FoMo: A Unifying Theory of Visual Foraging

Alasdair D. F. Clarke and Anna E. Hughes
Department of Psychology
University of Essex

Visual foraging lies at the intersection of visual perception, decision-making and action planning. An attractive feature of this paradigm is that it generates a rich stream of sequential decision data. However, this presents a number of challenges for analysis. To this end, we have developed FoMo, a robust and flexible generative model for spatial-sequential data. Building upon initial work by Clarke et al. (2022b), we present FoMo 2.0, which incorporates spatial structure allowing us to model organised spatial behaviours. We test the model over a range of secondary data sources and show that incorporating these parameters improve model accuracy, predominantly due to improvements for a subset of participants who use grid-following strategies. FoMo therefore allows us to explain results that would previously have been categorised as outliers. A real strength of this generative modelling approach is that it can account for individual differences across the wide range of summary statistics that have previously been used to explore human and non-human animal behaviour, providing a unified framework for analysing these data.

1 Introduction

When a person searches, they search through space. This is true whether we are talking about a classic cognitive psychology paradigm in which an observer searches for a pre-specified target on a computer screen (Treisman and Gelade, 1980); or more 'real-world' tasks, such as finding a specific wine bottle in a virtual store (Zhao et al., 2017). It even holds for abstract semantic spaces that are searched through when receiving a word from memory (Collins and Loftus, 1975; Hills et al., 2012). The effect of space is particularly relevant in foraging behaviours; a topic of study in human and non-human animal cognition (Bella-Fernández et al., 2021), as well as more widely in ecology and biology (Hills et al., 2015). We define foraging as a task in which participants collect multiple targets that are distributed throughout some (psyical or abstract) space. Targets are typically selected one at a time, with each target being selected at most once. The distance from the forager to prospective items is generally a strong predictor of behaviour (Clarke et al., 2022b).

While the above may seem obvious, the effect of space on search tasks is generally overlooked, at best treated as a nuisance parameter to be "averaged away". Instead of trying to incorporate spatial parameters into modelling efforts, it is more common to attempt to reduce their influence via experimental design. For example, when designing stimuli for visual search, a common approach is to arrange items in a circle around a central fixation cross in order to control for any effects of eccentricity. The effects of direction are averaged away by ensuring that there are multiple trials with the target in randomised positions in each one. Some illustrative examples include Wolfe and Myers (2010); Becker (2010); Chen and Zelinsky (2006); Buetti et al. (2019); Hughes et al. (2024a), but we could cite many hundreds of other papers, including many of our own. Similarly, some visual foraging paradigms use stimulus displays with constantly moving items (Grössle et al., 2023) in order to disrupt, remove or at least minmise the spatial factors. The hope here is that this allows researchers to experimentally isolate the particular factors that they are investigating, and for many research questions this approach has been extremely fruitful. However, we would argue that in building a full model of visual foraging, we need to grapple with spatial effects more directly. Incorporating spatial position explicitly in models allows us to not only partition out this variance, but also to understand how spatial position might interact with other factors, such as stimulus features.

This introduction is structured as follows. First, we discuss the effect of space in the visual atten-

tion literature, with a focus on how visual search models account for spatial effects. We then move to the specific case of visual foraging and introduce the key paradigms. Finally, we discuss the various summary statistics and modelling approaches that have been used to analyse data from visual foraging experiments. This general introduction is followed by Section 2 in which we introduce our generative model of visual foraging (FoMo, Clarke et al. (2022b)) and extend it to create a unifying model that incorporates target-based as well as spatial parameters. We purposely have kept this literature review relatively brief: For more in-depth recent reviews, see Bella-Fernández et al. (2021) and Kristjánsson and Kristjánsson (2025).

1.1 Eccentricity, Eye Movements & Attention

The importance of space in vision is well established. For a start, the density of rods and cones on the retina (and hence visual acuity) falls off steeply with increasing eccentricity from the fovea (Harvey Jr and Pöppel, 1972). The effect of this is further compounded by M-scaling, in which the amount of visual cortex devoted to processing each unit of visual space decreases systematically with eccentricity (Rovamo and Virsu, 1979; Virsu et al., 1987). The result is a dramatic gradient of spatial resolution across the visual field, with the highest acuity concentrated in a small central region and steep drop-offs toward the periphery.

One way in which the human visual system deals with these eccentricity effects is to make use of saccadic eye movements (Findlay and Gilchrist, 2003), and a number of spatial effects have been well documented over the years. For example, phenomena such as saccadic momentum (the tendency of saccades to continue in the same direction and with the same velocity as the previous saccade, (Smith and Henderson, 2009; Wang et al., 2011)) and inhibition of return (the reduced probability of a saccade returning to a previously fixated location (Klein, 2000; MacInnes et al., 2014)) play a role in making some regions of the stimulus more likely to be fixated than others. It is also possible in many cases to detect regularities in the direction of saccades. Gilchrist and Harvey (2006) demonstrated that scan paths during visual search contain more horizontal than vertical saccades, and that this ratio is sensitive to spatial regularities in the search array: participants made more horizontal saccades in trials where the stimuli had a stronger grid structure.

Finally, behavioural biases such as the central bias (Clarke and Tatler, 2014), where observers favour making fixations towards the centre of an image, regardless of content, and pseudoneglect, the tendency to be biased towards the left-hand side of space (Brooks et al., 2014) also affect fixational patterns. Despite these factors being well known, there are few attempts to account for them or even measure them in visual search tasks, although the saccadic flow model is one attempt to provide a relatively simple model of how people look around photographs of scenes which takes into account the central bias (Clarke et al., 2017; Mairon and Ben-Shahar, 2025).

In addition to eye movements, our visual system employs spatial attention to selectively prioritise different regions of the visual field, enhancing detection sensitivity and discrimination accuracy in attended areas. (Posner, 1980). However, despite the name emphasising the *spatial* nature of attention, much of the empirical work aims to investigate theoretical accounts of how attention may be deployed away from the fovea, without much explicit reference to the distance or direction. As mentioned above, the variables are often controlled for, in the hope that there the effects under investigation do not interact with direction or distance[1]. We discuss this in more depth in the context of visual search below.

1.2 Spatial Effects in Visual Search

The most influential models of visual search are generally relatively unconcerned with spatial features. The only mention in Feature Integration Theory (FIT) (Treisman and Gelade, 1980) is to argue that there is no requirement to spatially locate target features. More recent iterations of the Guided Search (GS) Model (Wolfe, 2021) highlight the importance of spatial information and acknowledge "*in classic GS, spatial factors were largely ignored*". However, the discussion is largely limited to the functional field of view, with little focus on the effects of spatial layout on the order in which items are inspected. Finally, ideal observer models of visual search (Najemnik and Geisler, 2008) are able to make predictions about the optimal distributions of saccadic directions and distances, but do not necessarily provide a good ac-

---

[1]Furthermore, it should be noted that the theoretical underpinnings of the whole attentional enterprise are *still* contested (Anderson, 2011; Hommel et al., 2019; Rosenholtz, 2024).
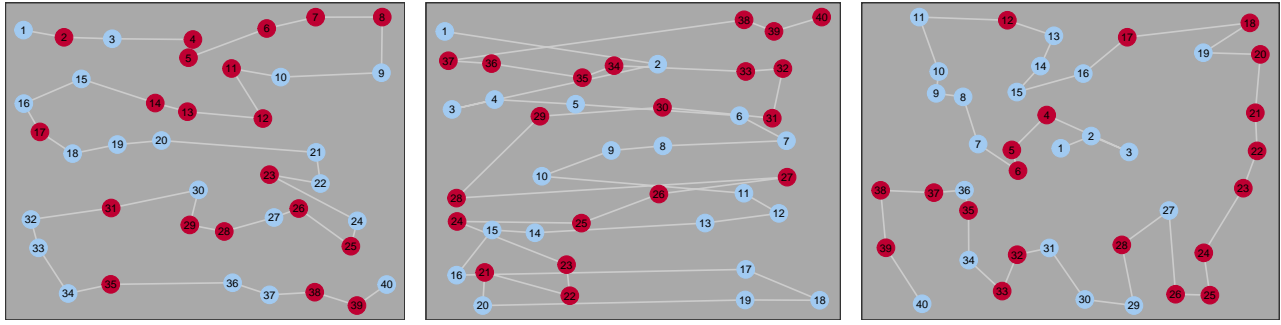
Figure 1

Examples of three trials from the Clarke et al. (2022c) dataset. Each trial is from the same condition (feature) but with a different participant. We can clearly see that there are some large individual differences. The first two participants appear to prioritise a horizontal scanning strategy, although the second implements this as two long runs (one colour, and then the next). The third participant de-emphasises a directional strategy and instead focuses on collecting nearby items of the same colour whenever possible.

count of individual spatial search behaviour (Morvan and Maloney, 2012; Nowakowska et al., 2017; Clarke et al., 2016).

One well-understood spatial effect is the fall off in visibility with increasing eccentricity: it is almost trivial to vision scientists that all other things being equal, a target that is further from fixation will take longer to find than one that is closer, even in the case where an observer is carrying out a parallel, 'pop out' search that is thought to occur simultaneously across the visual field (Buetti et al., 2019). However, the importance of spatial effects goes beyond this. For example, Hughes et al. (2024a) demonstrate that search slopes vary with eccentricity, particularly when the target is distinguished from the distractors by shape. There are also other studies that demonstrate broader spatial effects in visual search tasks. Amor et al. (2016) showed that participants often use a 'reading-like' strategy when completing a visual search task. Hooge and Erkelens (1996) showed that systematic scanning can happen around circular displays, suggesting that these biases do not simply reflect oculomotor biases but instead are related to the display structure. These types of scanning strategies can also have practical consequences: for example, when searching through medical images, radiologists who adopted a 'scanning' strategy, searching an entire 2D scan before moving in depth performed more poorly, making more search errors, compared to participants who had a 'drilling' strategy, restricting their eye movements to one portion of

the image and scrolling through in depth (Drew et al., 2013). However, these types of strategies are rarely considered in models of search, despite potentially explaining a high proportion of the variance in people's searching behaviours.

1.3 Visual Foraging

It is clear that spatial factors strongly affect visual search, and we therefore need a paradigm that allows us to examine the spatial components of search behaviour easily. We believe that the visual foraging paradigm, first developed by Kristjánsson et al. (2014), is particularly suitable for this. In this task, participants are directed to sequentially collect targets (of multiple classes e.g. red circles and green circles) by clicking on them, while ignoring distractors (e.g. blue circles and yellow circles). Thus, we are able to obtain from each trial a sequence of target selections that give us information about how participants spatially explore the environment. This paradigm has the advantage of being constrained and relatively tractable, as in the basic paradigm there are fewer than 100 simple items on the screen. In addition, a range of previous work has demonstrated that participant's strategies can vary. For example, Kristjánsson et al. (2014) showed that people's behaviour in a feature foraging task (where targets are defined from distractors by just one feature) differs compared to a conjunction foraging task (where the targets must be distinguished from distractors using a combination of factors). In particular, most people switch more frequently between

target types in the feature condition compared to the conjunction condition, generating longer 'run' lengths in the latter (although there are individual differences in this behaviour, and it appears to be less prominent in 3D environments (Kristjánsson et al., 2020; Prpic et al., 2019)). There is also good evidence that participants can plan ahead during foraging, preparing the next one or two targets that they will visit (Kosovicheva et al., 2020; Thornton and Horowitz, 2004). In our own work (Clarke et al., 2022c), we have also seen that participants clearly use a range of strategies, leading to interesting individual differences in target selection patterns (see Figure 1).

Spatial factors within the context of foraging have also been explored. In one VR study, human participants organised their search behaviour based on the layout of the area to be searched (De Lillo et al., 2014). Similarly, human participants often used highly systematic search paths when attempting to find small targets (coins) in an open terrain search area (Riggs et al., 2017) Interestingly, other primate species (such as capuchin monkeys) are also able to forage 'patchy' stimuli in an efficient manner (De Lillo et al., 1997), but only humans were able to efficiently search a less naturalistic matrix stimulus (De Lillo et al., 2014), indicating that some level of spatially organised foraging may be a highly evolutionarily conserved ability, perhaps driven by diet (Smith and De Lillo, 2022) (as primates eat a diet rich in fruit, which tends to be ephemeral and patchily distributed). However, humans seem to be able to flexibly adapt these strategies to serve a wider range of scenarios, and understanding which strategies they choose in different circumstances is critical for understanding foraging behaviour more fully.

1.4   Models and Measurements of Foraging

Foraging paradigms are a rich source of data and can be analysed in many different ways. In the animal cognition and ecology literature, foraging is often studied in terms of *patch leaving*: when should the local area be considered depleted, triggering a move to a new, hopefully, more bountiful, region? This has often been studied via the lens of the *marginal value theorem* (MVT) (Charnov, 1976), which states that an animal should leave a patch when the marginal capture rate dips below the average rate for its habitat. This is a specific example of the more general *optimal foraging theory* (Pyke et al., 1977). These frameworks have been used to study foraging behaviour in a wide range of species (Bella-Fernández et al., 2021), and MVT and patch-leaving statistics have also been applied to human visual foraging tasks, albeit with limited success: a recent review by Kristjánsson and Kristjánsson (2025) concluded that these methods provided a poor account of human visual behaviour.

Another concept from the ecology foraging literature is the idea of *lévy flights* which have been used to describe movement patterns in animals as varied as the albatross (Viswanathan et al., 1996) and polar bears (Auger-Méthé et al., 2016) (see Kamaruzaman et al. (2013) for a general review). In short, lévy flight analyses aim to describe foraging behaviour as a random walk, and assume that the distribution of inter-target distances are sampled from a power law. As with MVT, these ideas have also been taken up by researchers studying foraging in humans (Garg and Kello, 2021). However, such analysis is of limited use: as lévy flight analysis only attempts to summarise the distribution on inter-target distances, it is not able to offer predictions of which items are selected next. Furthermore, Pyke (2015) points out a specific weakness; that these methods do not incorporate directionality, a point we explore in the present manuscript.

Whereas foraging studies from ecology typically focus on explaining patch-leaving or how animals move around their environment, studies on human visual foraging have often focused on explaining behaviour *within* a trial. As the paradigm has gained in popularity, a wealth of summary statistics have been applied. As previously discussed, many of these experiments use two different target classes and focus on 'runs' (Dawkins, 1971; Kristjánsson et al., 2014): a condition can then be summarised by calculating the average number of runs, or the average run length. This work has inspired a model by Le et al. (2021) that aims to predict when a participant will break with the current run and switch to the other target type. Their classifier makes use of 11 features (such as the distance to the closest same target, and the ratio of remaining targets of one type to the other type) and aims to predict switches during both feature and conjunction foraging data. Their approach achieves impressive accuracy rates (92.1% for conjunction foraging) and can also predict rare switching events relatively well. However, this type of model cannot yet deal with the individual differences seen in this type of foraging task.

Beyond run statistics, a variety of other mea-

sures have been introduced with the aim of capturing something about spatial organisation during foraging (Bella-Fernández et al., 2024). The best-r statistic is based on correlating the x and y coordinates of the selected targets with the order in which they were selected. Trials in which targets were systematically collected from left-to-right or top-to-bottom receive high scores, while more stochastic, less organised behaviour receives lower scores. While this measure appears relatively well established in the literature (Woods et al., 2013; Ólafsdóttir et al., 2021; Johannesson et al., 2016) it has several weaknesses that we believe undermine its usefulness as a measure of spatial organisation. First of all, it can only capture horizontal and vertical patterns and is blind to diagonal zig-zag and circular spiral patterns. While such behaviours are not well-documented in the foraging literature, we should desire that our summary statistics have good validity. Secondly, this statistic is a summary statistic of the whole foraging path, and it is hard to see how it could be used as a causal predictor to help us understand item-to-item selections within a trial. Similarly, we also have the *percentage above optimal* (*PAO*) and *intersection rate* (*IR*) statistics, both of which are calculated from a completed foraging path. *PAO* compares the total distance travelled by the participant to the shortest possible path, while *IR* counts the number of times the path intersects itself. Again, while both statistics have been used to describe foraging organisation with some success (Ólafsdóttir et al., 2021; Wiener et al., 2007), they describe only a summary statistic of the whole foraging path and cannot easily be used to make predictions of participant behaviour on an item-by-item basis.

1.5 Individual Differences in Foraging

Visual foraging requires a range of cognitive abilities from low-level attention to higher-order decision making. As such, it is an interesting paradigm to study individual differences in. Early work by Kristjánsson et al. (2014) identified potential *super foragers* - a small subset of participants who appeared to be able to carry out the same strategy in both feature and conjunction foraging. However, later work (Tagu and Kristjánsson, 2020) demonstrated that participants who implement this strategy appear to make a larger number of errors, suggesting that these differences may reflect different strategies (perhaps different choices regarding a speed-accuracy trade-off) rather than

differences in performance on this task. This conclusion is supported by (Clarke et al., 2022c), where low correlations were found between an individual's performance on a visual foraging task when compared with other similar visual attention tasks, suggesting that individuals might display fairly consistent behaviours within a given task, but that these may reflect a chosen strategy rather than an innate cognitive skill that varies between people.

Despite this, there is clear evidence for foraging behaviours varying across the lifespan (Ólafsdóttir et al., 2020, 2021; Bella-Fernández et al., 2024), with younger participants displaying less organised foraging behaviour. Similarly, recent work has demonstrated that differences in foraging behaviour may be useful in a clinical context: Tagu et al. (2025) have shown that individuals with higher trait anxiety are more affected by negative emotional stimuli in an 'emotion foraging' task, and Lloyd et al. (2025) have shown that learning to adjust one's exploration-exploitation balance based on environmental volatility in a patch-leaving foraging task is affected by anxiety. Cancellation tasks are frequently used to assess visuospatial attention deficits in neurological populations (Checketts et al., 2021; Moore et al., 2022), and involve participants searching for multiple targets among distractors in a manner very similar to a foraging task. It therefore seems likely that visual foraging paradigms are sensitive to developmental differences, and that being able to model behaviour in this task could have benefits for clinical populations by potentially allowing earlier or more sensitive diagnoses.

2 FoMo: a Model of Visual Foraging

In our work, we have taken a different approach to understanding foraging within a patch by conceptualising it as a '*sampling without replacement*' problem (Clarke et al., 2022b). Our Foraging Model (FoMo) provides accurate predictions of target-by-target behaviour using relatively few parameters. Moreover, because our model is generative, we can simulate data to systematically evaluate how well it captures different aspects of human behaviour. This approach allows us to identify specific limitations in the model and use these insights to refine our theoretical understanding of visual foraging.

One weakness of the original version of FoMo (Clarke et al., 2022b) is that it made relatively poor predictions of absolute foraging directions
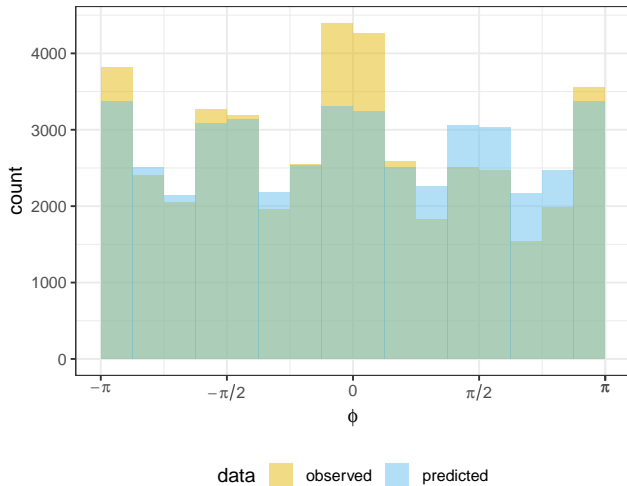
Figure 2

FoMo v1.0 inter-item selection results for the Clarke et al. (2022c) dataset. Histograms comparing the absolute directions (ϕ). Green indicates overlap of overlap of the two distributions. We can see that while the model favours the cardinal directions - presumably due to a proximity bias combined with grid-like stimuli - it fails to capture the strong left to right peak we see in the human data.

(see Figure 2), potentially leading to decreased accuracy in cases where participants are using an absolute direction strategy (e.g. follow cardinal directions) to structure their behaviour. In the current manuscript, we update FoMo to account for absolute direction biases, and hence characterise spatial behaviour and participant strategies more accurately. We use a flexible mixture model of directions and show that this improves overall model accuracy, predominantly by dramatically improving accuracy for a subset of individuals who use highly directional strategies to complete foraging tasks. We also demonstrate the flexibility of our model by showing how models with different parameter combinations can easily be compared in our framework, increasing the utility of our approach for a range of different foraging questions. Finally, we show how our modelling approach is able to act as a unifying framework, as it is able to accurately represent the individual differences present in a wide range of other foraging metrics that have been used in previous studies.

On the whole, we find that our model offers a compelling description of behaviour and individual differences in the visual foraging paradigm.

The rest of this section contains descriptions of our existing model (both general, Section 2.1, and mathematical, Section 2.2), followed by the new directional components in Section 2.3. Section 2.4 details how the model is implemented, while Section 2.5 gives a brief overview of the datasets we have used to test the model. Finally, Section 2.6 outlines the analysis plan, including the range of foraging metrics we test.

2.1 General Overview

Clarke et al. (2022b) modelled visual foraging as a weighted sampling-without-replacement process. The foraging model assumes that target items belong to one of two classes (A or B) and accounts for four influences on how participants decide which item to select next:

- *stick/switch rate* (b_s): values above 0 lead to increasingly long runs of the same target class while values below zero lead to a preference to alternate between target classes.

- *class weights* (b_a): This can be thought of as modelling the relative attractiveness of item class A over B. This attractiveness could be due to a number of different reasons such as low-level visual saliency or value. When b_a is above 0 the model favours selecting items of class A while values below 0 lead to a preference for items of class B.

- *proximity tuning* (ρ_δ): the extent to which participants select items that are in close proximity to the previously selected item.

- *direction tuning* (ρ_ψ): the extent to which participants select items that "ahead" of them. Negative values would indicate that items that are "behind" the participant are favoured.

These four features are illustrated in Figure 3. These are more formally defined in Section 2.1, and expanded to include absolute directions in Section 2.2.

2.2 FoMo v1.0: Model Specification

To formally define FoMo we first define some terms. Let the N target items in a stimulus be represented by $z_i = (x_i, y_i, c_i)$, where $i = 1, \ldots, N$; (x_i, y_i) is the spatial location of the target z_i and $c_i \in \{A, B\}$ represents its class.
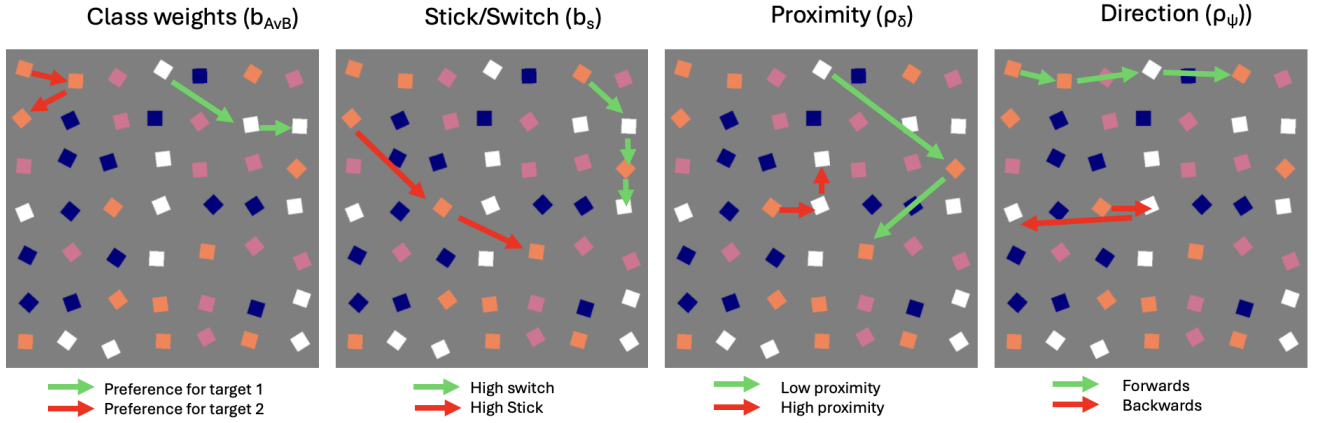
Figure 3

The original Clarke et al. (2022b) model parameters. From left to right: stick/switch rate, class weights, proximity tuning, and direction tuning.

Item class components

First, we define some simple functions that give us the weight contributions due to item class and stick/switching behaviour:

$$p_a(i) = p_{AvB}(z_i) = \begin{cases} \text{invlogit}(b_a) & \text{if } c_i = A \\ 1 - \text{invlogit}(b_a) & \text{if } c_i = B \end{cases} \quad (1)$$

$$p_s(i|j) = p_s(z_i|z_j) = \begin{cases} \text{invlogit}(b_s) & \text{if } c_i = c_j \\ 1 - \text{invlogit}(b_s) & \text{if } c_i \neq c_j \end{cases} \quad (2)$$

Spatial components

The spatial components of the model are based on the distance, δ, and direction, ϕ, of the vector from the previously selected item z_i to each remaining item z_j:

$$\delta(i,j) = \delta(z_i, z_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

$$\phi(i,j) = \phi(z_i, z_j) = \tan^{-1}(y_j - y_i, x_j - x_i) \quad (4)$$

We use *relative direction*, ψ, defined as the angular difference between the vector from z_{i-1} to z_i and the vectors from z_i to each remaining z_j:

$$\phi_d(i,j) = \phi(j,i) - \phi(i-1,i) \quad (5)$$

before a final transform to calculate the shortest angular difference:

$$\psi(i,j) = \min(|\phi_d(i,j) \% 2\pi|, |-\phi_d(i,j) \% 2\pi|) \quad (6)$$

This results in a value which equals 1 if the vectors $z_{i-1} \to z_i$ and $z_i \to z_j$ are parallel in the same direction. A value of $\psi = 0.5$ indicates that the vectors are perpendicular, while $\psi = 0$ corresponds to the situation in which $z_{i-1} \to z_i$ and $z_i \to z_j$ are parallel but in opposite directions.

Following Clarke et al. (2022b), we calculate the contribution of δ and ψ to the item-weights with a negative exponential: $e^{-\rho_\delta \delta(j) - \rho_\psi \psi(j)}$.

Likelihood

We can now use these functions to calculate the likelihood of selecting different items. Let z_i ($i = 1, 2, \ldots$) represent the sequence of items that we have selected so far in this trial, while z_j represents the remaining items that have not yet been selected. To simplify the mathematical notation we can define $\delta_j = \delta(z_j, z_i)$ and $\psi_j = \psi(z_{i-1}, z_i, z_j)$. Then weight for selecting each item z_j given our previous selections $z_1, z_2, \ldots, z_i$ is defined by:

$$\omega_{j,i} \sim p_a(i) p_s(j|i) e^{-\rho_\delta \delta_j - \rho_\psi \psi_j} \quad (7)$$

These can be converted into probabilities:

$$p(j|i) = \begin{cases} \frac{\omega(j,i)}{\sum_{z=1}^{N} \omega(z,i)} & \text{if } j \notin \{1, 2, \ldots, i\} \\ 0 & \text{if } j \in \{1, 2, \ldots, i\} \end{cases} \quad (8)$$

In brief, these formulae tell us that the probability of selecting a particular item next depends on its class, the class of the previously selected item, the distance between items, and the difference in directions.

2.3   Towards FoMo v2.0: Absolute Direction

One limitation of the FoMo v1.0 is that it only cares about relative direction, and is blind to the left-to-right scanning behaviour illustrated in Figure 1. To incorporate these behaviours we have to make use of *absolute direction, ϕ.* This is somewhat more complex than our relative direction parameter as we now have to make decisions in our modelling about *which* directions we think may be more heavily weighted by human participants, and how best to model these weights.

Directional Components

Our implementation of absolute directions for FoMo is based on a von Mises mixture model[2], illustrated in Figure 4. In short, we assume that items that can be selected by following one of the four cardinal directions $(0, \frac{\pi}{2}, \pi, \frac{3\pi}{2})$ receive greater weight, with items receiving less weight the further they are from these directions. This is done by defining a new function, f that computes a score for a given angle ϕ:

$$f_{\theta,\kappa}(\phi_j) = 1 + \sum_{k=1}^{4} \theta_\psi \texttt{vonMises}(\mu_k, \kappa_k) \tag{9}$$

I.e., for each direction we need to fit a θ_k, that gives the strength of preference for direction μ_k and a κ_k, that gives the concentration around the direction. A high values means that only directions very close to μ_k receive a large weight, while low values spread the weight over a larger range of directions. We set the mu_k values to be equal to the four cardinal directions.

We can now adopt our weights calculation to take this new feature into account:

$$\omega_{j,i} \sim p_{AvB}(i)p_s(j|i)e^{-\rho_\delta\delta_j - \rho_\psi\psi_j - f_{\theta,\kappa}(\phi_j)} \tag{10}$$

Parameters and Hyper-Parameters

The inclusion of this directional information adds a further two parameters (θ_k, κ_k) for $k = 1, 2, 3, 4$ per direction per condition to our model. I.e, before we even get to modelling individual differences we are looking at 16 new parameters. Furthermore, jointly fitting θ and κ to the data can be challenging: as θ approaches 0, κ becomes relatively unconstrained and vice versa. Therefore, we manually set the κ values as a hyper-parameter. For the results presented in this manuscript, we found that a value of 20 produced good results.

Model Versions

We will test a number of different versions of FoMo, allowing us to determine whether including absolute direction parameters improves the model, and whether absolute and relative direction parameters both need to be included in the model. The versions we will test are as follows:

- v1.0: This is the new and improved implementation of the model presented in Clarke et al. (2022b).

- v1.1: *(presented in supplementary material)* This was used in development to test an alternative way of computing distance.

- v1.2: This is the same as v1.0, but we have removed the relative direction component.

- v1.3: This is the same as v1.0, but adds in four new absolute direction components.

- v1.4: This is the same as v1.0, but adds in eight new absolute direction components.

- v1.5: this is the same as v1.3, but we have removed the relative direction component.

In order to rigorously test our models we split all tested datasets into separate training and testing partitions. As we are interested in how well the random effect structure of our models can capture individual differences between participants, we fit the model using the first $N/2$ trials-per-participant-per-condition, and test to the last $N/2$ trials.

2.4   Implementation Details

This model has been implemented in Stan (a probabilistic programming language) which allows us to estimate the posterior probability for each parameter given observed data. Specifically, we used cmdStan v2.36.0 (Stan Development Team, 2025), R v4.4.1 (R Core Team, 2024) and cmdStanr (Gabry et al., 2025).

Multi-Level Framework

To account for individual differences between participants, we implement a maximal multi-level framework in which each parameter is allowed to vary across conditions and from one participant to the next:

---

[2]the von Mises distribution is simply an extension of the normal distribution to work on the circle.
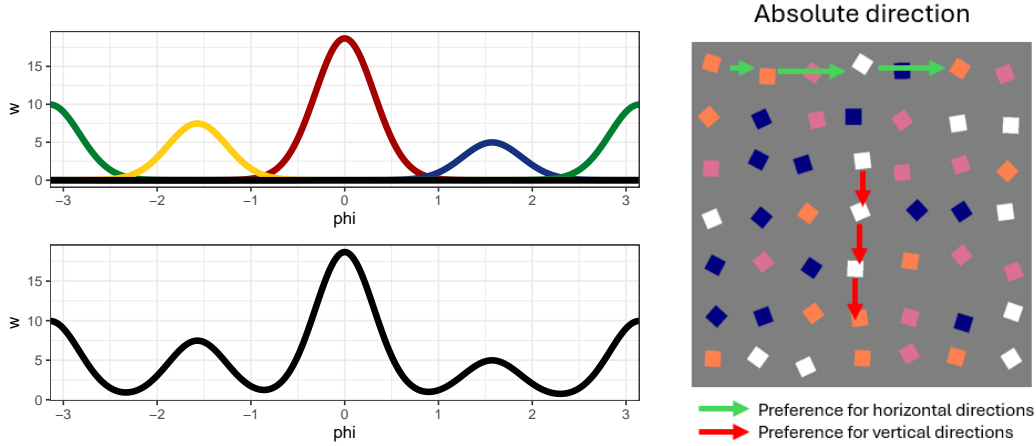
Figure 4

(left): Illustration of four von Mises distributions (top) that can be summed to give (bottom) a multi-model distribution that is able to fleixbly weight horizontal and vertical directions. (right): How these weights could encode different strategies.

$$b_a \sim \text{condition} + (\text{condition|person}) \quad (11)$$
$$b_s \sim \text{condition} + (\text{condition|person}) \quad (12)$$
$$\rho_\delta \sim \text{condition} + (\text{condition|person}) \quad (13)$$
$$\rho_\psi \sim \text{condition} + (\text{condition|person}) \quad (14)$$

A full variance-covariance matrix is estimated for these core parameters. As the experiments analysed in this manuscript all have two conditions, this leads to an 8×8 matrix.

For the directional components we also allow for θ to have a maximal random effects structure:

$$\theta_k \sim \text{condition} + (\text{condition|person}) \quad (15)$$

However, modelling the full covariance matrix is computationally challenging, especially when we move from four to eight directions, as this would produce a 16×16 matrix[3]. As such, we do not model the correlations between distance weights.

Rescaling distances

In our previous work we found that the ρ_δ parameter was typically an order of magnitude larger than the other parameters. This is simply an artefact of the scale that we have used to measure the distances between items in the display. As it is often computationally easier to fit models when the parameters are on similar scales, we have rescaled δ by dividing all distances by a hyper-parameter δ_0:

$$\delta(z_i, z_j) = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{\delta_0} \quad (16)$$

In the models presented in this paper we have used $\delta_0 = 20$.

Priors

Based on the results of Clarke et al. (2022b) we use the following weakly-informative priors for the fixed-effect components:

$$b_a \sim \mathcal{N}(0, 1.5) \quad (17)$$
$$b_s \sim \mathcal{N}(0, 1.5) \quad (18)$$
$$\rho_\delta \sim \mathcal{N}(1, 0.25) \quad (19)$$
$$\rho_\psi \sim \mathcal{N}(0, 1) \quad (20)$$

For the random-effects, we follow Lewandowski et al. (2009) and use an LKJ prior. We use an Exponential(5) distribution for the prior for all group-level variances. See supplementary materials for full model code.

Generated Quantities

To facilitate model comparison we have implemented a *generated quantities* block to our Stan model that computes a range of posterior predictions:

---

[3]Initalising highly constrained parameters, such as high dimensional variance-covariance matrices is computationally challenging

- `log_lik`: the log-likelihood that each item z_i was selected next, given the participant and previous foraging selections in the current trial. While this could in theory be computed for all remaining items in the stimulus, our implementation only computes it for the item that was selected next by the participant.

- *P*: the model's guess of which item will be selected next. This item is randomly sampled from all remaining items using the FoMo weights.

- *Q*: simulating a whole trial from start to finish.

- *F*: the same as *Q*, but with the intial item selection fixed to match the item that was selected in the empirical data.

These values allow us to implement a range of model summary and comparison measures, going beyond the simple posterior density plots used in Clarke et al. (2022b).

2.5 Data

Below we give an overview of the datasets we will make use of in this paper: a brief summary is included in Table 1.

Kristjánsson et al. (2014) was one of the first experiments to investigate multi-target search in humans, with an iPad based paradigm that involved participants 'foraging' for two different types of target (mixed in with two distractor types) by tapping on them sequentially. In the "feature" condition, the targets and distractors were discriminable using just colour as a cue (e.g. the targets were red and green, and the distractors were blue and yellow): in the "conjunction" condition, the categories were defined by both colour and shape (e.g. the targets were red squares and green triangles, and the distractors were red triangles and green squares).

Their experiment used 16 student participants, foraging for 40 targets on each of 20 trials per condition. The dataset used for the following analysis was taken directly from the supplementary material of the journal article: we used trials 6-25 for each participant in each condition, as these were assumed to be the ones used in the original manuscript based on the description in the methods.

Tagu and Kristjánsson (2022) is a paper using a foraging paradigm that also includes a manipulation involving value: in the 'value' blocks of the experiment, participants received more points for selecting a high-value target colour compared to low-value target colours. This condition thus manipulated participants' preference for a particular target type (compared to the 'no-value' block where all target types were equal).

We analysed only the mouse foraging blocks, and trials where the targets were randomly distributed (as opposed to in patches). Note that the paper had 3 different target types, but we have collapsed these into two target types for convenience. The dataset includes data from 24 participants, with 10 trials for each of the 'value' and 'no-value' blocks per participant, where they could find up to 54 targets on each trial (though they did not always have to find all the targets). For our reanalysis, we used 23 participants (we excluded one participant who did not complete all 10 trials).

Clarke et al. (2022c) contains a near-replicate of Kristjánsson et al. (2014), except using a larger number of participants: we used 58 participants in our re-analysis (the same number as in the original paper, except that we also removed one additional participant that was labelled as being a different version of the experiment).

Hughes et al. (2024b) is a dataset on a foraging paradigm that includes a scarcity manipulation: in some trials there were unequal ratios of the two target types, whereas in others there were equal ratios (similar to previous experiments such as Clarke et al. (2022c) and Kristjánsson et al. (2014)). The dataset includes data from 36 participants, with 10 trials for each of 6 conditions: two levels of difficulty (conjunction and feature) vs. three levels of scarcity (scarce target type A, scarce target type B and equal A & B). As no effects of scarcity were found in Hughes et al. (2024b), we have collapsed across different scarcity conditions in the current analyses.

2.6 Analysis Plan

An advantage of generative models is that they can be compared and contrasted with human participants and one another over a range of summary statistics that represent different aspects of behaviour. Not only does this allow us to identify areas in which FoMo preforms well, it also suggests aspects of human foraging behaviour that are not yet captured by our model.

We start with *accuracy*: for each item selection

after the first, we calculate how often our model correctly predicts which item the human participant will select next. Given the nature of the foraging task, baseline (chance) accuracy increases throughout the trial as there are fewer items remaining. We can summarise accuracy across trials to investigate the range of individual differences in predictability, and explore whether some experimental conditions lead to more predictable behaviour.

As well as predicting which specific item will be selected next, we can also summarise our model in terms of its ability to predict when a participant will start a new run of items. Blah.

Whereas our accurracy statistics are based on the item-wise (P) predictions, the rest of our summary statistics are based on the trial-wise Q predictions. A summary is given below:

- *Number of runs*: this is computed per trial, using the `rle()` function in R, which calculates the lengths and values of runs of a vector (in this case, the 'item class' label).

- *Maximum run length*: computed as above, but rather than counting the number of runs, we find the value of the longest continuous run in each trial.

- *Percentage Above Optimal (PAO)*: this is computed using the `ForagingOrg` package for R (Ólafsdóttir et al., 2021). The optimal path is calculated by estimating the shortest Hamiltonian path among the locations of the targets collected, and is compared with the actual path. The difference between the actual and optimal path is expressed as a percentage e.g. 0% would be where the actual path corresponds to the optimal path and 100% would be where the actual path was double the optimal.

- *Best-r*: defined as the best absolute Pearson correlation coefficient between the order in which targets are collected and their respective x and y values (Woods et al., 2013). As above, this is computed using the `ForagingOrg` package.

- *intersection rate*: the number of intersections of the foraging path divided by the ... check details.

- *Lévy flight metric*: This is calculated using the `poweRlaw` package (Clauset et al., 2009). Briefly, we calculate the exponent μ that describes the power-law tail of the movement lengths using a maximum likelihood-based approach (Edwards, 2008).

Our aim is to provide a comprehensive account of our model over as many of the commonly used metrics and summary statistics as possible. This naturally leads to *a lot* of results and comparisons between different datasets, models, and summary statistics. As such, we focus on a selection of the more interesting results in this paper, and provide more comprehensive accounts (including full uncertainty propagation?) in our supplementary materials available from git/osf.

3 Results

In this section we present the results from testing the different versions of FoMo over a range of datasets.

citation	participants	target items	trials	conditions
Kristjánsson et al. (2014)	16	40	20	feature v conjunction
Tagu and Kristjánsson (2022)	23	up to 54	10	value v no-value
Clarke et al. (2022c)	58	40	20	feature v conjunction
Hughes et al. (2024b)	36	20	10	feature v conjunction

Table 1

Table overviewing the data used in this paper. The 'trials' column indicates the number of trials in each condition, and the 'target items' column indicates how many targets the participant had to collect (all were exhaustive search tasks, apart from Tagu and Kristjánsson (2022)). The 'conditions' column indicates the comparisons that were looked at in our analysis, and do not necessarily reflect the full range of conditions that were tested in the original papers.

3.1 Accuracy for predicting items

Overall accuracy is reported in Table 2. We can see that over the course of a trial, all versions of FoMo have an accuracy of around 50% on the majority of datasets.

These headline figures hide a lot of variation within trials, between experimental conditions, and from one participant to the next. For example, if we compute the accuracy of each participant individually, we find there is considerable variation. For example, in Figure 5 (left, top row), we can see that the model can predict some participants with over 60% accuracy, whereas others are closer to 30% accuracy. These results suggest that while the overall accuracy improvements for Fomo v1.3 (with absolute spatial components included) are relatively modest, using only the averages hides some dramatic improvements for individual participants. For example, in the Hughes et al. (2024b) data, displayed in Figure 5 (left, bottom row), one participant who was one of the worst predicted in FoMo v1.0, with around 30% accuracy, turns out to be highly predictable, over 60%, when analysed with FoMo v1.3. Crucially, the inclusion of the new directional component does not decrease accuracy in any participants.

We can also see variation in predictive accuracy as a trial progresses (see Figure 5, right): as would be expected, accuracy increases as the number of items remaining decreases (though it is worth noting that accuracy is far above chance performance even at the beginning of the trial). It is also interesting to note that v1.3 is consistently more accurate, although the difference between the two models is more pronounced earlier in the trial, suggesting that the directional model is most useful when there are more items available in the display.

Our results suggest that, while the addition of a more complex representation of space only offers marginal benefits to the overall accuracy, for some individuals, they offer a large improvement in our ability to account for foraging behaviour. However, at least in these data, moving from four (v1.3) to eight (v1.4) directions in the mixture model appears to offer decreased accuracy. This likely reflects the fact that most foraging experiments have been conducted on a (slightly jittered) cardinal grid, and therefore four directions are often sufficient to describe behaviour: a strategy that involved diagonal directions would also involve longer distances, and proximity is generally a strong predictor of behaviour in these studies.

Interestingly, relative direction seems to offer only negligible improvements when absolute direction has been added to the model: model version 1.5 includes only the absolute direction parameter (with the relative direction parameter removed) and in three out of four datasets, performs equivalently to model version 1.3 in overall accuracy. The exception to this is Kristjánsson et al. (2014), which is the only dataset we are using that makes use of 'finger-based' foraging on an iPad (compared to the mouse-based foraging in the other paradigms). While clearly speculative, this may hint at relative vs. absolute directions having different importances based on modality.

3.2 Accuracy for predicting switches

We now evaluate our model in terms of predicting switches from a run of one class of item to another. As can be seen in Table 3 FoMo offers good performance - 70% – 80% in most conditions. FoMo compares favourably to the classifier by Le et al. (2021) on the Kristjánsson et al. (2014) dataset, especially when considering that this earlier model was trained specifically to detect changes in runs and makes use of a wider selection of features. For conjunction trials our model is within 1.5% accuracy of the classifier, outperforming Le et al. (2021) when classifying switches and underperforming when detecting repeated selections of the same item. The behaviour diverges more in the feature condition, and the comparison looks quite different depending on whether we compare to the initial or oversampled model from Le et al. (2021). Performance for other datasets is available in the Supplementary Material.

3.3 Posterior densities

Figure 6 shows the posterior densities for Fomo v1.3 applied to the Hughes et al. (2024a) dataset (see the Supplementary Materials for others). The parameters for b_a, b_s, ρ_δ and ρ_ψ are similar to those reported in v1.0 and show familiar patterns (e.g. the *conjunction* condition has a stronger stick bias but a weaker proximity bias compared to the *feature* condition). The bottom section of Figure 6 shows the strength of the absolute direction parameters in this dataset: generally, horizontal preferences seem to be stronger than vertical ones, and it seems that the biases are generally stronger in the *feature* compared to the *conjunction* condition. This is consistent with behaviour in the *feature* condition being more driven by spatial layout,

dataset	chance	FoMo version				
		1.0	1.2	1.3	1.4	1.5
Kristjánsson et al. (2014)	0.11	0.48	0.47	**0.50**	0.48	0.48
Tagu and Kristjánsson (2022)	0.08	0.40	0.40	**0.41**	0.40	**0.41**
Clarke et al. (2022c)	0.11	0.49	0.49	**0.51**	0.49	**0.51**
Hughes et al. (2024b)	0.18	0.48	0.47	**0.54**	0.48	**0.54**

Table 2

Mean test set accuracy for each dataset. The difference in the chance baselines reflects the fact that some experiments involved different numbers of target items. The prior model is based on FoMo v1.0. Note: these overall accuracy scores are relatively insensitive to the differences between models, as accuracy rates vary within trials, between conditions and from person to person, as illustrated in Figure 5.
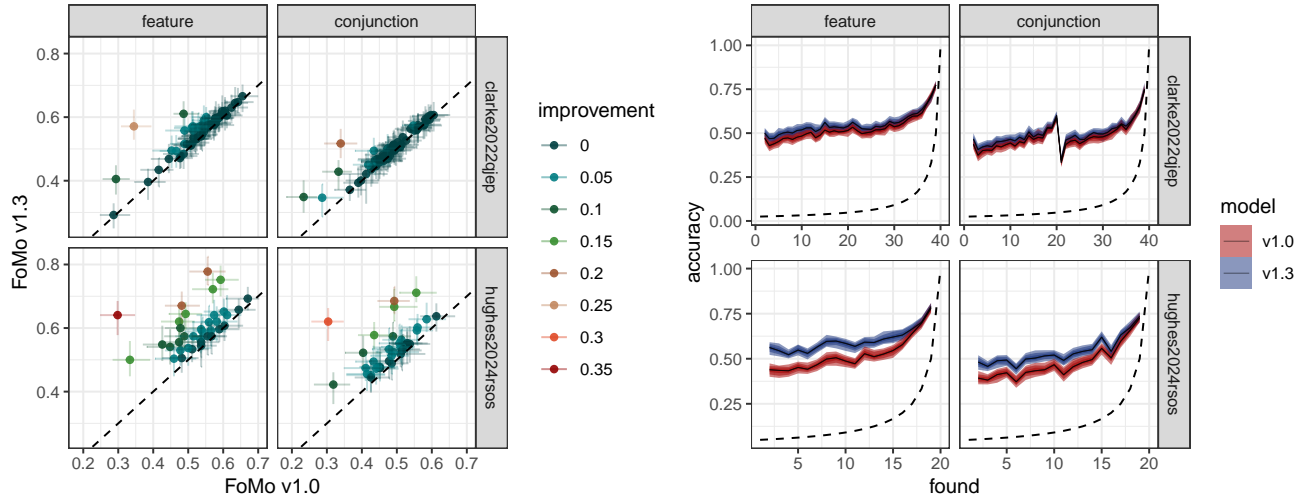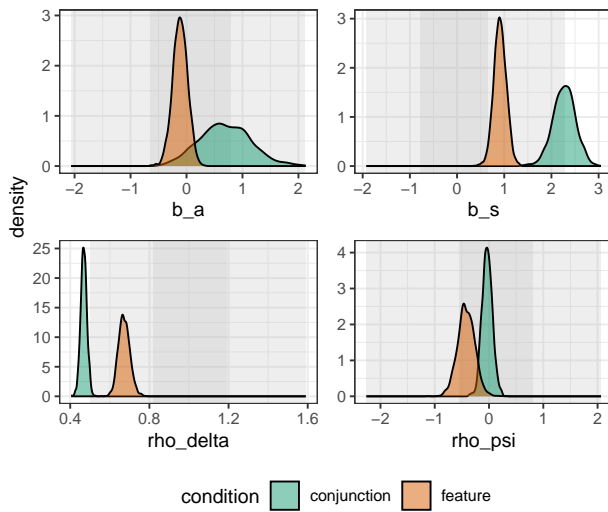

Figure 5

(left): Comparison in accuracy between FoMo v1.0 and 1.3 for the Clarke et al. (2022c) and Hughes et al. (2024b) data. Each point represents an individual participant, and the crosshairs provide the 95% HPDI. Colours indicate the improvement in accuracy between models. (right): 95% HPDI for accuracy throughout a trial for FoMo v1.3. The dotted line indicates chance performance.

condition	switch	Le et al. (2021)		FoMo version		difference
		initial	final	1.0	1.3	
feature	stick	88.0	60.7	74.4	75.2	14.5
feature	switch	51.3	88.1	69.3	70.5	-17.6
conjunction	stick	99.4	93.5	92.0	92.2	-1.3
conjunction	switch	46.7	75.7	76.4	76.8	1.1

Table 3

Comparison between FoMo and the classifier developed by Le et al. (2021) on the Kristjánsson et al. (2014) dataset. We report their final results, post feature pruning and oversampling.
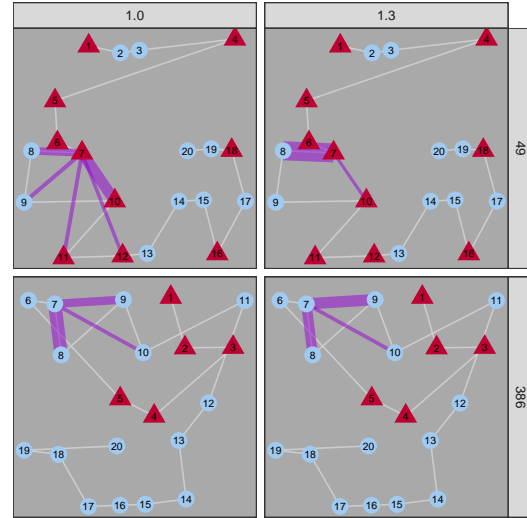
Figure 7

Comparison of simulated paths from FoMo v1.0 and v1.3 with observed data from a human participant from the Hughes et al. (2024b) dataset. The purple lines indicate the weightings that the model assigns to each possible next target (thicker means more weight).

Figure 6

(Posterior densities for FoMo v1.3, applied to the Hughes et al. (2024a) dataset.

whereas the *conjunction* trials are more driven by target properties.

3.4   Posterior Predictions

By using a generative modelling framework, we can also make predictions: i.e. at a given point in any trial, we can ask the model what it thinks that participant will select next. The Bayesian framework of the model also means that we can quantify the uncertainty in this prediction by taking multiple draws from the model posterior. Figure 7 shows some representative examples: the top row shows dramatic improvement for FoMo v1.3, which is clearly driven by the participant's 'left-right' strategy. The absolute direction component allows the model to account for this strategy, and the model makes quite different (and more accu-

rate) predictions from v1.0 accordingly. In the bottom row, we see an example where both v1.0 and v1.3 are relatively uncertain about the best option, although in this case model 1.0 is slightly better, perhaps because this participant seems to have a less consistent strategy.

As discussed above, an advantage of using a generative modelling framework is that we can assess the extent to which FoMo can account for foraging behaviour over a range of metrics commonly used in the literature. In Figure 8 we give an overview of how well FoMo v1.3 is able to predict these metrics across the range of datasets. In general, we can see that FoMo is doing a good job, with its outputs correlating well with all metrics (and in some cases lying almost directly on the identity line). Some of the poorer fits include the maximum run length for Tagu and Kristjánsson (2022): this is likely because this is an inexhaustive search paradigm, and the model currently assumes exhaustive foraging. The predictions are also often slightly biased for best-r and PAO metrics. While the predicted and observed values generally correlate well, the model systematically predicts lower best-r values and higher PAO values than are seen in the data, suggesting that humans

remain more 'optimal' than the model, perhaps through the use of more systematic strategies. We can also see that model v1.3 does a better job of predicting these metrics on aggregate than v1.0 (Figure 10, top: note we have not included Tagu and Kristjánsson (2022) as the model will predict beyond the observed data, confounding error metrics).

3.5 Fixing the first item selection

Clarke et al. (2022a) about first item selections
One factor that affects trialwise predictions is the initial starting point: clearly, if the model starts at a different item to human participants, the overall path is likely to be quite different. Figure 9 shows examples of this. When the starting position is fixed (top row), the model gets the first 8 selections completely correct in one prediction, and generally the top row of targets are selected early in the trial (similar to the participant's real behaviour). The increasing variability in the start position shown in the bottom row leads to more variability in the selections, meaning that while the overall patterns still look similar, trialwise accuracy is likely lower. Across all predictions, Figure 10 (bottom) suggests that the fixed start point also improves predictions of the foraging metrics, as if we take the scaled difference between observed and predicted measures (so that they are comparable) and then average, we can see that the fixed start point model has lower average error for most metrics.

3.6 Different κ values

κ is passed as a hyperparameter to our models with absolute directions included: for the basic version of model 1.3, its value was set as 20. However, to check that this is an appropriate value for κ, we re-ran the models for v1.3 on the Hughes et al. (2024b) dataset with a range of other κ values (κ = 10, 25 and 50). There were no strong differences in overall accuracy for these different values, and thus we used κ = 20 for the other datasets.

4 Discussion

FoMo aims to predict visual foraging behaviour on a target-by-target basis. In the current manuscript, we demonstrate that FoMo is remarkably successful in this aim, achieving an average accuracy of 40-50% for all the datasets we test (compared to a much lower average chance level of around 10-20%, depending on the exact dataset). Importantly, incorporating a parameter for absolute direction into the model improves predictions compared to the 'baseline' model with only relative direction included, and we show this is largely because a small subset of participants show a strong preference for absolute directions, and the model is therefore able to get a lot better at predicting these individuals. The modularity of our modelling approach means that we can also start to ask questions about the most parsimonious set of parameters to predict behaviour: for example, we show that for many (but not all) of our datasets, including absolute direction means that the relative direction parameter is redundant, at least in terms of improving average accuracy. Finally, we believe FoMo can act as a unifying framework for visual foraging research, as it is able to predict a wide range of summary statistics that have been used in previous work.

4.1 Outliers and Individual Differences

In general, we think that one strength of our approach is for characterising individual differences: our model can deal with people who show 'outlier' behaviour, and can actually tell us something about why those people are outliers. For example, in the Hughes et al. (2024b) dataset, we found a number of participants for whom including an absolute direction parameter starkly improved the model's prediction of their behaviour. We think this is particularly exciting because it does not seem to matter that this is relatively rare behaviour: the model predictions improve, without affecting predictions for participants for whom absolute direction matters less. Often modelling approaches deal poorly with outliers or 'rare' events, but FoMo seems to be able to incorporate them well. This modelling approach also provides a principled way to tackle the idea of 'strategies', which are often quite hard to define: for example, we could determine that participants whose model prediction accuracy increases more than a certain percentage when tested with model v1.3 (compared to model v1.0) are using an 'absolute direction' strategy. This has the potential to be very powerful for developmental and clinical work, where we may want to be able to group participants based on characteristic behaviours.

It is interesting to note that we see bigger improvements as we move from model version 1.0 to 1.3 for some datasets compared to others: for example, Clarke et al. (2022c) shows a smaller in-
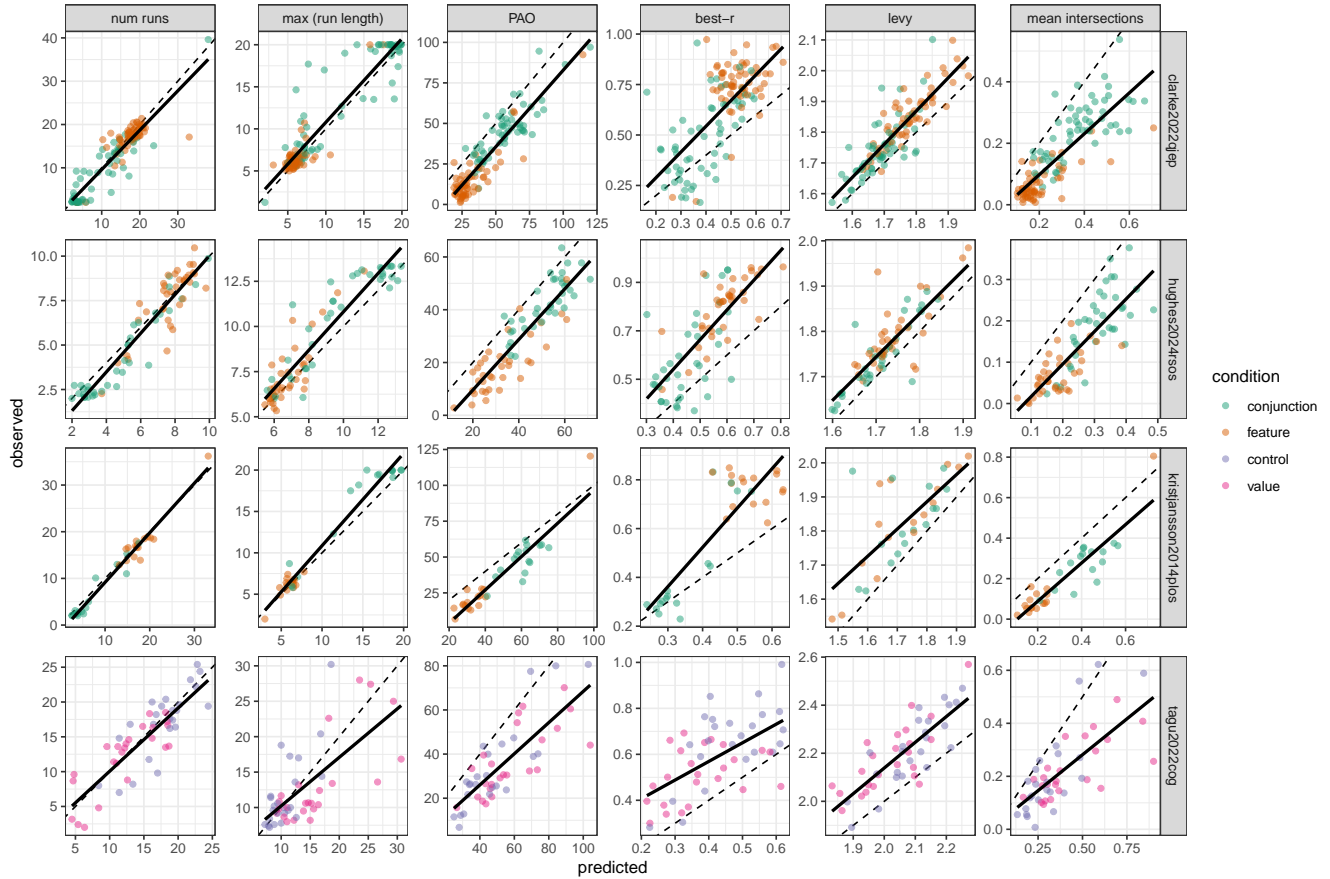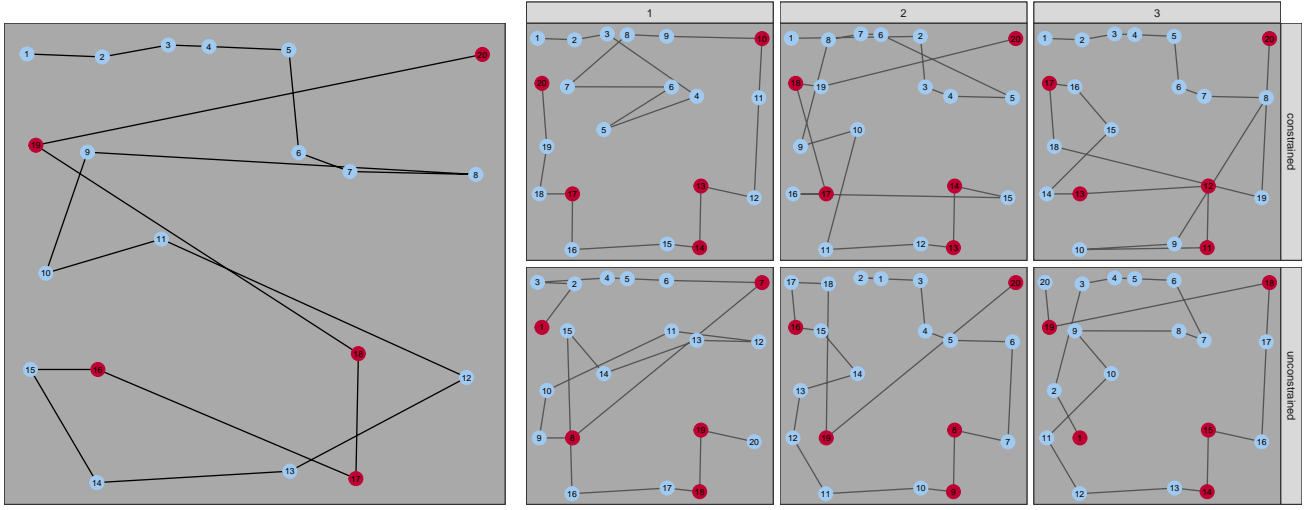
Figure 8

Comparison of predicted and observed path statistics (FoMo v1.3) over a range of datasets. Each dot represents an individual participant in a condition, the dotted line indicates the y = x identity line, while the solid black lines give the best fit linear regression. Note that the poorer performance for the Tagu and Kristjánsson (2022) is almost certainly due, at least in part, by our use of an exhaustive foraging model to account for an inexhastive paradigm.

crease than Hughes et al. (2024b). Hughes et al. (2024b) involved a 'scarcity' based task where on some trials, one target was less common than the other. We had originally hypothesised that participants might show a scarcity bias, preferring to collect the rarer targets first, which we did not find evidence for in our original manuscript, but it does seem that at least for a subset of participants, this manipulation may have pushed them into a scanning-based strategy to ensure they had found all the 'rare' targets. In addition, it is likely that the structure of the foraging stimulus will strongly affect behavioural patterns: to date, most visual foraging studies have used 'grid-like' 2D stimuli that are likely to encourage participants to use an absolute-direction-based strategy, which may be less evident if targets are found in patches, or if

participants must forage in a real-world environment where not all targets are immediately visible. Eye and head movements are also likely to become more important in paradigms that move away from screen-based foraging. We therefore think it may be unrealistic to expect the parameter values in our model to look similar for each foraging dataset, but with more data, we may be able to understand better which environmental conditions change the importance of each parameter, and which parameters are most likely to see large individual differences.

The foraging literature to date has used a wide range of different metrics for evaluating performance, from run-length statistics (e.g. Kristjánsson et al. (2014)) to scanpath metrics such as best-r (e.g. Ólafsdóttir et al. (2021)) or Lévy flight met-

Figure 9

Example of (left): a trial from the Hughes et al. (2024b) dataset and (right): six predictions from FoMo v1.3. The three plots in the top row represent simulations in which we have constrained the initial item selection to match the item selected by the human participant on this trial. The three plots on the bottom are simulations in which the initial item selection is unconstrained.
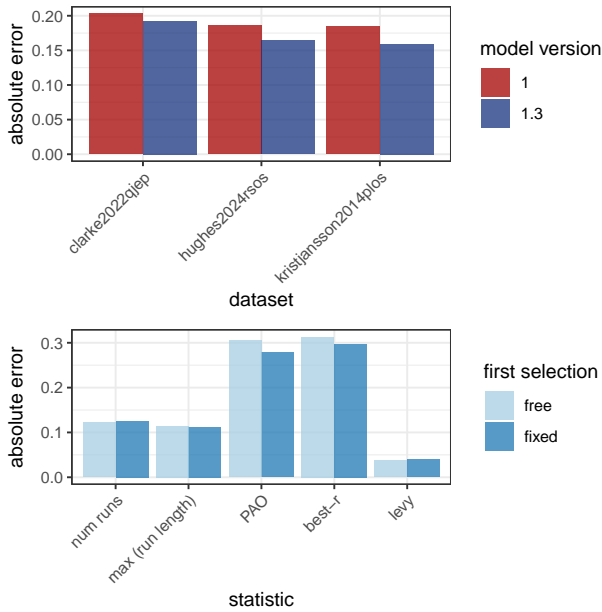
Figure 10

(top): estimate of scaled absolute error averaged across first selection variants and statistics, for each dataset and model version. (bottom): estimate of scaled absolute error averaged across datasets for model v1.3, for fixed and free first selection and for each statistic.

rics (e.g. Edwards (2008)). We show here that by using predictions from the FoMo model, we can generate very good approximations to the true empirical values for essentially all of these metrics. We would therefore argue that FoMo can act as a unifying theory: the model is not explicitly designed to account for these metrics, and yet it is able to make good estimations in most cases, suggesting that it is capturing something fundamental about how behaviour unfolds in this task.

4.2 Limitations and future directions

Although our model is capable of making good predictions of human behavior in a foraging task, there is clearly still room for improvement. One challenge is that some strategies on this task limit the observation space: for example, if a participant always works top-to-bottom to find targets, the model never gets a chance to 'observe' their behaviour when there are potential target items above there current (x, y) position in the display (as these have already all been selected). A consequence of this is that the model's corresponding θ parameter is now unconstrained, limiting our ability to make good predictions in novel situations. Similarly, one common strategy is a 'reading-type' strategy where participants scan targets left-to-right and then do a 'carriage return' back to the left side of the screen to begin again, which the

model cannot easily predict. While it would be possible to include ways to account for this strategy, it is more difficult to envisage a method that would also deal well with non-screen-based tasks, where we predict this behaviour would be less prevalent.

We used von Mises mixture models to implement absolute direction parameters into FoMo, which involves adding θ and κ parameters into our model. However, it can be challenging to fit these parameters simultaneously: in particular, you can end up with one parameter compensating for the other, essentially leading to κ being unconstrained (as if θ is small, κ can be large, and vice versa). We circumvent this issue with use a hyperparameter for κ. Our testing suggests that the exact value of κ has a relatively small effect and does not influence any of the conclusions we draw. However, it is possible that a more sophisticated modelling approach with more tailored priors could avoid the need to use hyperparameters in this way.

FoMo is currently designed to predict target-by-target selection in exhaustive foraging paradigms. It does a reasonable job at predicting inexhaustive search in Tagu and Kristjánsson (2020), but the accuracy is notably lower compared to the other datasets considered in this manuscript; though since we have only one inexhaustive dataset, it is unclear whether this lower accuracy is due to this, or other unique aspects of that dataset (such as its value vs. no value manipulation). However, broadening FoMo to be able to work better with inexhastive foraging by including a parameter for when to 'give up' on a given area of targets would also allow the model to make predictions for 'patch leaving' paradigms, helping to bridge the gap between these two types of foraging experiment that have normally been studied separately.

4.3 Conclusion

We have updated our model of foraging, FoMo, to incorporate absolute direction parameters and have shown that we can make good predictions of target-by-target over a range of datasets. Futhermore we have demonstrated that our model can account for human behaviour over a wide range of secondary metrics and summary statistics that have previously been used in the foraging literature. We would therefore argue that FoMo is the first visual foraging model that can be considered 'unifying', in that it can bring together data and approaches from different research groups and explain them using a single framework. This approach is exciting both for studying foraging and for wider cognitive psychology research for several reasons. Firstly, we believe that emphasising prediction in psychological theories is important as it provides a good way of truly measuring how successfully we are answering the question we actually want to know about i.e. how well can we predict behaviour? Furthermore, We have designed the model to be modular in order to facilitate hypothesis testing i.e. does a given parameter affect behaviour in my paradigm? Finally, we have aimed to make the model as accessible and open source as possible to encourage other researchers to attempt to use it - and break it! - with their own data. By finding situations in whiche FoMo fails to make good predictions, we can learn more about which aspects of foraging behaviour are still poorly explained and gain insight about how to improve our modelling efforts.

5   References

Amor, T. A., Reis, S. D., Campos, D., Herrmann, H. J., and Andrade Jr, J. S. (2016). Persistence in eye movement during visual search. *Scientific reports*, 6(1):20815.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in psychology*, 2:246.

Auger-Méthé, M., Derocher, A. E., DeMars, C. A., Plank, M. J., Codling, E. A., and Lewis, M. A. (2016). Evaluating random search strategies in three mammals from distinct feeding guilds. *Journal of Animal Ecology*, 85(5):1411–1421.

Becker, S. I. (2010). The role of target–distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General*, 139(2):247.

Bella-Fernández, M., Suero Suñé, M., and Gil-Gómez de Liaño, B. (2024). The time course of visual foraging in the lifespan: Spatial scanning, organization search, and target processing. *Psychonomic Bulletin & Review*, 31(1):325–339.

Bella-Fernández, M., Suñé, M. S., and de Liaño, B. G.-G. (2021). Foraging behavior in visual search: A review of theoretical and mathematical models in humans and animals. *Psychological Research*, pages 1–19.

Brooks, J. L., Della Sala, S., and Darling, S. (2014). Representational pseudoneglect: A review. *Neuropsychology review*, 24:148–165.

Buetti, S., Xu, J., and Lleras, A. (2019). Predicting how color and shape combine in the human visual system to direct attention. *Scientific reports*, 9(1):20258.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, 9(2):129–136.

Checketts, M., Mancuso, M., Fordell, H., Chen, P., Hreha, K., Eskes, G. A., Vuilleumier, P., Vail, A., and Bowen, A. (2021). Current clinical practice in the screening and diagnosis of spatial neglect post-stroke: Findings from a multidisciplinary international survey. *Neuropsychological Rehabilitation*, 31(9):1495–1526.

Chen, X. and Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision research*, 46(24):4118–4133.

Clarke, A. D., Green, P., Chantler, M. J., and Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of vision*, 16(7):4–4.

Clarke, A. D., Hunt, A. R., and Hughes, A. E. (2022a). A bayesian statistical model is able to predict target-by-target selection behaviour in a human foraging task. *Vision*, 6(4):66.

Clarke, A. D., Hunt, A. R., and Hughes, A. E. (2022b). Foraging as sampling without replacement: A bayesian statistical model for estimating biases in target selection. *PLOS Computational Biology*, 18(1):e1009813.

Clarke, A. D., Irons, J. L., James, W., Leber, A. B., and Hunt, A. R. (2022c). Stable individual differences in strategies within, but not between, visual search tasks. *Quarterly Journal of Experimental Psychology*, page 1747021820929190.

Clarke, A. D., Stainer, M. J., Tatler, B. W., and Hunt, A. R. (2017). The saccadic flow baseline: Accounting for image-independent biases in fixation behavior. *Journal of vision*, 17(11):12–12.

Clarke, A. D. and Tatler, B. W. (2014). Deriving an appropriate baseline for describing fixation behaviour. *Vision research*, 102:41–51.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

Collins, A. M. and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6):407.

Dawkins, M. (1971). Shifts of 'attention'in chicks during feeding. *Animal Behaviour*, 19(3):575–582.

De Lillo, C., Kirby, M., and James, F. C. (2014). Spatial working memory in immersive virtual reality foraging: path organization, traveling distance and search efficiency in humans (homo sapiens). *American Journal of Primatology*, 76(5):436–446.

De Lillo, C., Visalberghi, E., and Aversano, M. (1997). The organization of exhaustive searches in a patchy space by capuchin monkeys (cebus apella). *Journal of Comparative Psychology*, 111(1):82.

Drew, T., Vo, M. L.-H., Olwal, A., Jacobson, F., Seltzer, S. E., and Wolfe, J. M. (2013). Scanners and drillers: characterizing expert visual search through volumetric images. *Journal of vision*, 13(10):3–3.

Edwards, A. M. (2008). Using likelihood to test for lévy flight search patterns and for general power-law distributions in nature. *Journal of Animal Ecology*, 77(6):1212–1222.

Findlay, J. M. and Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Number 37. Oxford University Press.

Gabry, J., Češnovar, R., Johnson, A., and Bronder, S. (2025). *cmdstanr: R Interface to 'CmdStan'*. R package version 0.9.0, https://discourse.mc-stan.org.

Garg, K. and Kello, C. T. (2021). Efficient lévy walks in virtual human foraging. *Scientific reports*, 11(1):5242.

Gilchrist, I. D. and Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual cognition*, 14(4-8):704–715.

Grössle, I. M., Schubö, A., and Tünnermann, J. (2023). Testing a relational account of search templates in visual foraging. *Scientific Reports*, 13(1):12541.

Harvey Jr, L. O. and Pöppel, E. (1972). Contrast sensitivity of the human retina. *Optometry and Vision Science*, 49(9):748–753.

Hills, T. T., Jones, M. N., and Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological review*, 119(2):431.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., and Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, 19(1):46–54.

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., and Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, 81(7):2288–2303.

Hooge, I. T. C. and Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58:969–976.

Hughes, A. E., Nowakowska, A., and Clarke, A. D. (2024a). Bayesian multi-level modelling for predicting single and double feature visual search. *cortex*, 171:178–193.

Hughes, A. E., Statham, H. R., and Clarke, A. D. F. (2024b). The effect of target scarcity on visual foraging. *Royal Society: Open Science*.

Johannesson, O. I., Thornton, I. M., Smith, I. J., Chetverikov, A., and Kristjánsson, A. (2016). Visual foraging with fingers and eye gaze. *i-Perception*, 7(2):2041669516637279.

Kamaruzaman, A. F., Zain, A. M., Yusuf, S. M., and Udin, A. (2013). Levy flight algorithm for optimization problems-a literature review. *Applied mechanics and materials*, 421:496–501.

Klein, R. M. (2000). Inhibition of return. *Trends in cognitive sciences*, 4(4):138–147.

Kosovicheva, A., Alaoui-Soce, A., and Wolfe, J. M. (2020). Looking ahead: When do you find the next item in foraging visual search? *Journal of Vision*, 20(2):3–3.

Kristjánsson, Á., Jóhannesson, Ó. I., and Thornton, I. M. (2014). Common attentional constraints in visual foraging. *PloS one*, 9(6):e100752.

Kristjánsson, T., Draschkow, D., Pálsson, Á., Haraldsson, D., Jónsson, P. Ö., and Kristjánsson, Á. (2020). Moving foraging into three dimensions: Feature-versus conjunction-based foraging in virtual reality. *Quarterly Journal of Experimental Psychology*, page 1747021820937020.

Kristjánsson, T. and Kristjánsson, Á. (2025). Understanding human visual foraging: a review. *Biological Cybernetics*, 119(4):1–16.

Le, S. T. T., Kristjánsson, Á., and MacInnes, W. J. (2021). Bayesian approximations to the theory of visual attention (tva) in a foraging task. *Quarterly Journal of Experimental Psychology*, page 17470218221094572.

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001.

Lloyd, A., McKay, R., and Furl, N. (2025). Stochastic decisions support optimal foraging of volatile environments, and are disrupted by anxiety. *Cognitive, Affective, & Behavioral Neuroscience*, pages 1–18.

MacInnes, W. J., Hunt, A. R., Hilchey, M. D., and Klein, R. M. (2014). Driving forces in free visual search: An ethology. *Attention, Perception, & Psychophysics*, 76:280–295.

Mairon, R. and Ben-Shahar, O. (2025). The polar saccadic flow model: Re-modeling the center bias from fixations to saccades. *Vision Research*, 228:108546.

Moore, M., Milosevich, E., Beisteiner, R., Bowen, A., Checketts, M., Demeyere, N., Fordell, H., Godefroy, O., Laczó, J., Rich, T., et al. (2022). Rapid screening for neglect following stroke: A systematic search and european academy of neurology recommendations. *European Journal of Neurology*, 29(9):2596–2606.

Morvan, C. and Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS computational biology*, 8(2):e1002342.

Najemnik, J. and Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, 8(3):4–4.

Nowakowska, A., Clarke, A. D., and Hunt, A. R. (2017). Human visual search behaviour is far from ideal. *Proceedings of the Royal Society B: Biological Sciences*, 284(1849):20162767.

Ólafsdóttir, I. M., Gestsdóttir, S., and Kristjánsson, Á. (2020). Age differences in foraging and executive functions: A cross-sectional

study. *Journal of Experimental Child Psychology*, 198:104910.

Ólafsdóttir, I. M., Gestsdóttir, S., and Kristjansson, A. (2021). The development of foraging organization. *Attention, Perception, & Psychophysics*, 83(7):2891–2904.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32(1):3–25.

Prpic, V., Kniestedt, I., Camilleri, E., Maureira, M. G., Kristjánsson, Á., and Thornton, I. M. (2019). A serious game to explore human foraging in a 3d environment. *PloS one*, 14(7):e0219827.

Pyke, G. H. (2015). Understanding movements of organisms: it's time to abandon the lévy foraging hypothesis. *Methods in Ecology and Evolution*, 6(1):1–16.

Pyke, G. H., Pulliam, H. R., and Charnov, E. L. (1977). Optimal foraging: a selective review of theory and tests. *The quarterly review of biology*, 52(2):137–154.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Riggs, C. A., Cornes, K., Godwin, H. J., Liversedge, S. P., Guest, R., and Donnelly, N. (2017). The importance of search strategy for finding targets in open terrain. *Cognitive research: principles and implications*, 2(1):1–17.

Rosenholtz, R. (2024). Visual attention in crisis. *Behavioral and brain sciences*, pages 1–32.

Rovamo, J. and Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental brain research*, 37(3):495–510.

Smith, A. D. and De Lillo, C. (2022). Sources of variation in search and foraging: A theoretical perspective. *Quarterly Journal of Experimental Psychology*, 75(2):197–231.

Smith, T. J. and Henderson, J. M. (2009). Facilitation of return during scene viewing. *Visual Cognition*, 17(6-7):1083–1108.

Stan Development Team (2025). Stan modeling language users guide and reference manual, version 2.18.0.

Tagu, J. and Kristjánsson, Á. (2020). Dynamics of attentional and oculomotor orienting in visual foraging tasks. *Quarterly Journal of Experimental Psychology*, page 1747021820919351.

Tagu, J. and Kristjánsson, Á. (2022). The selection balance: Contrasting value, proximity and priming in a multitarget foraging task. *Cognition*, 218:104935.

Tagu, J., Robert, C., and Mathey, S. (2025). Foraging through emotions: emotional stimuli and participants' trait anxiety shape visual foraging. *Royal Society Open Science*, 12(5):242175.

Thornton, I. M. and Horowitz, T. S. (2004). The multi-item localization (milo) task: Measuring the spatiotemporal context of vision for action. *Perception & Psychophysics*, 66:38–50.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.

Virsu, V., Näsänen, R., and Osmoviita, K. (1987). Cortical magnification and peripheral vision. *Journal of the Optical Society of America A*, 4(8):1568–1578.

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., and Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature*, 381(6581):413–415.

Wang, Z., Satel, J., Trappenberg, T. P., and Klein, R. M. (2011). Aftereffects of saccades explored in a dynamic neural field model of the superior colliculus. *Journal of Eye Movement Research*, 4(2).

Wiener, J. M., Ehbauer, N. N., and Mallot, H. A. (2007). Path planning and optimization in the traveling salesman problem: nearest neighbor vs. region-based strategies. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Wolfe, J. M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28(4):1060–1092.

Wolfe, J. M. and Myers, L. (2010). Fur in the midst of the waters: Visual search for material type is inefficient. *Journal of Vision*, 10(9):8–8.

Woods, A. J., Göksun, T., Chatterjee, A., Zelonis, S., Mehta, A., and Smith, S. E. (2013). The development of organized visual search. *Acta psychologica*, 143(2):191–199.

Zhao, H., Huang, F., Spence, C., and Wan, X. (2017). Visual search for wines with a triangle on the label in a virtual store. *Frontiers in psychology*, 8:2173.