of risk and reward modelling (Fig. 1), the study tested the following hypotheses: (1) that Chimpanzees and humans share foundational cognitive processes for risk learning and strategy updating (De Petrillo and Rosati, 2021), the ABM replicates the follow/explore dynamics of humans and chimpanzees when agents are tuned to the initial following proportion observed in the gameplay session; (2) that the PEIRS model more accurately captures the follow dynamics of participants due to the dual roles of dopamine in encoding teaching signals and risk regulation have been suggested to interfere with each other (da Silva et al., 2018; Hamid et al., 2016); and (3) that models which best capture follow/explore behaviour also more closely reproduce empirical foraging patterns.



**Figure 1. Four models of risk cognition and learning related to the likelihood of revisiting fruit trees by following the avatar employed in the study; model details are provided in the Methods and Appendix.**

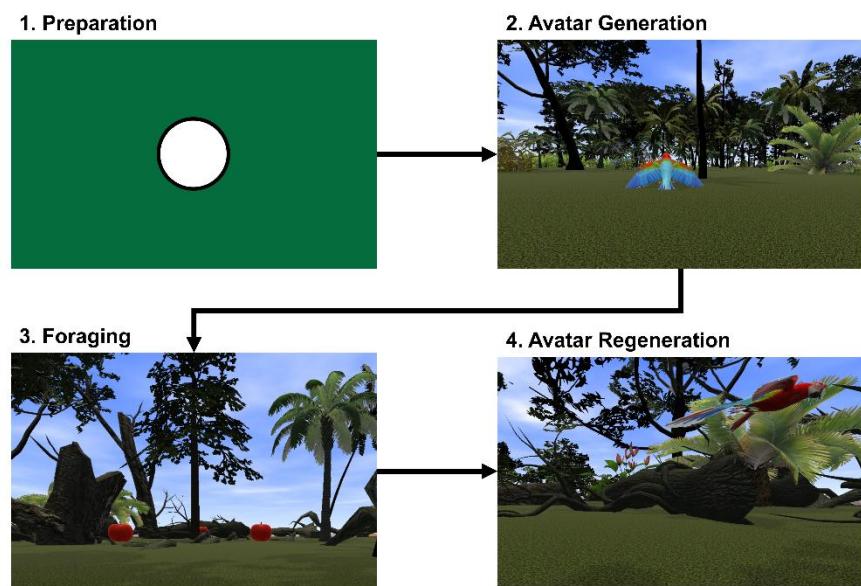
## 2. Methods

### 2.1. General Information of Experiments

The video game used for experiments were created by Bryndan van Pinxteren, Bram van der Perk and Pia den Braver. In this experiment, participants navigate a virtual forest containing seven trees capable of bearing fruit. Their objective is to collect fruit

by approaching these trees. Each participant, hereafter referred to as the forager, plays the VR game 10 times. The general experience during the game includes starting the game, following the avatar, searching for fruit and following the avatar again (Fig 2). At the beginning of each session, an avatar appears, which the forager can choose to follow. The image of the avatar is a red macaw (*Ara chloropterus*) to attract the attention of foragers by offering colorful stimulus. This avatar serves as a cue, representing a knowledgeable agent that moves directly toward a fruit-bearing tree. Initially, all fruit trees contain fruit, making it advantageous for the forager to follow the avatar. However, once a tree has been visited, its fruit is depleted. As the experiment progresses, the avatar becomes less reliable due to the greater probability of revisiting the fruit trees, increasingly leading the forager to empty fruit trees. The game will end when participant found all 7 fruits or it reached the 600s time limitation.

To facilitate spatial learning, each fruit tree is surrounded by a distinct set of landmarks, allowing foragers to associate specific locations with visual cues (Epstein and Vass, 2014; Vericel et al., 2024). The landmarks include trees, dead trees, pebbles, rocks, flowers, and plants, with rocks being the only impassable obstacles. The fruit trees could be visually identified from a distance by their narrow, straight trunks. However, a few morphologically similar tree, referred to as “Redwoods” in the game file, never bore fruit. Fruit-bearing trees are morphologically identical to depleted trees, meaning foragers must rely on associative memory to locate fruit-bearing trees. The spatial layout and fruit tree locations remain unchanged across all 10 game sessions, enabling foragers to use memory from previous games to balance their reliance on the avatar’s guidance with their own accumulated knowledge. This design allows for an assessment of how individuals transition from social information to independent decision-making as the reliability of external cues diminishes.



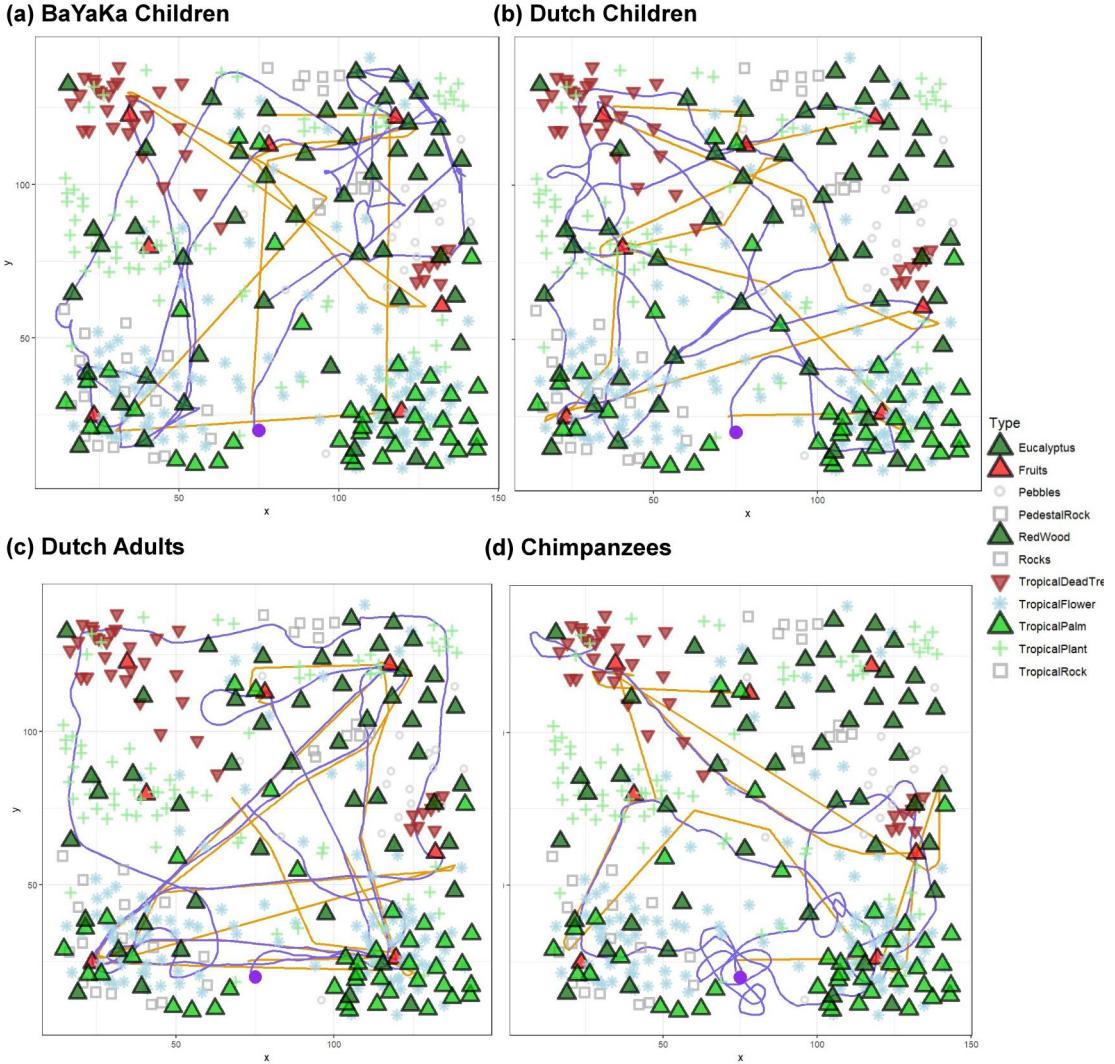
**Figure 2. Example of the process of each fruit-finding trial in the VR game**

## 2.2. General Description of the Agent-based Model

In this study, the ABM was developed to simulate the foraging game, the participants involved, and the events occurring within the gameplay. The detailed model description follows the updated ODD (Overview, Design concepts, Details) protocol suggested by Grimm et al., 2010, Grimm et al., 2020 (see Appendix 4). The ABM was based on the game design and outputs from initial experiments with the game (see Fig 3). The ABM consists of three core components: agents (representing foragers), avatars (representing public information through knowledgeable organism), and environmental modules (including foraging landscapes and landmarks). Agents use specific movement rules and possess both a cognitive map and mechanism for learning risk. The agents make movement decisions using Softmax, an activation function commonly used in neural networks for decision-making, with Bayesian updates (Xiong et al., 2017). This set enable agent to alternate between self-explore and following avatar based on experience during the game. Agents can detect landmarks in 40m, the maximum visible distance determined from the VR game, constrained to an angular orientation of 120 degrees. The calculation of detection distance is based on a sensitivity test. Researchers approached a selected landmark, then walked backward until it was no longer visible, pausing the game at that point. The pause location and distance to the landmark were recorded to compute the mean maximum visibility distance ( $39.5\text{m} \pm 7.2\text{m}$ ). Three instances of each landmark type were randomly selected for testing.

Each agent tracks its own state via position coordinates, movement velocity, and memory of prior experiences, dynamically updating selecting probability of move strategy to improve foraging success. The cognitive mapping module employs a variant of the Feature Association Matrix (FAM), a computationally lightweight graph-learning algorithm mimicking hippocampal auto- and hetero-associative circuits for episodic memory (Gribkova et al., 2024). This allows agents to assess environmental uncertainty using visual inputs, with uncertainty calculated via Shannon entropy. The risk learning module for following avatar is based on the Rescorla–Wagner rule, integrating learning rules for latent variables (e.g. value, expectancy) with choice rules that map these variables to selection probabilities, the activation patterns of dopaminoceptive areas in both human and nonhuman primate brain have been formalized in terms of such reinforcement learning models (Delamater and Oakeshott, 2007; Kobayashi and Okada, 2007). Agent movement speed was calibrated based on the speed setting of real game. Avatars replicate the in-game guide bird's behaviour: they are generated near the agent at the start of each trial and fly directly to a randomly selected fruit tree. After arrival, the avatar waits until the agent discovers the tree, then disappears and reappears to repeat the guidance behaviour. Avatar speed also mirrors parameters from the real-world game. The foraging environment comprises a  $200 \times 200$  plain, virtual forest containing seven fruit-bearing trees that visually resemble ordinary trees; once harvested, the fruits do not regenerate until another session starts. All environmental landmarks are consistent

with those in the game design.



**Figure 3. Examples for Movement Patterns of Foragers** (Bayaka children from hunter-gatherer society, Dutch children and adults from WEIRD (Western, Educated, Industrialized, Rich and Democratic) society, chimpanzees from ARTIS zoo). Each panel includes the performance of one participant in one session. This figure illustrates the movement patterns of the four groups of foragers. The purple line represents the movement of the forager, the purple point represents the start point, and the orange line represents the movement of the avatar.

### 2.3. Learning and Strategy Updating Rule of the Agent

#### 2.3.1. General logic of each strategy and their update

To simulate exploratory behaviour in the absence of explicit goal cues, a randomized movement module was implemented to incorporate both memory-guided and visually guided navigation components. The agent's movement was updated at each time step according to a stochastic process that integrated internal memory associations and perceptual input (see section 2.3.4). If no valid memory-based vector could be retrieved (e.g., due to the absence of associative history or navigational uncertainty),

the agent introduced a random perturbation to its current heading within a defined angular field of view. In contrast, the agent selects avatar as target and move forward to it each iteration when it under the avatar-targeted move. The avatar moves independently, providing an external cue about fruit tree locations. It moves randomly between fruit-bearing trees, following a unit vector normalization process to maintain efficiency in navigation. Upon reaching a fruit tree, the avatar temporarily disappears until the agent finds a fruit tree independently, at which point it reappears and resumes movement. This dynamic interaction allows agents to balance reliance on external cues with self-directed exploration.

The agent's movement strategy is determined using a SoftMax rule (Sutton and Barto, 1998), which provides a probabilistic framework for decision-making. The model incorporates two primary movement modes: avatar-targeted movement and random movement. The agent updates its position based on the direction and distance to the selected target avatar during avatar-targeted movement, while in random movement, its position is determined based on the memory and perception of landmarks. In the avatar-targeted movement, the agent follows an avatar that moves randomly between fruit trees. The strategy is updated when the agent either encounters a fruit tree while following the avatar or when the risk of following the avatar exceeds the reward of independently locating a fruit tree from the currently closest landmark. In contrast, the random movement parallel with memory-based guidance, the agent computed a visual gradient vector directed toward perceptually salient landmarks. Landmarks within the agent's field of view contributed to this vector based on their eligibility trace and relative distance. Each landmark's influence on the visual vector was inversely proportional to the square of its Euclidean distance from the agent, promoting attraction to nearby, eligible targets. The resulting gradient vector was normalized to obtain a visual heading.

### 2.3.2. Cognitive mapping with Feature Association Matrix

To learn the foraging map, agents integrate associative information from nearby landmarks of fruit-bearing trees. This process is mediated by a FAM, which encodes and replays sequential input patterns. Given a set of potential inputs, such as sensory cues and reward signals, FAM encodes the temporal structure of these sequences by modulating the strength and order of associations between paired inputs. Each associative link within the matrix is further assigned an expected reward value, thereby enabling the replay of previously acquired sequences. The matrix exhibits plasticity: sequences may be subject to decay and eventual forgetting if they cease to be reinforced by repeated presentation or reward, particularly in the presence of competing input sequences. Eligibility traces for each input are derived from sensory stimuli and serve as short-term temporal records of stimulus occurrence. Although reward signals may be transient, their associated eligibility traces decay more gradually, facilitating reinforcement learning for temporally proximal cues. In the present context, sensory inputs arise from visual stimuli generated by various landmarks (landmarks and fruit trees) within the virtual environment. The sensory

input strength of a given landmark  $n$  is computed using a model in which visual features, such as impassability, conspicuous coloration, and size, are weighted to reflect their relative salience in influencing perceptual input intensity:

$$S_n = p_n \times c_n \times s_n$$

where  $p_n$  represents the landmark that can be crossed through by the agent or not;  $c_n$  represents the conspicuity of the landmark  $n$  colours compared to the background;  $s_n$  represents the size of the landmark  $n$ .

At each iteration, the eligibility trace  $E_n$  for a given landmark  $n$  is updated with following equations:

$$E_n^{t+1} = \begin{cases} k_E + \frac{1 - k_E}{e^{a_E \cdot S_n}}, & \text{if the landmark } n \text{ is in sight} \\ E_{decay} \cdot E_n^t, & \text{if the landmark } n \text{ is not in sight} \end{cases}$$

$$E_{decay} = e^{1 - N},$$

where constants  $k_E$  is the constant for scaling  $E_n$ ,  $a_E$  is the constant for scaling  $S_n$ ,  $E_{decay}$  is the decay variable for eligibility trace, which is higher when there's more new sensory input,  $N$  is the number of landmarks agent can see in the current iteration. The eligibility trace also used in calculating the visual vector for each landmark.

Within the FAM, each pairwise combination of landmark eligibility traces is characterised by three key parameters: Strength, Order, and Reward. These variables respectively represent: (1) the degree of co-activation or overlap between two inputs; (2) the temporal order and relative latency with which the inputs are received; and (3) the expected reward associated with their joint activation. The simplified FAM used for this model only contains the strength and reward due to the dense distribution of landmarks weakens the effect of the order in which landmarks are encountered on associative learning, as similar landmarks are seen and memorized at the same time as agent move. The associative strength between a landmark  $i$  and its surrounding fruit-bearing tree  $j$  is ultimately determined by the co-occurrence and temporal overlap of their eligibility traces  $E_i$  and  $E_j$ :

$$C_{ij}^{t+1} = C_{ij}^t + \Delta C_{ij}^t,$$

$$\Delta C_{ij}^t = k_c(E_i \cdot E_j - E_{threshold} + k_a \cdot R_{ij})(E_i + E_j) + k_d(Strength_{ij} - C_{ij}),$$

$$Strength_{ij} = \frac{1}{1 + e^{a_s \cdot C_{ij}}},$$

where  $k_a, k_c, k_d, k_s$  denote constant parameters modulating the associative dynamics. The co-activation index  $C_{ij}$  increases in proportion to the degree of temporal overlap between the traces (i.e.  $E_i + E_j > 0$ ), and declines when such overlap is insufficient. The associative strength between landmark  $i$  and its surrounding fruit-bearing tree  $j$ , denoted as  $Strength_{ij}$ , is computed as a logistic transformation of  $C_{ij}$ , bounded within the interval [0, 1]. This measure quantifies the learned associative strength between the two entities. To prevent uncontrolled escalation of associative values, the dynamics of  $C_{ij}$  are further modulated by the associative reward signal  $R_{ij}$  and  $Strength_{ij}$ , thereby ensuring stable convergence within the learning system. The associative reward value assigned to each landmark  $i$  is derived from the reward output and its associative strength with its neighboring fruit-bearing tree  $j$ , as follows:

$$R_{ij} = Strength_{ij} \cdot R_j,$$

### 2.3.3. Path learning rule and the selection rule of landmarks

To enable navigation through non-overlapping spatial sequences, path integration (Etienne et al., 2004) is employed to compute the net direction and distance between pairs of landmarks and fruit trees. This module was triggered each time an agent arrived at a novel landmark and functioned as a foundation for encoding directed associations within a feature association memory. While the agent moved between landmarks (i.e., not within any arrival threshold), it continuously accumulated a path integration vector to approximate the relative displacement between successive landmarks. At each timestep  $t$ , the agent's displacement vector  $\vec{v}_t = (v_t^x, v_t^y)$  was added to an internal accumulator:

$$\vec{V}_{PI} \leftarrow \vec{V}_{PI} + \vec{v}_t$$

$$D_{PI} \leftarrow D_{PI} + \|\vec{v}_t\| = D_{PI} + \sqrt{(v_t^x)^2 + (v_t^y)^2}$$

where  $\vec{V}_{PI}$  is the cumulative path vector and  $D_{PI}$  is the total distance travelled since the last landmark arrival. Upon confirmed arrival at a new landmark  $j$ , if a valid previous landmark  $i$  was recorded and the integrated distance  $D_{PI} > 0$ , the path integration vector  $\vec{V}_{PI}$  was stored as a directional memory vector from landmark  $i$  to landmark  $j$ . These vectors encoded the relative direction between landmark pairs and served as internal guides for future memory-based navigation decisions. Reverse vectors were encoded symmetrically depending on the assumed model of bidirectional learning. Assume that the agent takes  $n$  steps, and at each step  $t \in \{1, \dots, n\}$ , the movement vector is  $\vec{v}_t = (v_t^x, v_t^y)$ . Then the stored memory vector is calculated as:

$$\vec{V}_{i \rightarrow j} = \left( \sum_{t=1}^n v_t^x, \sum_{t=1}^n v_t^y \right)$$

To support learning from delayed reinforcement, a list of previously visited landmark indices was maintained in a reward history buffer. Each time the agent arrived at a distinct landmark, the index of the prior landmark was appended to this buffer. This history was subsequently used for reward backpropagation through temporally ordered landmark transitions. Upon arrival at a new landmark, the path integrator was reset to zero to ensure that memory vectors represented discrete inter-landmark transitions rather than cumulative trajectories.

#### 2.3.4. The movement rule of avatar-based move and random move

Two principal types of agent movement were implemented for the agent: (1) goal-directed movement toward avatar, and (2) random exploratory movement, which included both memory-guided navigation and stochastic directional changes. When an agent selected a specific avatar, it moved deterministically toward that location using a normalized direction vector. The agent then updated its position and velocity based on the normalized direction vector and the step length.

Under the random move, the direction of agent was determined by both memory-based vector and visual gradient vector. To calculate the memory-based

vector if the agent was located at a known landmark  $i$ , a directional vector toward another landmark  $j$  was retrieved from the feature association memory, conditioned on learned reward and associative strength. The set of available candidate transitions is denoted:

$$\mathcal{T}_i = \{j | j \neq i\},$$

For each candidate  $j \in \mathcal{T}_i$ , the expected value of transitioning was based on the associative reward of the landmark. The target landmark with maximal learned reward was selected:

$$j^* = \operatorname{argmax}_{j \in \mathcal{T}_i} R_{ij}$$

The associated memory vector  $\vec{m}_{i \rightarrow j^*} \in \mathbb{R}^2$  was then normalized to unit length:

$$\vec{m}_{i \rightarrow j^*} = \frac{1}{\|\vec{v}\|} \cdot \begin{bmatrix} \vec{V}_x[i, j^*] \\ \vec{V}_y[i, j^*] \end{bmatrix}, \text{ where } \|\vec{v}\| = \sqrt{\vec{V}_x[i, j^*]^2 + \vec{V}_y[i, j^*]^2}$$

If no such memory vector was available, a fallback stochastic rotation (detailed below) was applied.

For visual gradient vector, The agent also computed a visual salience gradient based on currently visible landmarks. For each landmark  $n$  in view, a weighted vector was computed:

$$\vec{v}_n = \frac{1}{d_n^2} \cdot E_n \cdot \hat{r}_n, \text{ where } \hat{r}_n = \frac{(x_n - x_t, y_n - y_t)}{\|(x_n - x_t, y_n - y_t)\|}$$

Here,  $d_n$  is the Euclidean distance to landmark  $n$ , and  $E_n \in [0, 1]$  is an eligibility trace indicating attentional salience. The overall visual gradient vector is then:

$$\vec{v}_{\text{visual}} = \sum_k \vec{v}_k \text{ (normalized to unit length if } \|\vec{v}_{\text{visual}}\| > 0\text{)}$$

The final movement direction was computed as a weighted combination of the memory vector  $\vec{v}_{\text{mem}}$  and the visual vector  $\vec{v}_{\text{visual}}$  :

$$\vec{v}_{\text{nav}} = (1 - w) \cdot \vec{v}_{\text{mem}} + w \cdot \vec{v}_{\text{visual}}, w = \text{visual\_weight}$$

This vector was renormalized:

$$\hat{v}_{\text{nav}} = \frac{\vec{v}_{\text{nav}}}{\|\vec{v}_{\text{nav}}\|}$$

and used to update the velocity and position of agent:

$$\vec{v}_t = \ell \cdot \hat{v}_{\text{nav}}, \vec{x}_{t+1} = \vec{x}_t + \vec{v}_t$$

where  $\ell$  is the agent's fixed step length.

If both memory and visual vectors were unavailable or null (i.e., agent was uninformed and visual occluded), directional noise was introduced. A small angular perturbation  $\theta \sim \mathcal{U}\left(-\frac{\phi}{2}, \frac{\phi}{2}\right)$  was applied to the agent's previous velocity vector  $\vec{v}_{t-1}$  :

$$\vec{v}_t = \ell \cdot \mathbf{R}_\theta \vec{v}_{t-1}, \mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This preserved movement continuity while enabling exploration in the absence of directed cues.

In agent encounter a landmark cannot be passed, such as rocks, the agent performed a directional reversal, approximating a U-turn, by sampling a heading angle  $\theta \sim \mathcal{TN}(\mu = \pi, \sigma = \phi/5, [\pi - \phi/2, \pi + \phi/2])$ , where  $\mathcal{TN}$  is a truncated normal distribution centered at  $180^\circ$ . The resulting movement vector was:

$$\vec{v}_t = \ell \cdot \mathbf{R}_\theta \vec{v}_{t-1}$$

Position remained fixed in this step  $(\vec{x}_{t+1} = \vec{x}_t)$ , but the heading vector was updated for the subsequent movement cycle.

### 2.3.5. Strategy updating rule and SoftMax probability calculation

Bayesian reinforcement learning framework was adopted to simulate strategy adaptation based on prior experience and environmental uncertainty (Trimmer et al., 2011; Xiong et al., 2017). Specifically, for each movement strategy  $s \in \{\text{avatar, random}\}$ , the agent maintains Beta-distributed priors over success rates, updated trial-by-trial using weighted evidence. The agent updates its priors with new trial information as follows:

$$\begin{aligned}\alpha_s(t) &= \lambda \cdot \alpha_s(t-1) + n_s^{\text{success}}(t), \\ \beta_s(t) &= \lambda \cdot \beta_s(t-1) + n_s^{\text{fail}}(t),\end{aligned}$$

where  $n_s^{\text{success}}(t)$  is the number of successful decisions using strategy  $s$  in trial  $t$ , and  $n_s^{\text{fail}}(t)$  is the number of failed decisions using strategy  $s$  in trial  $t$ . The decay term  $\lambda$  captures temporal forgetting, ensuring recent experiences influence behavior more.

The posterior belief in success for each strategy is calculated using the Beta posterior:

$$p_s(t) = \frac{\alpha_s(t)}{\alpha_s(t) + \beta_s(t)},$$

This is combined with initial prior probabilities  $p_s^{(0)} \in [0, 1]$  to update the Bayesian belief, grounded in initial trust toward each strategy:

$$BayesianProb_s(t) = \frac{p_s(t) \cdot p_s^{(0)}}{p_s(t) \cdot p_s^{(0)} + (1 - p_s(t)) \cdot (1 - p_s^{(0)})},$$

Each strategy's value is then updated using a Rescorla-Wagner-like rule, scaled by its Bayesian success probability:

$$V_s(t+1) = V_s(t) + k_s \cdot (R_s(t) - V_s(t)) \cdot BayesianProb_s(t),$$

where  $R_s(t)$  is the trial outcome (1 for success, 0 for failure) using strategy  $s$  in trial  $t$ , ensuring agent can integrate rewards and cognitive costs to update expected results for each strategy.

Given the constraints of working memory (Sweller, 1988), learned uncertainty of environment was incorporated into the model to simulate the agent's limited capacity for memory utilization. As the primary focus of this study is the cognitive processes underlying foraging decision-making, uncertainty in environment was specifically modeled as the ambiguity imposed by working memory during decision-making. The uncertainty of decision outcomes in the random movement is primarily influenced by the availability of information within the association in this game. Conceptualizing information processing as a shift in information states allows for the application of Shannon entropy to quantify uncertainty in a computationally efficient manner (Frank, 2013; Ortega and Braun, 2013; Wilkes and Gallistel, 2017). Thus, in this study, the uncertainty in the associative rewards of certain landmark  $i$  is quantified using entropy, defined as:

$$P_i = \frac{R_i}{\sum_j^N R_j},$$

$$H = - \sum_{i=1}^N P_i \cdot \ln P_i,$$

where  $N$  represents the number of paths can be chose from the closet landmark in sight at the iteration  $t$ ;

Let  $Q_{\text{random}}$  denote the utility of self-exploration by random move at trial  $t$ , defined as a function of expected outcomes and the uncertainty associated with random move:

$$Q_{\text{random}} = V_{t, \text{random}} - b\sqrt{H_{\text{cog}}}$$

where  $b$  is the sensitivity parameter for the uncertainty in changing into the self-exploration.

The risk perception mechanism of the avatar-following movement was employed with two different model, one is the dopaminergic reward prediction frameworks described by Moeller et al. (2021). This model considered risk preferences emerge as side effects of reward prediction errors, matching with the risk-seeking behaviour caused by the hot hand effects. Another is based on the classic risk aversion model (d'Acremont et al., 2009). By integrating these elements with the SoftMax decision rule, the ABM captures how agents refine their foraging strategies over time, offering insights into the cognitive processes underlying movement decisions and the trade-offs between public and individual information use.

Predicted risk were updated following a Rescorla–Wagner rule (Preuschoff and Bossaerts, 2007). Let the reward  $r_t$  denote the payoff in trial t. Let the  $V_{t,avatar}$  be the expected outcome of following avatar. The reward prediction error at trial t can be calculated as:

$$\delta_t = \begin{cases} r_t - V_{t,avatar}, & \text{if model = 'PU' OR 'AU'} \\ r_t - R_{expect}, & \text{if model = 'PN' OR 'AN'} \end{cases}$$

$R_{expect}$  was defaulted as 1 due to the fruit reward in this game continuously equal to 1.

This design ensured that agents in the PU and AU models adjusted their expected rewards based on experience, whereas those in the PN and AN models did not. Then the risk prediction error  $\xi_t$  is denoted as the difference between the squared reward prediction error and the estimated predicted risk (d'Acremont et al., 2009):

$$\xi_t = \delta_t^2 - h_t$$

The Rescorla–Wagner rule is used to update predicted risk  $h_{t+1}$  for the next trial:

$$h_{t+1} = h_t + k_{risk} \xi_t$$

where  $k_{risk}$  is the risk learning rate.

Let  $Q_{avatar}$  denote the utility of avatar-targeted move at trial t, defined as a function of the expected outcome and the predicted risk associated with avatar-targeted move. When the agent perceive risk following the PEIRS model (Moeller et al., 2021), the utility updating rule will be:

$$Q_{avatar} = V_{t,avatar} + \tanh(a\delta_t) \cdot h_t,$$

When the agent perceive risk following the classic risk aversion model, the utility updating rule will be:

$$Q_{\text{avatar}} = V_{t, \text{avatar}} - \alpha \sqrt{h_t},$$

Where  $\alpha$  is the risk preference parameter. In this model,  $\delta = 0$  corresponds to steady state dopamine release,  $\delta < 0$  means that dopamine release is suppressed, and  $\delta > 0$  means that dopamine release is enhanced (Moeller et al., 2021).

The probability of selecting the specific strategy in the trial  $t$  was defined with the log-SoftMax-based probability (Banerjee et al., 2025):

$$\log p_s(t) = \beta \cdot Q_s - \log \left( \sum_{i=1}^N e^{\beta \cdot Q_i} \right),$$

To ensure numerical stability and prevent overflow, log-sum-exp trick was used for subtracting the maximum scaled value:

$$\log p_s(t) = \beta \cdot Q_s - \left[ \max_j (\beta \cdot Q_j) + \log \left( \sum_{i=1}^N e^{\beta \cdot Q_i - \max_j \beta \cdot Q_j} \right) \right],$$

Where,  $\beta$  is the parameter used for adjusting the cognitive flexibility of agent, higher  $\beta$  will let agent tend to select the decision with higher value.

## 2.4. Exploratory Data Analysis of Emerging Patterns

In this study, model predictions of ABMs were used to compare with empirical observations, refining model parameters to better capture the cognitive mechanisms underlying the observed behaviors. To minimise bias, we first conduct exploratory data analysis to inform model rules and parameter selection, ensuring that empirical patterns guide model development rather than preconceptions (Tukey, 1977). To mitigate data dredging bias (Erasmus et al., 2022), only 50% of the data within each subgroup (Dutch adults, Dutch children, Bayaka children, and adult chimpanzees) was used for exploratory data analysis. This analysis primarily focused on identifying patterns in the raw data, including movement trajectories, summary statistics of following behaviour, and potential correlations. However, to reduce the risk of overfitting in subsequent comparative analyses (Nilsen et al., 2020; Rubin and Donkin, 2024), only movement patterns and summary statistics for proportion of following behaviour were considered. This approach allows for a data-driven yet cautious development of the ABM.

Each group of foragers exhibited similar movement patterns, characterised by a combination of following the avatar and random navigation between landmarks (Fig. 1). Following behavior was quantified by analyzing deviations in the forager's trajectory relative to the avatar's, based on angular velocity and distance (Flack et al., 2012; Harel et al., 2017). Additionally, a forager was classified as following the avatar when its movement direction was directly oriented toward the avatar while within the defined detection range of 45 meters, this threshold reflects the average distance beyond which the avatar is no longer visible in the game tested by researcher. Changes in following behavior across number of fruits found were assessed by calculating mean values over 10 gameplay sessions. In line with the central research questions of this study, we identified several emerging patterns for comparison between the agent-based ABM and empirical data: (1) foraging efficiency based on number of fruits found; (2) the dynamics of following versus exploration; and (3) the transition sequences of independently found fruit-bearing trees. Foraging efficiency serves to evaluate whether the ABM successfully captures participants' real-world foraging success and the learning trajectory toward optimal foraging strategies. The dynamics of following versus exploration are examined to determine whether the ABM can simulate the emergent behavioral trade-offs observed in participants as they navigate between social and individual foraging strategies. Finally, the transition sequences of independently found fruit-bearing trees were used for quantifying the spatial learning possess of agents and participants, used for further understanding how follow-explore dynamics help agent optimise foraging efficiency. Data from BaYaka children were excluded from emerging pattern comparisons due to unclear counts of individual gameplay session.

Given the strong influence of early-game experiences on subsequent behaviour, the ABM data were filtered according to the observed range of initial proportion of following behaviour in selected participant group (Dutch adults, Dutch children, and adult chimpanzees). The resulting subset of model data was then subjected to exploratory analyses alongside empirical data for each corresponding group. The participant group is further generalized into 'Human' group and 'Chimpanzee' group for further analysis. Foraging efficiency was quantified using two complementary measures: the total number of fruits collected per session and the time taken to locate a predefined number of fruit-bearing trees within each session. Although path efficiency is frequently employed in foraging research as an indicator of individual optimization and overall foraging performance (Janmaat et al., 2021), this metric was excluded due to technical constraints inherent to the touchscreen-based VR environment, which leads participants exhibited erratic movement patterns due to touchscreen-induced noise. Poisson regression was selected to capture the relationship between number of fruits found and sessions due to number of fruits is non-negative integer values and the find of each fruit-bearing tree is independent. Poisson regression was performed separately for the aggregated data of each group (humans, chimpanzees, and correspond ABMs) and for each individual participant/agent. Regression coefficients at the individual level were calculated only for human

participants and their corresponding ABMs, as the models failed to successfully replicate the foraging behaviour observed in chimpanzees. Earth mover's distances between regression coefficient distribution of models and real participants were calculated by *transport* package in R (Schuhmacher et al. 2024) to find the best match.

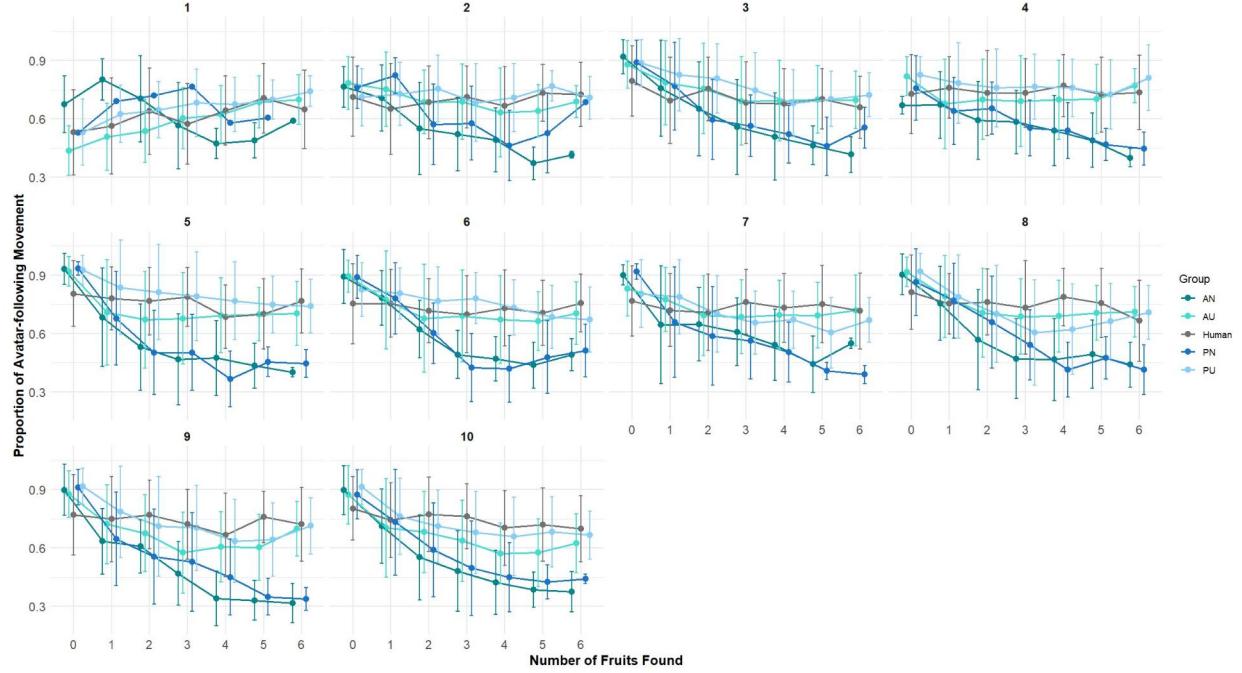
To characterise the dynamics between following and exploratory behaviours, we assessed the proportion of each behavioural strategy employed by individuals at varying stages of each game, defined by the number of fruit-bearing trees discovered. The Fréchet distance was calculated between the curves representing the proportion of following behaviour versus the number of fruits found in both empirical and simulated data, using the *SimilarityMeasures* package (Toohey K, 2015), to identify the model that best matched participant behaviour. Spatial learning patterns were examined in human participants to clarify the relationship between follow-explore dynamics and foraging efficiency; chimpanzees were excluded from this analysis due to mismatches with the agent-based model predictions. The transition matrix was computed based on the probability that an agent, having encountered a specific fruit tree, would next discover another specific tree through autonomous exploration. Jensen–Shannon divergence was employed to identify the model that best approximated human participant behavior, as it is commonly used for assessing similarity between transition matrices (Mao et al., 2025; Xue et al., 2025). To evaluate whether the most probable transition path (i.e. the tree-to-tree transition with the highest probability) represented an optimal (minimum length) foraging path, we referenced the spatial layout of trees in the game environment (Appendix 1), where the optimal path was defined as the shortest distance between two fruit-bearing trees. Additionally, Levenshtein distance, implemented via the *stringdist* package in R (van der Loo M, 2014), was used to quantify the dissimilarity between sequences of autonomously discovered tree IDs across successive gameplay sessions, in order to examine whether agents developed consistent foraging routes over time. All empirical data used for comparison were drawn from participants who completed the full game protocol and played all 10 sessions consecutively. All statistical analyses and model implementations were conducted in R (v4.3.2) and the graphics were made by *tidyverse* (Wickham et al., 2019) and *ggpubr* R package (Kassambara A, 2023).

### 3. Results

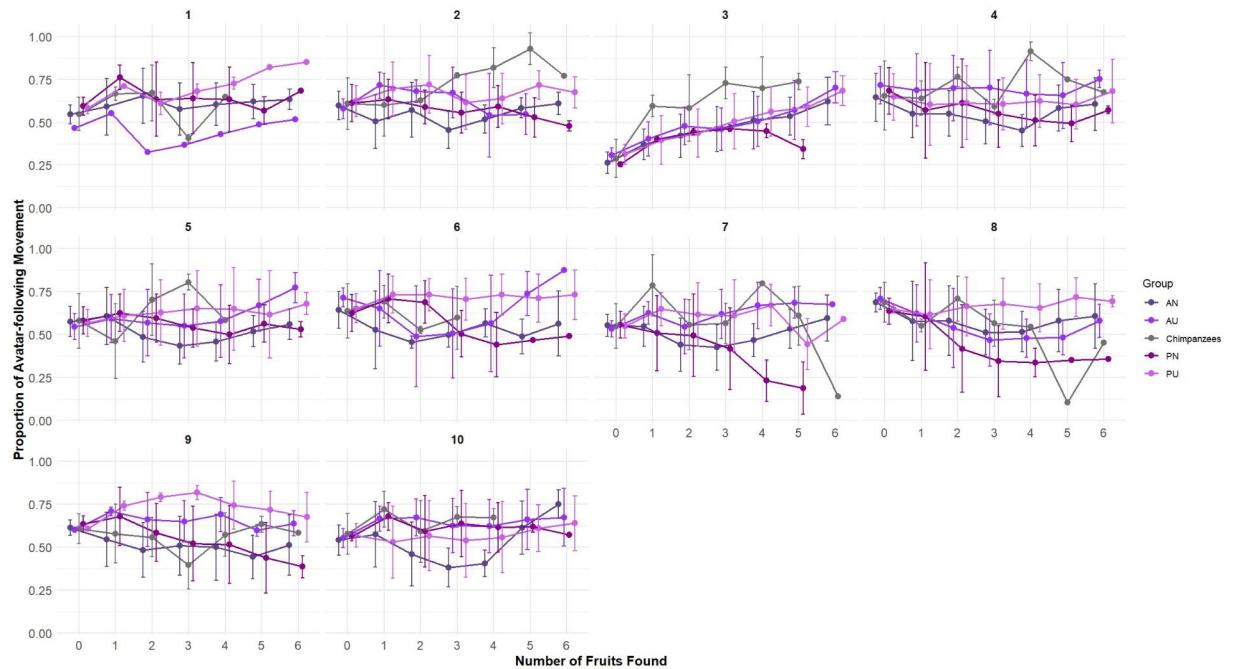
#### 3.1. Follow/Explore Dynamics

To assess the model capacity of reproducing the follow/explore dynamics observed in participants, the proportion of avatar-following behaviour changes with the number of fruits found was examined (Fig 4 & Fig 5). Both the PU and AU models closely tracked the following behavioural trends in human participants (Fig 4 & Fig 6) while PU models dropped slower than AU models. In contrast, the ABM did not adequately reproduce the following behaviour of chimpanzee participants (Fig 5). However, models with updating expected rewards exhibit lower Fréchet distances to both

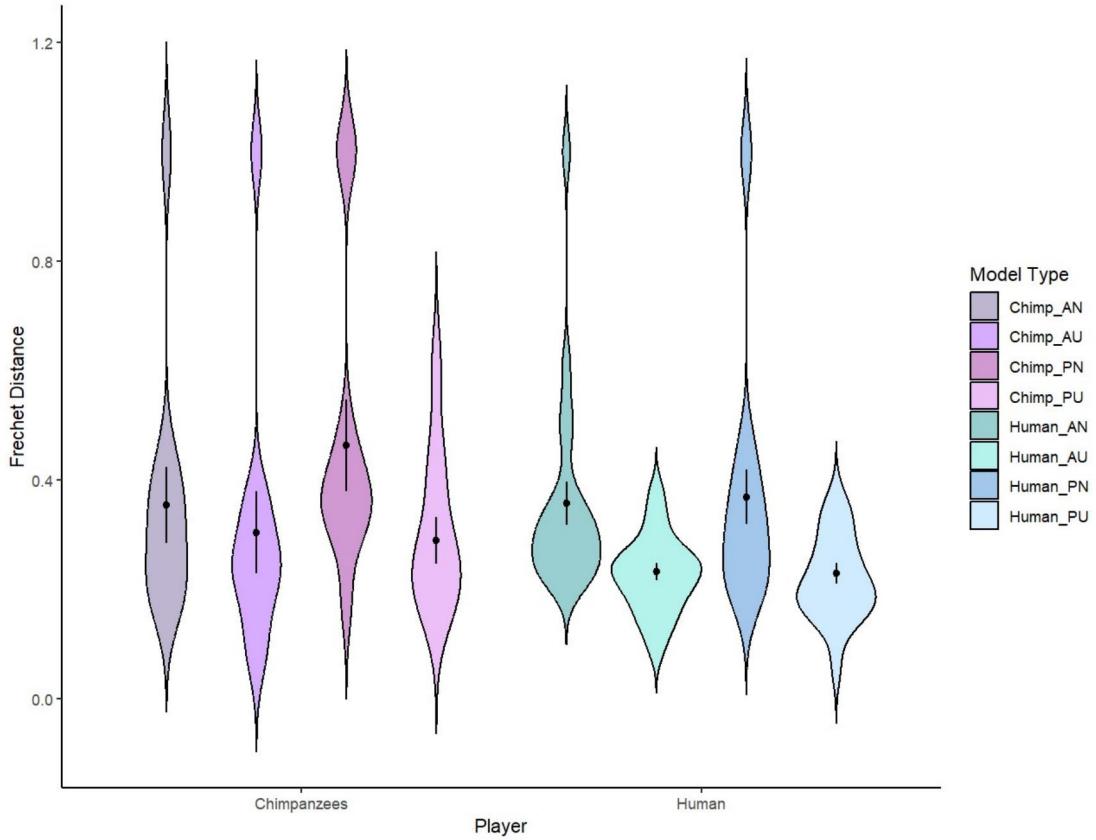
human and chimpanzee data (Fig. 6), characterised not only by fewer agents with large distances but also by smaller standard error ranges compared to other models.



**Figure 4. PU and AU models most accurately reproduce human patterns of initial social following:** Line plots show changes in the proportion of avatar-following behaviour across the number of fruits found, for both human participants and agents within  $\pm 1$  standard deviation of participants' initial following proportion (prior to locating the first fruit). Points and error bars indicate mean and standard deviation of following behaviour, respectively. Numeric labels on panels denote the number of game session.



**Figure 5. The ABM fails to replicate chimpanzee exploration–exploitation dynamics:** Line plots show changes in the proportion of avatar-following behaviour across the number of fruits found, for both chimpanzee participants and agents within  $\pm 1$  standard deviation of participants' initial following proportion (prior to locating the first fruit). Points and error bars indicate mean and standard deviation of following behaviour, respectively. Numeric labels on panels denote the number of game session.

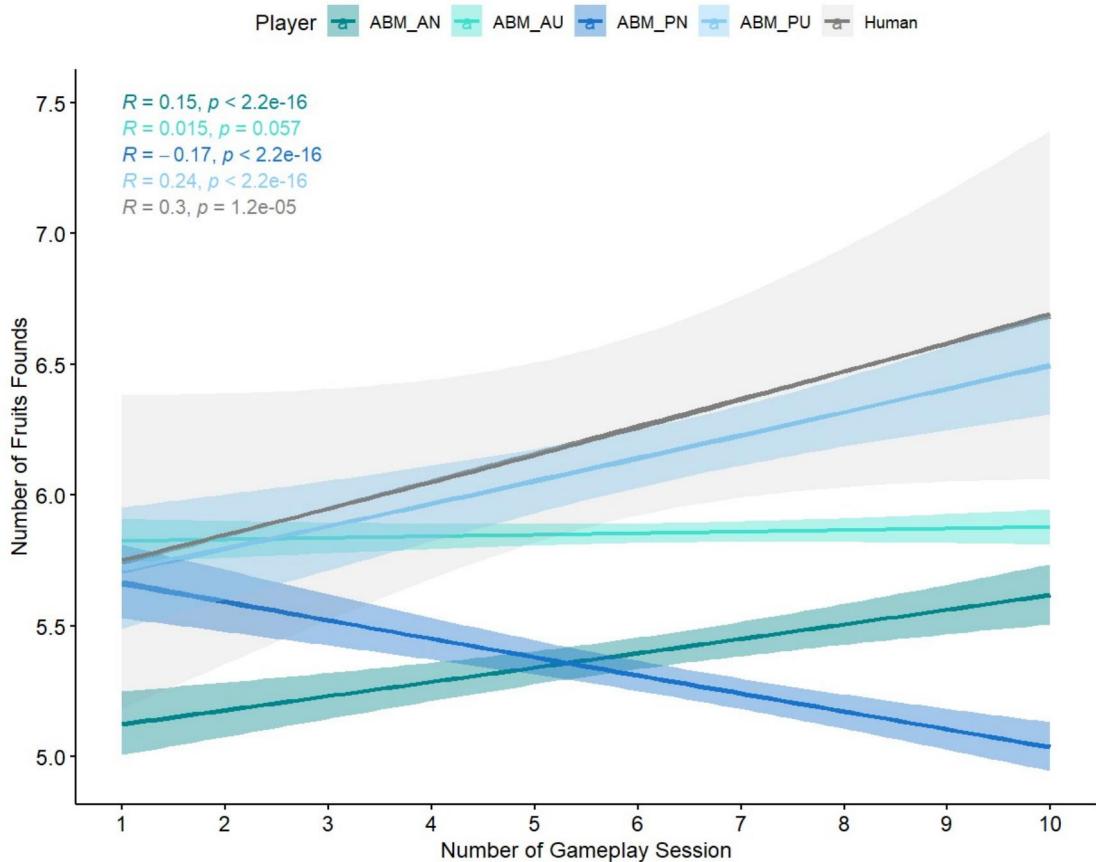


**Figure 6. Models incorporating updating expected reward of following avatar better approximate empirical following behaviour:** Violin plots display the Fréchet distances between observed and model-predicted trajectories of avatar-following as a function of fruit discoveries, for chimpanzee and human participants and their matched agents (based on  $\pm 1$  standard deviation of initial following behaviour). Overlaid dot-line plots show the mean and standard error of Fréchet distances for each model.

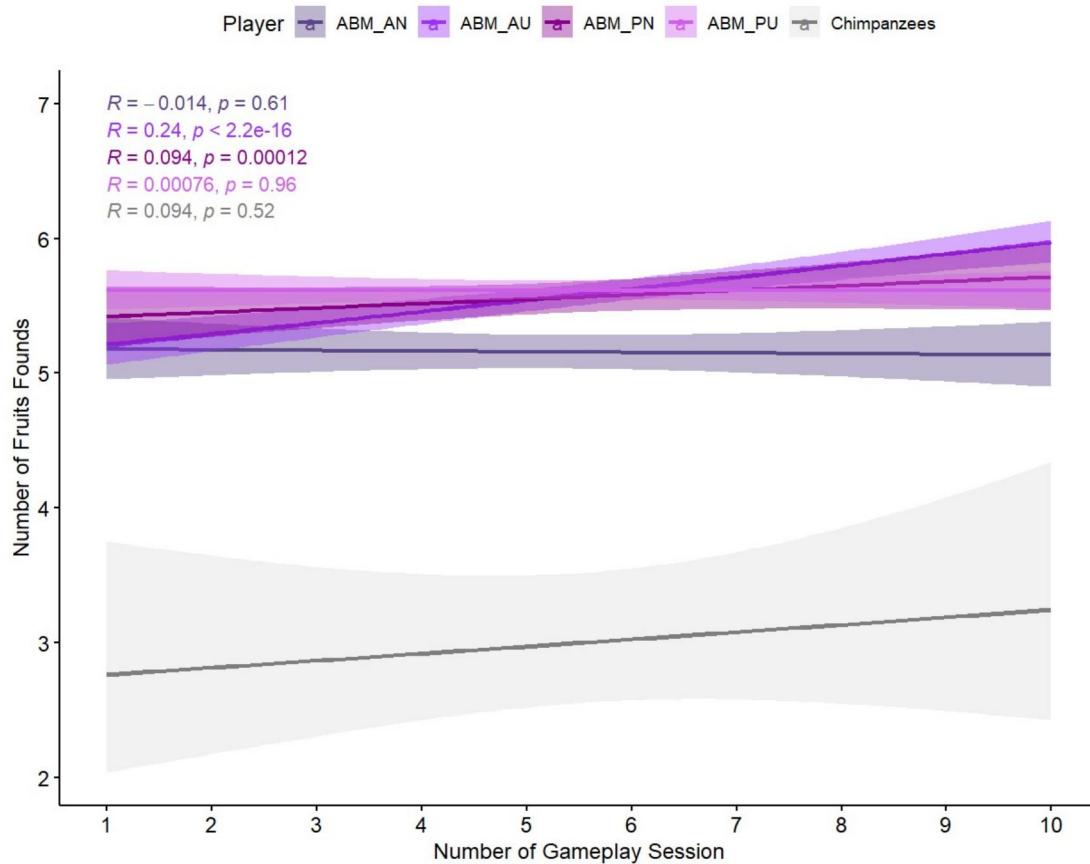
### 3.2. Foraging Performance

In human participants, the total number of fruits discovered increased across game sessions, suggesting an optimization in foraging performance. This pattern was best captured by the PU model, which showed a significant positive association between game number and fruit acquisition among agents mimicking human follow behaviour (Fig 7). Other models either underestimated the number of fruits found or failed to capture this learning pattern. By contrast, none of the models successfully reproduced the observed learning pattern in chimpanzees (Fig 8). Empirical data revealed no significant relationship between game session and fruit acquisition among

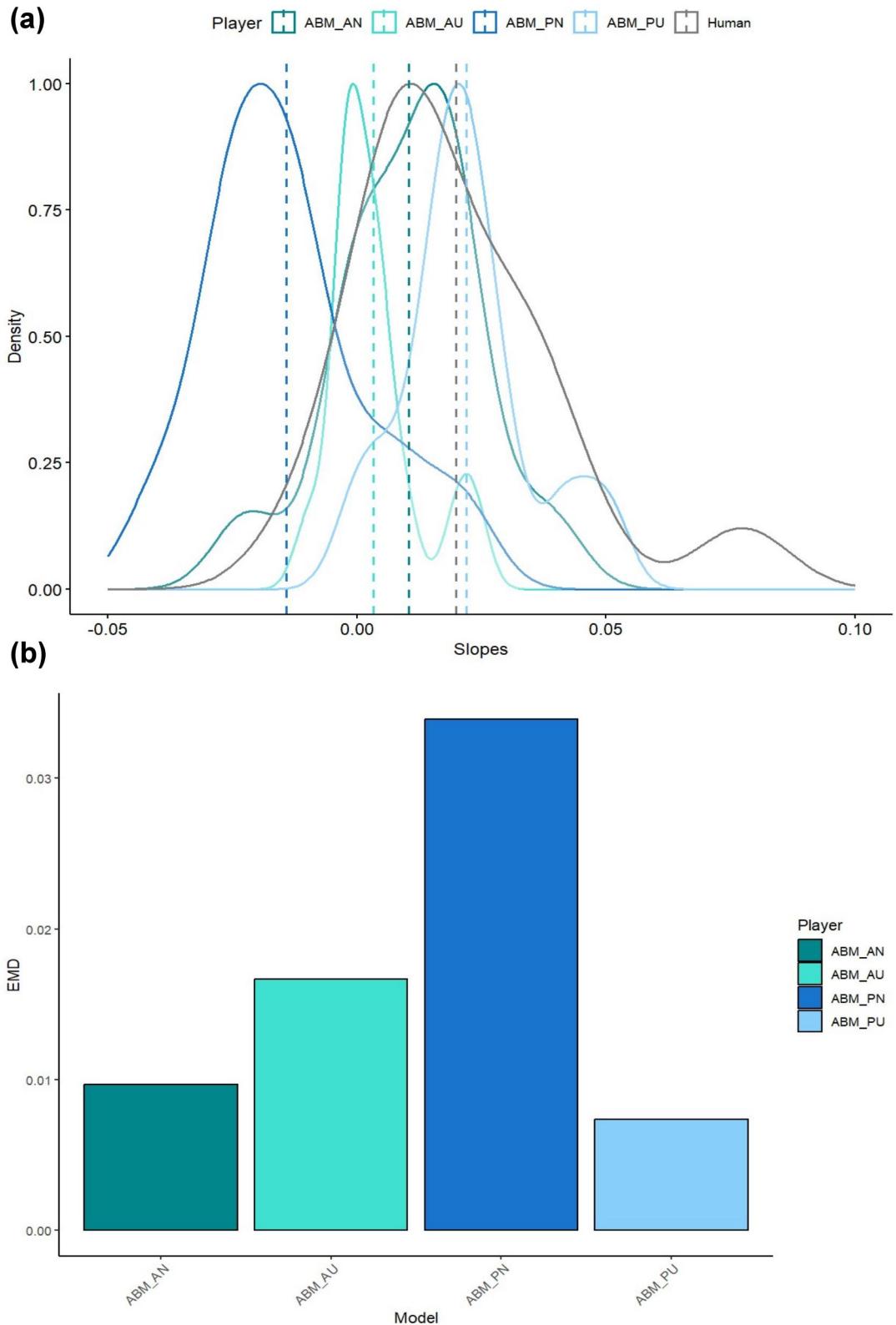
chimpanzees, and model agents failed to simulate any dynamic. Individual-level analysis further supported the superior fit of the PU model to human data. Distributions of Poisson regression coefficients quantifying the relationship between number of game session and fruits found were closely aligned between human participants and PU agents (Fig 9a). Moreover, Earth Mover's Distance values confirmed minimal divergence between the model and empirical coefficient distributions (Fig 9b). In addition, The distribution results show that the learning rates of most human participants are close to zero. Both the model that updates expected rewards and the slope distribution of human participants exhibit two peaks. Among them, the peak appearing on the right side of the x-axis for human participants corresponds to a larger slope value than those of both the PU model and the AU model. Additionally, the right-side peak of the PU model corresponds to a larger slope value than that of the AU model.



**Figure 7. The PU model best captures human fruit acquisition and learning dynamics:**  
 Regression plots show the relationship between total fruits acquired and number of gameplay session for human participants and simulated agents matched on initial avatar-following proportion (based on  $\pm 1$  standard deviation of initial following behaviour). Shaded areas indicate 95% confidence intervals. R means the correlation coefficient of regression model and the regression considered as significant when the p-value is smaller than 0.05.



**Figure 8. No model adequately captures chimpanzee fruit acquisition and foraging learning trajectories:** Regression plots show the relationship between total fruits discovery and number of game sessions for chimpanzee participants and matched agents (based on  $\pm 1$  standard deviation of initial following behaviour). Shaded bands represent 95% confidence intervals. R means the correlation coefficient of regression model and the regression considered as significant when the p-value is smaller than 0.05.

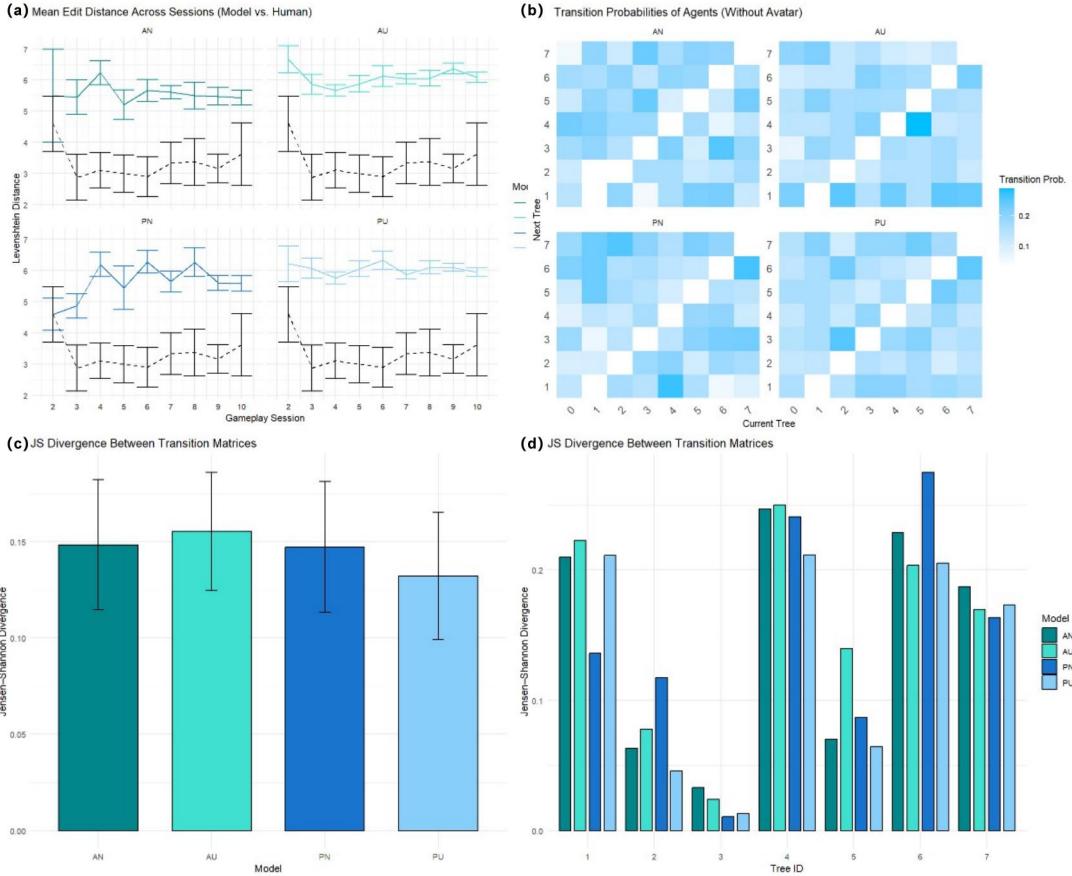


**Figure 9. The PU model more accurately captures individual-level foraging learning dynamics:** (a) Density plots show the distribution of Poisson regression coefficients (slopes) relating number of game sessions to fruits discovery for each human participant and their matched simulated agents (defined within  $\pm 1$  s.d. of each participant's initial avatar-following

behaviour prior to first fruit discovery). Dashed lines indicate medians of respective distributions. (b) Bar plot displays the Earth Mover’s Distance (EMD) between the distribution of model-derived regression coefficients and those observed in human participants.

### 3.3. Spatial Learning

To better understand how model successfully replicates the foraging pattern in human participants, whether agents could replicate human-like spatial learning were evaluated. None of models replicate the value of Levenshtein distances between self-explored fruit-bearing tree visitation sequences across gameplay session. However, PU agents and AU agents displayed a decreasing pattern similar to the real participant in the previous sessions (Figure 10a). For agents whose relationship between the number of games played and the number of fruits found is non-negative (AU, AN, and PU), the fruit tree sequences discovered through self-exploration become increasingly similar as the number of games increases. In contrast, for agents whose relationship between the number of games played and the number of fruits found is negative, the fruit tree sequences discovered through self-exploration become increasingly divergent with more gameplay. According to the transition matrix (Fig 10b), PU agents exhibited some path learning in selecting nearby fruit-bearing trees to visit ( $2 \rightarrow 3, 6 \rightarrow 5$ ). PN agents also exhibit higher probability in choosing closer fruit-bearing trees ( $2 \rightarrow 7$ ). These two model also performed a same but not the optimal path with real human participants ( $7 \rightarrow 6$ ; Appendix 1 & 2). The AU agents showed choice preference of selecting closer tree but only one path ( $5 \rightarrow 4$ ) while AN agents failed to reproduce any such optimal choice. The spatial transitions simulated by PU model were more closely matched human participants (Fig 10c, d), as indicated by lower Jensen–Shannon divergence values.



**Figure 10. The PU model better approximates navigation patterns observed in human participants compare to other models:** (a) Line plots show Levenshtein distances between sequences of fruit tree encounters in successive trials for each human participant and their matched agents. (b) Heatmaps illustrate the transition probability matrices of tree encounters, showing the likelihood of transitioning from a given fruit tree (rows; 0 = start location) to the next fruit-bearing tree visited by self-explore (columns), for each ABM. (c) Bar plots represent the Jensen–Shannon divergence between the agent- and participant-derived tree transition matrices, with error bars indicating standard errors. (d) Bar plots show the Jensen–Shannon divergence between the distributions of next-tree encounters (conditional on the last tree visited) for agents and their respective human participants.

## 4. Discussion

### 4.1. Dynamic Risk Perception Enhances Adaptive Foraging in Agents

Distinct cognitive models of risk not only exhibited clear divergences in follow-explore dynamics, but the differences in risk evaluation also extended into the process and outcomes of foraging learning. Models that did not incorporate predicted reward updating (AN and PN) performed comparatively fewer fruits discovery and displayed a steeper decline in following behaviour relative to expected reward-updating models. This indicates that updating expectations of reward supports adaptive flexibility in the use of public information, thereby improving learning in foraging contexts. Analyses of navigational behaviour further demonstrated that

models incorporating reward updating effectively acquired and retained stable foraging paths, suggesting that leveraging public information can substantially reduce the cognitive cost of self-exploration. This allows individuals to more rapidly learn the spatial distribution of resource patches in unfamiliar settings. The brain possesses the capacity of generalizing learned information to make predictions about future events (Yon et al., 2020). To improve predictive accuracy, it is essential that learning processes integrate new information to reflect changing environmental conditions (Kobayashi and Okada, 2007). Prediction error—the discrepancy between expected and actual outcomes—plays a crucial role in optimising these learning processes. Accurate computation of prediction error requires that individuals can estimate expected reward, ensuring that subsequent strategy updates remain adaptive. Findings from neuroeconomic research indicate that individuals update expected reward not only based on outcomes of their own choices but also according to beliefs about others' behaviours (Zhu et al., 2012). This is consistent with our results showing that participants revised their expectations of reward from following behaviour based on the avatar performance. Importantly, this implies that even when the outcome of a risky decision remains objectively constant, perceived reward can shift through learning. Such belief updating should be accounted for in experimental designs and analyses focused on information use. Future research paradigms should consider that the reward expectations of agents may vary with perceived reliability of public information, rather than treating reward perception as fixed and externally imposed. Recognising the dynamic nature of expected reward will facilitate the development of more nuanced experimental designs and analytical frameworks, leading to deeper insights into decision-making processes.

The observed differences between the PEIRS model and the risk aversion model highlight the critical role of prediction error induced risk-seeking behaviour in enhancing foraging efficiency. Although both models incorporate similar mechanisms for updating expected rewards, the PEIRS model not only showed a slower decline in following behaviour but also demonstrated superior performance in locating fruit. This trait may support agents in unfamiliar environments to exploit the full utility of available public information, thereby optimising foraging outcomes. In natural settings, such behaviour may further help agents minimise energy expenditure and cognitive load associated with private exploration and spatial memory retrieval (Longman et al., 2023; Reichle, 2020; Rosati, 2017), allowing for greater foraging gains under constrained costs. The observed reversal of risk preferences based on the magnitude of risk and potential reward contributes to behavioural plasticity, enabling individuals to more rapidly adapt their foraging strategies. Risk preference reversals have been documented across a range of ecological decision-making scenarios (Barrett and Fiddick, 2000), from energy budgeting in foraging options (Craft, 2016; Houston and Rosenström, 2024b) to group-beneficial decision strategies (Kusev et al., 2020; Wang, 1996), suggesting that the capacity to incorporate expected value, outcome variance, and personal expectations into risk decisions may have evolved across diverse socioecological contexts.

Previous research has often separated risk and ambiguity evaluations through distinct experimental designs or framing manipulations (Kühberger, 1998; Mishra and Fiddick, 2012). However, our findings suggest that processes previously considered context-specific, such as risk assessments in public information use versus in foraging decisions, may co-occur within the same behavioural setting to facilitate adaptive responses. While complex dynamics of risk-seeking and risk-aversion have been extensively explored and validated in economic contexts, discussions of risk in ecology, particularly foraging ecology, have largely remained within a linear risk evaluation framework. These perspectives frequently frame risk as a cost to be mitigated, such as avoiding predation (Bach et al., 2020; McNamara and Houston, 1992) or increasing risk tolerance under resource scarcity (Mobbs et al., 2018; Rosati and Hare, 2012), thereby overlooking the possibility that risk itself may serve as a motivational force (Li et al., 2019; Loewenstein et al., 2001). Given that many species leverage both intra-specific (Gil et al., 2018) and inter-specific public information (Jackson et al., 2020; Kane and Kendall, 2017) during foraging, dynamic risk-based information use may be key to optimising foraging strategies and enhancing environmental adaptability. Thus, a deeper understanding of the neurobiological underpinnings of agents' risk cognition is essential for elucidating the co-evolution of foraging strategies and decision-making mechanisms. Moving beyond a default view of risk as a cost and instead framing it as a dynamic component of decision-making plasticity will allow for more nuanced interpretations of decision dynamics in foraging and other contexts, ultimately shedding light on the emergence of collective decision-making and the evolution of complex social behaviours from ecological challenges.

The distribution of foraging performance among human participants and their corresponding ABMs indicates a population-level divergence in foraging strategies. This difference further highlights the importance of strategically using reliable public information for learning in unfamiliar environments — although it may result from some agents continuing to follow the avatar despite repeatedly encountering depleted fruit trees and consistently applying this strategy across different sessions. This behaviour resembles overreliance on external navigational aids, such as GPS (Ruginski et al., 2019), which has been shown to impair the formation of internal spatial representations. In our paradigm, such persistent avatar-following may have interfered with the development of landmark-based memory vectors, limiting agents' ability to navigate independently. Consequently, agents that disengaged from the avatar under higher-risk conditions often failed to relocate fruit trees using their own memory, reinforcing a dependence on the avatar and resulting in stagnant foraging success over time. However, the interpretation regarding overreliance on navigational aids is not fully supported by the data. The fact that the constant reward model with lower levels of following still shows slope values close to zero or even negative suggests that this effect is not solely due to following behavior. An alternative explanation lies in the formation of path preferences during early exploratory phases,

potentially leading agents to become locked-in to suboptimal routes with a relatively longer length. Given the abundance of landmarks and the lack of explicit navigational cues prior to fruit tree discovery, agents may have failed to encode efficient spatial representations. Randomized avatar trajectories further increased the likelihood that early encounters involved inefficient paths or fruit trees that were visually salient but poorly integrated within the broader spatial layout. These early experiences may have led agents to reinforce memory vectors associated with suboptimal, yet repeatedly rewarding, paths (e.g., route 7 → 6), thereby entrenching less effective search strategies. Each successful return to a fruit tree along a suboptimal path served as a positive reinforcement, increasing the likelihood of its future selection. In contrast, exploratory behaviour aimed at identifying globally optimal routes involved higher short-term uncertainty and risk of failure, rendering it less appealing. The distribution of individual-level learning outcomes may reflect divergent developmental trajectories shaped by stochastic differences in early exploratory experience: one group of agents fortuitously identified effective paths linking multiple resource sites and built robust spatial maps, while another group fixated on locally reinforcing but globally inefficient trajectories, thereby limiting their overall learning potential.

This interpretation also provides insight into the complex performance patterns observed across agents in foraging learning. The comparatively poorer learning performance of the AU model relative to the AN model suggests that while reliance on reliable public information can enhance initial food discovery in unfamiliar environments, it does not necessarily facilitate spatial learning. Although AN agents found fewer food items overall, partly due to consistently lower and more rapidly declining follow rates across all sessions, they may have benefited from increased opportunities for individual exploration. This exploratory behaviour likely allowed them to acquire a greater number of memory vectors, resulting in superior performance in learning foraging paths. Notably, AN agents did not discover any shortest paths; however, they exhibited the fastest consolidation of foraging-route composition, suggesting a pattern of gradual but persistent improvement in foraging efficiency. Furthermore, the emergence of a subset of both PU agents and human participants who exhibited accelerated learning despite being fewer in number points to the possibility that early access to particularly reliable public information may have coincidentally included optimal paths. This interpretation is further supported by the gradient in performance across human participants, PU agents, and AU agents, with PU agents and humans discovering a greater number of optimal routes. These results indicate that cognitive inertia in route selection, when interacting with varying degrees of information availability, can yield both adaptive and maladaptive outcomes in foraging contexts. Thus, although updating expected reward alone may improve the efficiency of information use in novel environments, achieving optimal foraging likely requires a more nuanced integration of public information and spatial learning. These findings underscore the importance of bounded rationality in shaping navigational decisions and highlight cognitive inertia as a key feature of foraging behaviour. Future research should investigate the influence of bounded rationality on

spatial learning of agents to better understand the emergence of cognitive inertia during foraging decision-making and its potential evolutionary origins.

#### **4.2. Social Implications of Dynamic Risk Perception in Public Information Use**

Beyond its implications for decision-making and foraging ecology, the risk-based cognition of information use demonstrated in this study offers an evolutionary perspective on pressing societal issues related to the spread of misinformation. With the rise of the internet and digital media, access to novel information has become unprecedentedly effortless. Concurrently, the dissemination of misinformation and disinformation has grown more pervasive than at any point in history. A 2024 global survey reported that approximately 86% of individuals had encountered false or misleading news (Press Room, 2025). This phenomenon is further exacerbated by a well-documented human tendency to preferentially attend to negative information, commonly referred to as the “negativity bias.”(Corns, 2018) While the origins of negativity bias have been examined from various angles, including emotional appraisal (Brown et al., 2017), social cognition (Fiske, 2018), and developmental psychology (Vaish et al., 2008), its evolutionary roots remain uncertain. Moreover, the presence of documented positivity biases in some contexts challenges the assumed universality of negativity bias (Kauschke et al., 2019). Our findings offer a novel framework for interpreting biases in information transmission as potentially contingent on domain-specific cost–benefit trade-offs. Negativity/positivity biases may arise from domain-general mechanisms of strategic information sampling, which evolved in the context of ancestral foraging environments.

In decision-making environments analogous to those explored in this study, agents engage with novel information under uncertainty, where the outcomes of using such information are probabilistic and must be learned through experience, rather than being fully known in advance. This is similar to the process of navigating unfamiliar digital information, where individuals must assess the reliability of new inputs and decide whether to incorporate them into their behavioural strategies. Many instances of misinformation are difficult to falsify in daily life, meaning that individuals may not receive adequate negative feedback to learn the risks associated with believing such content. The lack of corrective signals may underlie the persistence of false beliefs, which may lead to barriers in belief revision (Ecker et al., 2022). Evidence suggests that negativity bias may enhance loss aversion in decision-making (Molins et al., 2022), leading individuals to more rapidly internalise the risks of using accurate information when it produces negative outcomes, potentially eroding trust in credible sources. The low baseline expectation, caused by vague or poorly substantiated information (Williams-Ceci et al., 2024), may increase the likelihood of occasional positive surprises, thereby reinforcing belief in high-risk, unreliable content. Together, these insights suggest that differences in expected reward associated with information sources may shape belief formation via mechanisms akin to risk preference reversal. This provides a new angle for future research—spanning communication science, neuroscience, and behavioural ecology—to investigate how populations come to

distrust reliable information while placing unwarranted confidence in misinformation. Given the profound societal impacts of misinformation, ranging from threats to democratic integrity (Ecker et al., 2024) and public health (Sylvia Chou et al., 2020) to impairments in disaster response (Muhammed T and Mathew, 2022), incorporating interdisciplinary and evolutionary frameworks into the study of information transmission could improve current strategies for risk communication and information governance, which may ultimately support the development of more resilient and sustainable social systems.

#### **4.3. Limitations, Applications, and Conclusions**

The robust performance of ABMs in simulating human behaviour, as demonstrated in this study, underscores their potential to support cognitive experimental design by reducing trial-and-error costs and generating testable hypotheses. ABMs have long held promise for elucidating complex, emergent phenomena and thus play a valuable role in the development of experimental paradigms. In recent years, the social sciences have increasingly adopted generative agent-based modelling approaches to replicate large-scale societal systems (Bail, 2024), offering insights into group-level behaviours and informing the design of broader experimental interventions. These models not only enhance researchers' understanding of phenomena such as collective decision-making and social network dynamics but also offer predictive insights into system behaviour when empirical data are sparse or participant recruitment is constrained (Vezhnevets et al., 2023). In behavioural ecology, particularly in the emerging field of cognitive ecology, researchers face analogous challenges in designing experiments to explore complex behavioural systems. Field studies are often constrained by temporal, logistical, and environmental uncertainties, with high sunk costs incurred from erroneous assumptions or suboptimal experimental designs (Rafiq et al., 2024). As ecological crises intensify and the pressure for publication requirement (Ríos-Saldaña et al., 2018), the time costs of unsuccessful research endeavours are becoming increasingly consequential for both knowledge production and academic career progression. Integrating ABMs into behavioural ecological research may mitigate these costs by aiding in the pre-selection of empirically tractable hypotheses and fostering more robust experimental designs. In fact, ecological studies have begun to demonstrate the utility of ABMs for simulating real-world individual behaviour. For example, Weller et al. (2023) showed that ABMs could effectively model movement decisions in mallards (*Anas platyrhynchos*), thereby assisting in the identification of critical habitat features and environmental thresholds. However, few studies have yet leveraged contemporary insights from cognitive science and neuroscience to build neurocomputational ABM capable of capturing the emergence of behavioural patterns in ecological contexts. The present findings demonstrate that ABMs incorporating biologically grounded cognitive architectures can accurately reproduce multilevel cognitive traits and decision-making strategies, offering a promising avenue for future research. As computational capacity and neuroscientific understanding continue to advance, the development of more comprehensive neurobiological modelling frameworks will become increasingly

feasible. Consequently, future research in cognitive ecology should consider integrating ABMs not only to optimise field study designs but also to generate new, testable hypotheses grounded in mechanistic models of cognition.

While the integration of VR and ABM has proven effective in advancing our understanding of the cognitive mechanisms and learning processes underpinning foraging behaviour, particularly in elucidating human risk perception in information use and its potential evolutionary links with hominid development, the results from chimpanzee participants reveal important limitations in the current approach. Specifically, chimpanzees exhibited minimal evidence of foraging-related learning and performed markedly worse than humans in locating fruit. Furthermore, the ABM failed to capture their foraging strategies and follow-explore dynamics with comparable accuracy. Given the well-documented foraging competencies of chimpanzees in natural settings, including sophisticated spatiotemporal memory (Janmaat et al., 2014), use of social cues (Harrison et al., 2021), and route optimization (Green et al., 2020). These findings are unlikely to reflect an absence of cognitive ability. Instead, the discrepancy is more plausibly attributed to difficulties in interpreting and interacting with the VR environment. Observational data (Appendix 3) suggest that sometimes chimpanzee participants did not engage with the task according to its intended goal—maximising fruit collection within a constrained timeframe. Rather, individuals frequently engaged in repetitive movements around specific landmarks. While this might partially stem from challenges in touchscreen control, the observed capacity for straight-line navigation implies a broader attempt to explore and make sense of the virtual environment's affordances. As captive individuals housed in zoological institutions, these chimpanzees do not face immediate foraging pressures. Thus, their motivation may have been exploratory rather than effectively engaging in what could be described as a qualitatively different “game.” Elements such as fruit falling when approached or the navigational utility of dense landmark configuration, are not directly analogous to naturalistic foraging but instead reflect conventions common in video game design (Tekinbas and Zimmerman, 2003). This suggests that human participants may have benefited from prior conceptual familiarity, whereas chimpanzees lacked this experiential framework. Consequently, chimpanzees may require extended exposure to such tasks and a more gradual learning progression—beginning with simplified, ecologically grounded game mechanics and advancing toward more complex, naturalistic simulations—to appropriately engage with experimental objectives. Addressing the current limitations of VR-based methods is essential for developing more inclusive and ecologically valid research frameworks. Additionally, the exploratory and interactive behaviours exhibited by chimpanzees toward salient landmarks may suggest that strong visual stimuli inherently elicit dopaminergic activation (Dommett et al., 2005) and emotional reward (Kauvar et al., 2025) in captive individuals. Future research should further investigate chimpanzees' acclimation to video game-based tasks and their experiential engagement, to inform improvements in animal welfare and the development of effective cognitive enrichment strategies.

VR environments and ABMs inherently represent simplified abstractions of natural ecosystems and the organisms within them. As such, they are inevitably limited in their capacity to fully capture the complex, reciprocal interactions between landscape features and animal foraging behaviour. Nonetheless, the strong alignment between the PU model and empirical human data, particularly in the follow-explore dynamics and overall foraging efficiency, demonstrates the model's potential to accurately simulate key aspects of human foraging cognition. Importantly, examining areas where the model fails to replicate empirical patterns can yield critical insights into other potential drivers of animal movement and decision-making (Gallagher et al., 2021; Green et al., 2020). Certain discrepancies may stem from an incomplete understanding of the cognitive processes underpinning landmark salience perception. This limitation likely contributed to the relatively poor performance of the model in replicating navigation-related behaviours, as evidenced by the diffuse transition matrix and absence of strongly preferred paths. Moreover, constraints in computational capacity necessitated the simplification of the underlying FAM module, potentially impairing the model's ability to simulate the adaptive nature of spatial learning. This simplification may have restricted the agent's capacity to exhibit flexible navigation strategies in response to environmental contingencies. Improving the ecological and cognitive realism of such models will require more precise experimental assessment of landmark perception in study subjects. Understanding how different organisms evaluate environmental salience can inform the development of more accurate cognitive architectures within ABMs. Additionally, advances in high-throughput computing and the increasing accessibility of quantum computing (Woolnough et al., 2023) offer promising avenues for enhancing the fidelity of ABMs based on neurocomputational principles. As computational technologies continue to evolve, they are poised to open new frontiers in cognitive ecology, offering researchers unprecedented opportunities to simulate, manipulate, and predict complex cognitive-environmental interactions at scale.

How animals make decisions under uncertainty is a fundamental question in biology. Our findings reveal that risk perception is not a static shield against danger but a dynamic computational tool for optimising learning and resource acquisition in novel environments. By combining VR foraging game in humans and chimpanzees with neurocomputational ABM, the study demonstrates that the capacity to flexibly update expectations based on prediction error is critical for successful foraging. Notably, the results reveal that risk-seeking behavior, specifically triggered by positive prediction errors when outcomes exceed expectations, confers a significant adaptive advantage. This mechanism promotes the sustained use of public information, allowing individuals to more rapidly learn the resource landscape while minimizing the cognitive and energetic costs of autonomous exploration. Our results suggest that the dynamic reversal of risk preference is a deeply conserved adaptive trait that likely co-evolved with social learning and decision-making across diverse ecological contexts. The high fidelity of our neurocomputational ABM in simulating human

cognitive and behavioral patterns validates its use as a powerful tool for generating and pre-testing hypotheses, particularly valuable in cognitive ecology research where long-term fieldwork remains challenging and resource-intensive. This integrative approach reveals the risk evaluation of information potentially emerged from foraging context, providing interdisciplinary avenues for understanding the evolution of information use and tackling its modern societal consequences.

## References

- Allritz, M., Call, J., Schweller, K., McEwen, E.S., de Guinea, M., Janmaat, K.R.L., Menzel, C.R., Dolins, F.L., 2022. Chimpanzees (*Pan troglodytes*) navigate to find hidden fruit in a virtual environment. *Science Advances* 8, eabm4754. <https://doi.org/10.1126/sciadv.abm4754>
- Altmann, J., 1988. Primate Societies. Barbara B. Smuts, Dorothy L. Cheney, Robert M. Seyfarth, Richard W. Wrangham, and Thomas T. Struhsaker, Eds. University of Chicago Press, Chicago, IL, 1987. xii, 578 pp. *Science* 240, 1076–1078. <https://doi.org/10.1126/science.240.4855.1076.b>
- Apicella, C., Norenzayan, A., Henrich, J., 2020. Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior, Beyond Weird* 41, 319–329. <https://doi.org/10.1016/j.evolhumbehav.2020.07.015>
- Bach, D.R., Moutoussis, M., Bowler, A., Dolan, R.J., 2020. Predictors of risky foraging behaviour in healthy young people. *Nat Hum Behav* 4, 832–843. <https://doi.org/10.1038/s41562-020-0867-0>
- Bail, C.A., 2024. Can Generative AI improve social science? *Proceedings of the National Academy of Sciences* 121, e2314021121. <https://doi.org/10.1073/pnas.2314021121>
- Banerjee, K., C, V.P., Gupta, R.R., Vyas, K., H, A., Mishra, B., 2025. Exploring Alternatives to Softmax Function. Presented at the 2nd International Conference on Deep Learning Theory and Applications, pp. 81–86.
- Barceló, J.A., Bernal, F.D.C., Olmo, R. del, Mameli, L., Quesada, F.J.M., Poza, D., Vilà, X., 2014. Social Interaction in Hunter-Gatherer Societies: Simulating the Consequences of Cooperation and Social Aggregation. *Social Science Computer Review* 32, 417–436. <https://doi.org/10.1177/0894439313511943>
- Barrett, H.C., Fiddick, L., 2000. Evolution and risky decisions. *Trends in Cognitive Sciences* 4, 251–252. [https://doi.org/10.1016/S1364-6613\(00\)01495-9](https://doi.org/10.1016/S1364-6613(00)01495-9)
- Beauchamp, G., 2008. A spatial model of producing and scrounging. *Animal Behaviour* 76, 1935–1942. <https://doi.org/10.1016/j.anbehav.2008.08.017>
- Bracken, A.M., Christensen, C., O'Riain, M.J., Fürtbauer, I., King, A.J., 2022. Flexible group cohesion and coordination, but robust leader–follower roles, in a wild social primate using urban space. *Proceedings of the Royal Society B: Biological Sciences* 289, 20212141. <https://doi.org/10.1098/rspb.2021.2141>
- Brown, C.C., Raio, C.M., Neta, M., 2017. Cortisol responses enhance negative valence perception for ambiguous facial expressions. *Sci Rep* 7, 15107.

- <https://doi.org/10.1038/s41598-017-14846-3>
- Clark, L., Cools, R., Robbins, T.W., 2004. The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition, Development of Orbitofrontal Function* 55, 41–53. [https://doi.org/10.1016/S0278-2626\(03\)00284-7](https://doi.org/10.1016/S0278-2626(03)00284-7)
- Clay, Z., Smith, C.L., Blumstein, D.T., 2012. Food-associated vocalizations in mammals and birds: what do these calls really mean? *Animal Behaviour* 83, 323–330. <https://doi.org/10.1016/j.anbehav.2011.12.008>
- Corns, J., 2018. Rethinking the Negativity Bias. *Rev Philos Psychol* 9, 607–625. <https://doi.org/10.1007/s13164-018-0382-7>
- Couzin, I.D., 2006. Behavioral Ecology: Social Organization in Fission–Fusion Societies. *Current Biology* 16, R169–R171. <https://doi.org/10.1016/j.cub.2006.02.042>
- Craft, B.B., 2016. Risk-sensitive foraging: changes in choice due to reward quality and delay. *Animal Behaviour* 111, 41–47. <https://doi.org/10.1016/j.anbehav.2015.09.030>
- d’Acremont, M., Lu, Z.-L., Li, X., Van der Linden, M., Bechara, A., 2009. Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage* 47, 1929–1939. <https://doi.org/10.1016/j.neuroimage.2009.04.096>
- da Silva, J.A., Tecuapetla, F., Paixão, V., Costa, R.M., 2018. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* 554, 244–248. <https://doi.org/10.1038/nature25457>
- Dall, S.R.X., Giraldeau, L.-A., Olsson, O., McNamara, J.M., Stephens, D.W., 2005. Information and its use by animals in evolutionary ecology. *Trends in Ecology & Evolution* 20, 187–193. <https://doi.org/10.1016/j.tree.2005.01.010>
- Danchin, É., Giraldeau, L.-A., Valone, T.J., Wagner, R.H., 2004. Public Information: From Nosy Neighbors to Cultural Evolution. *Science* 305, 487–491. <https://doi.org/10.1126/science.1098254>
- De Petrillo, F., Rosati, A.G., 2021. Variation in primate decision-making under uncertainty and the roots of human economic behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, 20190671. <https://doi.org/10.1098/rstb.2019.0671>
- Delamater, A.R., Oakeshott, S., 2007. Learning about Multiple Attributes of Reward in Pavlovian Conditioning. *Annals of the New York Academy of Sciences* 1104, 1–20. <https://doi.org/10.1196/annals.1390.008>
- Dm, F., E, K., A, M., C, M., G, B., J, J.-P., W, H., 2009. Navigating two-dimensional mazes: chimpanzees (*Pan troglodytes*) and capuchins (*Cebus apella sp.*) profit from experience differently. *Animal cognition* 12. <https://doi.org/10.1007/s10071-008-0210-z>
- Dolins, F.L., Klimowicz, C., Kelley, J., Menzel, C.R., 2014. Using virtual reality to investigate comparative spatial cognitive abilities in chimpanzees and humans. *American Journal of Primatology* 76, 496–513. <https://doi.org/10.1002/ajp.22252>

- Dominic Schuhmacher, Björn Bähre, Nicolas Bonneel, Carsten Gottschlich, Valentin Hartmann, Florian Heinemann, Bernhard Schmitzer and Jörn Schrieber 2024. transport: Computation of Optimal Transport Plans and Wasserstein Distances. R package version 0.15-4. <https://cran.r-project.org/package=transport>
- Dommett, E., Coizet, V., Blaha, C.D., Martindale, J., Lefebvre, V., Walton, N., Mayhew, J.E.W., Overton, P.G., Redgrave, P., 2005. How Visual Stimuli Activate Dopaminergic Neurons at Short Latency. *Science* 307, 1476–1479. <https://doi.org/10.1126/science.1107026>
- Ecker, U., Roozenbeek, J., van der Linden, S., Tay, L.Q., Cook, J., Oreskes, N., Lewandowsky, S., 2024. Misinformation poses a bigger threat to democracy than you might think. *Nature* 630, 29–32. <https://doi.org/10.1038/d41586-024-01587-3>
- Ecker, U.K.H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L.K., Brashier, N., Kendeou, P., Vraga, E.K., Amazeen, M.A., 2022. The psychological drivers of misinformation belief and its resistance to correction. *Nat Rev Psychol* 1, 13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- Epstein, R.A., Vass, L.K., 2014. Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 20120533. <https://doi.org/10.1098/rstb.2012.0533>
- Erasmus, A., Holman, B., Ioannidis, J.P.A., 2022. Data-dredging bias. *BMJ Evidence-Based Medicine* 27, 209–211. <https://doi.org/10.1136/bmjebm-2020-111584>
- Etienne, A.S., Maurer, R., Boulens, V., Levy, A., Rowe, T., 2004. Resetting the path integrator: a basic condition for route-based navigation. *Journal of Experimental Biology* 207, 1491–1508. <https://doi.org/10.1242/jeb.00906>
- Fahimipour, A.K., Gil, M.A., Celis, M.R., Hein, G.F., Martin, B.T., Hein, A.M., 2023. Wild animals suppress the spread of socially transmitted misinformation. *Proceedings of the National Academy of Sciences* 120, e2215428120. <https://doi.org/10.1073/pnas.2215428120>
- Filippo Aureli, by, M. Schaffner, C., Boesch, C., K. Bearder, S., Call, J., A. Chapman, C., Connor, R., Di Fiore, A., I. M. Dunbar, R., Peter Henzi, S., Holekamp, K., H. Korstjens, A., Layton, R., Lee, P., Lehmann, J., H. Manson, J., Ramos-Fernandez, G., B. Strier, K., P. van Schaik, C., 2008. Fission-Fusion Dynamics. *Current Anthropology*. <https://doi.org/10.1086/586708>
- Fiske, S., 2018. Social Cognition: Selected Works of Susan Fiske. Routledge.
- Flack, A., Pettit, B., Freeman, R., Guilford, T., Biro, D., 2012. What are leaders made of? The role of individual experience in determining leader–follower relations in homing pigeons. *Animal Behaviour* 83, 703–709. <https://doi.org/10.1016/j.anbehav.2011.12.018>
- Frank, S.L., 2013. Uncertainty Reduction as a Measure of Cognitive Load in Sentence Comprehension. *Topics in Cognitive Science* 5, 475–494. <https://doi.org/10.1111/tops.12025>
- Gallagher, C.A., Chudzinska, M., Larsen-Gray, A., Pollock, C.J., Sells, S.N., White,

- P.J.C., Berger, U., 2021. From theory to practice in pattern-oriented modelling: identifying and using empirical patterns in predictive models. *Biological Reviews* 96, 1868–1888. <https://doi.org/10.1111/brv.12729>
- Garber, P.A., Dolins, F.L., 2014. Primate spatial strategies and cognition: Introduction to this special issue. *American Journal of Primatology* 76, 393–398. <https://doi.org/10.1002/ajp.22257>
- Garcia, C., Bouret, S., Druelle, F., Prat, S., 2021. Balancing costs and benefits in primates: ecological and palaeoanthropological views. *Philos Trans R Soc Lond B Biol Sci* 376, 20190667. <https://doi.org/10.1098/rstb.2019.0667>
- Garg, K., Kello, C.T., Smaldino, P.E., 2022. Individual exploration and selective social learning: balancing exploration-exploitation trade-offs in collective foraging. *Journal of The Royal Society Interface* 19, 20210915. <https://doi.org/10.1098/rsif.2021.0915>
- Gil, M.A., Hein, A.M., Spiegel, O., Baskett, M.L., Sih, A., 2018. Social Information Links Individual Behavior to Population and Community Dynamics. *Trends in Ecology & Evolution* 33, 535–548. <https://doi.org/10.1016/j.tree.2018.04.010>
- Gilby, I.C., Machanda, Z.P., 2022. Advanced cognition in wild chimpanzees: lessons from observational studies. *Current Opinion in Behavioral Sciences* 46, 101183. <https://doi.org/10.1016/j.cobeha.2022.101183>
- Giraldeau, L.-A., Caraco, T., 2000. Social Foraging Theory. Princeton University Press.
- Green, S.J., Boruff, B.J., Bonnell, T.R., Grueter, C.C., 2020. Chimpanzees Use Least-Cost Routes to Out-of-Sight Goals. *Current Biology* 30, 4528-4533.e5. <https://doi.org/10.1016/j.cub.2020.08.076>
- Gribkova, E.D., Chowdhary, G., Gillette, R., 2024. Cognitive mapping and episodic memory emerge from simple associative learning rules. *Neurocomputing* 595, 127812. <https://doi.org/10.1016/j.neucom.2024.127812>
- Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: A review and first update. *Ecological Modelling* 221, 2760–2768. <https://doi.org/10.1016/j.ecolmodel.2010.08.019>
- Grimm, V., Railsback, S.F., Vincenot, C.E., Berger, U., Gallagher, C., DeAngelis, D.L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A.S.A., Milles, A., Nabe-Nielsen, J., Polhill, J.G., Radchuk, V., Rohwäder, M.-S., Stillman, R.A., Thiele, J.C., Ayllón, D., 2020. The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *JASSS* 23, 7.
- Grove, M., Pearce, E., Dunbar, R.I.M., 2012. Fission-fusion and the evolution of hominin social systems. *Journal of Human Evolution* 62, 191–200. <https://doi.org/10.1016/j.jhevol.2011.10.012>
- Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J., Berke, J.D., 2016. Mesolimbic dopamine signals the value of work. *Nat Neurosci* 19, 117–126. <https://doi.org/10.1038/nn.4173>
- Harel, R., Spiegel, O., Getz, W.M., Nathan, R., 2017. Social foraging and individual

- consistency in following behaviour: testing the information centre hypothesis in free-ranging vultures. *Proc Biol Sci* 284, 20162654. <https://doi.org/10.1098/rspb.2016.2654>
- Harrison, R.A., van Leeuwen, E.J.C., Whiten, A., 2021. Chimpanzees' behavioral flexibility, social tolerance, and use of tool-composites in a progressively challenging foraging problem. *iScience* 24, 102033. <https://doi.org/10.1016/j.isci.2021.102033>
- Herd, S., Krueger, K., Nair, A., Mollick, J., O'Reilly, R., 2021. Neural Mechanisms of Human Decision-Making. *Cogn Affect Behav Neurosci* 21, 35–57. <https://doi.org/10.3758/s13415-020-00842-0>
- Hesslow, G., 2012. The current status of the simulation theory of cognition. *Brain Research, The Cognitive Neuroscience of Thought* 1428, 71–79. <https://doi.org/10.1016/j.brainres.2011.06.026>
- 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, 46–54. <https://doi.org/10.1016/j.tics.2014.10.004>
- Houston, A.I., Rosenström, T.H., 2024a. A critical review of risk-sensitive foraging. *Biological Reviews* 99, 478–495. <https://doi.org/10.1111/brv.13031>
- Houston, A.I., Rosenström, T.H., 2024b. A critical review of risk-sensitive foraging. *Biological Reviews* 99, 478–495. <https://doi.org/10.1111/brv.13031>
- Jackson, C.R., Maddox, T., Mbise, F.P., Stokke, B.G., Belant, J.L., Bevanger, K., Durant, S.M., Fyumagwa, R., Ranke, P.S., Røskaft, E., May, R., Fossøy, F., 2020. A dead giveaway: Foraging vultures and other avian scavengers respond to auditory cues. *Ecol Evol* 10, 6769–6774. <https://doi.org/10.1002/ece3.6366>
- Janmaat, K.R.L., 2019. What animals do not do or fail to find: A novel observational approach for studying cognition in the wild. *Evolutionary Anthropology: Issues, News, and Reviews* 28, 303–320. <https://doi.org/10.1002/evan.21794>
- Janmaat, K.R.L., Guinea, M. de, Collet, J., Byrne, R.W., Robira, B., Loon, E. van, Jang, H., Biro, D., Ramos-Fernández, G., Ross, C., Presotto, A., Allritz, M., Alavi, S., Belle, S.V., 2021. Using natural travel paths to infer and compare primate cognition in the wild. *iScience* 24. <https://doi.org/10.1016/j.isci.2021.102343>
- Janmaat, K.R.L., Polansky, L., Ban, S.D., Boesch, C., 2014. Wild chimpanzees plan their breakfast time, type, and location. *Proceedings of the National Academy of Sciences* 111, 16343–16348. <https://doi.org/10.1073/pnas.1407524111>
- Janson, C.H., Byrne, R., 2007. What wild primates know about resources: opening up the black box. *Anim Cogn* 10, 357–367. <https://doi.org/10.1007/s10071-007-0080-9>
- Janssen, M.A., Hill, K., 2014. Benefits of Grouping and Cooperative Hunting Among Ache Hunter-Gatherers: Insights from an Agent-Based Foraging Model. *Hum Ecol* 42, 823–835. <https://doi.org/10.1007/s10745-014-9693-1>
- Kane, A., Kendall, C.J., 2017. Understanding how mammalian scavengers use information from avian scavengers: cue from above. *Journal of Animal Ecology* 86, 837–846. <https://doi.org/10.1111/1365-2656.12663>

- Kassambara, A., 2023. *ggsnubr: 'ggplot2' Based Publication Ready Plots* . R package version 0.6.0, <<https://CRAN.R-project.org/package=ggsnubr>>.
- Kauschke, C., Bahn, D., Vesker, M., Schwarzer, G., 2019. The Role of Emotional Valence for the Processing of Facial and Verbal Stimuli-Positivity or Negativity Bias? *Front Psychol* 10, 1654. <https://doi.org/10.3389/fpsyg.2019.01654>
- Kauvar, I., Richman, E.B., Liu, T.X., Li, C., Vesuna, S., Chibukhchyan, A., Yamada, L., Fogarty, A., Solomon, E., Choi, E.Y., Mortazavi, L., Chau Loo Kung, G., Mukunda, P., Raja, C., Gil-Hernández, D., Patron, K., Zhang, X., Brawer, J., Wrobel, S., Lusk, Z., Lyu, D., Mitra, A., Hack, L., Luo, L., Grosenick, L., van Roessel, P., Williams, L.M., Heifets, B.D., Henderson, J.M., McNab, J.A., Rodríguez, C.I., Buch, V., Nuyujukian, P., Deisseroth, K., 2025. Conserved brain-wide emergence of emotional response from sensory experience in humans and mice. *Science* 388, eadt3971. <https://doi.org/10.1126/science.adt3971>
- Kobayashi, Y., Okada, K.-I., 2007. Reward Prediction Error Computation in the Pedunculopontine Tegmental Nucleus Neurons. *Annals of the New York Academy of Sciences* 1104, 310–323. <https://doi.org/10.1196/annals.1390.003>
- Koops, M.A., Giraldeau, L.-A., 1996. Producer-scrounger foraging games in starlings: a test of rate-maximizing and risk-sensitive models. *Animal Behaviour* 51, 773–783. <https://doi.org/10.1006/anbe.1996.0082>
- Kühberger, A., 1998. The Influence of Framing on Risky Decisions: A Meta-analysis. *Organ Behav Hum Decis Process* 75, 23–55. <https://doi.org/10.1006/obhd.1998.2781>
- Kummer, H., 1971. Primate societies: group techniques of ecological adaptation. Arlington Heights, Ill.: AHM Pub.
- Kusev, P., van Schaik, P., Martin, R., Hall, L., Johansson, P., 2020. Preference reversals during risk elicitation. *J Exp Psychol Gen* 149, 585–589. <https://doi.org/10.1037/xge0000655>
- Li, M., Lauharatanahirun, N., Steinberg, L., King-Casas, B., Kim-Spoon, J., Deater-Deckard, K., 2019. Longitudinal link between trait motivation and risk-taking behaviors via neural risk processing. *Dev Cogn Neurosci* 40, 100725. <https://doi.org/10.1016/j.dcn.2019.100725>
- Lichtenberg, E.M., Heiling, J.M., Bronstein, J.L., Barker, J.L., 2020. Noisy communities and signal detection: why do foragers visit rewardless flowers? *Philosophical Transactions of the Royal Society B: Biological Sciences* 375, 20190486. <https://doi.org/10.1098/rstb.2019.0486>
- Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N., 2001. Risk as feelings. *Psychol Bull* 127, 267–286. <https://doi.org/10.1037/0033-2909.127.2.267>
- Longman, D.P., Wells, J.C.K., Stock, J.T., 2023. Human energetic stress associated with upregulation of spatial cognition. *American Journal of Biological Anthropology* 182, 32–44. <https://doi.org/10.1002/ajpa.24820>
- Madsen, A., de Silva, S., 2024. Societies with fission-fusion dynamics as complex adaptive systems: the importance of scale. *Philosophical Transactions of the*

- Royal Society B: Biological Sciences 379, 20230175.  
<https://doi.org/10.1098/rstb.2023.0175>
- Mao, X., Xia, C., Liu, J., Zhang, H., Ding, Y., Yao, Y., Liu, Z., 2025. A novel similarity measure based on dispersion-transition matrix and Jensen–Fisher divergence and its application on the detection of rail short-wave defects. Chaos, Solitons & Fractals 192, 115988.  
<https://doi.org/10.1016/j.chaos.2025.115988>
- McLinn, C.M., Stephens, D.W., 2006. What makes information valuable: signal reliability and environmental uncertainty. Animal Behaviour 71, 1119–1129.  
<https://doi.org/10.1016/j.anbehav.2005.09.006>
- McNamara, J.M., Houston, A.I., 1992. Risk-sensitive foraging: A review of the theory. Bulletin Mathematical Biology 54, 355–378. <https://doi.org/10.1007/BF02464838>
- Menzel Jr., E.W., Menzel, C.R., 2007. Do Primates Plan Routes? Simple Detour Problems Reconsidered, in: Primate Perspectives on Behavior and Cognition. American Psychological Association, Washington, DC, US, pp. 175–206.  
<https://doi.org/10.1037/11484-014>
- Mishra, S., Fiddick, L., 2012. Beyond gains and losses: The effect of need on risky choice in framed decisions. Journal of Personality and Social Psychology 102, 1136–1147. <https://doi.org/10.1037/a0027855>
- Mobbs, D., Trimmer, P.C., Blumstein, D.T., Dayan, P., 2018. Foraging for foundations in decision neuroscience: insights from ethology. Nat Rev Neurosci 19, 419–427. <https://doi.org/10.1038/s41583-018-0010-7>
- Moeller, M., Grohn, J., Manohar, S., Bogacz, R., 2021. An association between prediction errors and risk-seeking: Theory and behavioral evidence. PLoS Comput Biol 17, e1009213. <https://doi.org/10.1371/journal.pcbi.1009213>
- Molins, F., Martínez-Tomás, C., Serrano, M.Á., 2022. Implicit Negativity Bias Leads to Greater Loss Aversion and Learning during Decision-Making. Int J Environ Res Public Health 19, 17037. <https://doi.org/10.3390/ijerph192417037>
- Muhammed T, S., Mathew, S.K., 2022. The disaster of misinformation: a review of research in social media. Int J Data Sci Anal 13, 271–285.  
<https://doi.org/10.1007/s41060-022-00311-6>
- Murthy, V.N., 2024. Olfactory navigation in fluctuating environments. Current Biology 34, R1013–R1018. <https://doi.org/10.1016/j.cub.2024.07.049>
- Nilsen, E.B., Bowler, D.E., Linnell, J.D.C., 2020. Exploratory and confirmatory research in the open science era. Journal of Applied Ecology 57, 842–847.  
<https://doi.org/10.1111/1365-2664.13571>
- Normand, E., Boesch, C., 2009. Sophisticated Euclidean maps in forest chimpanzees. Animal Behaviour 77, 1195–1201.  
<https://doi.org/10.1016/j.anbehav.2009.01.025>
- Nunn, C.L., Barton, R.A., 2001. Comparative methods for studying primate adaptation and allometry. Evolutionary Anthropology: Issues, News, and Reviews 10, 81–98. <https://doi.org/10.1002/evan.1019>
- Ortega, P.A., Braun, D.A., 2013. Thermodynamics as a theory of decision-making with information-processing costs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 379, 20230175.  
<https://doi.org/10.1098/rsta.2023.0175>

- Mathematical, Physical and Engineering Sciences 469, 20120683. <https://doi.org/10.1098/rspa.2012.0683>
- Pelé, M., Broihanne, M.H., Thierry, B., Call, J., Dufour, V., 2014. To bet or not to bet? Decision-making under risk in non-human primates. *J Risk Uncertain* 49, 141–166. <https://doi.org/10.1007/s11166-014-9202-3>
- Penar, W., Magiera, A., Klocek, C., 2020. Applications of bioacoustics in animal ecology. *Ecological Complexity* 43, 100847. <https://doi.org/10.1016/j.ecocom.2020.100847>
- Portugal, S.J., White, C.R., 2018. Miniaturization of biologgers is not alleviating the 5% rule. *Methods in Ecology and Evolution* 9, 1662–1666. <https://doi.org/10.1111/2041-210X.13013>
- Post, D.J. van der, Semmann, D., 2011. Local Orientation and the Evolution of Foraging: Changes in Decision Making Can Eliminate Evolutionary Trade-offs. *PLOS Computational Biology* 7, e1002186. <https://doi.org/10.1371/journal.pcbi.1002186>
- Press Room, 2025. The Impact of Fake News: Statistics by Country, Artificial Intelligence Influence, and Misinformation Trends | DISA. URL <https://disa.org/the-impact-of-fake-news-statistics-by-country-artificial-intelligence-influence-and-misinformation-trends/> (accessed 6.6.25).
- Pretelli, I., Ringen, E., Lew-Levy, S., 2022. Foraging complexity and the evolution of childhood. *Science Advances* 8, eabn9889. <https://doi.org/10.1126/sciadv.abn9889>
- Preuschoff, K., Bossaerts, P., 2007. Adding Prediction Risk to the Theory of Reward Learning. *Annals of the New York Academy of Sciences* 1104, 135–146. <https://doi.org/10.1196/annals.1390.005>
- Rafiq, K., Jordan, N.R., McNutt, J.W., Neelo, J., Attias, N., Boersma, D., Palmer, M.S., Ruesink, J., Abrahms, B., 2024. The value of field research in academia. *Science* 384, 855–856. <https://doi.org/10.1126/science.ad06937>
- Ramos-Fernandez, G., King, A.J., Beehner, J.C., Bergman, T.J., Crofoot, M.C., Di Fiore, A., Lehmann, J., Schaffner, C.M., Snyder-Mackler, N., Zuberbühler, K., Aureli, F., Boyer, D., 2018. Quantifying uncertainty due to fission–fusion dynamics as a component of social complexity. *Proceedings of the Royal Society B: Biological Sciences* 285, 20180532. <https://doi.org/10.1098/rspb.2018.0532>
- Reichle, D.E., 2020. Chapter 6 - Species adaptations to their energy environment, in: Reichle, D.E. (Ed.), *The Global Carbon Cycle and Climate Change*. Elsevier, pp. 79–93. <https://doi.org/10.1016/B978-0-12-820244-9.00006-8>
- Ríos-Saldaña, C.A., Delibes-Mateos, M., Ferreira, C.C., 2018. Are fieldwork studies being relegated to second place in conservation science? *Global Ecology and Conservation* 14, e00389. <https://doi.org/10.1016/j.gecco.2018.e00389>
- Robira, B., Benhamou, S., Masi, S., Llaurens, V., Riotte-Lambert, L., 2021. Foraging efficiency in temporally predictable environments: is a long-term temporal memory really advantageous? *Royal Society Open Science* 8, 210809. <https://doi.org/10.1098/rsos.210809>

- Rosati, A.G., 2017. Foraging Cognition: Reviving the Ecological Intelligence Hypothesis. *Trends in Cognitive Sciences* 21, 691–702. <https://doi.org/10.1016/j.tics.2017.05.011>
- Rosati, A.G., Hare, B., 2012. Chimpanzees and bonobos exhibit divergent spatial memory development. *Developmental Science* 15, 840–853. <https://doi.org/10.1111/j.1467-7687.2012.01182.x>
- Rubin, M., Donkin, C., 2024. Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests. *Philosophical Psychology* 37, 2019–2047. <https://doi.org/10.1080/09515089.2022.2113771>
- Ruginski, I.T., Creem-Regehr, S.H., Stefanucci, J.K., Cashdan, E., 2019. GPS use negatively affects environmental learning through spatial transformation abilities. *Journal of Environmental Psychology* 64, 12–20. <https://doi.org/10.1016/j.jenvp.2019.05.001>
- Schultz, W., Dayan, P., Montague, P.R., 1997. A Neural Substrate of Prediction and Reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>
- Sellers, W.I., Hill, R.A., Logan, B.S., 2007. An agent-based model of group decision making in baboons. *Philos Trans R Soc Lond B Biol Sci* 362, 1699–1710. <https://doi.org/10.1098/rstb.2007.2064>
- Semple, S., Higham, J.P., 2013. Primate Signals: Current Issues and Perspectives. *American Journal of Primatology* 75, 613–620. <https://doi.org/10.1002/ajp.22139>
- Strandburg-Peshkin, A., Farine, D.R., Couzin, I.D., Crofoot, M.C., 2015. Shared decision-making drives collective movement in wild baboons. *Science* 348, 1358–1361. <https://doi.org/10.1126/science.aaa5099>
- Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks* 9, 1054–1054. <https://doi.org/10.1109/TNN.1998.712192>
- Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12, 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Sylvia Chou, W.-Y., Gaysinsky, A., Cappella, J.N., 2020. Where We Go From Here: Health Misinformation on Social Media. *Am J Public Health* 110, S273–S275. <https://doi.org/10.2105/AJPH.2020.305905>
- Tekinbas, K.S., Zimmerman, E., 2003. Rules of Play: Game Design Fundamentals. MIT Press.
- Toohey, K., 2015. SimilarityMeasures: Trajectory Similarity Measures. R package version 1.4, <<https://CRAN.R-project.org/package=SimilarityMeasures>>.
- Trimmer, P.C., Houston, A.I., Marshall, J.A.R., Mendl, M.T., Paul, E.S., McNamara, J.M., 2011. Decision-making under uncertainty: biases and Bayesians. *Anim Cogn* 14, 465–476. <https://doi.org/10.1007/s10071-011-0387-4>
- Tukey, J.W. (John W., 1977. Exploratory data analysis. Reading, Mass.: Addison-Wesley Pub. Co.
- Turchin, P., 1998. Quantitative Analysis of Movement: Measuring and Modeling

- Population Redistribution in Animals and Plants. Sinauer.
- Tversky, A., Kahneman, D., 1971. Belief in the law of small numbers. *Psychological Bulletin* 76, 105–110. <https://doi.org/10.1037/h0031322>
- Vaish, A., Grossmann, T., Woodward, A., 2008. Not all emotions are created equal: The negativity bias in social-emotional development. *Psychol Bull* 134, 383–403. <https://doi.org/10.1037/0033-2909.134.3.383>
- van der Loo M., 2014. “The stringdist package for approximate string matching.” *The R Journal*, \*6\*, 111-122. <<https://CRAN.R-project.org/package=stringdist>>.
- Vericel, M.E., Baraduc, P., Duhamel, J.-R., Wirth, S., 2024. Organizing space through saccades and fixations between primate posterior parietal cortex and hippocampus. *Nat Commun* 15, 10448. <https://doi.org/10.1038/s41467-024-54736-7>
- Vezhnevets, A.S., Agapiou, J.P., Aharon, A., Ziv, R., Matyas, J., Duéñez-Guzmán, E.A., Cunningham, W.A., Osindero, S., Karmon, D., Leibo, J.Z., 2023. Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia. <https://doi.org/10.48550/arXiv.2312.03664>
- Wang, X.T., 1996. Domain-specific rationality in human choices: violations of utility axioms and social contexts. *Cognition* 60, 31–63. [https://doi.org/10.1016/0010-0277\(95\)00700-8](https://doi.org/10.1016/0010-0277(95)00700-8)
- Weller, F.G., Webb, E.B., Fogelnburg, S., Beatty, W.S., Kesler, D., Blenk, R.H., Ringelman, K.M., Miller, M.L., Arzel, C., Eadie, J.M., 2023. An agent-based model to quantify energetics, movement and habitat selection of mid-continent mallards in the Mississippi Alluvial Valley. *Ecological Modelling* 485, 110488. <https://doi.org/10.1016/j.ecolmodel.2023.110488>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4, 1686. <https://doi.org/10.21105/joss.01686>
- Wilkes, J.T., Gallistel, C.R., 2017. Information Theory, Memory, Prediction, and Timing in Associative Learning, in: Computational Models of Brain and Behavior. John Wiley & Sons, Ltd, pp. 481–492. <https://doi.org/10.1002/9781119159193.ch35>
- Williams-Ceci, S., Macy, M.W., Naaman, M., 2024. Misinformation does not reduce trust in accurate search results, but warning banners may backfire. *Sci Rep* 14, 1–16. <https://doi.org/10.1038/s41598-024-61645-8>
- Wilson, R.P., Sala, J.E., Gómez-Laich, A., Ciancio, J., Quintana, F., 2015. Pushed to the limit: food abundance determines tag-induced harm in penguins. *Animal Welfare* 24, 37–44. <https://doi.org/10.7120/09627286.24.1.037>
- Wise, T., Charpentier, C.J., Dayan, P., Mobbs, D., 2023. Interactive cognitive maps support flexible behavior under threat. *Cell Reports* 42. <https://doi.org/10.1016/j.celrep.2023.113008>
- Woolnough, A.P., Hollenberg, L.C.L., Cassey, P., Prowse, T.A.A., 2023. Quantum

- computing: a new paradigm for ecology. *Trends in Ecology & Evolution* 38, 727–735. <https://doi.org/10.1016/j.tree.2023.04.001>
- Wren, C.D., Botha, S., De Vynck, J., Janssen, M.A., Hill, K., Shook, E., Harris, J.A., Wood, B.M., Venter, J., Cowling, R., Franklin, J., Fisher, E.C., Marean, C.W., 2020. The foraging potential of the Holocene Cape south coast of South Africa without the Palaeo-Agulhas Plain. *Quaternary Science Reviews*, *The Palaeo-Agulhas Plain: a lost world and extinct ecosystem* 235, 105789. <https://doi.org/10.1016/j.quascirev.2019.06.012>
- Wright, T.F., 2006. The Evolution of Animal Communication: Reliability and Deception in Signaling Systems. *The Condor* 108, 989–991. <https://doi.org/10.1093/condor/108.4.989>
- Xiong, F., Liu, Z., Yang, X., Sun, B., Chiu, C., Qiao, H., 2017. A Bayesian Posterior Updating Algorithm in Reinforcement Learning, in: Liu, D., Xie, S., Li, Y., Zhao, D., El-Alfy, E.-S.M. (Eds.), *Neural Information Processing*. Springer International Publishing, Cham, pp. 418–426. [https://doi.org/10.1007/978-3-319-70139-4\\_42](https://doi.org/10.1007/978-3-319-70139-4_42)
- Xue, W., Sun, J., Liang, F., Hou, J., Yang, Y., Shang, W., Chen, X., Gradoni, G., Huang, Y., 2025. Jensen-Shannon Divergence Hypothesis Test for Determining Reverberation Chamber Field Distribution. *IEEE Transactions on Antennas and Propagation* 1–1. <https://doi.org/10.1109/TAP.2025.3574921>
- Yon, D., Heyes, C., Press, C., 2020. Beliefs and desires in the predictive brain. *Nat Commun* 11, 1–4. <https://doi.org/10.1038/s41467-020-18332-9>
- Zappala, J., Logan, B., 2010. Effects of resource availability on consensus decision making in primates. *Comput Math Organ Theory* 16, 400–415. <https://doi.org/10.1007/s10588-010-9080-4>
- Zedadra, O., Jouandeau, N., Seridi, H., Fortino, G., 2017. Multi-Agent Foraging: state-of-the-art and research challenges. *COMPLEX ADAPTIVE SYSTEMS MODELING* 5. <https://doi.org/10.1186/s40294-016-0041-8>
- Zhu, L., Mathewson, K.E., Hsu, M., 2012. Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings of the National Academy of Sciences* 109, 1419–1424. <https://doi.org/10.1073/pnas.1116783109>

## **Ethics Approval**

The study was part of the general enrichment program of ARTIS zoo. The study provided the animals with autonomy and voluntary control over part of their own food supply. All enrichment trials were fully integrated into the daily routine of the study animals and required no manipulation of individuals. The chimpanzees were never deprived of food or water at any stage of the study. Our study did not meet the definition of an animal experiment as mentioned in Article 1 of the Dutch “Experiments on Animals Act”, due to the absence of potential discomfort for the study animals and the non-invasive character of the study. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All activities to conduct the enrichment procedures were approved by ARTIS Zoo, Amsterdam.

## **Acknowledgements**

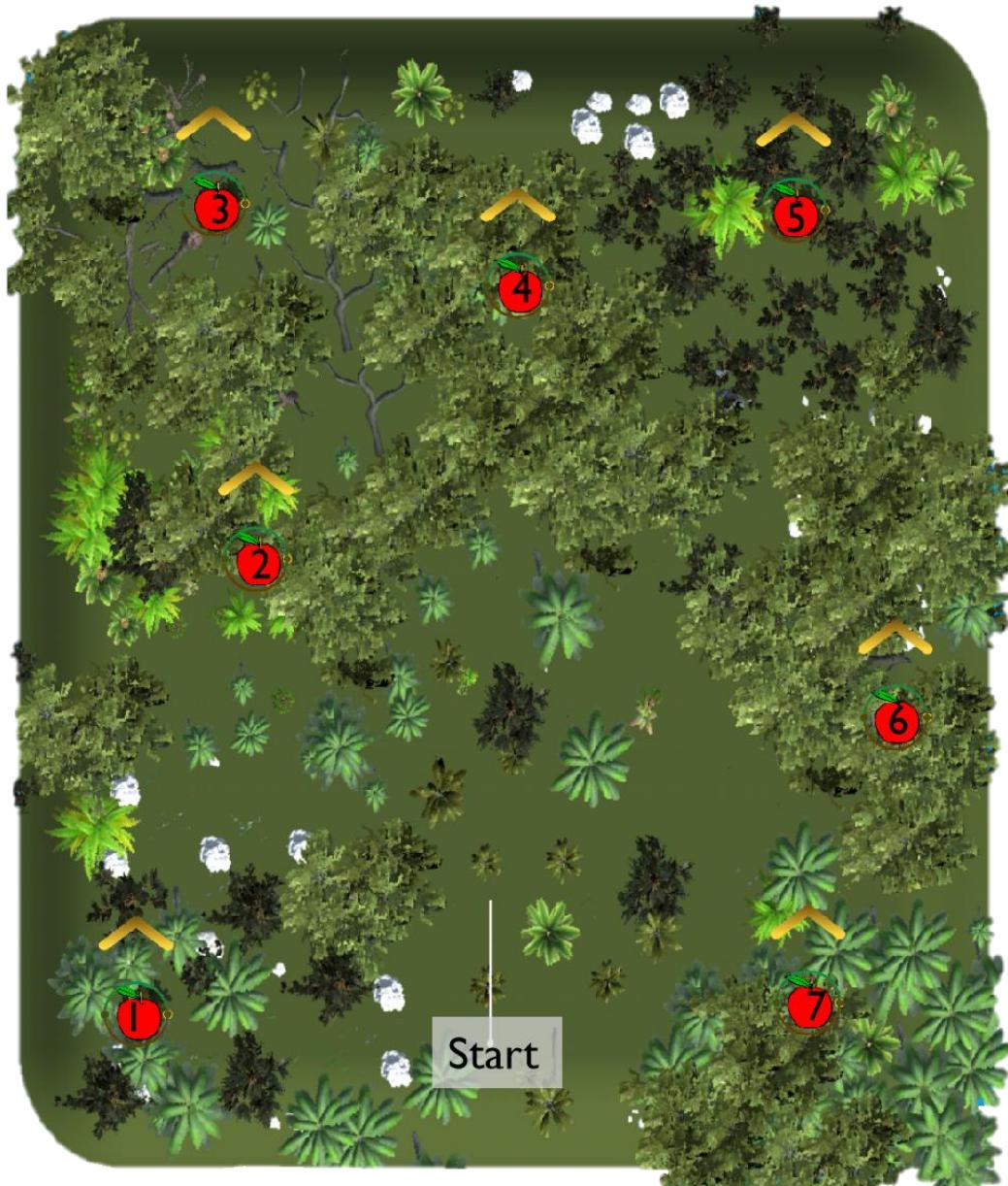
I’m deeply grateful to Emiel van Loon for his invaluable guidance throughout this project. In particular, I thank him for fostering rigorous coding and data processing practices, as well as for his efforts in enhancing my understanding of statistical methodologies and agent-based modelling. I’m also sincerely appreciative of the academic opportunities he has provided; it has been a privilege to work under his supervision, and our discussions have been both enlightening and formative.

Thank Karline Janmaat for her essential guidance during this research, particularly for her expertise in primate ecology and foraging cognition. Beyond this project, I’m deeply thankful for the numerous academic opportunities she has made available—from teaching assistant to research project—which have allowed for continuous growth and self-reflection in my academic development. It has been an honour to have her as a supervisor, and I’m more than proud to remain part of the CoBEco group.

I also thank Matthias Allritz, Rafa Bidara, and Evy van Berlo for their contributions to the development of the VR foraging game, and Bryndan van Pinxteren, Bram van der Perk, and Pia den Braver for designing the Forest Mind environment, which was instrumental to the implementation of this research.

I’m grateful to Anne van Dijk for coordinating the research at ARTIS Zoo, and to the zookeepers whose dedicated care of the chimpanzees made this study possible. Finally, I thank all human and chimpanzee participants whose voluntary involvement was fundamental to the success of this project.

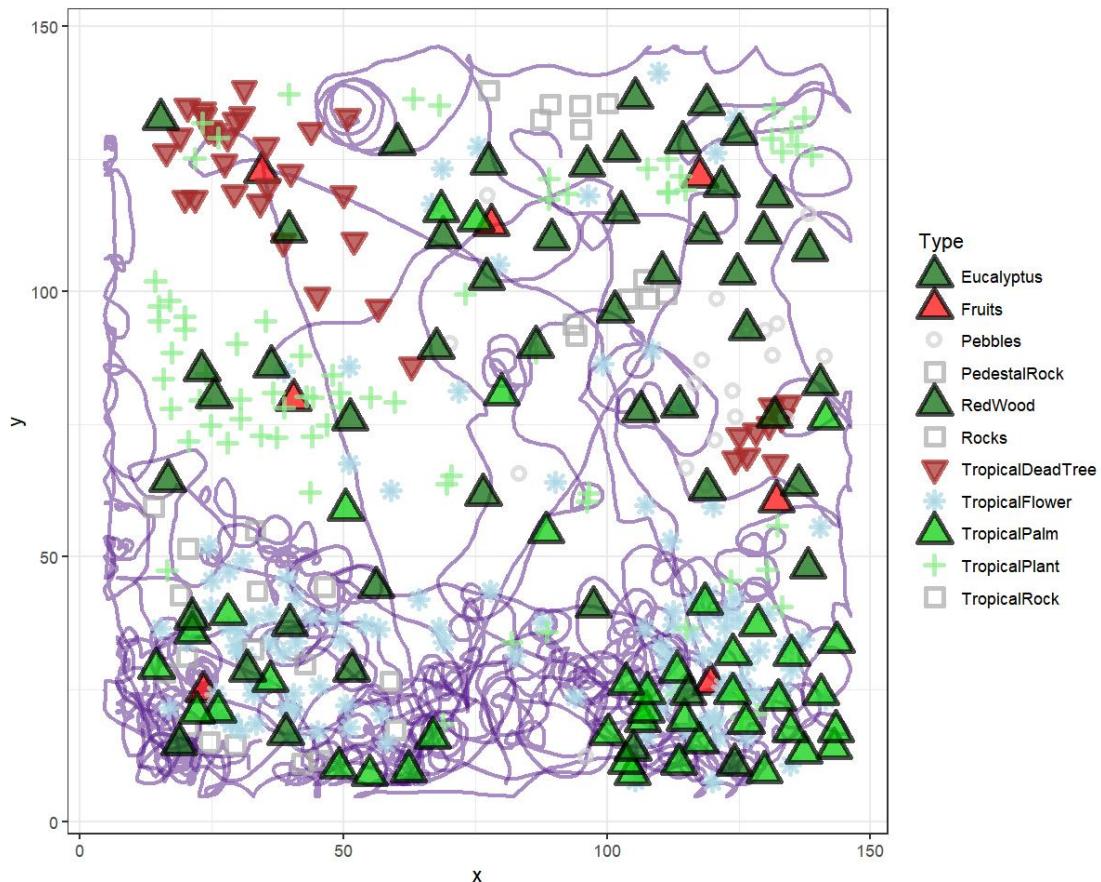
## Appendices



**Appendix 1.** A schematic overview of the virtual foraging environment. Red apple icons indicate the locations of fruit-bearing trees, with numerical labels denoting their respective tree IDs. The point marked 'start' represents the initial position of the participant.



**Appendix 2.** Heatmap illustrates the transition probability matrices of tree encounters, showing the likelihood of transitioning from a given fruit tree (rows; 0 = start location) to the next fruit-bearing tree visited by self-explore (columns), for Dutch participants.



**Appendix 3. Trajectory plots illustrate an example of game performance by a chimpanzee participant, showing movement patterns across all ten foraging sessions.**

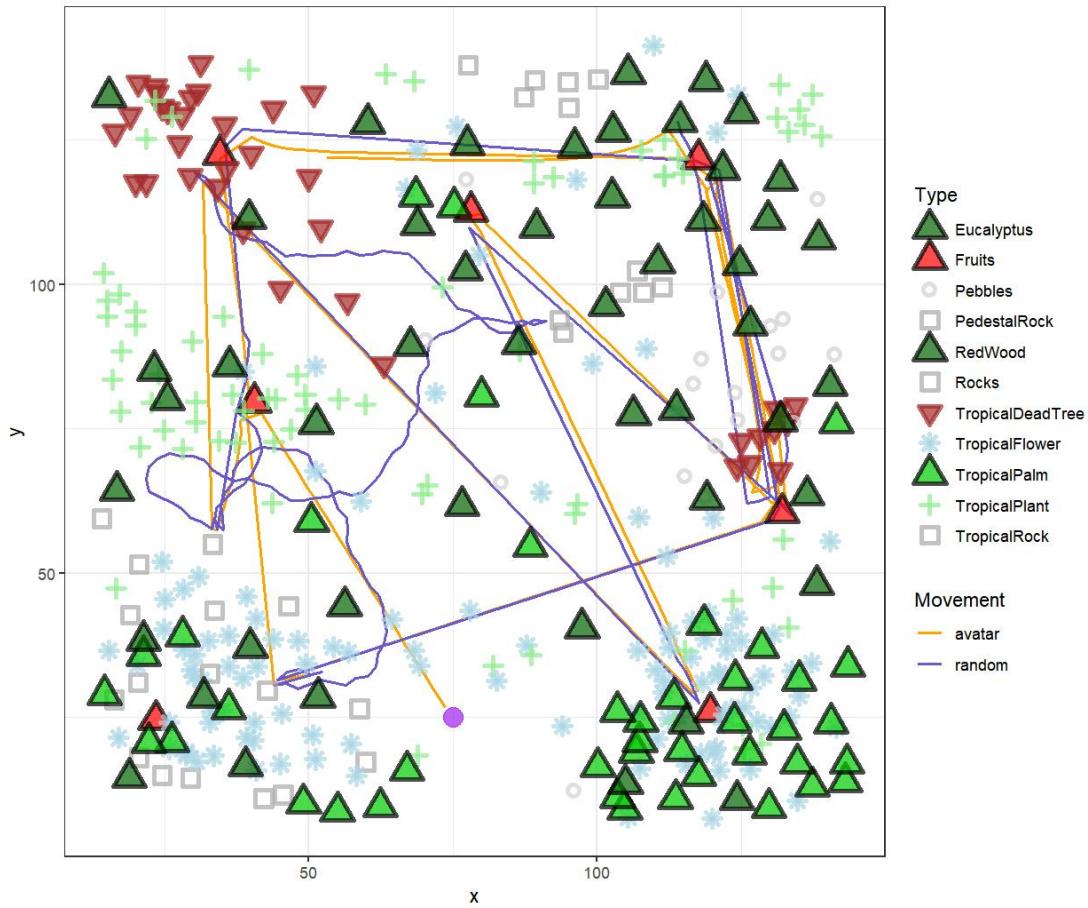
## Appendix 4: ODD protocol

The ODD protocol (Grimm et al., 2020) presented here includes elements from the model descriptions of the agent-based modelling (ABM) for simulating the foraging decision-making of primates. This ABM simulates the behavioural dynamics of foragers participating in an ecologically structured foraging task, designed to emulate a virtual reality (VR) game environment. The model was developed to investigate experience-dependent foraging strategies, integrating public and personal information sources in a dynamic spatial setting. It draws upon behavioural data and design principles from initial experimental trials with human and chimpanzee participants and is intended to explore how agents adaptively balance exploration, memory, and social cues to optimise resource acquisition.

### 1. Purpose and patterns

The model was developed to investigate how primates balance the use of private and public information during foraging under ecologically realistic assumptions. During foraging, individuals must respond to local environmental cues to navigate toward

resource patches, while also utilizing public information to efficiently explore unfamiliar environments—together giving rise to emergent behavioural patterns (Post and Semmann, 2011). Accordingly, the model was structured around two core functional components: spatial learning, which supports navigation based on environmental memory, and decision-making, which governs the dynamic allocation of reliance between public information and individual experience. Particular emphasis is placed on the modulatory role of risk perception in this trade-off, as well as the emergence of cognitive mapping. The model specifically simulates the movement dynamics and follow-explore decisions of participants in the Forest Mind game, with hypothesis testing centred on three primary behavioural metrics: foraging efficiency, the proportion of time spent following avatars, and the sequence in which participants encounter fruit-bearing trees (Fig 1.1).



**Figure 1.1. Snapshot of a Forest Mind game simulation. The purple point is the start point of simulation and the color shows different movement strategy represent by the agent.**

## 2. Entities, state variables and scales

The model comprises three principal components: (i) agents representing foraging individuals; (ii) knowledgeable avatars that act as carriers of public information by guiding agents toward potential resource sites; and (iii) a spatially explicit environment incorporating static foraging patches (fruit-bearing trees) and discrete

environmental landmarks. Each agent maintains a dynamic internal state comprising spatial coordinates, velocity vectors, movement modality, history of tree encounters, memory structures, perceived environmental risk, and intrinsic cognitive load (Tab 2.1). This state includes a cognitive map and a value estimation mechanism based on reinforcement learning. Agent decision-making is governed by a softmax rule combined with Bayesian updating, allowing agents to adaptively balance alternative foraging strategies, landmark-based navigation versus avatar-following, based on past experiences.

Avatars, which emulate the guide organism from the Forest Mind game, are randomly assigned to a fruit tree (irrespective of its fruiting status) and possess a state defined by their location, velocity, visibility to agents, tree encounter history, and current target. The state of fruit trees and other landmarks is defined by attributes including position, colour, size, penetrability, visibility duration, encounter frequency, eligibility traces, local co-occurrence with fruit trees (within 35 m), associative strength to neighbouring fruit trees, and associated reward value. The distance range used to determine neighbouring situation is based on the landmark distribution in the game.

The simulation is implemented on a 200 m × 200 m virtual landscape, designed to reflect empirical tree distributions observed in tropical forest environments (Fig. 2.1). This landscape includes 373 static landmarks, of which seven are fruit-bearing trees, corresponding to typical daily foraging targets reported for wild chimpanzees. Each fruit tree is surrounded by a unique constellation of co-occurring landmarks. Landmarks are taxonomically and morphologically categorised into broad classes including trees, rocks, pebbles, tropical flowers, and tropical plants. Tree types include *Eucalyptus*, Redwood, and tropical palm species, with fruit trees visually indistinguishable from *Eucalyptus* trees. The sole additional environmental parameter is simulation duration. Each simulation proceeds in one-second time steps for a maximum of 10 minutes. If all seven fruit trees are discovered before the allotted time, the simulation ends early and proceeds to the next iteration, reflecting the structure of the real-world game.

**Table 2.1. State variables used for describing the agent, avatar, and landmarks**

Symbol	Code	Description
<i>State variables of all entities:</i>		
$g_s$	sim_seed	Current gameplay session
<i>State variables of both agent and avatar:</i>		
x	x	x position on the grid
y	y	y position on the grid
$v_x$	x_vel	Velocity x component
$v_y$	y_vel	Velocity y component
$t_v$	trees_visited	ID of Trees visited by agent/avatar

---

*State variables of agent:*

$D_a$	dist_to_knowledgeable	Distance to the knowledgeable avatar
$s$	movement_mode	movement mode chosen for a given time step
$s_m$	movement_mode_mod	movement mode modifier (e.g. for "random" movement mode, "FAM" modifier)
$L_c$	current_landmark_id	ID of the current visited landmark
$L_l$	last_landmark_id	ID of the last visited landmark
$D_{PI}$	path_integrator_distance	Cumulative path length from the last visited landmark to the current visited landmark
$\vec{V}_{x, PI}$	path_integrator_vector_x	Total velocity of direction x from the last visited landmark to the current visited landmark
$\vec{V}_{y, PI}$	path_integrator_vector_y	Total velocity of direction y from the last visited landmark to the current visited landmark
$\vec{V}_{x, i \rightarrow j}$	memory_vectors_x	Memory vector of direction x from the landmark $i$ to the landmark $j$
$\vec{V}_{y, i \rightarrow j}$	memory_vectors_y	Memory vector of direction y from the landmark $i$ to the landmark $j$
$h$	risk	Learned risk about revisiting depleted fruit trees of following avatar
$H_{cog}$	intrinsic_load	Intrinsic cognitive load of agent to make decision from following avatar to self-explore

---

*State variables of avatar:*

$t_{id}$	target_id	ID of selected target on the grid
$x_t$	target_x	x position of selected target on the grid
$y_t$	target_y	y position of selected target on the grid

---

*State variables of avatar and landmarks:*

	in_dd	Whether an avatar/landmark is in agent's detection distance (0 if not, 1 if yes)
	in_fov	Whether an avatar/landmark is in agent's field of view (0 if not, 1 if yes)
	in_sight	whether an avatar/landmark is in agent's sight (in_dd and in_fov both equal to 1)

---

*State variables of landmarks:*

$p_{v_n}$	pv	Number of previous visits to the landmark $n$
$p_n$	is_block	Landmark $n$ will block the move of agent or not (0 if not, 1 if yes)
$c_n$	color	Color vibrancy of the landmark $n$ compare to the background
$s_n$	size	Size of the landmark $n$
$T_n$	time_insight	Number of iterations the landmark $n$ has

		been in sight
$S_n$	sensory	Sensory input of the landmark $n$ to agent
$E_n$	eligibility	Eligibility trace of agent to the landmark $n$
$C_{ij}$	coccurrence	Co-occurrence of fruit tree $j$ and its nearby landmark $i$
$S_{ij}$	association	Association strength between the landmark $i$ and its nearby fruit tree $j$
$R_{ij}$	reward	Expected reward of the landmark $i$ associate with its nearby fruit tree $j$

### 3. Process overview and scheduling

At each simulation timestep, agents move according to their currently selected strategy and consume any food resources encountered en route. Agents retain memory of the distances traversed between encountered landmarks, enabling them to construct navigational paths via path integration—a mechanism inspired by mammalian spatial cognition. In exploration mode, agent movement is additionally influenced by eligibility traces generated by sensory cues emitted from landmarks, primarily visual features such as colour and size. The strength of these eligibility traces modulates the magnitude of vector-based movement. Upon locating a fruiting tree, these memory vectors are further modulated by the agent’s tendency to return to previously rewarding patches. Overall, exploratory movement is governed by a combination of memory-based vectors and vector biases toward salient patches, either previously encountered or distinguished by strong sensory cues. In the absence of established memory vectors, agent locomotion defaults to a correlated random walk (CRW) process (Turchin, 1998). During follower mode, agents navigate by aligning their movement direction with the visible avatar, following it directly. Avatars are initialized at a random position within a 1-metre radius of the agent at the beginning of each trial. They are then assigned to travel toward a randomly selected fruit tree, irrespective of fruit presence. Upon reaching the tree, the avatar disappears until the agent also encounters the same tree, at which point it reappears near the agent and repeats its guidance behaviour. Visibility of both avatars and landmarks is updated at every timestep. In addition, each landmark updates its associative memory values using the Feature Association Matrix (FAM) algorithm (Gribkova et al., 2024; see Section 7.1), which operates at each timestep. The associative reward signal derived from this process is used to compute the agent’s internal cognitive load during decision-making, quantified via Shannon entropy.

A trial is defined as the path between either the starting location or the previous tree encounter and the next encountered tree. If the agent successfully reaches a fruit-bearing tree, the trial is marked as successful and one fruit is added to the foraging count. The encountered tree is then removed from the set of unseen trees and no longer treated as fruiting in subsequent trials. If no fruiting tree is found, the trial is

recorded as a failure. Beyond binary success or failure, agents also evaluate whether the empty fruit tree was reached via avatar-following, allowing them to update their estimate of the risk associated with this strategy. Agents continuously adjust the value estimates of their respective movement strategies based on trial outcomes, avatar-associated risk, and the cognitive load experienced during exploratory decisions. These adjustments are implemented via reinforcement learning algorithms (see Section 7.3), allowing agents to iteratively refine their decision policies. The simulation concludes either when all seven fruit trees are discovered or when the 600-timestep limit is reached. Agents retain their learned landmark associations between runs. Each simulation is repeated 11 times, corresponding to the number of game sessions played by human participants in the empirical study. One complete model run consists of these 11 repeated simulations, from which performance and behavioural data are extracted for analysis (Fig. 3.1).

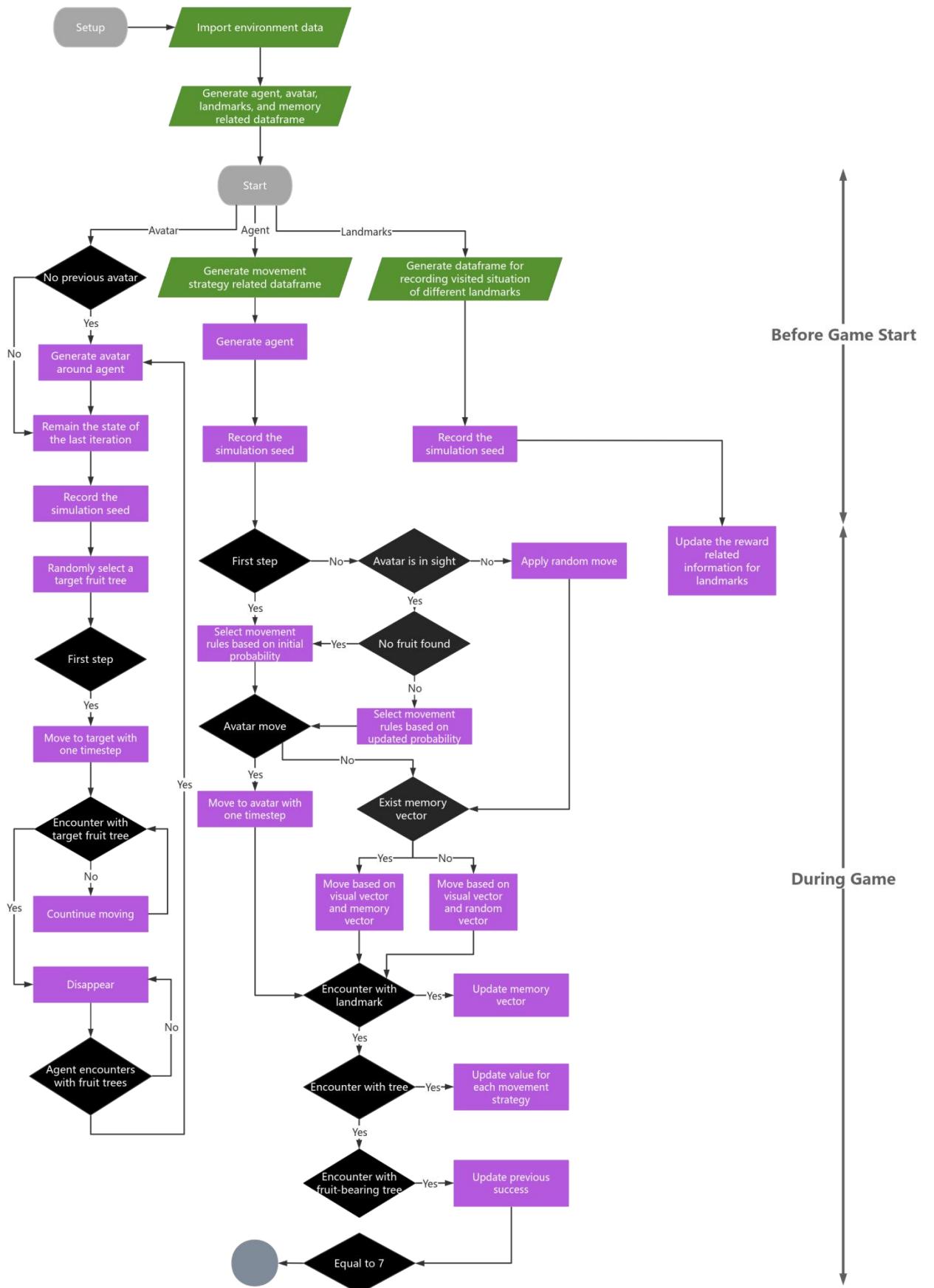


Figure 3.1. Flowchart describing the key operational processes of the

**agent-based model. Diamond-shaped symbols indicate logical, parallelograms indicate model input and rectangles indicate calculations. Circle indicate the end of simulation. A detailed description of all elements of this flow diagram can be found in the following submodel descriptions.**

## 4. Design concepts

### ***Basic principles:***

This model is founded on the hypothesis that agents foraging in unfamiliar environments exhibit a tendency to follow more knowledgeable guide (avatar) when locating food resources. As agents accumulate personal spatial knowledge over time, they begin to evaluate the relative reliability of personal versus public information and adapt their decision-making accordingly, thereby optimising their foraging efficiency.

### ***Emergence:***

Foraging efficiency emerges from two interacting processes: (i) the effective utilization of reliable public information, and (ii) accurate navigation based on spatial memory of resource locations. The utility of public information depends on the agent's capacity to assess the risk associated with relying on others, which in turn is shaped by risk sensitivity and the ability to update expectations based on failed outcomes. Navigation accuracy relies on the fidelity of path memory, itself governed by associative and reinforcement learning. Agents evaluate their memory confidence via internal uncertainty estimates. Associative learning reinforces spatial memory through repeated co-occurrence of landmarks and reward outcomes. The emergent foraging strategies reflect a balance between avatar-following and self-directed exploration.

### ***Adaptation:***

Agents adaptively shift their behavioural strategies based on the perceived reliability of information sources. When public information becomes unreliable (e.g., due to avatars increasingly selecting empty trees) agents favour self-guided exploration. Conversely, when agents' memory is imprecise or uncertain, they exhibit a greater propensity to follow knowledgeable guides.

### ***Objectives:***

Over the course of a game session, as fruit trees are successively encountered and avatars are more likely to guide agents to depleted (non-fruiting) trees, agents are expected to reduce reliance on avatar-following and instead prioritize decisions guided by accumulated spatial memory to maximize foraging returns. In some cases, however, agents may persist in following avatars despite previous failures. If such a strategy unexpectedly results in a reward, this may trigger a disproportionate dopaminergic response (De Petrillo and Rosati, 2021), potentially reinforcing suboptimal decision patterns (i.e., continued avatar-following despite high risk).

***Learning:***

Agents incrementally learn the spatial relationships between landmarks and food rewards by constructing cognitive maps during gameplay. Learning is both associative through repeated pairings of features and outcomes, and reflective, as agents revise their decision strategies based on trial-level successes and failures to optimise future foraging performance.

***Prediction:***

The model predicts that agents, through repeated gameplay, will increasingly optimise their decision-making strategies. Consequently, the total number of fruit trees located should increase across successive games, and the latency to find each tree should decrease. If agents exhibit risk-averse behaviour, the proportion of avatar-following is expected to decline as more fruit trees are located. Conversely, if agents follow the Prediction Errors Induce Risk Seeking (PEIRS) model (Moeller et al., 2021), avatar-following may drop slowly or even increase near the end of a session due to high positive reward errors resulting from the discovery of fruit with avatar in later trials.

***Sensing:***

Agents can perceive landmarks and avatars within a limited visual range, processing their colour, size, and whether the objects are traversable. This perceptual input enables agents to decide whether to follow an avatar or engage in autonomous exploration, using visual salience to reinforce associative memories. Agents also monitor the perceived risk associated with avatar guidance and the uncertainty of their own memory, facilitating dynamic switching between movement strategies.

***Interaction:***

The only direct interaction between agents and avatars is the option to follow the avatar's trajectory. Both agents and avatars interact with the environment by encountering landmarks, which serve as anchors for memory and navigational learning.

***Stochasticity:***

Stochastic elements within the model include exploratory movement patterns, avatar selection of fruit trees, strategy switching, the number of fruit trees located, encounters with non-fruiting trees, and the learned spatial paths themselves.

***Collectives:***

This model does not incorporate social structures or group dynamics; agents act independently with no collective behaviour.

***Observation:***

The simulation records agent-level data at one-second intervals, including movement trajectories, movement mode (exploratory or following), landmark encounters,

learned risk values, memory uncertainty, and decision values associated with competing movement strategies.

## 5. Initialization

The model initializes a virtual environment comprising 373 spatially distributed landmarks, including both fruit-bearing and non-fruiting trees, consistent with the spatial layout and visual design of the Forest Mind game. Agent entities representing primate participants are initialized at fixed coordinates ( $x = 75$ ,  $y = 25$ ), corresponding to the participant starting location in the actual game. Avatars, acting as carriers of public information, are initialized at a random position within a 1 m radius from the agent's location at the start of each trial. Cognitive and learning-related parameters are set prior to model execution. These include cognitive flexibility, sensitivity to cognitive load, risk learning rate, risk sensitivity, strategy learning rate, memory decay rate, and the type of risk cognition model employed (Tab 7.3.1). Initial strategy probabilities for avatar-following versus independent exploration are calibrated based on empirical behavioural distributions observed in real participants, comprising chimpanzees, hunter-gatherers from BaYaka society, and Dutch participants, prior to encountering their first fruit tree.

## 6. Input data

The sole input dataset used by the model is the Forest Mind game map, comprising the coordinates and unique identifiers of all landmarks. Additional attribute columns were appended to each landmark to support agent perception and decision-making. These attributes include colour, size, passability (i.e., whether the landmark can be traversed), and classification as a tree or non-tree object. Colour values were assigned a salience score ranging from 1 to 5 based on the contrast between landmark colour and background terrain. Size values were assigned based on their relative object dimensions (Tab 6.1).

**Table 6.1. Assigned visual features for the different type of landmarks**

Landmarks	Type	Color	Size
Fruit Trees	Trees	1	4
Redwoods	Trees	1	4
Eucalyptus	Trees	1	4
Tropical Dead Trees	Dead Trees	3	3
Tropical Palm	Palm Trees	2	4
Rocks	Rocks	5	3
Pedestal Rocks	Rocks	5	3
Tropical Rocks	Rocks	5	3
Tropical Flowers	Flowers	4	2
Tropical Plants	Plants	2	2
Pebbles	Pebbles	2	1

## 7. Submodels

As previously described, the model comprises two principal components—navigation and decision-making. To facilitate transparency, replicability, and broader understanding, a nested ODD protocol is employed to delineate the model’s internal logic. The architecture is subdivided into three interrelated submodels: spatial learning, movement, and strategy updating. For each submodel, the objectives, operational procedures, and theoretical underpinnings are described in detail. While a comprehensive list of model parameters is provided in this section, parameters specific to particular processes are organized in corresponding subsection tables for clarity.

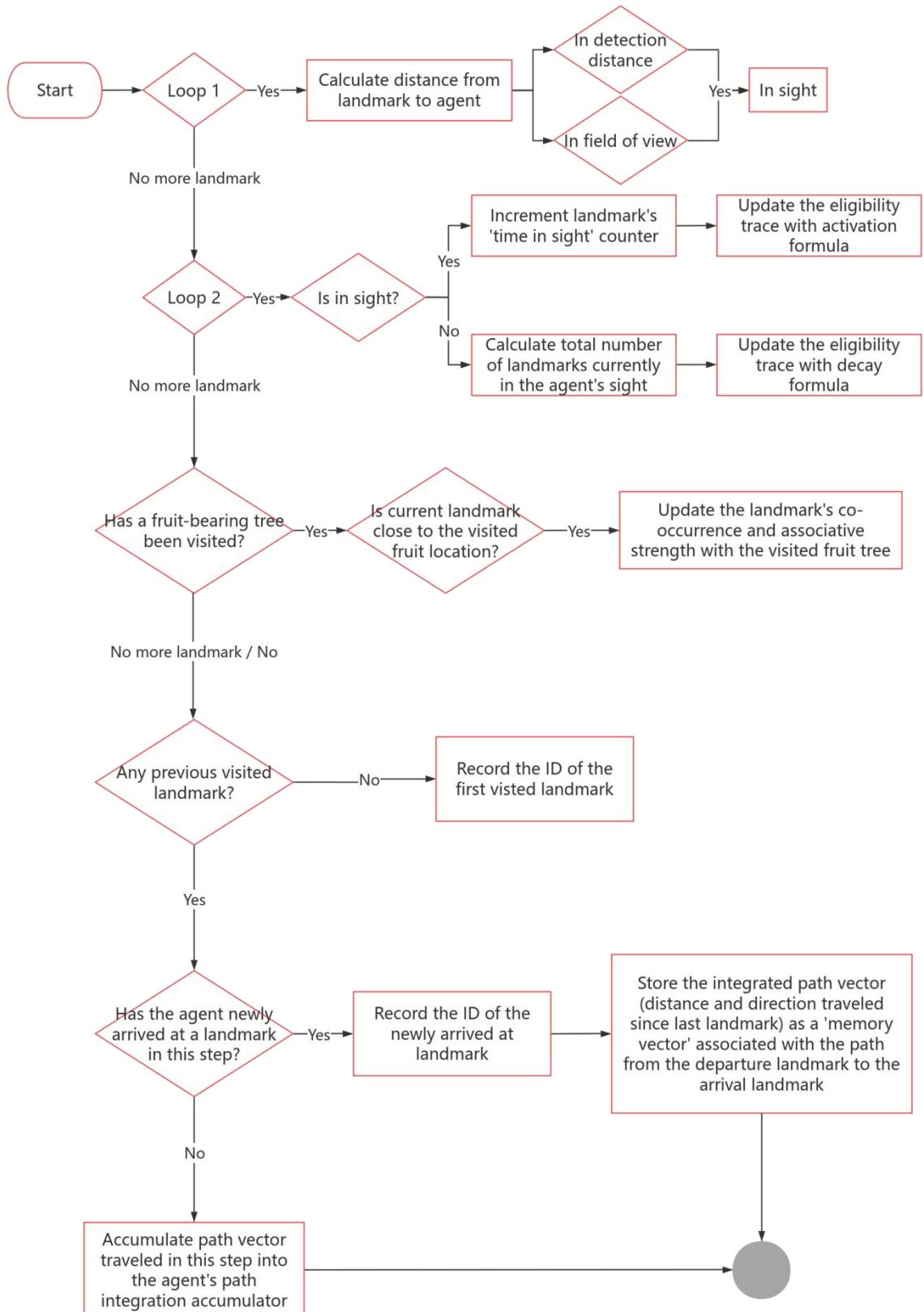
### 7.1. Spatial learning model

#### 7.1.1. Purpose

The spatial learning module enables agents to form associative links between environmental landmarks and resource rewards, facilitating the acquisition of foraging routes. This capacity allows agents to adaptively adjust their foraging decisions based on the perceived status and features of encountered landmarks.

#### 7.1.2. Process overview & scheduling

State variables associated with spatial learning are stored across two core data structures: a landmark dataset and a memory vector matrix. The landmark dataset encodes associative memory values linking each landmark to fruit-bearing trees, while the memory vector matrix maintains directional and distance information for navigational paths between landmarks. These spatial learning state variables are dynamically updated by the function `update_landmark_and_fam_simple`, which is invoked at each simulation timestep. This function operates conditionally based on the agent’s current state, executing the necessary computations to maintain and refine spatial memory representations (Fig. 7.1.2.1).



**Fig 7.1.2.1** Flowchart describing the key operational processes of the spatial model. Diamond-shaped symbols indicate logical, rectangle indicate the calculation and the circle indicate the end of simulation.

### 7.1.3. Associative learning

To learn the foraging map, agents integrate associative information from nearby landmarks of fruit-bearing trees. This process is mediated by a FAM, which encodes and replays sequential input patterns. Given a set of potential inputs, such as sensory cues and reward signals, FAM encodes the temporal structure of these sequences by modulating the strength and order of associations between paired inputs. Each associative link within the matrix is further assigned an expected reward value, thereby enabling the replay of previously acquired sequences. The matrix exhibits plasticity: sequences may be subject to decay and eventual forgetting if they cease to be reinforced by repeated presentation or reward, particularly in the presence of competing input sequences. Eligibility traces for each input are derived from sensory stimuli and serve as short-term temporal records of stimulus occurrence. Although reward signals may be transient, their associated eligibility traces decay more gradually, facilitating reinforcement learning for temporally proximal cues. In the present context, sensory inputs arise from visual stimuli generated by various landmarks (landmarks and fruit trees) within the virtual environment. The sensory input strength of a given landmark  $n$  is computed using a model in which visual features, such as impassability, conspicuous coloration, and size, are weighted to reflect their relative salience in influencing perceptual input intensity:

$$S_n = p_n \times c_n \times s_n$$

where  $p_n$  represents the landmark that can be crossed through by the agent or not;  $c_n$  represents the conspicuity of the landmark  $n$  colours compared to the background;  $s_n$  represents the size of the landmark  $n$ .

At each iteration, the eligibility trace  $E_n$  for a given landmark  $n$  is updated with following equations:

$$E_n^{t+1} = \begin{cases} k_E + \frac{1 - k_E}{e^{a_E \cdot S_n}}, & \text{if the landmark } n \text{ is in sight} \\ E_{decay} \cdot E_n^t, & \text{if the landmark } n \text{ is not in sight} \end{cases}$$

$$E_{decay} = e^{1 - N},$$

where constants  $k_E$  is the constant for scaling  $E_n$ ,  $a_E$  is the constant for scaling  $S_n$ ,  $E_{decay}$  is the decay variable for eligibility trace, which is higher when there's more new sensory input,  $N$  is the number of landmarks agent can see in the current iteration.

The eligibility trace also used in calculating the visual vector for each landmark.

Within the FAM, each pairwise combination of landmark eligibility traces is characterised by three key parameters: Strength, Order, and Reward. These variables respectively represent: (1) the degree of co-activation or overlap between two inputs; (2) the temporal order and relative latency with which the inputs are received; and (3) the expected reward associated with their joint activation. The simplified FAM used for this model only contains the strength and reward due to the dense distribution of landmarks weakens the effect of the order in which landmarks are encountered on associative learning, as similar landmarks are seen and memorized at the same time as agent move. The associative strength between a landmark  $i$  and its surrounding fruit-bearing tree  $j$  is ultimately determined by the co-occurrence and temporal overlap of their eligibility traces  $E_i$  and  $E_j$ :

$$C_{ij}^{t+1} = C_{ij}^t + \Delta C_{ij}^t,$$

$$\Delta C_{ij}^t = k_c(E_i \cdot E_j - E_{threshold} + k_a \cdot R_{ij})(E_i + E_j) + k_d(Strength_{ij} - C_{ij}),$$

$$Strength_{ij} = \frac{1}{1 + e^{a_s \cdot C_{ij}}},$$

where  $k_a, k_c, k_d, k_s$  denote constant parameters modulating the associative dynamics. The co-activation index  $C_{ij}$  increases in proportion to the degree of temporal overlap between the traces (i.e.  $E_i + E_j > 0$ ), and declines when such overlap is insufficient. The associative strength between landmark  $i$  and its surrounding fruit-bearing tree  $j$ , denoted as  $Strength_{ij}$ , is computed as a logistic transformation of  $C_{ij}$ , bounded within the interval  $[0, 1]$ . This measure quantifies the learned associative strength between the two entities. To prevent uncontrolled escalation of associative values, the dynamics of  $C_{ij}$  are further modulated by the associative reward signal  $R_{ij}$  and  $Strength_{ij}$ , thereby ensuring stable convergence within the learning system. The associative reward value assigned to each landmark  $i$  is derived from the reward output and its associative strength with its neighboring fruit-bearing tree  $j$ , as follows:

$$R_{ij} = Strength_{ij} \cdot R_j,$$

**Table 7.1.3.1. Parameters used in the Feature Association Matrix.**

Symbol	Value	Code	Description
$k_E$	0.2	k_E	Constant for eligibility trace
$a_E$	-0.25	a_E	Exponential constant for eligibility trace
$k_a$	1	k_a	Constant for scaling changes in co-concurrence
$k_c$	0.8	k_c	Constant for reward-based changes in co-concurrence
$k_d$	0.001	k_d	Constant for decay of co-concurrence
$a_s$	-0.25	a_s	Exponential constant for Strength
$E_{threshold}$	0.04	E_threshold	Eligibility product threshold for Co-occurrence

#### 7.1.4. Path integration and memory vectors

To enable navigation through non-overlapping spatial sequences, path integration (Etienne et al., 2004) is employed to compute the net direction and distance between pairs of landmarks and fruit trees. This module was triggered each time an agent arrived at a novel landmark and functioned as a foundation for encoding directed associations within a feature association memory. While the agent moved between landmarks (i.e., not within any arrival threshold), it continuously accumulated a path integration vector to approximate the relative displacement between successive landmarks. At each timestep  $t$ , the agent's displacement vector  $\vec{v}_t = (v_t^x, v_t^y)$  was added to an internal accumulator:

$$\vec{V}_{PI} \leftarrow \vec{V}_{PI} + \vec{v}_t$$

$$D_{PI} \leftarrow D_{PI} + \|\vec{v}_t\| = D_{PI} + \sqrt{(v_t^x)^2 + (v_t^y)^2}$$

where  $\vec{V}_{PI}$  is the cumulative path vector and  $D_{PI}$  is the total distance travelled since the last landmark arrival. Upon confirmed arrival at a new landmark  $j$ , if a valid previous landmark  $i$  was recorded and the integrated distance  $D_{PI} > 0$ , the path integration vector  $\vec{V}_{PI}$  was stored as a directional memory vector from landmark  $i$  to landmark  $j$ . These vectors encoded the relative direction between landmark pairs and

served as internal guides for future memory-based navigation decisions. Reverse vectors were encoded symmetrically depending on the assumed model of bidirectional learning. Assume that the agent takes  $n$  steps, and at each step  $t \in \{1, \dots, n\}$ , the movement vector is  $\vec{v}_t = (v_t^x, v_t^y)$ . Then the stored memory vector is calculated as:

$$\vec{V}_{i \rightarrow j} = \left( \sum_{t=1}^n v_t^x, \sum_{t=1}^n v_t^y \right)$$

To support learning from delayed reinforcement, a list of previously visited landmark indices was maintained in a reward history buffer. Each time the agent arrived at a distinct landmark, the index of the prior landmark was appended to this buffer. This history was subsequently used for reward backpropagation through temporally ordered landmark transitions. Upon arrival at a new landmark, the path integrator was reset to zero to ensure that memory vectors represented discrete inter-landmark transitions rather than cumulative trajectories.

## 7.2. Movement model

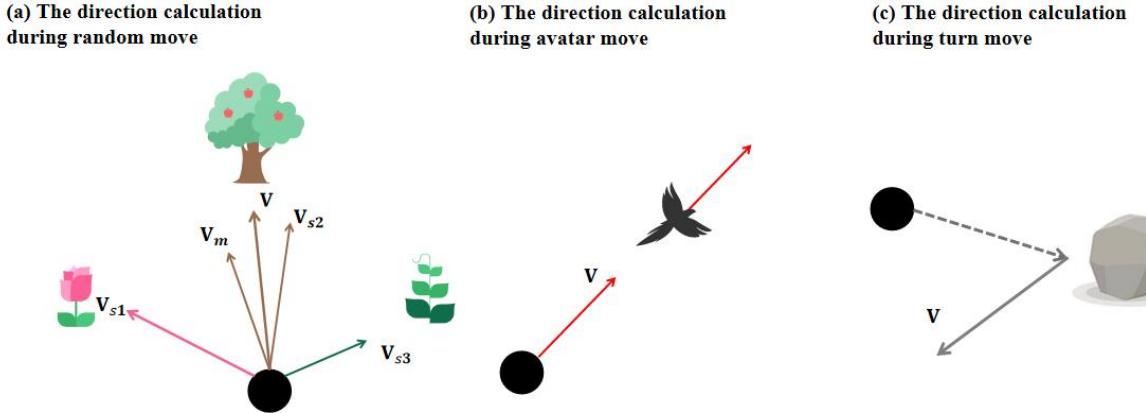
### 7.2.1. Purpose

The movement model enables agents to make adaptive locomotor decisions informed by spatial memory and current strategic states, while simultaneously perceiving and responding to environmental cues.

### 7.2.2. Process overview & scheduling

At the onset of each trial, the agent stochastically selects a movement strategy based on the current probability distribution across available strategies. If the selected strategy is to follow the avatar, the agent further evaluates whether to maintain or switch strategies within the trial, based on its learned estimate of risk. Specifically, if the estimated risk of re-encountering a non-fruited tree while following the avatar exceeds the expected failure probability of self-directed exploration (i.e. 1 minus the expected reward), the agent will re-sample its movement strategy according to the prevailing selection probabilities. During self-exploration, movement direction is jointly governed by visual stimulus vectors from all landmarks within perceptual range and memory vectors pointing toward the landmark associated with the highest expected reward (Fig 7.2.2.1a). In the absence of a memory vector, stochastic perturbations are introduced by generating vectorized random values drawn from a univariate truncated normal distribution. In contrast, during avatar-following, the agent moves directly towards the avatar's position (Fig 7.2.2.1b). If the agent collides with a rock, it switches to a turning-mode correlated random walk (CRW) pattern (Fig 7.2.2.1c); such movement episodes are excluded from retrospective strategy updating. At each time step, the model records the agent position, orientation, chosen strategy, current time step, and encounter status with environmental features. Encounters are

defined as follows: a landmark is encountered if the agent is within 0.5 m; a tree or fruiting tree is encountered if within 5 m (reflecting the fruit-drop radius in the original game); and a rock is encountered if within 1.5 m (based on the average radius of rocks in the game environment).



**Figure 7.2.2.1. Illustration of the different movement strategy used in the agent-based model**

### 7.2.3. The movement rule of avatar-based move and random move

Two principal types of agent movement were implemented for the agent (Fig 7.2.1): (1) goal-directed movement toward avatar, and (2) random exploratory movement, which included both memory-guided navigation and stochastic directional changes. When an agent selected a specific avatar, it moved deterministically toward that location using a normalized direction vector. The agent then updated its position and velocity based on the normalized direction vector and the step length.

Under the random move, the direction of agent was determined by both memory-based vector and visual gradient vector. To calculate the memory-based vector if the agent was located at a known landmark  $i$ , a directional vector toward another landmark  $j$  was retrieved from the feature association memory, conditioned on learned reward and associative strength. The set of available candidate transitions is denoted:

$$\mathcal{T}_i = \{j | j \neq i\},$$

For each candidate  $j \in \mathcal{T}_i$ , the expected value of transitioning was based on the associative reward of the landmark. The target landmark with maximal learned reward was selected:

$$j^* = \operatorname{argmax}_{j \in \mathcal{T}_i} R_{ij}$$

The associated memory vector  $\vec{m}_{i \rightarrow j^*} \in \mathbb{R}^2$  was then normalized to unit length:

$$\vec{m}_{i \rightarrow j^*} = \frac{1}{\|\vec{v}\|} \cdot \begin{bmatrix} \vec{V}_x[i, j^*] \\ \vec{V}_y[i, j^*] \end{bmatrix}, \text{ where } \|\vec{v}\| = \sqrt{\vec{V}_x[i, j^*]^2 + \vec{V}_y[i, j^*]^2}$$

If no such memory vector was available, a fallback stochastic rotation (detailed below) was applied.

For visual gradient vector, The agent also computed a visual salience gradient based on currently visible landmarks. For each landmark  $n$  in view, a weighted vector was computed:

$$\vec{v}_n = \frac{1}{d_n^2} \cdot E_n \cdot \hat{r}_n, \text{ where } \hat{r}_n = \frac{(x_n - x_t, y_n - y_t)}{\|(x_n - x_t, y_n - y_t)\|}$$

Here,  $d_n$  is the Euclidean distance to landmark  $n$ , and  $E_n \in [0,1]$  is an eligibility trace indicating attentional salience. The overall visual gradient vector is then:

$$\vec{v}_{\text{visual}} = \sum_k \vec{v}_k \text{ (normalized to unit length if } \|\vec{v}_{\text{visual}}\| > 0 \text{)}$$

The final movement direction was computed as a weighted combination of the memory vector  $\vec{v}_{\text{mem}}$  and the visual vector  $\vec{v}_{\text{visual}}$  :

$$\vec{v}_{\text{nav}} = (1 - w) \cdot \vec{v}_{\text{mem}} + w \cdot \vec{v}_{\text{visual}}, w = \text{visual\_weight}$$

This vector was renormalized:

$$\hat{v}_{nav} = \frac{\vec{v}_{nav}}{\|\vec{v}_{nav}\|}$$

and used to update the velocity and position of agent:

$$\vec{v}_t = \ell \cdot \hat{v}_{nav}, \quad \vec{x}_{t+1} = \vec{x}_t + \vec{v}_t$$

where  $\ell$  is the agent's fixed step length.

If both memory and visual vectors were unavailable or null (i.e., agent was uninformed and visual occluded), directional noise was introduced. A small angular perturbation  $\theta \sim \mathcal{U}\left(-\frac{\phi}{2}, \frac{\phi}{2}\right)$  was applied to the agent's previous velocity vector

$$\vec{v}_{t-1} :$$

$$\vec{v}_t = \ell \cdot \mathbf{R}_\theta \vec{v}_{t-1}, \quad \mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This preserved movement continuity while enabling exploration in the absence of directed cues.

In agent encounter a landmark cannot be passed, such as rocks, the agent performed a directional reversal, approximating a U-turn, by sampling a heading angle  $\theta \sim \mathcal{T}\mathcal{N}(\mu = \pi, \sigma = \phi/5, [\pi - \phi/2, \pi + \phi/2])$ , where  $\mathcal{T}\mathcal{N}$  is a truncated normal distribution centered at  $180^\circ$ . The resulting movement vector was:

$$\vec{v}_t = \ell \cdot \mathbf{R}_\theta \vec{v}_{t-1}$$

Position remained fixed in this step ( $\vec{x}_{t+1} = \vec{x}_t$ ) , but the heading vector was updated for the subsequent movement cycle.

**Table 7.2.3.1. Parameters used in the Movement Model.**

Symbol	Value	Code	Description
$\ell$	2.15	step_length	Step length (in m), set up based on the velocity extracted from game
$w$	0.3	visual_weight	Weight for scaling the visual stimulus

### 7.3. Strategy updating model

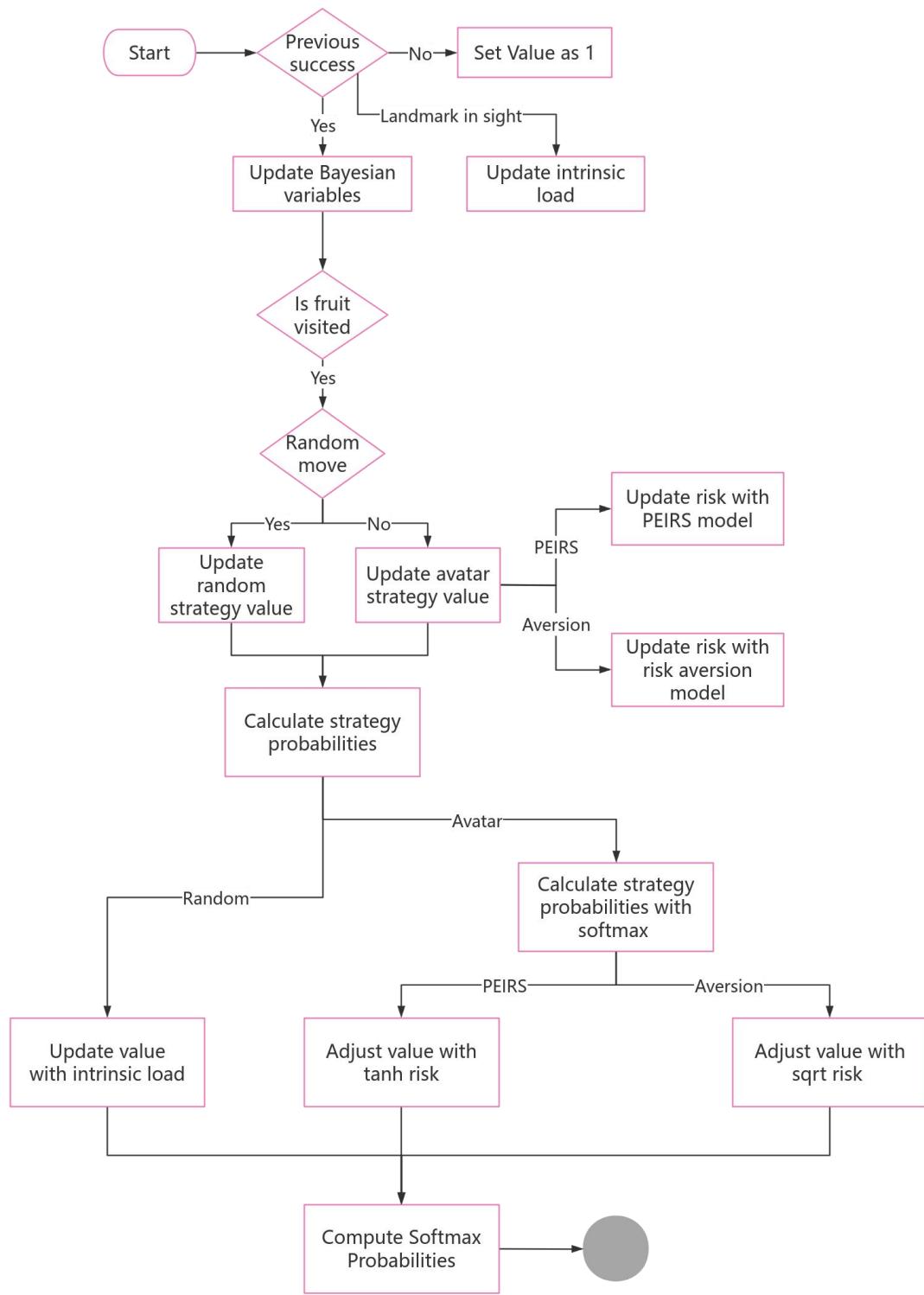
#### 7.3.1. Purpose

The strategy update model is designed to capture the cognitive and ecological processes by which foraging agents dynamically adapt their movement decisions in response to environmental uncertainty, prior experience, and evolving estimates of risk and reward. Specifically, it aims to formalize how agents balance exploitation of known resource locations with exploration of novel areas, leveraging both individual memory and socially available cues.

#### 7.3.2. Process overview & scheduling

The model implements an agent-based decision architecture in which individual foragers dynamically update their movement strategies through a hierarchical Bayesian reinforcement learning process. Strategy selection is influenced by spatial memory, perceived risk, and cognitive load. The process is executed iteratively at each simulation time step, consisting of the following core stages (Fig 7.3.2.1). At the beginning of each trial, the agent assesses its history of successful foraging experiences. If no prior success is recorded, strategy values are reset to uniform priors (default value = 1), enabling unbiased exploration in early learning phases. Cognitive load associated with random move are calculated based on the visual input of agent. If the detected targets were associated with reward, the intrinsic cognitive load of applying random move is computed as Shannon entropy of learned reward (see 7.3.3), representing the uncertainty of finding fruit reward through self-exploration. In the absence of discernible rewards, load is estimated using a logarithmic function of the number of visible trees, which is the maximum entropy under current situation. This load value penalizes the utility of the random movement strategy, discouraging disorganized search behaviour. If prior success exists, historical trials are partitioned into successful and failed outcomes, and the posterior belief in success of each strategy is updated via exponentially discounted parameters. The updated belief will be used for further calculation of the expected outcome of each strategy, the outcome update only happened when agent encountered fruit reward with correspond strategy. Risk was updated when the success or failure of trial was caused by avatar following behaviour. The calculation formula is based on the setup of risk update model. Intrinsic load and risk-related penalties are incorporated at the stage when all strategy values are converted to choice probabilities using a softmax transformation. Each of these steps is executed per time step as part of the agent decision-making cycle. The process ensures continuous adaptation to environmental feedback, balancing exploration and exploitation under constraints of memory, perception, and risk

evaluation.



**7.3.2.1. Flowchart describing the key operational processes of the spatial model.**  
**Diamond-shaped symbols indicate logical, rectangle indicate the calculation and the circle indicate the end of simulation.**

### 7.3.3. Strategy updating rule and softmax probability calculation

Bayesian reinforcement learning framework was adopted to simulate strategy

adaptation based on prior experience and environmental uncertainty (Trimmer et al., 2011; Xiong et al., 2017). Specifically, for each movement strategy  $s \in \{\text{avatar, random}\}$ , the agent maintains Beta-distributed priors over success rates, updated trial-by-trial using weighted evidence. The agent updates its priors with new trial information as follows:

$$\alpha_s(t) = \lambda \cdot \alpha_s(t-1) + n_s^{\text{success}}(t),$$

$$\beta_s(t) = \lambda \cdot \beta_s(t-1) + n_s^{\text{fail}}(t),$$

where  $n_s^{\text{success}}(t)$  is the number of successful decisions using strategy  $s$  in trial  $t$ , and  $n_s^{\text{fail}}(t)$  is the number of failed decisions using strategy  $s$  in trial  $t$ . The decay term  $\lambda$  captures temporal forgetting, ensuring recent experiences influence behavior more.

The posterior belief in success for each strategy is calculated using the Beta posterior:

$$p_s(t) = \frac{\alpha_s(t)}{\alpha_s(t) + \beta_s(t)},$$

This is combined with initial prior probabilities  $p_s^{(0)} \in [0, 1]$  to update the Bayesian belief, grounded in initial trust toward each strategy:

$$\text{BayesianProb}_s(t) = \frac{p_s(t) \cdot p_s^{(0)}}{p_s(t) \cdot p_s^{(0)} + (1 - p_s(t)) \cdot (1 - p_s^{(0)})},$$

Each strategy's value is then updated using a Rescorla-Wagner-like rule, scaled by its Bayesian success probability:

$$V_s(t+1) = V_s(t) + k_s \cdot (R_s(t) - V_s(t)) \cdot \text{BayesianProb}_s(t),$$

where  $R_s(t)$  is the trial outcome (1 for success, 0 for failure) using strategy  $s$  in trial  $t$ , ensuring agent can integrate rewards and cognitive costs to update expected results for each strategy.

Given the constraints of working memory (Sweller, 1988), learned uncertainty of environment was incorporated into the model to simulate the agent's limited capacity for memory utilization. As the primary focus of this study is the cognitive processes underlying foraging decision-making, uncertainty in environment was specifically modeled as the ambiguity imposed by working memory during decision-making. The

uncertainty of decision outcomes in the random movement is primarily influenced by the availability of information within the association in this game. Conceptualizing information processing as a shift in information states allows for the application of Shannon entropy to quantify uncertainty in a computationally efficient manner (Frank, 2013; Ortega and Braun, 2013; Wilkes and Gallistel, 2017). Thus, in this study, the uncertainty in the associative rewards of certain landmark  $i$  is quantified using entropy, defined as:

$$P_i = \frac{R_i}{\sum_j^N R_j},$$

$$H = - \sum_{i=1}^N P_i \cdot \ln P_i,$$

where  $N$  represents the number of paths can be chose from the closet landmark in sight at the iteration  $t$ ;

Let  $Q_{\text{random}}$  denote the utility of self-exploration by random move at trial  $t$ , defined as a function of expected outcomes and the uncertainty associated with random move:

$$Q_{\text{random}} = V_{t,\text{random}} - b\sqrt{H_{\text{cog}}}$$

where  $b$  is the sensitivity parameter for the uncertainty in changing into the self-exploration.

The risk perception mechanism of the avatar-following movement was employed with two different model, one is the dopaminergic reward prediction frameworks described by Moeller et al. (2021). This model considered risk preferences emerge as side effects of reward prediction errors, matching with the risk-seeking behaviour caused by the hot hand effects. Another is based on the classic risk aversion model (d'Acremont et al., 2009). By integrating these elements with the softmax decision rule, the ABM captures how agents refine their foraging strategies over time, offering insights into the cognitive processes underlying movement decisions and the trade-offs between social and individual information use.

Predicted risk were updated following a Rescorla–Wagner rule (Preuschoff and Bossaerts, 2007). Let the reward  $r_t$  denote the payoff in trial  $t$ . Let the  $V_{t,\text{avatar}}$  be the expected outcome of following avatar. The reward prediction error at trial  $t$  can be calculated as:

$$\delta_t = \begin{cases} r_t - V_{t,avatar}, & \text{if model} = 'PU' \text{ OR } 'AU' \\ r_t - R_{expect}, & \text{if model} = 'PN' \text{ OR } 'AN' \end{cases}$$

$R_{expect}$  was defaulted as 1 due to the fruit reward in this game continuously equal to 1.

This design ensured that agents in the PU and AU models adjusted their expected rewards based on experience, whereas those in the PN and AN models did not. Then the risk prediction error  $\xi_t$  is denoted as the difference between the squared reward prediction error and the estimated predicted risk (d'Acremont et al., 2009):

$$\xi_t = \delta_t^2 - h_t$$

The Rescorla–Wagner rule is used to update predicted risk  $h_{t+1}$  for the next trial:

$$h_{t+1} = h_t + k_{risk} \xi_t$$

where  $k_{risk}$  is the risk learning rate.

Let  $Q_{avatar}$  denote the utility of avatar-targeted move at trial  $t$ , defined as a function of the expected outcome and the predicted risk associated with avatar-targeted move. When the agent perceive risk following the PEIRS model (Moeller et al., 2021), the utility updating rule will be:

$$Q_{avatar} = V_{t,avatar} + \tanh(a\delta_t) \cdot h_t,$$

When the agent perceive risk following the classic risk aversion model, the utility updating rule will be:

$$Q_{avatar} = V_{t,avatar} - a\sqrt{h_t},$$

Where  $a$  is the risk preference parameter. In this model,  $\delta = 0$  corresponds to steady state dopamine release,  $\delta < 0$  means that dopamine release is suppressed, and  $\delta > 0$  means that dopamine release is enhanced (Moeller et al., 2021).

The probability of selecting the specific strategy in the trial  $t$  was defined with the log-softmax-based probability (Banerjee et al., 2025):

$$\log p_s(t) = \beta \cdot Q_s - \log \left( \sum_{i=1}^N e^{\beta \cdot Q_i} \right),$$

To ensure numerical stability and prevent overflow, log-sum-exp trick was used for

subtracting the maximum scaled value:

$$\log p_s(t) = \beta \cdot Q_s - \left[ \max_j (\beta \cdot Q_j) + \log \left( \sum_{i=1}^N e^{\beta \cdot Q_i - \max_j \beta \cdot Q_j} \right) \right],$$

Where,  $\beta$  is the parameter used for adjusting the cognitive flexibility of agent, higher  $\beta$  will let agent tend to select the decision with higher value.

**Table 7.3.1. Parameters used in the Strategy Updating Model.**

Symbol	Value	Code	Description
$\beta$	2	beta	Parameters regulate exploration of options when making a decision; the higher beta is, the more privileged will be the options with higher perceived value
$k_{risk}$	0.8	k	Learning rate of the perceived risk of following avatar
$k_s$	0.75	ks	Learning rate of the new experience for each strategy
$a$	0.9	a	Sensitivity of the learned risk
$b$	0.1	b	Sensitivity of the intrinsic load for changing into exploration
$\lambda$	0.8	E_threshold	Discounting factor for decay of past experiences

## 8. References

- Banerjee, K., C, V.P., Gupta, R.R., Vyas, K., H, A., Mishra, B., 2025. Exploring Alternatives to Softmax Function. Presented at the 2nd International Conference on Deep Learning Theory and Applications, pp. 81–86.
- d'Acremont, M., Lu, Z.-L., Li, X., Van der Linden, M., Bechara, A., 2009. Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage* 47, 1929–1939. <https://doi.org/10.1016/j.neuroimage.2009.04.096>
- De Petrillo, F., Rosati, A.G., 2021. Variation in primate decision-making under uncertainty and the roots of human economic behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, 20190671. <https://doi.org/10.1098/rstb.2019.0671>
- Etienne, A.S., Maurer, R., Boulens, V., Levy, A., Rowe, T., 2004. Resetting the path integrator: a basic condition for route-based navigation. *Journal of Experimental Biology* 207, 1491–1508. <https://doi.org/10.1242/jeb.00906>
- Frank, S.L., 2013. Uncertainty Reduction as a Measure of Cognitive Load in Sentence Comprehension. *Topics in Cognitive Science* 5, 475–494. <https://doi.org/10.1111/tops.12025>
- Gribkova, E.D., Chowdhary, G., Gillette, R., 2024. Cognitive mapping and episodic memory emerge from simple associative learning rules. *Neurocomputing* 595, 127812. <https://doi.org/10.1016/j.neucom.2024.127812>
- Grimm, V., Railsback, S.F., Vincenot, C.E., Berger, U., Gallagher, C., DeAngelis, D.L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A.S.A., Milles,

- A., Nabe-Nielsen, J., Polhill, J.G., Radchuk, V., Rohwäder, M.-S., Stillman, R.A., Thiele, J.C., Ayllón, D., 2020. The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *JASSS* 23, 7.
- Moeller, M., Grohn, J., Manohar, S., Bogacz, R., 2021. An association between prediction errors and risk-seeking: Theory and behavioral evidence. *PLoS Comput Biol* 17, e1009213. <https://doi.org/10.1371/journal.pcbi.1009213>
- Ortega, P.A., Braun, D.A., 2013. Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469, 20120683. <https://doi.org/10.1098/rspa.2012.0683>
- Post, D.J. van der, Semmann, D., 2011. Local Orientation and the Evolution of Foraging: Changes in Decision Making Can Eliminate Evolutionary Trade-offs. *PLOS Computational Biology* 7, e1002186. <https://doi.org/10.1371/journal.pcbi.1002186>
- Preuschoff, K., Bossaerts, P., 2007. Adding Prediction Risk to the Theory of Reward Learning. *Annals of the New York Academy of Sciences* 1104, 135–146. <https://doi.org/10.1196/annals.1390.005>
- Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12, 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Trimmer, P.C., Houston, A.I., Marshall, J.A.R., Mendl, M.T., Paul, E.S., McNamara, J.M., 2011. Decision-making under uncertainty: biases and Bayesians. *Anim Cogn* 14, 465–476. <https://doi.org/10.1007/s10071-011-0387-4>
- Turchin, P., 1998. Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. Sinauer.
- Wilkes, J.T., Gallistel, C.R., 2017. Information Theory, Memory, Prediction, and Timing in Associative Learning, in: Computational Models of Brain and Behavior. John Wiley & Sons, Ltd, pp. 481–492. <https://doi.org/10.1002/9781119159193.ch35>
- Xiong, F., Liu, Z., Yang, X., Sun, B., Chiu, C., Qiao, H., 2017. A Bayesian Posterior Updating Algorithm in Reinforcement Learning, in: Liu, D., Xie, S., Li, Y., Zhao, D., El-Alfy, E.-S.M. (Eds.), Neural Information Processing. Springer International Publishing, Cham, pp. 418–426. [https://doi.org/10.1007/978-3-319-70139-4\\_42](https://doi.org/10.1007/978-3-319-70139-4_42)