A Rational Information Gathering Account of Infant Habituation

Gili Karni
Princeton Neuroscience Institute
Princeton University
gili@princeton.edu

Marcelo G. Mattar
Department of Psychology
New York University
marcelo.mattar@nyu.edu

Lauren Emberson
Department of Psychology
University of British Columbia
emberson@psych.ubc.ca

Nathaniel D. Daw
Princeton Neuroscience Institute
and Department of Psychology
Princeton University
ndaw@princeton.edu

Abstract

Gaze is one of the primary experimental measures for studying cognitive development, especially in preverbal infants. However, the field is only beginning to develop a principled explanatory framework for making sense of the various factors affecting gaze. We approach this issue by addressing infant gaze from first principles, using rational information gathering. In particular, we revisit the influential descriptive account of Hunter and Ames (1988) (H&A), which posits a set of regularities argued to govern how gaze preference for a stimulus changes with experience and other factors. When H&A's model is reconsidered from the perspective of rational information gathering (as recently also proposed by other authors), one feature of the model emerges as surprising: that preference for a stimulus is not monotonic with exposure. This claim, which has at least some empirical support, is in contrast to most statistical measures of informativeness, which strictly decline with experience. We present a normative, computational theory of visual exploration that rationalizes this and other features of the classic account. Our account suggests that H&A's signature nonmonotonic pattern is a direct manifestation of a ubiquitous principle of the value of information in sequential tasks, other consequences of which have recently been observed in a range of settings including deliberation, exploration, curiosity, and boredom. This is that the value of information gathering, putatively driving gaze, depends on the interplay of a stimulus' informativeness (called *Gain*, the focus of other rationally motivated accounts) with a second factor (called *Need*) reflecting the estimated chance that information will be used in the future. This computational decomposition draws new connections between infant gaze and other cases of exploration, and offers novel, quantitative interpretations and predictions about the factors that may impact infant exploratory attention.

Introduction

Visual attention has long been a primary tool in the study of cognitive processing and development, especially in preverbal infants (Fantz & Ordy, 1959). Gaze (looking time) has been used to examine language, memory, object recognition, and intuitive physics, among many other topics (Damon, Quinn, & Pascalis, 2021; Frank, Vul, & Johnson, 2009; E. K. Johnson & Jusczyk, 2001; Kidd, Piantadosi, & Aslin, 2012; Kosie et al., 2023; Rose, Feldman, & Jankowski, 2004; Thiessen, Hill, & Saffran, 2005). However, the design of such experiments and the interpretation of the large body of existing results are hampered by the relative lack of a principled explanatory framework or predictive model. Such a framework would ideally serve as a linking hypothesis formally grounding gaze in underlying cognitive processes of interest (Aslin, 2007; Aslin & Fiser, 2005; Poli, Serino, Mars, & Hunnius, 2020). Here we aim to address this gap by revisiting gaze preference — and rederiving the key features of a canonical descriptive model of it — from first principles, as an instance of rational information gathering. In this, we build upon and extend another recent model (Cao, Raz, Saxe, & Frank, 2022; Raz, Cao, Saxe, & Frank, 2024) that connects gaze and habituation to learning.

Classically, a primary descriptive framework for gaze is the model of Hunter and Ames (1988) (H&A), which aims to systematize a set of factors argued to directionally affect infant visual attention. H&A envisioned that preference between stimuli arises from a hypothetical, covert response to each stimulus, varying with familiarization time. The key feature of the model is that the response has a nonmonotonic, inverted U-shape: initially rising, then sinking. When comparing some pre-exposed target stimulus to a novel foil, this produces a progression, with exposure, from a transient, early familiarity preference (favoring moderately pre-exposed stimuli over entirely novel ones) to a later novelty preference (disfavoring extensively exposed stimuli). H&A further posit that the overall speed of these dynamics (hence the trade-off between familiarity and novelty preference, all else equal) is moderated by other factors, notably age and stimulus complexity. Although these general dynamics have not been directly tested in a single study (Kosie et al., 2023), there is much empirical evidence for novelty preferences (Colombo & Bundy, 1983; Fantz, 1964; Rose & Feldman, 1987), and some reports of familiarity preferences (S. P. Johnson et al. 2009; Lew-Williams and Saffran 2012; Sonne, Kingo, and Krøjgaard 2023, albeit alongside other studies failing to detect such preferences; Raz et al. 2024).

Although H&A provides an influential descriptive taxonomy for cataloging patterns of directional tendencies, it does not fully serve the ideal of grounding the observable measures in underlying cognitive processes. Its posited U-shaped habituation dynamics are taken as axiomatic, motivated on loose empirical grounds, and not clearly related to functional considerations — particularly cognitive ones like exploration or curiosity. We aim to fill this gap by addressing H&A-style gaze preference dynamics in light of computational models of information gathering, taking the view that gaze is allocated, primarily, to investigate the world. Recent work has used reinforcement learning and statistical decision-theoretic models to investigate a set of common principles motivating informa-

tion gathering in a range of settings, from covert (deliberation) to overt (exploration in trial-and-error tasks, trivia questions) and has argued that these same underlying principles play out in a range of observable measures including choice behavior, neural recordings, and self-reported boredom and curiosity (Agrawal, Mattar, Cohen, & Daw, 2021; Dubey & Griffiths, 2020; Mattar & Daw, 2018). Here, we build on these theories to argue that the same basic model provides a new, rational underpinning for H&A and the data it describes. In turn, this account proposes a concrete, formal grounding for the abstract, covert variables at the center of the H&A program, which may help to guide, constrain, and interpret experimental work aimed at testing the theory (Kosie et al., 2023).

From a pure information gathering perspective, H&A's model has one particularly puzzling feature: the signature nonmonotonic form of its habituation function and the resulting transient familiarity preference. In general, statistical informativeness (as measured by changes in uncertainty, entropy, or sampling error, and formalized as the main engine driving gaze in the recent model of Cao et al. 2022) declines monotonically with repeated samples, meaning that on average each new sample resolves less uncertainty than its predecessors. Why, then, should preference for a stimulus, at least according to H&A, increase with exposure, before decreasing?

Here we suggest that this effect can be understood as a manifestation of a feature common to decision-theoretic accounts of many other exploratory domains: While informativeness generally declines strictly with experience, the *value* of information (e.g., for increasing future reward gathering) does not. In particular, in most sequential decision tasks, information about some situation or stimulus is only valuable to the extent that stimulus is likely to be encountered in the future (Anderson & Schooler, 1991; Mattar & Daw, 2018). This common feature of value-of-information calculations has different consequences in different domains: in deliberation, it means we should focus consideration on imminent decisions (Mattar & Daw, 2018); in memory, we should retain the items most likely to be retrieved again (Anderson & Schooler, 1991); in trial-and-error "bandit" tasks, it implies that the value of exploration is modulated by the number of choices left to make (Agrawal et al., 2021; Gittins, 1979; Wilson, Geana, White, Ludvig, & Cohen, 2014); and similar considerations may explain why curiosity and boredom are also impacted by expectations about future events (Agrawal et al., 2021; Dubey & Griffiths, 2020). In the case of gaze, transient familiarity preferences, then, may reflect the expectation that frequently encountered stimuli are likely to be encountered still more in the future.

In what follows, we extend the value of information approach from Mattar and Daw (2018; Agrawal et al. 2021) to build a novel, computational foundation for gaze dynamics and in particular to rationalize the functions proposed by Hunter and Ames (1988). We then use the model to examine current empirical findings and to put forward new predictions for how visual exploratory behavior may vary between situations and individuals, and to clarify the underlying mechanisms driving it.

Results

A *need* and *gain* model for visual preference

Our aim is to model the dynamics of infant visual habituation by adopting models of optimal information gathering. In particular, this type of model quantifies the value of information-gathering actions in terms of the expected future reward they may enable. Even actions like eye movements that do not directly obtain primary reward may nevertheless ultimately increase rewards gathered, if the information changes decisions the agent later makes in a way that allows them to obtain more reward. For instance, in multi-step trial-and-error reward-gathering tasks, both exploratory choices (Agrawal et al., 2021; Gittins, 1979; White et al., 2019) and steps of planning (Mattar & Daw, 2018) have value in units of reward that can be quantified in terms of how they increase the expected reward earned through later choices. Here we apply this same notion of value of information to infant visual fixations. Unlike for trial-and-error tasks with experimentally controlled rewards, we cannot ground these computations in the specific later decisions and rewards being optimized. We can nevertheless, by analogy, make specific predictions about the sign and direction of many effects.

In particular, we adopt the formulation of Mattar and Daw (2018; Agrawal et al. 2021), who show that in a generic class of sequential decision tasks known as Markov Decision Problems, the value of obtaining information about some "state" (e.g., stimulus) can, suggestively, be decomposed into the product of two terms, which they call *need* and *gain*. *Gain(s)* captures the increment in cumulative expected future reward, based on the change in the decision that would be made each time state s is encountered again: i.e., the return of the updated choice after obtaining the information, compared to before. But in sequential tasks of this sort, these rewards are only realized if the state is visited in the future; *gain(s)* is thus multiplied by *need(s)*, formally the expected future occupancy of s (the expected, time-discounted number of future encounters with s). It is then proposed that exploratory actions for each particular option s be evaluated by computing (perhaps approximately) its *need* and *gain*, and selected between or prioritized on this basis. (This is also known as "greedily optimal control," since the proposed prioritization is "greedy" in the sense of not taking account, at each step, of how exploratory actions may also change the value of other subsequent exploratory actions.)

To analogize experiments on infant visual fixation and habituation to this setting, we assume (similar to Cao et al. 2022; Raz et al. 2024) that infants act as greedy ideal observers encountering a series of stimuli s IID according to some initially unknown multinomial probabilities p. The goal of fixating is statistical sampling: Each stimulus also has some initially unknown hidden variable(s) associated with it, which can be noisily sampled (or not) by fixating it when encountered. Finally — since we lack a full account of how information gained via sampling will drive later decisions for experimentally controlled rewards — we instead assume that this estimation is itself the goal, i.e. that *gain* for this sampling is taken as ultimately proportional, at each encounter, to the information gained about the hidden

quantity by sampling (Kaelbling, 1993).

Accordingly, let us associate the value of fixating some stimulus S_t on some trial t with the value of information:

$$\mathrm{V}(S_t) = \mathrm{gain}(S_t) \times \mathrm{need}(S_t) \tag{1}$$

Even before elaborating the specifics of *need* and *gain* (see Methods), this framework very generally captures H&A's proposed dynamics. The main points concern the qualitative behavior of *need* and *gain* as any particular stimulus s is repeatedly encountered. Because sampling has diminishing returns as the estimates converge, *gain* declines monotonically, and stimulus preference habituates (Cao et al., 2022). For instance, if a random variable is sampled n_s times, the central limit theorem entails that the standard deviation of the sample mean declines proportionally to $\frac{1}{\sqrt{n_s}}$.

If this were the whole story, then repeatedly sampling a stimulus would have strictly diminishing returns; however, *need* has opposite effects, since expectations about how likely a stimulus is to be encountered in the future depend on how frequently it has been encountered previously. In the simplest case, stimuli $S_t = s$ are encountered according to a multinomial distribution with unknown probabilities p_s; if the agent estimates these probabilities using standard Dirichlet inference, then they estimate $p_s = \frac{n_s+1}{\sum_j (n_j+1)}$, where j indices all stimuli. Since *need* (the discounted number of future encounters) grows proportionally to p_s, it grows with repeated encounters n_s.

Finally, to connect these quantities to H&A's thought experiments, we consider two-stimulus forced choice. Following H&A, we take preference as the difference between the values of two alternatives, and specifically, novelty preference as the difference between a novel baseline stimulus b encountered $n_b = 0$ times and a familiarized stimulus s encountered n_s times:

$$\text{novelty preference} = V(b) - V(s) \tag{2}$$

Novelty and familiarity preferences in the model

Figure 1 shows how the product of *need* and *gain* recapitulates H&A's classic U-shaped dynamics. With repeated exposures to some stimulus, *need* increases and saturates, while *gain* decreases monotonically to zero. Their product, the value of sampling the familiarized item, rises then falls, mimicking H&A's proposed U-shaped function. Finally, novelty preference is the difference between the (constant) value of a novel baseline item and the varying value of a progressively familiarized item, resulting in an inverted U-shape function in which the familiar stimulus is first preferred, then the novel one.

While several empirical results support this theoretical model (Kinney & Kagan, 1976; Perone & Spencer, 2014; Roder, Bushnell, & Sasseville, 2000; Rose et al., 2004), perhaps the most direct echo of these dynamics is the study of Kidd et al. (2012). In their first experiment, the authors measured infants' visual attention to target stimuli that appeared, or not, in a sequence with varying levels of appearance frequency vs. a blank background stimulus. Infants were least likely to look away from rare or common stimuli, and more likely to disengage when the appearance frequency was intermediate (Figure 2B). Since stimuli

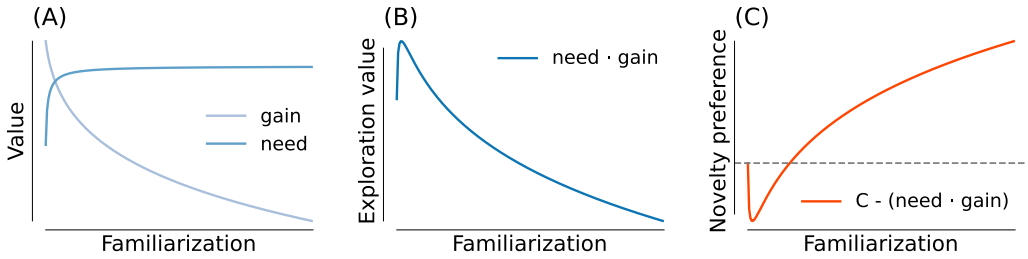


Figure 1: A *need* · *gain* model of novelty preference. (A) *need* and *gain* components, (B) The value of information through familiarization - guiding exploration, (C) Novelty preference as the difference between a constant value of a novel item to the varying one of a familiarized item

that appear more frequently in the fixed-length series are consequently encountered more times, we can view this result as reflecting a range of familiarization levels analogous to the nonmonotonic preference function from Figure 1. Accordingly, our model captures the result (Figure 2A). (Since this is a one-stimulus experiment, we stylize it, for each probability level, as choice between a target stimulus familiarized a number of times proportional to the appearance frequency, and a fixed baseline value reflecting the rest of the environment.)

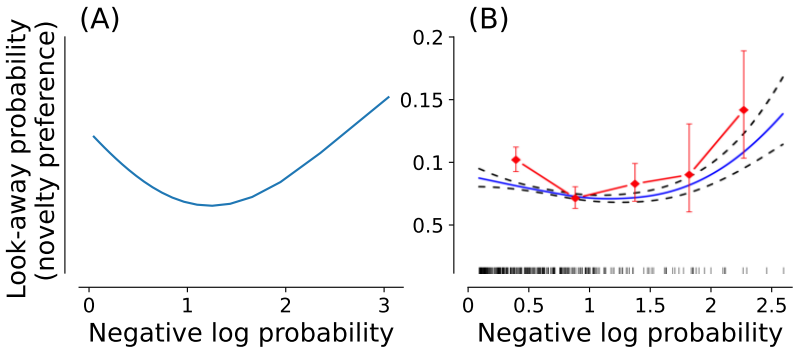


Figure 2: Our model explains empirical results, showing that look-away probability has a U-shaped curve as a function of stimulus appearance frequency. (A) Look away probability resulting from the product of *need* and *gain* (B) Empirical results depicting a U-shape curve of looking away probability as a function of negative log probability. Mean data, with CI, in red (Kidd et al., 2012)

Modulators of habituation

Another aspect of the H&A program was the identification of a few factors (notably age and stimulus complexity) hypothesized to affect the progression of the stimulus familiarization dynamics and thus, all else equal, the predominance of novelty vs. familiarity preference in any particular situation. Here we explore how different factors underlying *need* and/or *gain* might, in the current model, affect the dynamics of gaze preference. These

factors may capture individual differences, cognitive development (or decline), or experimentally manipulable factors. These factors are free parameters of the model, but, importantly (since the model is itself just a generative description of the statistics of stimuli), they reflect features of events that are interpretable and may be experimentally manipulable.

In our model (following Cao et al. 2022; Goodman, Tenenbaum, Feldman, and Griffiths 2008), stimuli are formalized via binary vectors, each element of which is independently drawn from a binomial distribution. Inference then involves estimating the vector of stimulus probabilities from samples. This process drives *gain*, which is modeled as the expected information gain (KL divergence from prior to posterior) upon receiving each sample.

Modulators of *Gain*

We first consider stimulus dimensionality, which we use to formalize the notion of stimulus complexity. *Gain* scales with the length of the stimulus vector, such that a higher dimensionality yields a larger initial value of *gain*. These differences, in turn, drive changes in the exploration dynamics 3A. Since waning *gain* drives the shift toward novelty preference, larger stimulus dimensionality results in slower habituation: a slower shift from familiarity to novelty preference 3B, as in H&A's model.

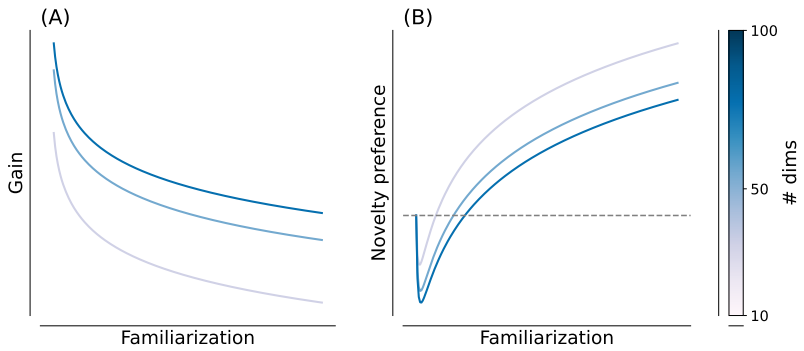


Figure 3: Our model predicts that habituation varies with the number of stimulus dimensions. (A) Larger # of dimensions (f) drives a higher initial *gain* value (B) Larger # of dimensions (f) results in slower habituation

A related parameter is the model's prior belief distribution over the stimulus probability parameters. Bayesian updating famously involves the precision-weighted trade-off between the prior distribution and the data likelihood. Thus, the speed of Bayesian updating reflects the precision of the prior beliefs, such that more diffuse priors change more nimbly upon sampling data, while more precisely focused priors (whether accurate or inaccurate) update more sluggishly. This means that a stronger prior results in a lower initial gain, and also a flatter, slower decay, compared to a weaker one. Stronger prior precision, in turn, drives flatter, slower habituation, whereas weaker prior precision yields faster, more pronounced habituation (Figure 4B).

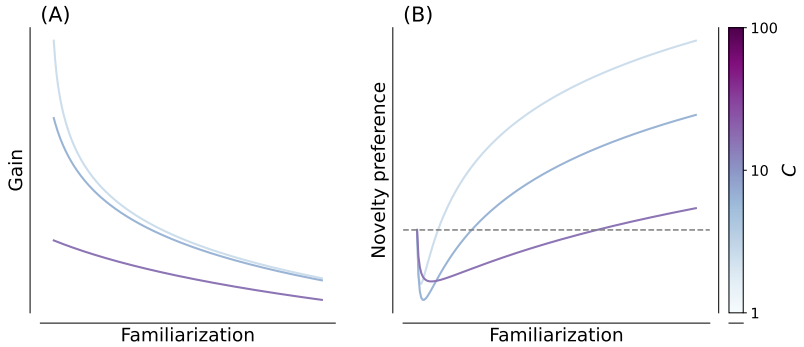


Figure 4: Our model predicts that habituation varies with the precision of prior knowledge. (A) More precise prior belief (higher C) drives a flatter decline of *gain* value, as well as a lower initial value *gain* value (B) More precise prior belief results in flatter and slower habituation

Modulators of *Need*

In the model, we assume the infant takes stimuli to be generated according to a distance-dependent Chinese restaurant process. The Chinese restaurant process is a distribution capturing the idea that an unknown number of stimuli will occur with unknown probabilities, like a weighted die with an unknown number of sides. Furthermore, in a distance-dependent Chinese restaurant process, these stimuli are grouped in time; there is some characteristic timescale over which stimuli are likely to repeat.

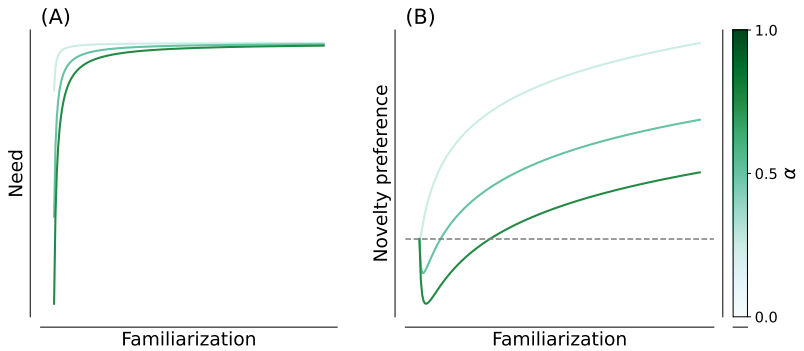


Figure 5: Our model predicts that habituation varies with concentration. (A) Larger concentration (higher α) drives a lower initial *need* value (B) Larger concentration results in slower habituation

The key parameter governing the Chinese restaurant process is the concentration or dispersion parameter α, which captures how often new stimuli arrive that have never been observed before. A high value of α will result in fewer repeated stimuli (in an extreme case, every new stimulus is novel), whereas a low value will yield a denser distribution, where some stimuli repeat many times. Accordingly, a higher dispersion α yields *need* that starts lower and saturates more slowly (Figure 5A). Since the saturation of *need* drives the asymptotic dominance of diminishing *gain* (and the shift to novelty preference), these

differences drive changes in the exploration dynamics such that higher dispersion α results in a slower shift from familiarity to novelty preference (Figure 5B).

Finally, a temporal decay parameter λ affects the timescale over which stimuli are temporally grouped. A stronger decay permits a smaller window in which stimuli tend to recur, while a weaker decay enlarges this window. Accordingly, stronger decay leads to a faster saturation of *need* with time (Figure 6A). In turn, a stronger decay blunts the impact of *need* yielding more transient familiarity preference and a faster transition to novelty preference (Figure 6B).

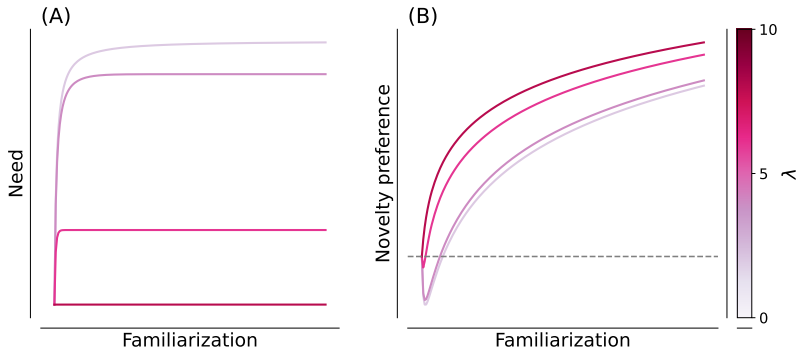


Figure 6: Our model predicts that exploration dynamics vary with temporal decay. (A) Steeper decay (larger λ) drives a faster saturation of the *need* value (B) Steeper decay results in faster exploration dynamics.

Discussion

We propose that significant features of infant exploratory behavior can be captured by a computational model that has previously been used to explain information gathering in other settings in psychology and neuroscience. The model offers a novel rational basis for a predominant account of gaze in developmental psychology, which was not previously expressed in formal terms. A key aspect of the model is that it explains the putatively nonmonotonic preference for stimuli over exposures as the result of the product of two factors: *need* and *gain*, which provide a new theoretical understanding of the rational basis of these phenomena and set the stage for further manipulations testing the theory.

Our model demonstrates the importance of adding *need* to classic information gain considerations to understand the putative nonmonotonicity of exploration and novelty preference. The idea of a decomposing value of information as *need* $\times$ *gain* is not unique to this setting — indeed a major merit of the current proposal is that it provides a unified perspective with other domains including memory (Anderson & Schooler, 1991), deliberation (Mattar & Daw, 2018), exploration (Agrawal et al., 2021), and (in perhaps the closest argument to the current one) curiosity (Dubey & Griffiths, 2020). The use of a common framework suggests that manipulations and measures from these other domains (for instance, manipulating a stimulus' *need* relative to its competitors while holding absolute

experience fixed (Liu, Mattar, Behrens, Daw, & Dolan, 2021) might be brought to bear on infant exploratory behaviors as well.

The *gain* portion of our model draws closely from another recent model of infant gaze, by Cao et al. (2022; Raz et al. 2024), which views moment-to-moment gaze decisions as driven by expected information gain in a similar statistical estimation setting. We adopt as our starting point a similar formalism and rationale, but use a slightly abstracted setting: In particular, we consider sampling as a binary decision at an encounter-by-encounter level, rather than more granular ongoing gaze-or-look away to capture per-stimulus looking times. This simplification is driven by our central goal of augmenting the logic of information gain with the idea that the value of learning about some stimulus or concept is not purely a function of information, but additionally depends on the rate of discrete future encounters with it, i.e. *need* (Anderson & Schooler, 1991).

In our theory, the growth of *need* over multiple encounters drives a transient familiarity preference, as also suggested by H&A. Importantly, this effect is theoretically driven by the target stimulus' encounter rate relative to other stimuli (Mattar & Daw, 2018) — and not by total exposure time, which is (depending on the situation) only imperfectly correlated with it. We suspect this distinction may contribute to the general sense that H&A's putative familiarity preference is, empirically, relatively difficult to observe. For instance, the planned ManyBabies 5 experiment (Kosie et al., 2023) modulates familiarization via controlling the length of a single prolonged exposure. In contrast, the U-shaped response measured by Kidd et al. (2012), which we model here, arises as a function of a stimulus disappearing and reappearing repeatedly. We suggest that the distinction between exposure time and encounter rate may help clarify previous and future experimental results. It is less clear, on statistical learning models, to what extent *gain* should depend on the number of encounters vs total looking time, though here we assume the former for consistency and simplicity. This ultimately depends on the assumed noise model by which concept learning is abstracted as noisy sampling of a prototype (Goodman et al., 2008; Raz et al., 2024), i.e., to what extent new information is resampled within vs. between encounters.

Our model also suggests two parameters that primarily affect gain, and two that primarily affect *need*. *Gain* is scaled by stimulus complexity, which has effects consistent with H&A's proposal that more complex stimuli undergo slower habituation. *Gain* is also affected by the strength of prior information about a stimulus, which might relate to empirical reports of enhanced familiarity preference for common categories like faces (Houston-Price & Nakai, 2004; Matsuda, Okamoto, Ida, Okanoya, & Myowa-Yamakoshi, 2012).

Importantly, one further aspect of *gain* that simple statistical learning models like ours do not directly capture is that some stimuli are objectively more important, and more valuable to learn about, than others. In common with much other work on novelty preferences and exploration (Brafman & Tennenholtz, 2002; Cao et al., 2022; Kaelbling, 1993), our theory assumes the infant takes learning per se as valuable (at least, learning about stimuli likely to be encountered again), notionally as an approximation to a fuller computation of how particular information helps to obtain later rewards. Nevertheless, an important

question for future work is deriving a better formal understanding of how information gain is connected to the objective later reward. Effects of this sort might also contribute to different processing for reward-relevant categories like faces, as well as for other incentives like family members or food.

We also showed how *need* — and in turn, the dynamics of familiarization — is affected by model parameters that capture the (infant's implied assumptions about the) clustering of stimulus encounters in the environment. These parameters might present interesting candidates for experimental manipulation, insofar as infants might be expected to learn the relevant time horizons by which they expect stimuli to be re-encountered and adapt their *need* computations accordingly. Similar timescale effects are seen in animal and human learning experiments (Piray & Daw, 2021) but have not, to our knowledge, been studied in this setting.

Of course, statistical expectations might also differ between individuals, and could change systematically with an individual's characteristics, such as age. In this respect, it is worth saying that neither of our *need*-related parameters (those reflecting cluster reliability and the temporal scale) seem likely to coincide with H&A's proposal that the speed of familiarization dynamics increases with age (Perone & Spencer, 2014; Rose et al., 2004; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). For instance, we might plausibly expect older infants to have longer temporal horizons and to form finer stimulus groupings (larger dispersion parameter; Bornstein and Arterberry 2010), but these would actually both promote slower familiarization. Instead, while these effects might indeed exist, we suspect that the overall net effect of age is dominated by gain: i.e., faster, more effective learning, similar to that arising from effectively less complex stimuli. Separate from the explanation for the effect, the possibility that the overall habituation dynamics speed up with age may be another explanation for why familiarity preferences (anyway transient) have been inconsistently detected in experiments. In particular, as exemplified in the recent study of Raz et al. (2024), adults may habituate too quickly for familiarity preferences to be observed, whereas infant gaze is noisy such that experiments may be underpowered for detecting (or, crucially from our perspective, rejecting) a transient effect.

The model presented has a few limitations. First, the underlying model of stimulus or concept learning via binary feature vectors is highly stylized, and surely fails to capture many aspects of how infants learn about the world. We also neglect aspects of stimuli (including value-relevant ones) while considering only familiarization level. Finally, the formal model is abstract and does not speak to the details how quantities like *gain* and *need* might be approximated. Although future work could apply and expand the model to consider these additions, the current model does demonstrate a rational, straightforward account for a previously perplexing phenomenon.

Methods

Model definition

The proposed model utilizes Bayesian learning to formalize novelty preference by calculating two quantities: namely *need* and *gain* to determine the exploratory value of observed stimuli. In our discrete temporal setting, the model calculates the exploratory value per target stimulus at every encounter, as a function of the stimuli encountered previously, yielding a metric allowing to contrast multiple stimuli across time. That is, we can now consider the simulated subject's preference (difference in exploratory value) as akin to the inverse of their "looking away" tendencies as seen in experimental studies (Cao et al., 2022; Kidd et al., 2012; Poli et al., 2020).

Let us formulate a detailed model of *need* and *gain*. *Need* is formalized as distance-dependent Chinese Restaurant Process (dd-CRP) with exponential decay, reflecting a clustering process for non-exchangeable data (i.e., time-dependent), such that stimuli that have been rarely observed are weighted less (Blei and Frazier 2011; Pitman et al. 2002; see equations 3 and 5). This model accounts for the idea that as a stimulus is repeatedly encountered, it becomes increasingly likely that it will be re-encountered in the future; thus (all else equal), it is more worthwhile to learn about. Let X_t denote the assignment of the stimulus presented at time t, assuming such assignment occupies the set $\mathbf{K}$ of categories of familiar stimuli. In a traditional CRP, the probability $p(X_{t+1} = k|X_{(1:t)}, \alpha)$ is proportional to the number of stimuli previously assigned to category k, with a scaling parameter α governing the probability of initiating a new category, that is, a 'novel' one. Unlike the traditional CRP, the distance-dependent form addresses the inherent order of appearance invariance, allowing us to consider the sequential nature of our setting.

Here, α is the Chinese Restaurant Process dispersion or concentration parameter, reflecting the tendency to assign a stimulus as 'novel'; the higher it is, the more concentrated the stimuli's categorizations will be. λ is a decay parameter, allowing the model to consider the time difference between stimuli, where a larger value favors more contemporaneous stimuli. Importantly, the *need* model is unconcerned with the form of the stimulus X.

$$\text{Need}_{t+1} \propto p(X_{t+1} = k|X_{(1:t)}, \alpha) \tag{3}$$

$$= \frac{\sum \phi(x_t)}{\sum_{k=1}^{K}(\phi(X_t)) + \alpha} \tag{4}$$

$$\phi(t) = \exp(-\lambda t) \tag{5}$$

The setup for the *gain* model follows a sequential hierarchical Bayesian inference process. We describe the generative model by which stimuli are actually produced in our simulations, and the corresponding inference model (which corresponds to Bayesian inference using the true generative model, except where noted). The agent's goal is to use a series of per-trial noisy observations to estimate a stimulus vector, $\boldsymbol{\theta}$. The per-trial obser-

vations are binary vectors $\mathbf{X}$ of length f (for features), with each feature element drawn independently from a Bernoulli distribution with probability given by the corresponding element of the stimulus vector $\boldsymbol{\theta}$ (see Equation 6):

$$\mathbf{X} \sim \text{Bernoulli}(\boldsymbol{\theta}) \tag{6}$$

In turn, the vector of probabilities corresponding to a stimulus, $\boldsymbol{\theta}$ is itself generated from a prior distribution representative of some category of stimuli. In particular, a category is given by a vector of means $\boldsymbol{\mu}$, where for generative purposes each element of the stimulus $\boldsymbol{\theta}$ is independently drawn from a beta distribution centered around the corresponding element of the mean $\boldsymbol{\mu}$. In particular, for we draw each element of $\boldsymbol{\theta}$ from a distribution $\beta(a, b)$, where a is equal to a scale parameter $C = 10$ times the corresponding element of μ and $b = C - a$. (We assume that $.1 <= \mu <= .9$ in order to ensure that the prior distribution is unimodal.)

For each experiment, we simulate $N = 200$ instances of $S = 10$ stimuli in $K = 5$ categories (where the elements of $\boldsymbol{\mu}$ are each randomly drawn from uniform distributions over [.1,.9]), such that $\mathbf{X}c, i, t$ — the instance at the tth trial of the ith stimulus in the cth category — is drawn from a Bernoulli distribution with probability $\boldsymbol{\theta}c, i$ corresponding to the ith stimulus in the cth category (Figure 7 summarizes the generative process).

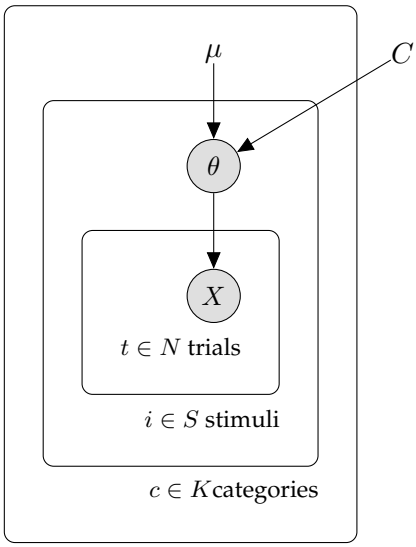
The gain model includes two free parameters. The first is the number of features, f, and the second is the strength of the prior. For most of the experiments (all those except Figure 4), the prior distribution is taken as maximally uninformative and category-nonspecific: $\beta(1, 1)$ for each element, i.e. uniform. We also (in Figure 4) consider different prior variances, in this case assumed to be category-specific (i.e. centered around the true category means μ) but with a range of variances. In this case, we parameterize each element of the prior as $\beta(a, b)$ with a equaling C times the corresponding true category mean from μ. Thus we assume the category mean μ is known a priori to the subject, and they still aim to infer the stimulus-specific θ. However, for the prior variance we consider a range of inferential C (e.g, $1, 10, 100, \ldots$) both larger and smaller than the generating distribution ($C = 10$), so as to produce prior distributions more or less concentrated around the true category means.

The model's belief at trial t for stimulus i in category c, $p(\boldsymbol{\theta}_{c,i,t})$, is given by Equation 7.

$$\begin{aligned} p(\boldsymbol{\theta}_{c,i,t}|\mathbf{X}_{c,i,t}; C, \mathbf{a}, \mathbf{b}) &= p(\mathbf{X}_{c,i,t}|\boldsymbol{\theta}_{c,i,t-1})p(\boldsymbol{\theta}_{c,i,t-1}|\mathbf{a}, \mathbf{b}) \\ &= \beta(C\mathbf{a} + \sum \mathbf{X}_{c,i,t}, C\mathbf{b} + t - \sum \mathbf{X}_{c,i,t}) \end{aligned} \tag{7}$$

Gain is then calculated as the expected information gain (EIG) over the next observation, a metric from Information Theory that has been previously used to measure surprise in cognitive science studies (Cao et al., 2022; Liquin, Callaway, & Lombrozo, 2021; Nelson, McKenzie, Cottrell, & Sejnowski, 2010; Poli et al., 2020; Yin, Chen, Pan, & Tschiatschek, 2021). EIG is approximated as the Kullback-Leibler (KL) divergence between the estimated

Figure 7: Generative model



probability density of some inferred random variable ($\boldsymbol{\theta}$) (true for every cth category and ith stimulus) before vs. after sampling another datum ($\mathbf{X}$), in expectation over the set of possible observation $0, 1$. In Bayesian terms, EIG sequentially models the distance between a prior distribution ($p(\boldsymbol{\theta})$) and the expectations of the posterior distribution ($p(\boldsymbol{\theta}|\mathbf{X})$). Equation 8 depicts the EIG over future observation $\mathbf{X}$ at time $t+1$ summed over all features f. It depicts that a larger update of one's belief would yield a large value of *gain*. Note that the *gain* is unbounded in its upper limit, as it can be any nonnegative number. To make it compatible with *need* (units of probability), we log-transformed it.

$$\begin{aligned}\text{Gain}_{t+1} &= EIG(\mathbf{X}_{t+1}) \\ &= \mathbb{E}_{\mathbf{X}}[D_{\text{KL}}(p(\boldsymbol{\theta}_{t+1})||p(\boldsymbol{\theta}_t))] \\ &= \sum_f \sum_{\mathbf{X}\in\{0,1\}} p(\mathbf{X}_{t+1}|\boldsymbol{\theta}_t) * D_{\text{KL}}(p(\boldsymbol{\theta}_{t+1}|\mathbf{X}_{t+1})||p(\boldsymbol{\theta}_t))\end{aligned} \tag{8}$$

Simulations

The model, described above, is then used to draw multiple simulations under different conditions. The *gain* component involves repeated inference of the underlying parameter $\boldsymbol{\theta}$ governing the series of observed binary vectors $\mathbf{X}$. Due to the stochastic nature of the model, we present the mean estimate over all simulations.

While manipulating each parameter, we held others constant. Baseline values are uniform prior ($\beta(1,1)$), $\alpha = .5$, $\gamma = 0$, $f = 10$, $\lambda = 0.001$.

References

Agrawal, M., Mattar, M. G., Cohen, J. D., & Daw, N. D. (2021). The temporal dynamics of opportunity costs: A normative account of cognitive fatigue and boredom. *Psychological review*.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological science*, *2*(6), 396–408.

Aslin, R. N. (2007). What's in a look? *Developmental science*, *10*(1), 48–53.

Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in cognitive sciences*, *9*(3), 92–98.

Blei, D. M., & Frazier, P. I. (2011). Distance dependent chinese restaurant processes. *Journal of Machine Learning Research*, *12*(8).

Bornstein, M. H., & Arterberry, M. E. (2010). The development of object categorization in young children: hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses. *Developmental psychology*, *46*(2), 350.

Brafman, R. I., & Tennenholtz, M. (2002). R-max-a general polynomial time algorithm for near-optimal reinforcement learning. *Journal of Machine Learning Research*, *3*(Oct), 213–231.

Cao, A., Raz, G., Saxe, R., & Frank, M. C. (2022). Habituation reflects optimal exploration over noisy perceptual samples.

Colombo, J., & Bundy, R. S. (1983). Infant response to auditory familiarity and novelty. *Infant Behavior and Development*, *6*(2-3), 305–311.

Damon, F., Quinn, P. C., & Pascalis, O. (2021). When novelty prevails on familiarity: Visual biases for child versus infant faces in 3.5-to 12-month-olds. *Journal of Experimental Child Psychology*, *210*, 105174.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review*, *127*(3), 455.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*(3644), 668–670.

Fantz, R. L., & Ordy, J. (1959). A visual acuity test for infants under six months of age. *The Psychological Record*, *9*, 159.

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *41*(2), 148–164.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive science*, *32*(1), 108–154.

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development: An International Journal of Research and Practice*, *13*(4), 341–348.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in infancy research*.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of memory and language*, *44*(4), 548–567.

Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract rule learning for visual sequences in 8-and 11-month-olds. *Infancy*, *14*(1), 2–18.

Kaelbling, L. P. (1993). *Learning in embedded systems*. MIT press.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.

Kinney, D. K., & Kagan, J. (1976). Infant attention to auditory discrepancy. *Child Development*, 155–164.

Kosie, J., Zettersten, M., Abu-Zhaya, R., Amso, D., Babineau, M., Baumgartne, H., ... others (2023). Manybabies 5: A large-scale investigation of the proposed shift from familiarity preference to novelty preference in infant looking time.

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *122*(2), 241–246.

Liquin, E. G., Callaway, F., & Lombrozo, T. (2021). Developmental change in what elicits curiosity. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43).

Liu, Y., Mattar, M. G., Behrens, T. E., Daw, N. D., & Dolan, R. J. (2021). Experience replay is associated with efficient nonlocal learning. *Science*, *372*(6544), eabf1357.

Matsuda, Y.-T., Okamoto, Y., Ida, M., Okanoya, K., & Myowa-Yamakoshi, M. (2012). Infants prefer the faces of strangers or mothers to morphed faces: an uncanny valley between social novelty and familiarity. *Biology letters*, *8*(5), 725–728.

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature neuroscience*, *21*(11), 1609–1617.

Nelson, J. D., McKenzie, C. R., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters: Information acquisition optimizes probability gain. *Psychological science*, *21*(7), 960–969.

Perone, S., & Spencer, J. P. (2014). The co-development of looking dynamics and discrimination performance. *Developmental Psychology*, *50*(3), 837.

Piray, P., & Daw, N. D. (2021). A model for learning based on the joint estimation of stochasticity and volatility. *Nature communications*, *12*(1), 6587.

Pitman, J., et al. (2002). Combinatorial stochastic processes. *Lecture notes in mathematics*, *1875*, 7–24.

Poli, F., Serino, G., Mars, R., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science advances*, *6*(39), eabb5053.

Raz, G., Cao, A., Saxe, R., & Frank, M. C. (2024). A stimulus-computable rational model of habituation in infants and adults. *bioRxiv*, 2024–08.

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, *1*(4), 491–507.

Rose, S. A., & Feldman, J. F. (1987). Infant visual attention: Stability of individual differ-

ences from 6 to 8 months. *Developmental Psychology*, *23*(4), 490.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2004). Infant visual recognition memory. *Developmental Review*, *24*(1), 74–100.

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, *18*(5), 704.

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2023). 6-, 10-, and 12-month-olds remember complex dynamic events across 2 weeks. *Journal of Experimental Child Psychology*, *229*, 105627.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53–71.

White, J. K., Bromberg-Martin, E. S., Heilbronner, S. R., Zhang, K., Pai, J., Haber, S. N., & Monosov, I. E. (2019). A neural network for information seeking. *Nature communications*, *10*(1), 5168.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074.

Yin, H., Chen, J., Pan, S. J., & Tschiatschek, S. (2021). Sequential generative exploration model for partially observable reinforcement learning. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 35, pp. 10700–10708).