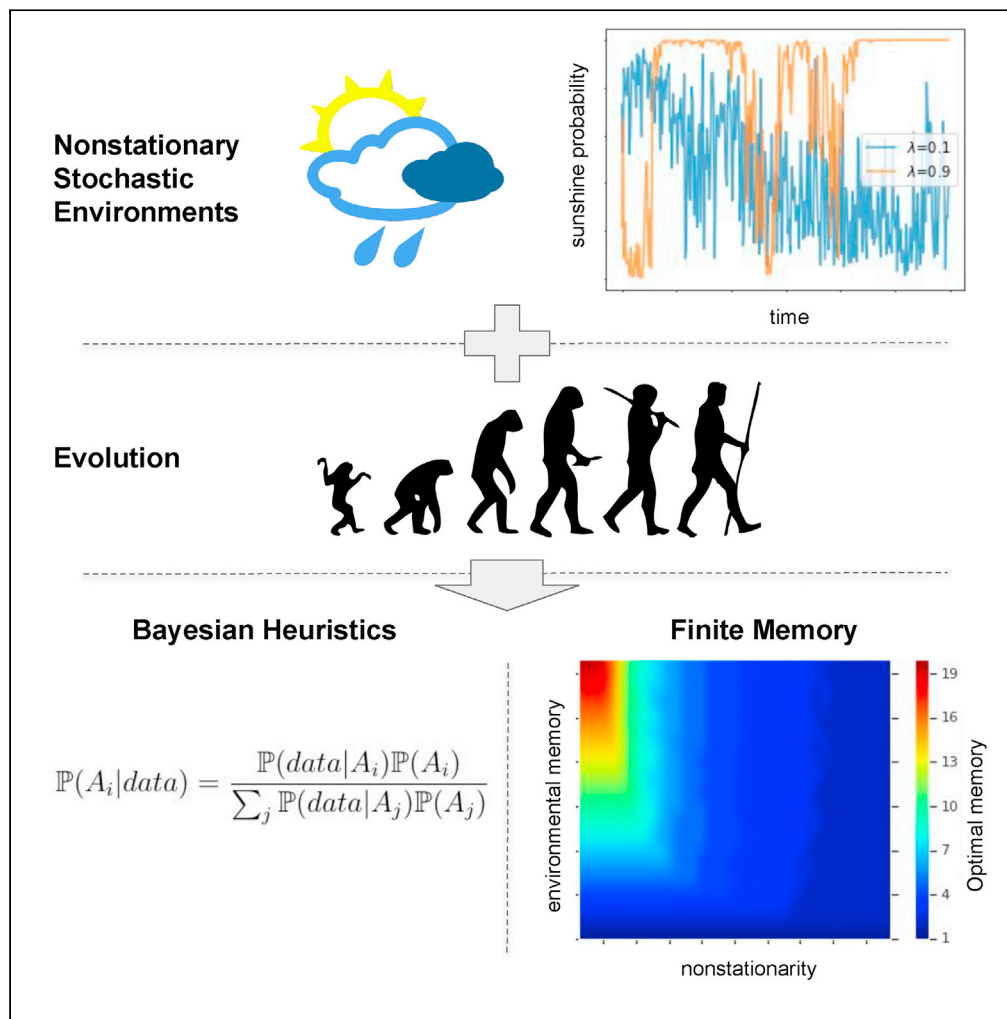


Article

The evolutionary origin of Bayesian heuristics and finite memory



Andrew W. Lo,
Ruixun Zhang

rxzhang@alum.mit.edu

Highlights

Evolution leads to behaviors that resemble Bayesian inference

Finite memory arises in nonstationary environments beyond resource constraints

Intelligence is what advances fitness from the evolutionary lens

This may provide insights for the quest toward artificial general intelligence



Article

The evolutionary origin of Bayesian heuristics and finite memory

Andrew W. Lo^{1,2,3,4} and Ruixun Zhang^{1,5,*}

SUMMARY

Bayes' rule is a fundamental principle that has been applied across multiple disciplines. However, few studies have addressed its origin as a cognitive strategy or the underlying basis for generalization from a small sample. Using a simple binary choice model subject to natural selection, we derive Bayesian inference as an adaptive behavior under certain stochastic environments. Such behavior emerges purely through the forces of evolution, despite the fact that our population consists of mindless individuals without any ability to reason, act strategically, or accurately encode or infer environmental states probabilistically. In addition, three specific environments favor the emergence of finite memory—those that are Markov, nonstationary, and environments where sampling contains too little or too much information about local conditions. These results provide an explanation for several known phenomena in human cognition, including deviations from the optimal Bayesian strategy and finite memory beyond resource constraints.

INTRODUCTION

Bayes' rule is a powerful inferential paradigm widely used in disciplines ranging from biology and cognitive science to signal processing and information theory to economics and finance. Cognitive scientists have applied Bayesian models to understand human inductive learning and generalization (Tenenbaum et al., 2006), visual scene perception and concepts (Lake et al., 2015), language processing and acquisition (Chater and Manning, 2006; Xu and Tenenbaum, 2007), and causal learning and inference (Steyvers et al., 2003; Griffiths and Tenenbaum, 2005), among other topics (Griffiths et al., 2008). Financial economists have also applied Bayes' rule to portfolio analysis (Black and Litterman, 1992; Pástor and Stambaugh, 2000; Avramov and Zhou, 2010). Evolutionary biologists and behavioral ecologists have studied the role of Bayes' theorem in biological evolution (Greaves and Wallace, 2006; Leitgeb and Pettigrew, 2010a, b; Okasha, 2013; Castellano, 2014; Campbell, 2016). One striking result from this line of inquiry is that the Bayes "policy"—that is, the choice of an organism's behavior that maximizes its conditional expected reproductive output—is evolutionarily optimal under a simple model of deterministic strategy when individuals face risks independent of each other (Okasha, 2013). More generally, Suchow et al. (2017) have pointed out the deep correspondence between certain evolutionary dynamics and Bayesian inference, leading to the reinterpretation of evolutionary processes as Bayesian inference algorithms.

Despite its successful applications across multiple domains, few studies have addressed two fundamental questions about the origin of Bayesian inference: how did humans acquire the cognitive capabilities for making Bayesian inferences, and why do these capabilities differ so much from one individual to the next? It is difficult to reconcile the perfect Bayesian strategy with known deviations in actual human behavior. Two examples of these deviations are the result that human subjects follow Bayesian strategies on average, but not individually (Goodman et al., 2008; Vul et al., 2014), and subjects make decisions based on a small number of samples, instead of computing the fully Bayesian answer (Griffiths and Tenenbaum, 2006; Mozer et al., 2008; Vul and Pashler, 2008; Xu and Tenenbaum, 2007; Sanborn and Griffiths, 2008).

In this article, we provide an evolutionary explanation for behaviors that are consistent with Bayesian inference, which shows that such inferences are, in fact, an emergent product of natural selection. Using a binary choice model in which individuals make decisions based on observed data, we first show that decision-making emerges based on the posterior distribution of environmental states, implying that behaviors and heuristics that resemble Bayesian inferences are a natural outcome of evolution. Here, we use the

¹MIT Laboratory for Financial Engineering, Cambridge, MA 02142, USA

²MIT Sloan School of Management, Cambridge, MA 02142, USA

³MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA 02142, USA

⁴Santa Fe Institute, Santa Fe, NM 87501, USA

⁵Lead contact

*Correspondence: rxzhang@alum.mit.edu
<https://doi.org/10.1016/j.isci.2021.102853>



term “resemble” intentionally because the Bayesian heuristics we derive are conditional on an evolutionarily endogenous utility function and do not include an explicit probabilistic representation of the environment. In other words, evolution leads to behaviors and heuristics with the appearance—but not the reality in the strict mathematical sense—of Bayesian inference. (We thank an anonymous reviewer for pointing out this subtle but important point.)

In addition, we identify three factors that constitute intelligence from an evolutionary perspective, including the ability to increase average fitness, to positively correlate with variations in fitness, and to reduce risk when fitness has high variance. The Bayesian heuristics we derive combine all three of these characteristics.

Finally, we demonstrate that finite memory naturally emerges in certain environments, including Markov environments, nonstationary environments, and environments where sampling contains too little or too much information about local conditions. This provides an evolutionary justification for decision-making with small sample sizes and offers a new rationale for limitations on memory that go beyond biological resource constraints (Anderson and Milson, 1989; Lieder and Griffiths, 2020). Our results underscore the importance of environmental conditions in shaping human behavior.

Our model consists of an initial population of individuals that live for one period of unspecified length and engage in a single binary decision that has implications for the random number of offspring they will generate. To the extent that their behavior is linked to fecundity, only the most reproductively successful behaviors will flourish due to the forces of natural selection. The link between different behaviors and their reproductive success, described by a probability distribution function Φ , is crucial in our framework. Embedded in Φ is the biological machinery that is fundamental to evolution, i.e., genetics, but this is of less direct interest to economists than the link between behavior and reproductive success. In this sense, the behaviors we derive are reflective of this link, Φ , rather than of the true underlying probabilities of different states in the world. This is consistent with Hoffman (2016)’s interface theory of perception, which argues that our perceptions of the world do not necessarily represent the world accurately, serving as an interface that allows us to interact with objective reality, while we remain blissfully ignorant of the complexity of that reality. (We thank an anonymous reviewer for pointing out this important connection to Hoffman’s interface theory of perception.)

Darwin’s ground-breaking work on natural selection (Darwin, 1859) has had an impact far beyond studies of biological adaptation. Evolutionary principles have been successfully used to understand human psychology (Buss, 1995) and economic behavior (Lo and Zhang, 2018), including altruism (Hirshleifer, 1977), cooperation (Nowak, 2006), the origin of utility functions (Robson, 1996; Zhang et al., 2014), and the dynamics of financial markets (Lo, 2017). By studying the impact of selection on behavior rather than genes, we are able to derive evolutionary consequences for behavior that cut across details of species, physiology, and genetic origin. In the same way that different levels of magnification of a microscope reveal different details of a specimen, applying evolutionary principles to behavioral variation leads to insights that may be of greater relevance for economics, psychology, and cognitive sciences. Our focus on behavior as the subject of selection serves as an alternate lens through which the effects of evolution may be studied.

THE BINARY CHOICE MODEL

Consider a population of individuals that live for one period, produce a random number of offspring asexually, and then die. Each individual makes only one decision during its life, a choice between one of two possible actions, a and b , which results in one of two corresponding random numbers of offspring, x_a and x_b . We use s_t to denote the state of its environment in period t . It determines the joint distribution of x_a and x_b . The state, s_t , is observable to individuals after period t . We use $\mathcal{S}_t = \{s_{t-1}, s_{t-2}, \dots\}$ to denote the information available to individuals in period t .

In our notation, the state, s_t , is an abstract representation of any information relevant to the environment. For example, when a tribe has a decision to build a shelter either by a river or on a plateau, s_t might be information related to the local weather, used to determine which option will be safer against the elements. When a squirrel decides to forage for food either on the ground or in the trees, s_t might be information about the location of the food relative to the location of predators. When a driver decides whether to

turn left for the tunnel or right for the bridge at a crossing, s_t might be the traffic conditions for both alternatives.

We assume that, once conditioned on state s_t , (x_a, x_b) are independently and identically distributed over time and identical for all individuals in a given generation. In other words, if two individuals choose the same action a , both will produce the same number of random offspring x_a . In such environments, the reproductive risks are systematic, in the sense that all individuals share the same risk. The environment is jointly determined by both s_t and (x_a, x_b) , which are described by a well-behaved probability distribution function $\Phi(\cdot)$ (see Assumptions 1 and 2 in [method details](#)).

After observing all available information, \mathcal{S} , each individual chooses action a with a probability $f(\mathcal{S}) \in [0, 1]$ and b with a probability $1 - f(\mathcal{S})$, where $f(\cdot)$ is a function that maps information into a probability. We shall henceforth refer to $f(\cdot)$ as an individual's "behavior" since it completely determines the individual choice between a and b given the available information, \mathcal{S} . (We drop the subscript t for notational convenience.) Offspring behave in a manner identical to their parents, i.e., they choose between a and b according to the same function $f(\cdot)$. Therefore, the population may be viewed as being segmented into groups of distinct types $f(\cdot)$.

The role of Φ is critical in our framework, as it represents the entirety of the implications of an individual's actions for reproductive success. If action a leads to higher fecundity than action b for individuals in a given population, the particular set of genes that predispose individuals to select a over b will be favored by natural selection, in which case these genes will survive and flourish, implying that the behavior "choose a over b " will flourish as well. This abstraction essentially defines two equivalent classes of actions that have different implications for reproduction, i.e., all actions yielding the same reproductive fitness are considered equivalent in our framework.

The specification of Φ also captures the fundamental distinction between traditional models of population genetics and more recent applications of evolution to behavior; the former focuses on the natural selection of traits (determined by genetics), whereas the latter focuses on the natural selection of behavior. Although behavior is obviously linked to genetics, the specific genes involved, their loci, and the mechanisms by which they are transmitted from one generation to the next are of less relevance to economic analysis than the ultimate implications of behavior for reproduction, which is captured by Φ .

Our binary choice model can be more easily understood through a simple concrete environment that we will use throughout this article. Consider an environment where the weather can either be sunny or rainy, and individuals will need to decide whether to build shelter by the river or on the plateau. The reproductive success of each action is specified as follows:

| Action | Sunny | Rainy |
|---------------|-----------|-----------|
| a (river) | $x_a = c$ | $x_a = 0$ |
| b (plateau) | $x_b = 0$ | $x_b = c$ |

This is a special case of the general environment in our model. At time t , with some probability p_t , the environment is sunny, and choice a yields $c > 0$ offspring, while choice b yields none. With probability $1 - p_t$, on the other hand, the environment is rainy, in which case the reproductive outcome is reversed. The probability of a sunny day, p_t , summarizes the impact on the environment in the state, s_t . We can think of s_t as any variable that affects weather, such as season or temperature.

In this article, we impose various assumptions on the state s_t and determine the type of behavior that emerges in each environment.

BAYESIAN BEHAVIORS IN STATIONARY ENVIRONMENTS

We first consider the case in which the state, s_t , is stationary (see Assumption 4 of [method details](#)). In this case, we can explicitly characterize how fast the population grows as a function of the behavior. Denote by $n_t^{f(\cdot)}$ the total number of offspring of type $f(\cdot)$ in generation t . It is easy to show that $T^{-1} \log n_T^{f(\cdot)}$, the log

population divided by time, converges to the log geometric average growth rate (see [method details](#) (Equation 13):

$$\alpha(f(\cdot)) \stackrel{\text{def}}{=} \mathbb{E}_{(x_a, x_b, \mathcal{S})} [\log(f(\mathcal{S})x_a + (1 - f(\mathcal{S}))x_b)] \quad (\text{Equation 1})$$

as time T increases without bound, where $\mathbb{E}_{(x_a, x_b, \mathcal{S})}$ is the expected value taken over the joint distribution of the environment. Maximizing (1) yields the evolutionarily dominant behavior in this environment, which we refer to as the “growth-optimal” behavior:

$$f^*(\mathcal{S}) = \begin{cases} 1 & \text{if } \mathbb{E}_{(x_a, x_b, \mathcal{S})} [x_b/x_a] < 1 \\ \text{solution to (3)} & \text{if } \mathbb{E}_{(x_a, x_b, \mathcal{S})} [x_a/x_b] \geq 1 \text{ and } \mathbb{E}_{(x_a, x_b, \mathcal{S})} [x_b/x_a] \geq 1 \\ 0 & \text{if } \mathbb{E}_{(x_a, x_b, \mathcal{S})} [x_a/x_b] < 1, \end{cases} \quad (\text{Equation 2})$$

where $f^*(\mathcal{S})$ is defined implicitly in the second case of (2) as follows:

$$\mathbb{E}_{(x_a, x_b, \mathcal{S})} \left[\frac{x_a - x_b}{f^*(\mathcal{S})x_a + (1 - f^*(\mathcal{S}))x_b} \right] = 0. \quad (\text{Equation 3})$$

The expected values in (2) and (3) are with respect to the joint distribution of (x_a, x_b) , given all available information \mathcal{S} .

This growth-optimal behavior is, in fact, what is implied by Bayesian inference. We can see this from two different perspectives. First, the population dynamics between two generations in evolution (also called the “replicator equation” in evolutionary biology) is given as follows:

$$f_t^{f(\cdot)} \stackrel{\text{def}}{=} \frac{f_{t-1}^{f(\cdot)} \cdot \text{Fitness}_t^{f(\cdot)}}{\sum_{g(\cdot)} f_{t-1}^{g(\cdot)} \cdot \text{Fitness}_t^{g(\cdot)}}$$

where $f_t^{f(\cdot)}$ is the relative frequency of behavior $f(\cdot)$ in the population in generation t , and $\text{Fitness}_t^{f(\cdot)} = (f(\mathcal{S}_t)x_{a,t} + (1 - f(\mathcal{S}_t))x_{b,t})$ represents the fitness (that is, the random reproductive success) of behavior $f(\cdot)$ in generation t . This is identical to Bayes formula, where we treat $f_t^{f(\cdot)}$ as the posterior, $f_{t-1}^{f(\cdot)}$ as the prior, and $\text{Fitness}_t^{f(\cdot)}$ as the likelihood (see further details in (Equation 8) and (Equation 9) of [method details](#)), an observation also shared by the recent literature in cognitive neuroscience (Suchow et al., 2017; Czégel et al., 2020).

Second, we can view the population growth rate for behavior $f(\cdot)$ in (1) as the expected value of a logarithmic utility function, where the utility is with respect to the reproductive success of this particular behavior: $f(\mathcal{S})x_a + (1 - f(\mathcal{S}))x_b$. We would like to emphasize that log utility is endogenous in our model and results from the systematic reproductive risks in the environment (all individuals share the same risk; see also Assumption 2 in [method details](#)). More generally, different forms of utility functions and different levels of risk aversion emerge from evolution in environments with different compositions of systematic and idiosyncratic reproductive risks (Zhang et al., 2014).

From this perspective, the dominant behavior in (2) corresponds to maximizing the endogenous expected utility, where the expected value is taken over the posterior distribution of the environment, conditioned on all available information, \mathcal{S} . (We impose no regularity conditions on the behavior $f^*(\cdot)$, other than it being bounded between 0 and 1. In fact, $f^*(\cdot)$ does not even need to be continuous. See also Proposition 2 in [method details](#).) This phenomenon is known as conditionalization in the evolutionary biology literature (Okasha, 2013), and this behavior is what Bayesian inference with a particular utility function would imply because individuals will appear to behave in a manner in which decisions are made in each generation to maximize the growth rate based on all available information at the time. (If the environment is independent over time, there is no benefit from information about past realized states. As a result, intelligent behaviors are not useful, and $f^*(\cdot)$ will simply be a constant (Brennan and Lo, 2011). In the case of our sunny/rainy example, the expected value of the probability of a sunny day, p_t , is constant when the environment is stationary. It is easy to show that the evolutionarily dominant behavior $f^* = \mathbb{E}[p_t]$, which is to build a shelter by the river with the same probability as the expected value of a sunny day. This is known as “probability matching” or “bet hedging” in the evolutionary biology literature (Frank, 2011).)

We use the phrase “appear to behave” in the previous paragraph because our evolutionary framework does not require individuals to possess any Bayesian reasoning ability. In fact, we do not require individuals to be sentient, much less having the capability of probabilistically encoding the environment or

manipulating those probabilities according to the Bayes' rule. In this sense, the dominant behavior may well be simply a heuristic rather than deliberate Bayesian inference. Nonetheless, our framework shows that even purely mindless individuals can adapt to show behaviors that are identical to Bayesian inference through the forces of natural selection. The claim that human minds learn and reason according to Bayesian principles is not a claim that the human mind can implement any Bayesian inference because this can be computationally costly and intractable (Gigerenzer, 2008; Mozer et al., 2008; Gershman et al., 2015). Rather, Bayesian inference gives individuals a rational framework for updating beliefs about latent variables in generative models given observed data (Tenenbaum et al., 2011).

In addition, in our framework, Bayesian-inferential behavior is with respect to the environment only to the extent that it matters for reproductive success. It does not necessarily track the true underlying probabilities of states of the world. In fact, as Hoffman (2016)'s interface theory of perception argues, our perceptions of the world are not "veridical," meaning that they do not necessarily represent the world accurately. (Hoffman (2016) provides a metaphor of computer interface where files and folders are just icons and abstractions of the true underlying transistors, voltages, magnetic fields, and megabytes of a system. In our model, the abstraction of the environment—the joint distribution of (x_a, x_b, \mathcal{S}) —precisely defines such an interface.) The perceptual systems with which we have been endowed by natural selection are a species-specific interface that allows us to interact adaptively and successfully with objective reality, while remaining blissfully ignorant of the complexity of that objective reality (Hoffman, 2016). This is consistent with our abstraction of the environment that matters for reproductive success: the joint distribution of (x_a, x_b, \mathcal{S}) . In this sense, the dominant behavior that resembles Bayesian inference is with respect to the fitness for individuals to reproduce—not necessarily to the actual underlying environment. The problem is not that veridical perceptions are necessarily counter adaptive but rather that veridicality is irrelevant to adaptation (Hoffman, 2016).

The three possible behaviors described in (2) reflect the relative reproductive success of the two choices. Choosing *a* deterministically will be optimal if choice *a* exhibits an unambiguously higher expected relative reproductive success, while choosing *b* deterministically will be optimal if the opposite is true. Randomizing between *a* and *b*, however, will be optimal if neither choice has a clear-cut reproductive advantage. This last outcome is perhaps the most counterintuitive because it is suboptimal from an individual's perspective (Brennan and Lo, 2011).

Because the dominant behavior is non-deterministic in the general case, $f^*(\cdot)$ cannot be observed for each individual based on one instance of the binary choice. It will therefore be difficult, if not impossible, to tell if any given individual exhibits perfect Bayesian behavior. However, at the population level, $f^*(\cdot)$ can be interpreted as the proportion of individuals that take action *a*, which can be easily observed. This is consistent with the observation described earlier that perfectly Bayesian behavior is only present as an average behavior, not in individual people nor in individual trials (Goodman et al., 2008; Vul et al., 2014).

WHAT IS INTELLIGENCE?

Our results also allow us to reexamine what constitutes intelligence from an evolutionary perspective. We can define intelligent behavior as behavior that adapts to the environment and accelerates population growth. The dominant behavior described in (2) is determined by the environment through the relative fitness of the two available choices.

With $r \stackrel{\text{def}}{=} x_a / x_b - 1$ denoting the relative fitness, we are able to rewrite the population growth rate in (1), $\alpha(f(\cdot))$, as follows:

$$\alpha(f(\cdot)) \approx \mathbb{E}[f(\mathcal{S})]\mathbb{E}[r] + \text{Cov}(f(\mathcal{S}), r) - \frac{1}{2}\mathbb{E}[(f(\mathcal{S})r)^2] + \text{constant} \quad (\text{Equation 4})$$

using a second-order Taylor approximation and additional regularity conditions (see [method details \(Equation 22\)](#)). In this notation, intelligent behavior $f(\cdot)$ increases the population growth rate given a particular environment as described by the relative fitness, r .

Each of the three terms in (4) highlights a different factor in determining intelligence. The first term in (4) is the average relative fitness, $\mathbb{E}[r]$. This states that intelligent behavior amplifies average relative fitness. The second term in (4) is the covariance of the behavior, $f(\mathcal{S})$, with relative fitness, r . In other words, intelligent behavior is correlated with the environment, moving in the same direction as relative fitness. The third term

is the second moment of $f(S)r$, its variance—intelligent behavior reduces risk most when relative fitness has high variability.

The Bayesian heuristics we derived are precisely the combination of these three characteristics which, in turn, define intelligence within our model's evolutionary perspective. This is complementary to earlier evolutionary definitions of intelligence (Brennan and Lo, 2012), where it is defined as behavior that is correlated with reproductive success across populations in environments with idiosyncratic risks. Here, intelligence is defined as behavior correlated with reproductive success over time.

FINITE MEMORY IN NONSTATIONARY ENVIRONMENTS

Our analysis has so far focused on stationary environments. In nonstationary environments, there is a trade-off between collecting more information and waiting too long to make a decision, as the past becomes increasingly irrelevant to the future. A balance between continuous learning and forgetting is thus necessary to cope with such environments (Lughofer and Sayed-Mouchaweh, 2015). By extending our model to nonstationary environments, we show that the degree of nonstationarity of the environment plays a critical role in the emergence of finite memory as an adaptive trait.

In nonstationary environments, it is easy to show that $T^{-1} \log n_T^{f^*}$, the log population divided by time, still converges (see [method details](#) (S.20)):

$$\begin{aligned} \mu(f(\cdot)) &= \overline{\mathbb{E}_{(x_{a,t}, x_{b,t}, s_t)} [\log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})]}_{t=1}^{\infty} \\ &\stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{(x_{a,t}, x_{b,t}, s_t)} [\log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})] \end{aligned} \quad (\text{Equation 5})$$

where we use $\overline{\cdot}_{t=1}^{\infty}$ to denote the average over t . (Assumption 5 in [method details](#) discusses the additional conditions of regularity needed for this result (states $\{s_t\}$ are α -mixing). We make the intuitive assumption that the environment forgets about the past asymptotically.) Compared to the growth rate in stationary environments in (1), the expected log fitness is replaced by its average over time. (In nonstationary environments that are independent over time, the dominant behavior is a constant, its explicit formula given in Proposition 3 of [method details](#). In terms of our sunny/rainy example, the dominant behavior, $f^* = \overline{\mathbb{E}[p_t]}_{t=1}^{\infty}$, is the average expected value of a sunny day over time, which still corresponds to probability matching in the average sense.)

We can now ask what is the dominant behavior in the population and what amount of information should $f^*(\cdot)$ use in nonstationary environments? Continuing with our sunny/rainy example, consider an environment where the states follow a linear combination of a simple random walk (SRW) and an autoregressive (AR) process: $s_t = \lambda \cdot SRW_t + (1 - \lambda) \cdot AR_t(k)$. Here, λ is a constant between 0 and 1, SRW_t is a simple random walk defined by $SRW_0 = 0$ and $SRW_t = SRW_{t-1} + z_t$, and $AR_t(k)$ is an AR process of order k defined by $AR_0 = 0$ and $AR_t(k) = 0.9 \cdot k^{-1} \sum_{i=1}^k AR_{t-i}(k) + \varepsilon_t$. (z_t is an IID Bernoulli random variable with an equal probability of being -1 or 1 , and ε_t is an IID standard normal random variable following $N(0,1)$.) The probability of a sunny day is the logistic transformation of the states that bounds it between 0 and 1: $p_t = \text{logistic}(s_t) = (1 + \exp(-s_t))^{-1}$. There are two important parameters that determine this environment: the order of the AR process (k), which determines its degree of memory over time, and the weight for the SRW (λ), which determines the degree of nonstationarity in the environment. [Figure 1A](#) shows two simulated paths of p_t in environments with different degrees of nonstationarity.

Individuals only observe the state of the environment, s_t , but not the realization of the subprocesses, SRW_t and $AR_t(k)$. We assume that individuals use the average of the previous m observed states as their best estimate of the next state and behave accordingly: $f^m(S_t) = \text{logistic}(m^{-1} \sum_{i=1}^m s_{t-i})$. (Strictly speaking, individuals do not have to follow this functional form. However, this behavior is a good approximation of a simple heuristic, and it suffices to illustrate our main point.)

[Figure 1B](#) shows the optimal amount of memory, m^* , that an individual should use in a given environment to maximize the population growth rate as functions of k and λ , based on an evolutionary simulation of 30,000 generations. (To reduce noise, we run the simulation 50 times and take the median values of m^* in each environment.)

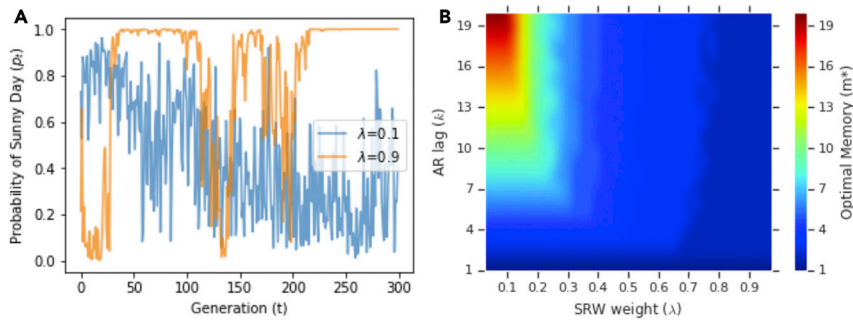


Figure 1. The adaptation of the optimal amount of memory as a function of environmental conditions

(A) Two sample paths of p_t (y axis) from a simulation of 300 generations (x axis) in environments with different degrees of nonstationarity. The blue path is generated from $\lambda = 0.1$ (low nonstationarity), and the orange path is generated from $\lambda = 0.9$ (high nonstationarity), both with an autoregressive lag of $k = 10$.

(B) The optimal amount of individual memory (shown in the color bar) in environments with a degree of memory as parameterized by the autoregressive lag k (y axis) and a degree of nonstationarity as parameterized by the simple random walk (SRW) weight λ (x axis).

Unsurprisingly, given a fixed degree of stationarity in the environment (λ), the optimal amount of memory, m^* , increases with the order of the AR process. In other words, the greater the degree of memory exhibited by the environment, the more memory individuals should adapt to use. In particular, when the environment is mostly driven by the AR process (i.e., when λ is very low), the optimal amount of memory, m^* , equals the maximum AR lag—that is, individuals adapt to the environment to match its memory.

Conversely, given a fixed amount of memory in the AR process, the optimal amount of memory, m^* , decreases as the degree of stationarity (λ) increases. This highlights another reason for finite memory and, more generally, for decision-making based on limited samples: the nonstationarity of the environment. When the environment is nonstationary and its distribution changes over time, it becomes costly for individuals to use out-dated data.

Finally, it should be observed that only an environment with high stationarity *and* high memory (the upper left quadrant of Figure 1B) warrants that an individual possess a large memory, highlighting the fact that finite memory can emerge in environments with either low memory or high nonstationarity. Indeed, in the vast majority of the area in Figure 1B, the optimal behavior only uses one or two observed states in the past.

This result is consistent with the well-documented phenomenon in cognitive science that human subjects tend to make decisions based only on a small number of data points instead of computing the fully Bayesian answer (Griffiths and Tenenbaum, 2006; Mozer et al., 2008; Vul and Pashler, 2008; Xu and Tenenbaum, 2007; Sanborn and Griffiths, 2008; Vul et al., 2014). The small-sample phenomenon is often explained as the consequence of costly data acquisition and Bayesian inference, whereas our model provides an alternate—and arguably more fundamental—explanation. Nonstationary environments induce adaptation to rapid generalization and decision-making based on a short memory, even if there is no cost for using an infinite amount of data.

This result also provides an explanation for the recency effect documented in a wide range of disciplines (Kahneman, 2011), which is often considered to be a bias relative to a fully rational agent. However, our results show that in nonstationary environments, the recency effect is, in fact, desirable, highlighting the key role of the environment in shaping behavior.

SAMPLING WITH LIMITED INFORMATION

We now turn to the context of cognition with limited computational resources (Lieder and Griffiths, 2020). We show that if obtaining and using data for inference is costly, the information contained in each data point also determines the optimal number of data points to use. Both too much or too little information can lead to inferences based on a small number of data points.

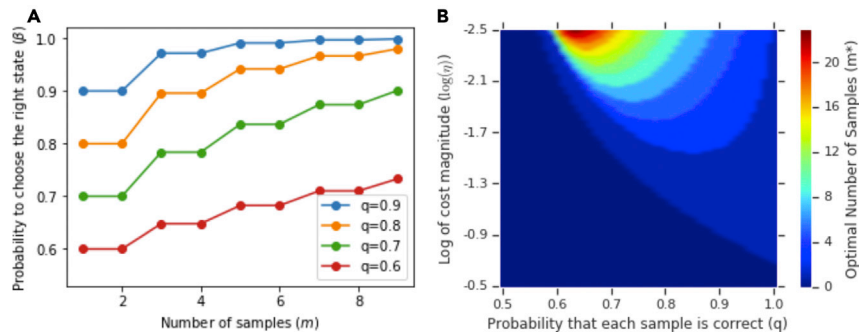


Figure 2. The optimal sample size as a function of the information contained in each sample

(A) Probability, β , that the majority of samples predicts the correct state (y axis), as a function of the number of samples, m (x axis), and the probability, q , that each sample is correct.

(B) Optimal sample sizes (shown in the color bar) given different levels of information contained in one sample (x axis) and the magnitude of the cost, $\log_{10}(\eta)$ (y axis). Greater absolute values of $\log_{10}(\eta)$ indicate a lower magnitude of cost.

Continuing with our sunny/rainy day example, we assume that the probability of a sunny day, p_t , is either $p = 0.8$ or $1 - p = 0.2$ with equal probability in each period. Once again, there are only two states in this environment. In each period, individuals receive imperfect information about their state. The information comes in the form of samples, $D = \{d_1, d_2, \dots\}$. Each sample, d_j , tells an individual its state with an accuracy $q > 0.5$: $\mathbb{P}(d_j = \text{correct state}) = q$. We can think of D as a weather forecast, and individuals are able to use data from D for their decisions. We assume that the cost of obtaining m samples is $C(m) = 1 + \eta m$, where $\eta \geq 0$ determines the magnitude of the cost. The cost is used as a multiplicative factor in reproductive success. (Method details (Equation 30) provides further details about how the cost of obtaining samples is incorporated into calculating the growth rate. The cost is multiplicative to the number of offspring of an individual. For example, when $\eta = 0$, there is no cost; when $\eta = 1$, $C(1) = 2$, which means that one sample cuts the reproductive success by half; when $\eta = 0.1$, $C(10) = 2$, which means that 10 samples also cut the reproductive success by half.)

We can now calculate the optimal number of data points individuals should use to determine their behavior $f(\cdot)$. Given m samples, if an individual follows a simple majority rule, the probability that the majority indicates the correct state is: $\beta = 1 - \text{CDF}(m/2; m, q)$, where $\text{CDF}(m/2; m, p)$ is the cumulative binomial distribution function. (When the data indicate a tie between the two states, we assume individuals choose randomly.) Figure 2A shows β as a function of m and q . When the information in each sample, q , is high, the confidence of the aggregate prediction of multiple samples quickly approaches 1 as the sample size increases, indicating that the marginal benefit of an extra sample quickly diminishes.

We now demonstrate that the optimal sample size depends on the amount of information contained in one sample. (Method details (Equation 32) provides the analytical solution of the optimal behavior given m samples, and (Equation 33) provides the corresponding population growth rate. The optimal number of samples, m^* , can be obtained by maximizing the growth rate over m .) Figure 2B shows the optimal sample size as a function of the information contained in each sample (q) and the magnitude of the cost (η in log scale). Given a fixed cost, finite memory appears when information contained in each sample is either too low or too high. When the amount of information in each sample is low (i.e., the q is close to 0.5), there is little to gain from each sample. As a result, it is not worth paying the cost for any sample. On the other hand, when the amount of information in each sample is very high (i.e., the q is close to 1), one sample is sufficient to infer the state reliably. As a result, it is not worth paying the cost beyond just one sample.

This provides an additional reason for the emergence of finite memory beyond nonstationary environments. It is only desirable to obtain a large number of samples when there is partial information. Otherwise, we adapt to generalize from a small number of samples. (Vul et al. (2014) considered a similar model, which also showed that individuals should make guesses based on very few samples).

DISCUSSION

In summary, we have provided a framework to explore the evolutionary origins of decision-making that is simple enough to uncover the most primitive forms of behavior, yet general enough to derive consequences that likely apply across evolutionary lineages. The same principles and insights apply equally to the evolution of mammalian foraging behavior and human economic decisions in financial markets.

Given its similarity to biological evolution, the implications of our framework can be tested through controlled laboratory experiments, such as those described for experimental evolution (Burnham et al., 2015). Central to this idea is the creation of test and control environments that vary in payoffs—or in a biological context, fitness. Various species, ranging from bacteria to *Drosophila* (fruit flies), have been used to design experiments to better understand decision-making under uncertainty (Mery and Kawecki, 2002; Yang et al., 2008; Beaumont et al., 2009; Dunlap and Stephens, 2014). [Method details](#) provides further details on specific designs of these experiments.

Finite memory emerges as an evolutionary adaptation to Markov environments, nonstationary environments, and environments in which samples contain either too little or too much information. We therefore conclude that we are hardwired to make quick decisions based on limited information precisely because evolution taught us to do so. From a policy perspective, our conclusion also highlights the importance of maintaining stability in public policy and financial markets. Otherwise, critical information about the past is doomed to be quickly forgotten.

Bayes' rule is now widely used in the field of machine learning, from the classical Bayes optimal classifier and multi-arm bandit problems to the more recent applications in deep learning (Kendall and Gal, 2017; Wang and Yeung, 2020; Wilson and Izmailov, 2020) and reinforcement learning (Strens, 2000; Ghavamzadeh et al., 2015). Therefore, understanding the evolutionary and cognitive origins of Bayesian heuristics can shed light on what machine learning algorithms derive more naturally from human intelligence and what kinds of inference (or algorithms) are more suitable for different environmental conditions in terms of their stochastic nature. Just like the "learning machine" Alan Turing proposed seven decades ago to parallel the principles of evolution (Turing, 1950), the fact that machine learning can benefit from nature's bounty for both inspiration and mechanism is more relevant today than ever in our quest for artificial general intelligence (Goertzel and Pennachin, 2007).

Limitations of the study

We have developed a formal theory of the evolution of Bayesian heuristics in stochastic environments in the context of the simplest possible nontrivial choice problem, a binary choice. The model developed here is by no means a complete description of reality. We have deliberately made minimal assumptions, allowing us to determine the most fundamental links between stochastic environments and adaptive behavior. However, more insights about, for example, strategic behavior, can be derived from a generalization of our framework to incorporate the interplay between sophisticated human behavior and highly complex environments with multiple unknown factors.

Our model also provides a theoretical framework to balance memorization versus generalization in nonstationary stochastic environments. Empirical validations of these insights require future work, and potential approaches include experimental evolution (see "Testable Implications" of [method details](#)). Another promising direction is to develop machine learning algorithms that are adaptive with respect to the three environmental characteristics we identified for finite memory (Markov environments, nonstationary environments, and environments where sampling contains too little or too much information about local conditions), which may lead to smarter utilization of past data and yield better trade-offs between memorization and generalization.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
 - Lead contact

- Materials availability
- Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
 - Related literature
 - Binary choice model
 - Stationary environments
 - Nonstationary environments
 - Environments with cost
 - Idiosyncratic reproductive risks
 - Testable implications
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.102853>.

ACKNOWLEDGMENTS

Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged. We also thank two anonymous reviewers for their insightful comments and Jayna Cummings for editorial assistance. The views and opinions expressed in this article are those of the authors only and do not necessarily represent the views and opinions of any institution or agency, any of their affiliates or employees, or any of the individuals acknowledged above.

AUTHOR CONTRIBUTIONS

Conceptualization and supervision, A.W.L.; methodology, formal analysis, writing - original draft, writing - review & editing, visualization, R.Z. and A.W.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 20, 2021

Revised: June 23, 2021

Accepted: July 9, 2021

Published: August 20, 2021

REFERENCES

- Abbott, J.T., and Griffiths, T.L. (2011). Exploring the influence of particle filter parameters on order effects in causal learning. *Proc. Annu. Meet. Cogn. Sci. Soc.* 33, 2950–2955.
- Alexander, R.D. (1974). The evolution of social behavior. *Annu. Rev. Ecol. Syst.* 5, 325–383.
- Almenberg, J., and Dreber, A. (2013). Economics and evolution: complementary perspectives on cooperation. In *Evolution, Games, and God*, M. Coakley and M. Nowak, eds. (Harvard University Press), pp. 132–149.
- Anderson, J.R., and Milson, R. (1989). Human memory: an adaptive perspective. *Psychol. Rev.* 96, 703.
- Anderson, J.R., and Schooler, L.J. (1991). Reflections of the environment in memory. *Psychol. Sci.* 2, 396–408.
- Avramov, D., and Zhou, G. (2010). Bayesian portfolio analysis. *Annu. Rev. Financial Econ.* 2, 25–47.
- Baker, C.L., Tenenbaum, J.B., and Saxe, R.R. (2007). Goal inference as inverse planning. *Proc. Annu. Meet. Cogn. Sci. Soc.* 29, 779–784.
- Beaumont, H.J., Gallie, J., Kost, C., Ferguson, G.C., and Rainey, P.B. (2009). Experimental evolution of bet hedging. *Nature* 462, 90.
- Becker, G.S. (1976). Altruism, egoism, and genetic fitness: economics and sociobiology. *J. Econ. Lit.* 14, 817–826.
- Black, F., and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts J.* 48, 28–43.
- Blume, L., and Easley, D. (1992). Evolution and market behavior. *J. Econ. Theor.* 58, 9–40.
- Brennan, T.J., and Lo, A.W. (2011). The origin of behavior. *Q. J. Finance* 1, 55–108.
- Brennan, T.J., and Lo, A.W. (2012). An evolutionary model of bounded rationality and intelligence. *PLoS One* 7, e50310.
- Burnham, T.C., Dunlap, A., and Stephens, D.W. (2015). Experimental evolution and economics. *SAGE Open* 5, 2158244015612524.
- Buss, D.M. (1995). Evolutionary psychology: a new paradigm for psychological science. *Psychol. Inq.* 6, 1–30.
- Campbell, J.O. (2016). Universal darwinism as a process of bayesian inference. *Front. Syst. Neurosci.* 10, 49.
- Castellano, S. (2014). Bayes' rule and bias roles in the evolution of decision making. *Behav. Ecol.* 26, 282–292.
- Chater, N., and Manning, C.D. (2006). Probabilistic models of language processing and acquisition. *Trends Cogn. Sci.* 10, 335–344.
- Cooper, W.S., and Kaplan, R.H. (1982). Adaptive "coin-flipping": a decision-theoretic examination of natural selection for random individual variation. *J. Theor. Biol.* 94, 135–151.

- Courville, A.C., Daw, N.D., and Touretzky, D.S. (2006). Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* 10, 294–300.
- Czégel, D., Giaffar, H., Zachar, I., Tenenbaum, J.B., and Szathmáry, E. (2020). Evolutionary Implementation of Bayesian Computations (bioRxiv), p. 685842.
- Czégel, D., Zachar, I., and Szathmáry, E. (2018). Major Evolutionary Transitions as Bayesian Structure Learning (bioRxiv), p. 359596.
- Darwin, C. (1859). *On the Origin of Species* (Routledge).
- Davis, E.S., and Marcus, G.F. (2020). Computational limits don't fully explain human cognitive limitations. *Behav. Brain Sci.* 43, e7.
- Del Moral, P. (1997). Nonlinear filtering: interacting particle resolution. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics* 325, 653–658.
- Ditzler, G., Roveri, M., Alippi, C., and Polikar, R. (2015). Learning in nonstationary environments: a survey. *IEEE Comput. Intelligence Mag.* 10, 12–25.
- Dunlap, A.S., and Stephens, D.W. (2014). Experimental evolution of prepared learning. *Proc. Natl. Acad. Sci.* 111, 11750–11755.
- Frank, S.A. (2011). Natural selection. i. variable environments and uncertain returns on investment. *J. Evol. Biol.* 24, 2299–2309.
- Frank, S.A., and Slatkin, M. (1990). Evolution in a variable environment. *Am. Naturalist* 136, 244–260.
- Frydman, C., and Jin, L.J. (2018). Efficient Coding and Risky Choice (SSRN), p. 3270773.
- Furnham, A. (1986). The robustness of the recency effect: studies using legal evidence. *J. Gen. Psychol.* 113, 351–357.
- Gershman, S.J., Horvitz, E.J., and Tenenbaum, J.B. (2015). Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science* 349, 273–278.
- Ghavamzadeh, M., Mannor, S., Pineau, J., and Tamar, A. (2015). Bayesian reinforcement learning: a survey. *Foundations Trends Machine Learn.* 8, 359–492.
- Gigerenzer, G. (2008). *Rationality for Mortals: How People Cope with Uncertainty* (Oxford University Press).
- Gigerenzer, G., and Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox* (MIT press).
- Glenberg, A.M., and Swanson, N.G. (1986). A temporal distinctiveness theory of recency and modality effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 3.
- Goertzel, B., and Pennachin, C. (2007). *Artificial General Intelligence, volume 2* (Springer).
- Goodman, N.D., Tenenbaum, J.B., Feldman, J., and Griffiths, T.L. (2008). A rational analysis of rule-based concept learning. *Cogn. Sci.* 32, 108–154.
- Greaves, H., and Wallace, D. (2006). Justifying conditionalization: conditionalization maximizes expected epistemic utility. *Mind* 115, 607–632.
- Greene, R.L. (1986). Sources of recency effects in free recall. *Psychol. Bull.* 99, 221.
- Griffiths, T.L., Kemp, C., and Tenenbaum, J.B. (2008). Bayesian models of cognition. In *The Cambridge Handbook of Computational Psychology*, R. Sun, ed. (Cambridge University Press), pp. 59–100.
- Griffiths, T.L., and Tenenbaum, J.B. (2005). Structure and strength in causal induction. *Cogn. Psychol.* 51, 334–384.
- Griffiths, T.L., and Tenenbaum, J.B. (2006). Optimal predictions in everyday cognition. *Psychol. Sci.* 17, 767–773.
- Guiral-Contreras, A., Gonzalo-Angulo, J.A., and Rodgers, W. (2007). Information content and recency effect of the audit report in loan rating decisions. *Account. Finance* 47, 285–304.
- Harper, M. (2009). The Replicator Equation as an Inference Dynamic (arXiv), p. 0911.1763.
- Hartono, J. (2004). The recency effect of accounting information. *Gadjah Mada Int. J. Business* 6, 85–116.
- Hirshleifer, D., and Teoh, S.H. (2009). Thought and behavior contagion in capital markets. In *Handbook of Financial Markets: Dynamics and Evolution*, Handbooks in Finance, pp. 1–46.
- Hirshleifer, J. (1977). Economics from a biological viewpoint. *J. L. Econ.* 20, 1–52.
- Hoffman, D.D. (2016). The interface theory of perception. *Curr. Dir. Psychol. Sci.* 25, 157–161.
- Houston, A.I., McNamara, J.M., and Steer, M.D. (2007). Do we expect natural selection to produce rational behaviour? *Philos. Trans. R. Soc. B: Biol. Sci.* 362, 1531–1543.
- Howes, A., Warren, P.A., Farmer, G., El-Deredy, W., and Lewis, R.L. (2016). Why contextual preference reversals maximize expected value. *Psychol. Rev.* 123, 368.
- Kahneman, D. (2011). *Thinking, Fast and Slow* (Macmillan).
- Kameda, T., and Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a nonstationary uncertain environment: an evolutionary simulation and an experiment with human subjects. *Evol. Hum. Behav.* 23, 373–393.
- Kendall, A., and Gal, Y. (2017). What uncertainties do we need in bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems*, pp. 1–12.
- Kenrick, D.T., Griskevicius, V., Sundie, J.M., Li, N.P., Li, Y.J., and Neuberg, S.L. (2009). Deep rationality: the evolutionary economics of decision making. *Social Cogn.* 27, 764.
- Kogan, L., Ross, S.A., Wang, J., and Westerfield, M.M. (2006). The price impact and survival of irrational traders. *J. Finance* 61, 195–229.
- Körding, K.P., and Wolpert, D.M. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* 10, 319–326.
- Lake, B.M., Salakhutdinov, R., and Tenenbaum, J.B. (2015). Human-level concept learning through probabilistic program induction. *Science* 350, 1332–1338.
- Leitgeb, H., and Pettigrew, R. (2010a). An objective justification of bayesianism i: measuring inaccuracy. *Philos. Sci.* 77, 201–235.
- Leitgeb, H., and Pettigrew, R. (2010b). An objective justification of bayesianism ii: the consequences of minimizing inaccuracy. *Philos. Sci.* 77, 236–272.
- Lieder, F., and Griffiths, T.L. (2020). Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behav. Brain Sci.* 43, e1.
- Liu, J.S., and Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *J. Am. Stat. Assoc.* 93, 1032–1044.
- Lo, A.W. (2017). *Adaptive Markets: Financial Evolution at the Speed of Thought* (Princeton University Press).
- Lo, A.W., and Zhang, R. (2018). *Biological Economics* (Edward Elgar Publishing Limited).
- Lughofer, E., and Sayed-Mouchaweh, M. (2015). Adaptive and on-line learning in non-stationary environments. *Evolving Syst.* 6, 75–77.
- Marcus, G. (2009). *Kluge: The Haphazard Evolution of the Human Mind* (Houghton Mifflin Harcourt).
- McDermott, R., Fowler, J.H., and Smirnov, O. (2008). On the evolutionary origin of prospect theory preferences. *J. Polit.* 70, 335–350.
- Mery, F., and Kawecki, T.J. (2002). Experimental evolution of learning ability in fruit flies. *Proc. Natl. Acad. Sci.* 99, 14274–14279.
- Mozer, M.C., Pashler, H., and Homaei, H. (2008). Optimal predictions in everyday cognition: the wisdom of individuals or crowds? *Cogn. Sci.* 32, 1133–1147.
- Nowak, M.A. (2006). Five rules for the evolution of cooperation. *Science* 314, 1560–1563.
- Oaksford, M., and Chater, N. (2001). The probabilistic approach to human reasoning. *Trends Cogn. Sci.* 5, 349–357.
- Obuchowicz, A., and Wawrzyniak, D. (2005). Evolutionary adaptation in non-stationary environments: a case study. In *International Conference on Parallel Processing and Applied Mathematics* (Springer), pp. 439–446.
- Okasha, S. (2013). The evolution of bayesian updating. *Philos. Sci.* 80, 745–757.
- Okasha, S., and Binmore, K. (2012). *Evolution and Rationality: Decisions, Co-operation and Strategic Behaviour* (Cambridge University Press).
- Pástor, L., and Stambaugh, R.F. (2000). Comparing asset pricing models: an investment perspective. *J. Financial Econ.* 56, 335–381.

- Robson, A.J. (1996). A biological basis for expected and non-expected utility. *J. Econ. Theor.* 68, 397–424.
- Rogers, A.R. (1994). Evolution of time preference by natural selection. *Am. Econ. Rev.* 84, 460–481.
- Ross, C., and Wilke, A. (2011). Past and present environments. *J. Evol. Psychol.* 9, 275–278.
- Samuelson, L. (2001). Introduction to the evolution of preferences. *J. Econ. Theor.* 97, 225–230.
- Sanborn, A., and Griffiths, T.L. (2008). Markov chain Monte Carlo with people. In *Advances in Neural Information Processing Systems*, pp. 1265–1272.
- Sanborn, A.N., Griffiths, T.L., and Navarro, D.J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychol. Rev.* 117, 1144.
- Sayed-Mouchaweh, M., and Lughofer, E. (2012). *Learning in Non-stationary Environments: Methods and Applications* (Springer Science & Business Media).
- Schulz, A.W. (2020). The evolutionary foundations of resource-rational analysis. *Behav. Brain Sci.* 43, e23.
- Shalizi, C.R. (2009). Dynamics of bayesian updating with dependent data and misspecified models. *Electron. J. Stat.* 3, 1039–1074.
- Shefrin, H. (2007). Behavioral finance: biases, mean-variance returns, and risk premiums. *CFA Inst. Conf. Proc. Q.* 24, 4–12.
- Simon, H.A. (1955). A behavioral model of rational choice. *Q. J. Econ.* 69, 99–118.
- Sims, C.R. (2016). Rate–distortion theory and human perception. *Cognition* 152, 181–198.
- Sims, C.R., Jacobs, R.A., and Knill, D.C. (2012). An ideal observer analysis of visual working memory. *Psychol. Rev.* 119, 807.
- Smith, A. (2013). *Sequential Monte Carlo Methods in Practice* (Springer Science & Business Media).
- Steyvers, M., Griffiths, T.L., and Dennis, S. (2006). Probabilistic inference in human semantic memory. *Trends Cogn. Sci.* 10, 327–334.
- Steyvers, M., Tenenbaum, J.B., Wagenmakers, E.J., and Blum, B. (2003). Inferring causal networks from observations and interventions. *Cogn. Sci.* 27, 453–489.
- Strens, M. (2000). A bayesian framework for reinforcement learning. *ICML 2000*, 943–950.
- Suchow, J.W., Bourgin, D.D., and Griffiths, T.L. (2017). Evolution in mind: evolutionary dynamics, cognitive processes, and bayesian inference. *Trends Cognitive Sciences* 21, 522–530.
- Suchow, J.W., and Griffiths, T.L. (2016). Deciding to remember: memory maintenance as a Markov decision process. *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2063–2068.
- Tenenbaum, J.B., Griffiths, T.L., and Kemp, C. (2006). Theory-based bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.* 10, 309–318.
- Tenenbaum, J.B., Kemp, C., Griffiths, T.L., and Goodman, N.D. (2011). How to grow a mind: statistics, structure, and abstraction. *science* 331, 1279–1285.
- Turing, A.M. (1950). Computing machinery and intelligence. *Mind* 59, 433–460.
- Vul, E., Goodman, N., Griffiths, T.L., and Tenenbaum, J.B. (2014). One and done? optimal decisions from very few samples. *Cogn. Sci.* 38, 599–637.
- Vul, E., and Pashler, H. (2008). Measuring the crowd within: probabilistic representations within individuals. *Psychol. Sci.* 19, 645–647.
- Waksberg, A.J., Smith, A.B., and Burd, M. (2009). Can irrational behaviour maximise fitness? *Behav. Ecol. Sociobiol.* 63, 461–471.
- Waldman, M. (1994). Systematic errors and the theory of natural selection. *Am. Econ. Rev.* 84, 482–497.
- Wang, H., and Yeung, D.Y. (2020). A survey on bayesian deep learning. *ACM Comput. Surv. (CSUR)* 53, 1–37.
- White, H. (2001). *Asymptotic Theory for Econometricians*, revised edition (Academic press).
- Wilson, A.G., and Izmailov, P. (2020). Bayesian deep learning and a probabilistic perspective of generalization. In *Advances in Neural Information Processing Systems*, pp. 1–12.
- Xu, F., and Tenenbaum, J.B. (2007). Word learning as bayesian inference. *Psychol. Rev.* 114, 245.
- Yang, Ch, Belawat, P., Hafen, E., Jan, L.Y., and Jan, Y.N. (2008). *Drosophila* egg-laying site selection as a system to study simple decision-making processes. *Science* 319, 1679–1683.
- Yuille, A., and Kersten, D. (2006). Vision as bayesian inference: analysis by synthesis? *Trends Cogn. Sci.* 10, 301–308.
- Zhang, H. (2013). Evolutionary justifications for non-bayesian beliefs. *Econ. Lett.* 121, 198–201.
- Zhang, R., Brennan, T.J., and Lo, A.W. (2014). The origin of risk aversion. *Proc. Natl. Acad. Sci.* 111, 17777–17782.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|-------------------------------|----------------------------|---|
| Software and algorithms | | |
| Python version 3.6.9 | Python Software Foundation | www.python.org |
| Colab notebooks | Google | research.google.com/colaboratory |
| Code for plots in the article | GitHub via Zenodo | https://doi.org/10.5281/zenodo.5035558 |
| Other | | |
| Model parameters | None | None |

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to the lead contact, Ruixun Zhang (rxzhang@alum.mit.edu).

Materials availability

This study did not generate any materials.

Data and code availability

- This article does not involve data.
- All original code has been deposited at github.com/raytheman/Bayesian and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study does not include experiments or subjects.

METHOD DETAILS

Related literature

The Bayesian paradigm has been widely used across a broad spectrum of disciplines, including cognitive science, evolutionary biology, computer science, and computational finance. [Griffiths et al. \(2008\)](#) provide a review of the application of Bayesian models in computational models of cognition. Examples include animal learning ([Courville et al., 2006](#)), visual scene perception and concepts ([Yuille and Kersten, 2006](#); [Lake et al., 2015](#)), motor control ([Körding and Wolpert, 2006](#)), semantic memory ([Steyvers et al., 2006](#)), symbolic reasoning ([Oaksford and Chater, 2001](#)), and social cognition ([Baker et al., 2007](#)).

Evolutionary biologists and behavioral ecologists have studied the role of the Bayes' rule in the context of biological evolution ([Greaves and Wallace, 2006](#); [Leitgeb and Pettigrew, 2010a, b](#); [Okasha, 2013](#); [Castellano, 2014](#); [Campbell, 2016](#)). Building on the known mathematical equivalence between Bayesian inference and replicator dynamics in evolutionary biology ([Harper, 2009](#); [Shalizi, 2009](#)), [Czégel et al. \(2018\)](#) argue that, "in order to devise seemingly-engineered complex organisms, evolution, on Earth or anywhere, utilized comparable hierarchical learning mechanisms as we humans do to make sense of the world around us."

Financial economists have also adopted the Bayes' rule for portfolio analysis (see [Avramov and Zhou \(2010\)](#) for a review and further references). [Frydman and Jin \(2018\)](#) argue that the perception of risky payoffs is imperfect, and is estimated through Bayesian inference, a concept known as the efficient coding hypothesis in neuroscience. In a simple market where only one asset generates positive returns in each period, [Blume and Easley \(1992\)](#) show that investors following Bayesian learning will dominate the market. [Zhang](#)

(2013) argues that perfect Bayesian updating is evolutionarily optimal only for a risk-neutral agent, and is not optimal for any other type of agent.

Resource-rational analysis, the approach to human cognition as the optimal use of limited computational resources, has been used to explain seemingly suboptimal behavior (Lieder and Griffiths, 2020). Finite memory arises from the trade-off between the cost of memory storage and the benefits of a long memory (Anderson and Milson, 1989; Anderson and Schooler, 1991; Sims et al., 2012; Howes et al., 2016; Sims, 2016; Suchow and Griffiths, 2016). Others argue that computational limits do not fully explain human cognitive limitations (Davis and Marcus, 2020), and that memory is a very clear case of a suboptimal system (Marcus, 2009). In fact, the evolutionary perspective benefits the resource-rational analysis by providing insights into the origins of the trade-off between accuracy and effort (Schulz, 2020).

Our study of evolutionary dynamics in nonstationary environments provides a new rationale for finite memory beyond constraints on computational resources. A balance between continuous learning and forgetting is necessary to cope with nonstationary environments (Lughofer and Sayed-Mouchaweh, 2015). Learning in nonstationary environments has been studied both in evolutionary biology (Kameda and Nakanishi, 2002) and computer science (Obuchowicz and Wawrzyniak, 2005; Sayed-Mouchaweh and Lughofer, 2012; Ditzler et al., 2015). Our results also provide an explanation for the recency effect documented in many widely different fields, including psychology (Glenberg and Swanson, 1986; Greene, 1986), accounting (Hartono, 2004; Guiral-Contreras et al., 2007), behavioral economics and finance (Shefrin, 2007; Kahneman, 2011), and even legal analysis (Furnham, 1986).

Recently, Suchow et al. (2017) point out a deep connection between evolutionary dynamics and Bayesian inference that can be used to join these two perspectives. They also call for new work on “finding equivalencies between particular evolutionary processes and algorithms for inference drawn from statistics and machine learning; creating process models of other cognitive capacities using the framework of evolutionary dynamics; and exploring the rich array of phenomena studied in modern evolutionary biology and what connection, if any, they may have to the similarly rich array of processes studied in the cognitive sciences.” For example, sequential Monte Carlo methods such as particle filters (Del Moral, 1997; Liu and Chen, 1998; Smith, 2013), partially inspired by evolutionary biology and originally proposed for state estimation, have also been used to model cognitive processes in other settings (Sanborn et al., 2010; Abbott and Griffiths, 2011). (We thank an anonymous reviewer for pointing out this connection to us.)

Our work contributes to this literature, but is different from Suchow et al. (2017) in that we make this connection with a particular focus on the stochastic nature of environmental risks that is applicable to economic behavior. In addition to establishing the equivalence between the population dynamics in evolution and the Bayes’ rule, we also derive a growth-optimal behavior that resembles Bayesian inference under a particular utility function determined by the stochastic nature of the reproductive success risks. By doing so, we provide a natural definition of intelligent behaviors from the evolutionary perspective. This is natural from the angle of evolutionary reproduction, but may provide new insights into why certain behaviors are more persistent than others. Moreover, our framework highlights the role of the environment in determining the specific form of behavior. Specifically, we identify three specific environments that favor the adaptation of finite memory, beyond cognitive resource constraints.

Finally, our article contributes to a growing literature that uses evolutionary principles to understand human economic behavior. See Lo and Zhang (2018) for a review of the literature on the interface between economics and biology. These ideas have proven themselves to be useful in understanding altruism (Alexander, 1974; Becker, 1976; Hirshleifer, 1977; Almenberg and Dreber, 2013), rationality (Houston et al., 2007; McDermott et al., 2008; Kenrick et al., 2009; Waksberg et al., 2009; Ross and Wilke, 2011; Brennan and Lo, 2012; Okasha and Binmore, 2012), the biological origin of utility functions and time preferences (Rogers, 1994; Waldman, 1994; Robson, 1996; Samuelson, 2001; Zhang et al., 2014), and the dynamics of financial markets (Blume and Easley, 1992; Kogan et al., 2006; Hirshleifer and Teoh, 2009; Lo, 2017).

Binary choice model

We make the following assumptions throughout this article.

Assumption 1. s_t has finite moments up to order 2 for all t . (x_a, x_b) and $\log(fx_a + (1 - f)x_b)$ have finite moments up to order 2 for all $f \in [0, 1]$, both marginally and conditioned on any state s_t .

Assumption 2. Conditioned on state s_t , (x_a, x_b) are IID over time and identical for all individuals in a given generation.

Individual behavior. Individuals in a given generation t are indexed by i , and generations are indexed by $t=1, \dots, T$. In all cases, we assume independently and identically distributed randomness over time, or equivalently, over generations. On occasion, we will omit the t subscript, unless we wish to emphasize the temporal ordering of the variables, such as in a recursive relation between two successive generations.

After observing all available information, \mathcal{S} , each individual i chooses action a with some probability $f(\mathcal{S}) \in [0, 1]$ and b with probability $1 - f(\mathcal{S})$, denoted by the Bernoulli variable $I_i^{f(\mathcal{S})}$. As a result, i 's offspring, $x_i^{f(\mathcal{S})}$, is given by:

$$x_i^{f(\mathcal{S})} = I_i^{f(\mathcal{S})}x_a + (1 - I_i^{f(\mathcal{S})})x_b, \quad I_i^{f(\mathcal{S})} = \begin{cases} 1 & \text{with probability } f(\mathcal{S}) \\ 0 & \text{with probability } 1 - f(\mathcal{S}) \end{cases} \quad (\text{Equation 6})$$

We assume that:

Assumption 3. $f(\cdot)$ is a function bounded between 0 and 1 (inclusive). (Strictly speaking, we require $f(\cdot)$ to be a \mathcal{F} -measurable function where (Ω, \mathcal{F}, P) is the same probability measure space on which \mathcal{S} is defined. For simplicity we omit the measure-theoretic definitions in the main body of this article.)

In contrast to the binary choice model of [Brennan and Lo \(2011\)](#), an individual's behavior here is a function $f(\cdot)$ of the observed information, \mathcal{S} , instead of a constant.

Population growth. We use the superscript $f(\cdot)$ to denote the particular type of individual as defined by the decision rule in [Equation 6](#). With this convention in mind, we denote by $n_t^{f(\cdot)}$ the total number of offspring of type $f(\cdot)$ in generation t , which is simply the sum of all the offspring from the type- $f(\cdot)$ individuals of the previous generation:

$$\begin{aligned} n_t^{f(\cdot)} &= \sum_{i=1}^{n_{t-1}^{f(\cdot)}} x_{i,t}^{f(\mathcal{S}_t)} = \left(\sum_{i=1}^{n_{t-1}^{f(\cdot)}} I_i^{f(\mathcal{S}_t)} \right) x_{a,t} + \left(\sum_{i=1}^{n_{t-1}^{f(\cdot)}} (1 - I_i^{f(\mathcal{S}_t)}) \right) x_{b,t} \\ &= n_{t-1}^{f(\cdot)} \left(\left(\frac{\sum_{i=1}^{n_{t-1}^{f(\cdot)}} I_i^{f(\mathcal{S}_t)}}{n_{t-1}^{f(\cdot)}} \right) x_{a,t} + \left(\frac{\sum_{i=1}^{n_{t-1}^{f(\cdot)}} (1 - I_i^{f(\mathcal{S}_t)})}{n_{t-1}^{f(\cdot)}} \right) x_{b,t} \right). \end{aligned}$$

Applying the Law of Large Numbers, we determine the population dynamics between two generations:

$$n_t^{f(\cdot) \text{ a.s.}} \stackrel{\text{a.s.}}{=} n_{t-1}^{f(\cdot)} (f(\mathcal{S}_t)x_{a,t} + (1 - f(\mathcal{S}_t))x_{b,t}) \quad (\text{Equation 7})$$

where the notation $\stackrel{\text{a.s.}}{=}$ in [Equation 7](#) denotes "almost sure" convergence as $n_{t-1}^{f(\cdot)}$ increases without bound.

We can rewrite [Equation 7](#) in terms of the relative frequency of each behavior in the population. Let $I_t^{f(\cdot)}$ be the relative frequency of behavior $f(\cdot)$ in the population in generation t . We have:

$$I_t^{f(\cdot)} \stackrel{\text{a.s.}}{=} \frac{n_{t-1}^{f(\cdot)} (f(\mathcal{S}_t)x_{a,t} + (1 - f(\mathcal{S}_t))x_{b,t})}{\sum_{g(\cdot)} n_{t-1}^{g(\cdot)} (g(\mathcal{S}_t)x_{a,t} + (1 - g(\mathcal{S}_t))x_{b,t})} = \frac{I_{t-1}^{f(\cdot)} \cdot \text{Fitness}_t^{f(\cdot)}}{\sum_{g(\cdot)} I_{t-1}^{g(\cdot)} \cdot \text{Fitness}_t^{g(\cdot)}} \quad (\text{Equation 8})$$

where $\text{Fitness}_t^{f(\cdot)} = (f(\mathcal{S}_t)x_{a,t} + (1 - f(\mathcal{S}_t))x_{b,t})$ is the fitness (random reproductive success) of behavior $f(\cdot)$ in generation t .

Note that [Equation 8](#) has the same mathematical form as the Bayes formula:

$$\mathbb{P}(A_i | \text{data}) = \frac{\mathbb{P}(\text{data} | A_i) \mathbb{P}(A_i)}{\sum_j \mathbb{P}(\text{data} | A_j) \mathbb{P}(A_j)} \quad (\text{Equation 9})$$

if we treat $f_t^{f(\cdot)}$ (the frequency in the current generation) as the posterior $\mathbb{P}(A_t | \text{data})$, $f_{t-1}^{f(\cdot)}$ (the frequency in the previous generation) as the prior $\mathbb{P}(A_t)$, and $\text{Fitness}_t^{f(\cdot)}$ (the fitness in the current generation) as the likelihood $\mathbb{P}(\text{data} | A_t)$.

(Equation 7) and (Equation 8) give the basic relationship between populations in two consecutive generations. We wish to characterize the growth of $n_t^{f(\cdot)}$ with respect to the environment, x_a , x_b , and S_t , and individual behavior, $f(\cdot)$. Different assumptions about the environment lead to different growth dynamics. We consider stationary and nonstationary environments in the next two sections, respectively.

Stationary environments

We consider stationary environments in this section, assuming that:

Assumption 4. The set of environmental states $\{s_t\}$ is stationary and ergodic. In particular, s_t is ergodic if $\{s_t\}$ is stationary with a finite second moment, and its covariance $\text{cov}(s_0, s_t) \rightarrow 0$ as $t \rightarrow \infty$.

Through backward recursion based on (Equation 7), the population size of type- $f(\cdot)$ individuals in generation T is given by

$$n_T^{f(\cdot) \text{ a.s.}} \equiv \prod_{t=1}^T (f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t}) \quad (\text{Equation 10})$$

$$\frac{1}{T} \log n_T^{f(\cdot) \text{ a.s.}} \equiv \frac{1}{T} \sum_{t=1}^T \log (f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t}) \quad (\text{Equation 11})$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}_{(x_a, x_b, S)} [\log (f(S)x_a + (1 - f(S))x_b)] \quad (\text{Equation 12})$$

where “a.s.” in (Equation 12) denotes almost sure convergence as T increases without bound. (This follows from the ergodic theorem (e.g. (White, 2001, Theorem 3.34)). Specifically, Assumptions 1 and 3 together imply that $\log (f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})$ has finite moments up to order 2. Assumptions 2, 3, and 4 together imply that $\log (f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})$ is stationary and ergodic (e.g. (White, 2001, Theorem 3.35)). The expected value is taken over the joint distribution of (x_a, x_b, S) , and we have assumed that $n_0 = 1$ without loss of generality.

Since the behavior, $f(\cdot)$, that maximizes the population size, $n_T^{f(\cdot)}$, is also the $f(\cdot)$ that maximizes $T^{-1} \log n_T^{f(\cdot)}$, (Equation 12) implies that this value converges almost surely to the maximum of the following expected value:

$$\alpha(f(\cdot)) \stackrel{\text{def}}{=} \mathbb{E}_{(x_a, x_b, S)} [\log (f(S)x_a + (1 - f(S))x_b)]. \quad (\text{Equation 13})$$

Expression (Equation 13) is simply the expected value of the log-geometric-average growth rate of the population, given a particular behavior $f(\cdot)$. Note that behavior $f(\cdot)$ is a function of the observed data S .

Next, we consider the case that $f(\cdot)$ is a constant, i.e., the individuals in our model are totally mindless. We then consider the general case that $f(\cdot)$ uses past information S_t .

Constant behavior. If the individuals in our model do not have intelligence, and only have constant behavior f , the following result (Brennan and Lo, 2011) gives the growth-optimal behavior f^* .

Proposition 1 (Stationary environments and constant behavior). Under Assumptions 1–4, the growth-optimal constant behavior f^* that maximizes (Equation 13) is:

$$f^* = \underset{f}{\text{argmax}} \mathbb{E}[\log (fx_a + (1 - f)x_b)]. \quad (\text{Equation 14})$$

In particular, f^* is explicitly given by:

$$f^* = \begin{cases} 1 & \text{if } \mathbb{E}[x_b/x_a] < 1 \\ \text{solution to } (S \cdot 11) & \text{if } \mathbb{E}[x_a/x_b] \geq 1 \text{ and } \mathbb{E}[x_b/x_a] \geq 1 \\ 0 & \text{if } \mathbb{E}[x_a/x_b] < 1, \end{cases} \quad (\text{Equation 15})$$

where f^* is defined implicitly in the second case of (Equation 15) by:

$$\mathbb{E} \left[\frac{x_a - x_b}{f^* x_a + (1 - f^*) x_b} \right] = 0 \quad (\text{Equation 16})$$

and the expected values in (Equation 14)–(Equation 16) are with respect to the joint distribution, $\Phi(x_a, x_b)$.

In terms of the sunny/rainy example, when s_t is stationary, the expected value of p_t is a constant. The growth rate of individuals with behavior f is:

$$\begin{aligned} \alpha(f) &= \mathbb{E}_{(x_a, x_b)} [\log(fx_a + (1 - f)x_b)] \\ &= \mathbb{E}_{p_t} [\mathbb{E}_{(x_a, x_b) | p_t} [\log(fx_a + (1 - f)x_b) | p_t]] \\ &= \mathbb{E}_{p_t} [p_t \log(fc) + (1 - p_t) \log((1 - f)c)] \\ &= \mathbb{E}[p_t] \log(f) + (1 - \mathbb{E}[p_t]) \log(1 - f) + \log(c). \end{aligned}$$

Note that we used the law of total expectation to break down the expected value over (x_a, x_b) into the conditional distribution of $(x_a, x_b | p_t)$ and the expected value over the state p_t . We will use this method again when we consider general intelligent behavior. Maximizing $\alpha(f)$, we get $f^* = \mathbb{E}[p_t]$. (The last term $\log(c)$ is a constant, and does not play a role in determining the optimal f .) The optimal behavior in our sunny/rainy example is to build shelter by the river with the same probability as the expectation of a sunny day. This is called “probability matching” or “bet hedging” in the evolutionary biology literature (Cooper and Kaplan, 1982; Frank and Slatkin, 1990; Frank, 2011; Brennan and Lo, 2011).

Intelligent behavior. In the general case, individual behavior is allowed to be a function of all available information at time t . In other words, $f(\cdot)$ is a function that maps all observed states, \mathcal{S}_t , to a value in $[0, 1]$. The pertinent evolutionary question is what forms of $f(\cdot)$ maximize growth.

We first provide more intuition to understand the population growth rate (Equation 13). By conditioning on the observed states \mathcal{S} (or mathematically, by the law of total expectation), (Equation 13) can be rewritten as

$$\alpha(f(\cdot)) = \mathbb{E}_{\mathcal{S}} [\mathbb{E}_{(x_a, x_b) | \mathcal{S}} [\log(f(\mathcal{S})x_a + (1 - f(\mathcal{S}))x_b) | \mathcal{S}]]. \quad (\text{Equation 17})$$

Given observed states, \mathcal{S} , the behavior $f(\mathcal{S})$ is fixed, and the growth rate (Equation 17) is simply an integration over the distribution of \mathcal{S} . It therefore suffices to find the function $f(\cdot)$ that maximizes the conditional expected value $\mathbb{E}_{(x_a, x_b) | \mathcal{S}} [\log(f(\mathcal{S})x_a + (1 - f(\mathcal{S}))x_b) | \mathcal{S}]$ for every possible value of \mathcal{S} . The next result (included in the main article) summarizes this function.

Proposition 2 (Stationary environments; intelligent behavior). Under Assumptions 1-4, the growth-optimal behavior $f^*(\cdot)$ that maximizes (Equation 13) is:

$$f^*(\mathcal{S}) = \operatorname{argmax}_{f(\mathcal{S})} \mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\log(f(\mathcal{S})x_a + (1 - f(\mathcal{S}))x_b) \middle| \mathcal{S} \right]. \quad (\text{Equation 18})$$

In particular, given observed states \mathcal{S} , $f^*(\cdot)$ is explicitly given by:

$$f^*(\mathcal{S}) = \begin{cases} 1 & \text{if } \mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\frac{x_b}{x_a} \right] < 1 \\ \text{solution to } (\mathcal{S} \cdot 15) & \text{if } \mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\frac{x_a}{x_b} \right] \geq 1 \text{ and } \mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\frac{x_b}{x_a} \right] \geq 1 \\ 0 & \text{if } \mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\frac{x_a}{x_b} \right] < 1, \end{cases} \quad (\text{Equation 19})$$

where $f^*(\mathcal{S})$ is defined implicitly in the second case of (Equation 19) by:

$$\mathbb{E}_{(x_a, x_b) | \mathcal{S}} \left[\frac{x_a - x_b}{f^*(\mathcal{S})x_a + (1 - f^*(\mathcal{S}))x_b} \right] = 0 \quad (\text{Equation 20})$$

and the expected values in (Equation 18)–(Equation 20) are with respect to the joint distribution, $\Phi(x_a, x_b | \mathcal{S})$.

Proposition 2 shows that in the case of systematic environmental risks, the growth-optimal behavior maximizes the expected log utility, where the expected value is taken over the posterior distribution of the environment conditioned on all available information \mathcal{S} .

What is intelligence? Proposition 2 characterizes the optimal behavior $f^*(\cdot)$. Here we provide more intuition behind the relationship between the optimal behavior, $f^*(\cdot)$, and the environment, highlighting the type of behavior that is considered intelligent.

Let $r \stackrel{\text{def}}{=} x_a/x_b - 1$ be the relative reproductive success. We can rewrite the population growth rate $\alpha(f(\cdot))$ in (Equation 13) as

$$\alpha(f(\cdot)) = \mathbb{E}[\log(f(S)r + 1)] + \mathbb{E}[\log(x_b)], \quad (\text{Equation 21})$$

provided that the expected value for both terms in (Equation 21) exists. (This is a more restrictive condition than Assumption 1. We do not make this distributional assumption in our article except here.) The last term, $\mathbb{E}[\log(x_b)]$, does not depend on $f(\cdot)$. Therefore, maximizing $\alpha(f(\cdot))$ over $f(\cdot)$ is equivalent to maximizing the first term in (Equation 21). We further approximate (Equation 21) by a Taylor expansion around $r=0$: (When the fitnesses of the two actions x_a and x_b are similar, r varies around 0. This is another assumption that is specific to this result, and not used in the remainder of this article).

$$\begin{aligned} \alpha(f(\cdot)) &\approx \mathbb{E}[f(S)r] - \frac{1}{2}\mathbb{E}[(f(S)r)^2] + \text{constant} \\ &= \mathbb{E}[f(S)]\mathbb{E}[r] + \text{Cov}(f(S), r) - \frac{1}{2}\mathbb{E}[(f(S)r)^2] + \text{constant} \end{aligned} \quad (\text{Equation 22})$$

up to the second-order Taylor approximation. This expansion is the basis for the three factors pertaining to intelligence in the main article.

Markov environments. Intelligent behavior only emerges in a time-dependent environment. Here we consider an example of a Markov environment to explicitly illustrate the relationship between optimal behavior and the amount of memory in the environment.

Suppose the environmental states $\{s_t\}$ are k -th order Markov variables. In other words, in period t , given the observed states in the past k periods: $\{s_{t-1}, \dots, s_{t-k}\}$, s_t is independent of states further away in the past: $\{s_{t-k-1}, s_{t-k-2}, \dots\}$. In such environments, the optimal behavior should only use information from the past k periods: $f^*(s_{t-1}, \dots, s_{t-k})$. In fact, using information older than k periods may hurt growth.

For example, suppose s_t follows an AR process of order k (AR(k)):

$$s_t = w_0 + w_1 s_{t-1} + \dots + w_k s_{t-k} + \varepsilon_t$$

where w_0, \dots, w_k are weights for different lags, and ε_t is white noise. In this case, the optimal behavior will be a function of the weighted observed states in the past k periods:

$$f^* \left(w_0 + \frac{1}{k} \sum_{i=1}^k w_i s_{t-i} \right).$$

In general, if there are no constraints to the functional forms of $f(\cdot)$, behaviors will emerge that learn the underlying process of the environment. However, if there are limitations to intelligence, (See, for example, the notion of "satisficing" and bounded rationality (Simon, 1955), and the notion of ecological rationality (Gigerenzer and Selten, 2002).) $f^*(\cdot)$ may not be able to learn the precise dynamics of the underlying environmental process. For example, if s_t is a complicated nonlinear stochastic process, and intelligence is constrained to linear functions of past information, $f^*(\cdot)$ can never perfectly learn the nonlinear relationship between s_t and past states.

Nonstationary environments

In nonstationarity environments, the Law of Large Numbers does not hold, and additional regularity conditions are needed to describe the evolutionary dynamics in such environments. We make the following assumption:

Assumption 5. The environmental states $\{s_t\}$ is α -mixing.

In general, $\{s_t\}$ can be nonstationary. (White, 2001, Definition 3.42) provides the formal measure-theoretic definition for mixing stochastic processes. In this section, we use a simple necessary condition for α -mixing to illustrate the intuition behind the concept. α -mixing stochastic processes are always asymptotically uncorrelated: there exist constants $\{\rho_\tau\}$ such that $0 \leq \rho_\tau \leq 1$, $\sum_{\tau=0}^{\infty} \rho_\tau < \infty$, and $\text{corr}(s_t, s_{t+\tau}) \leq \rho_\tau$ for all $\tau > 0$. In other words, Assumption 5 asserts that the environment forgets about the past asymptotically. We will see that the average log population size will still converge to some constant in these environments.

We return to the fundamental relationship between two generations, (Equation 7), to derive the population dynamics of nonstationary environments. Through backward recursion based on (Equation 7), the population size of type- $f(\cdot)$ individuals in generation T is given by

$$n_T^{f(\cdot) a.s.} \equiv \prod_{t=1}^T (f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t}) \quad (\text{Equation 23})$$

$$\frac{1}{T} \log n_T^{f(\cdot) a.s.} \equiv \frac{1}{T} \sum_{t=1}^T \log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t}) \quad (\text{Equation 24})$$

$$\xrightarrow{a.s.} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{(x_{a,t}, x_{b,t}, S_t)} [\log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})] \quad (\text{Equation 25})$$

where “ $a.s.$ ” in (Equation 25) denotes almost sure convergence as T increases without bound. (This follows from (White, 2001, Theorem 3.47). Specifically, Assumptions 1 and 3 together imply that $\log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})$ has finite moments up to order 2. Assumptions 2, 3, and 5 together imply that $\log(f(S_t)x_{a,t} + (1 - f(S_t))x_{b,t})$ is α -mixing (e.g., see (White, 2001, Theorem 3.49)). The expected value is taken over the joint distribution of (x_a, x_b, S_t) , and we have assumed that $n_0 = 1$ without loss of generality.

Constant behavior. When the optimal behavior is “mindless,” we can explicitly characterize it mathematically.

Proposition 3 (Nonstationary environments; constant behavior). Under Assumptions 1-3 and 5, the growth-optimal constant behavior f^* that maximizes (Equation 25) is:

$$f^* = \operatorname{argmax}_f \overline{\mathbb{E}[\log(fx_{a,t} + (1 - f)x_{b,t})]_{t=1}^{\infty}} \quad (\text{Equation 26})$$

In particular, f^* is explicitly given by:

$$f^* = \begin{cases} 1 & \text{if } \overline{\mathbb{E}[x_{b,t}/x_{a,t}]_{t=1}^{\infty}} < 1 \\ \text{solution to (S.23)} & \text{if } \overline{\mathbb{E}[x_{a,t}/x_{b,t}]_{t=1}^{\infty}} \geq 1 \text{ and } \overline{\mathbb{E}[x_{b,t}/x_{a,t}]_{t=1}^{\infty}} \geq 1 \\ 0 & \text{if } \overline{\mathbb{E}[x_{a,t}/x_{b,t}]_{t=1}^{\infty}} < 1, \end{cases} \quad (\text{Equation 27})$$

where f^* is defined implicitly in the second case of (Equation 27) by:

$$\overline{\mathbb{E}\left[\frac{x_{a,t} - x_{b,t}}{f^*x_{a,t} + (1 - f^*)x_{b,t}}\right]_{t=1}^{\infty}} = 0 \quad (\text{Equation 28})$$

and the expected values in (Equation 26)-(Equation 28) are with respect to the joint distribution $\Phi(x_{a,t}, x_{b,t})$.

Compared to Proposition 1 for stationary environments, the average expected values over time are needed here because the environment is nonstationary. In the sunny/rainy example, the probability of a sunny day, p_t , no longer has a constant expected value. The growth rate of individuals with behavior f is therefore:

$$\begin{aligned} \mu(f) &= \overline{\mathbb{E}_{(x_{a,t}, x_{b,t}, p_t)} [\log(fx_a + (1 - f)x_b)]_{t=1}^{\infty}} \\ &= \overline{\mathbb{E}_{p_t} [\mathbb{E}_{(x_a, x_b | p_t)} [\log(fx_a + (1 - f)x_b) | p_t]]_{t=1}^{\infty}} \\ &= \overline{\mathbb{E}_{p_t} [p_t \log(fc) + (1 - p_t) \log((1 - f)c)]_{t=1}^{\infty}} \\ &= \overline{\mathbb{E}[p_t]_{t=1}^{\infty}} \log(f) + \left(1 - \overline{\mathbb{E}[p_t]_{t=1}^{\infty}}\right) \log(1 - f) + \log(c). \end{aligned}$$

We get $f^* = \overline{\mathbb{E}[p_t]_{t=1}^{\infty}}$ by maximizing $\mu(f)$. Compared to the optimal behavior in stationary environments, f^* is the average expected value of a sunny day over time, which nevertheless still corresponds to probability matching in the average sense.

Intelligent behavior. For behaviors that use past information, there is no explicit solution for the optimal behavior $f^*(\cdot)$ for the general nonstationary environment. Although each term in the growth rate for the nonstationary environment, (Equation 25), is mathematically equivalent to the growth rate for the stationary environment, (Equation 13), the optimal behavior $f^*(\cdot)$ that maximizes (Equation 25) is fundamentally different from $f^*(\cdot)$ in stationary environments (Proposition 2). This is because the state in period t is conditioned on information in period t ($s_t | \mathcal{S}_t$, or equivalently, $s_t | s_{t-1}, s_{t-2}, \dots$). As a result, the conditional distribution $(x_{a,t}, x_{b,t} | \mathcal{S}_t)$ depends on t , but individual behavior $f^*(\cdot)$ is not time-aware (i.e., $f^*(\cdot)$ is not a function of time t). Therefore, it is impossible for $f^*(\cdot)$ to be optimal with respect to the distribution of $(x_{a,t}, x_{b,t} | \mathcal{S}_t)$ for any period t .

Proposition 4 (Nonstationary environments; intelligent behavior). Under Assumptions 1-3 and 5, the growth-optimal behavior $f^*(\cdot)$ that maximizes (Equation 25) is:

$$f^*(\mathcal{S}) = \operatorname{argmax}_{f(\mathcal{S})} \mu(f(\cdot)) \\ = \operatorname{argmax}_{f(\mathcal{S})} \mathbb{E}_{(x_{a,t}, x_{b,t}, s_t)} [\log(f(\mathcal{S}_t)x_{a,t} + (1 - f(\mathcal{S}_t))x_{b,t})]_{t=1}^{\infty}.$$

A simple example of a nonstationary environment. One simple example of a nonstationary environment is a time-varying deterministic state series. Following our sunny/rainy example, suppose $\{s_t\}$ is a binary series that encodes the season at time t , and the season changes every 10 periods. The probability of a sunny day, p_t , is 0.8 in one season, and 0.2 in another season. In other words, $\{p_t\}$ is a deterministic seasonal time series that changes its value every 10 periods:

$$\{p_1 = 0.8, \dots, p_{10} = 0.8, p_{11} = 0.2, \dots, p_{20} = 0.2, p_{21} = 0.8, \dots\}.$$

This environment is nonstationary because the distribution of (x_a, x_b) is not the same for all time periods t . The optimal constant behavior, that is, for “mindless” individuals, according to Proposition 3, is $f^* = 0.5$, and the optimal growth rate $\mu(f^*) = \log(0.5) + \log(c)$.

On the other hand, individuals with memory can use the last 10 observations of s_t to determine the current probability of “sunny”. The optimal behavior is then to match this probability in each period: $f^*(\cdot) = p_t$. This leads to the optimal growth rate $\mu(f^*(\cdot)) = \log(0.8) + \log(c)$, which is higher than the optimal growth rate for constant behavior.

This simple example shows that intelligent behavior can emerge in nonstationary environments. A more general example with stochastic states is given in the main article.

Environments with cost

We assume that all available information in period t , \mathcal{S}_t , contains a large number of data points. There is a cost for each individual to obtain and use this data, $C(m)$, where m is the number of data points used for inferential purposes. Instead of making decisions based on all available information, individuals make decisions based on this limited number of data points, and their behavior is denoted by $f(\mathcal{S}(m))$. When the reproductive risk is systematic, the number of individual i 's offspring $x_i^{f(\mathcal{S}(m))}$ is given by:

$$x_i^{f(\mathcal{S}(m))} = \frac{f_i^{f(\mathcal{S}(m))} x_a + (1 - f_i^{f(\mathcal{S}(m))}) x_b}{C(m)}, \quad f_i^{f(\mathcal{S}(m))} = \begin{cases} 1 & \text{with prob. } f(\mathcal{S}(m)) \\ 0 & \text{with prob. } 1 - f(\mathcal{S}(m)) \end{cases} \quad (\text{Equation 29})$$

where $C(m) > 0$ is an increasing function of m . (Equation 29) adjusts the individual reproductive success by a factor of $C(m)^{-1}$ relative to (Equation 6) to reflect this cost.

Through a similar transformation of the population dynamics described in (Equation 6)-(Equation 13), it is not hard to see that the population growth rate for individuals using limited samples of data points in stationary environments is:

$$\alpha(f(\mathcal{S}(m))) \stackrel{\text{def}}{=} \mathbb{E}_{(x_a, x_b, \mathcal{S})} [\log(f(\mathcal{S}(m))x_a + (1 - f(\mathcal{S}(m)))x_b)] - \log C(m). \quad (\text{Equation 30})$$

Note that in (Equation 30), the expected value is over the distribution of (x_a, x_b, \mathcal{S}) , that is, all available information \mathcal{S} , but individual behavior is restricted to only limited samples, $\mathcal{S}(m)$.

In this setting, an evolutionarily optimal behavior is achieved by maximizing the growth rate $\alpha(f(S(m)))$ over both the number of samples, m , and the specific functional form, $f(\cdot)$. As m increases, the first term in (Equation 30) increases, while the second term in (Equation 30) decreases due to the increase in the cognitive cost. The optimal number of samples, m^* , depends on the relative magnitude of the marginal change of an additional sample.

Without sampling. The sunny/rainy example in the main article takes place in a stationary environment whose states are independent over time. Therefore, the optimal behavior with no additional information is a constant $f^* = 0.5$, given by Proposition 1. The optimal growth rate is then $\alpha(f^*) = \log(0.5 \cdot c)$. The optimal behavior is to build a shelter by the river or on the plateau with equal probability.

No-cost and infinite sampling. We briefly consider the case of no-cost sampling. An infinite number of samples will give an individual the correct state of the environment with probability 1, by the Law of Large Numbers. An individual's behavior will now be dependent on all the data it receives, and is therefore a function $f(D)$. Because individuals will have perfect information about the two states, the optimal behavior $f^*(D)$ is to match the probability of a sunny day in the current state:

$$f^*(D) = \begin{cases} p, & \text{if the probability of a sunny day is } p, \\ 1 - p, & \text{if the probability of a sunny day is } 1 - p. \end{cases}$$

Therefore, the growth rate of individuals with behavior $f^*(D)$ is:

$$\begin{aligned} \alpha(f^*(D)) &= \mathbb{E}_D[\mathbb{E}_{(x_a, x_b|D)}[\log(f(D)x_a + (1 - f(D))x_b)]] \\ &= 0.5 \cdot [p \log(pc) + (1 - p) \log((1 - p)c)] \\ &\quad + 0.5 \cdot [(1 - p) \log((1 - p)c) + p \log(pc)] \\ &= p \log(pc) + (1 - p) \log((1 - p)c). \end{aligned}$$

Compared with the optimal solution in the case of no information, perfect information allows individuals to follow a probability matching strategy at a granular level, in each state of the environment. This greatly improves the growth rate of the population.

Limited sampling with cost. Using the cost given in the main article, we solve for the optimal behavior and growth rate in two steps. First, we fix the number of samples m , and solve for the optimal behavior $f^*(D(m))$. We then maximize the growth rate by optimizing over m .

Given m samples, the probability that the majority indicates the correct state is:

$$\beta = 1 - CDF(m/2; m, q)$$

where $CDF(m/2; m, p)$ is the binomial cumulative distribution function. The growth rate for a behavior using m samples can therefore be written as:

$$\begin{aligned} \alpha(f(D(m))) &= \mathbb{E}_D[\mathbb{E}_{(x_a, x_b|D)}[\log(f(D(m))x_a + (1 - f(D(m)))x_b)]] - \log(C(m)) \\ &= p(\beta \log(fc) + (1 - \beta) \log((1 - f)c)) \\ &\quad + (1 - p)(\beta \log((1 - f)c) + (1 - \beta) \log(fc)) - \log(C(m)) \\ &= (1 - p - \beta + 2p\beta) \log(fc) + (p + \beta - 2p\beta) \log((1 - f)c) - \log(C(m)). \end{aligned} \tag{Equation 31}$$

Given m samples, we can derive the optimal $f^*(D(m))$ by maximizing $\alpha(f(D(m)))$:

$$f^*(D(m)) = \begin{cases} \beta p + (1 - \beta)(1 - p), & \text{if } D(m) \text{ implies summer,} \\ \beta(1 - p) + (1 - \beta)p, & \text{if } D(m) \text{ implies winter.} \end{cases} \tag{Equation 32}$$

This is the weighted average of the state probabilities p and $1 - p$, where the probability that the majority indicates the correct state, β , is used as the weight. (Equation 32) can be treated as an approximation to probability matching, where the approximation is determined by the confidence of the data.

Substituting (Equation 32) into (Equation 31), we can determine the optimal growth rate given m samples, $\alpha(f^*(D(m)))$:

$$\begin{aligned} \alpha(f^*(D(m))) &= (1 - p - \beta + 2p\beta) \log((\beta p + (1 - \beta)(1 - p))c) \\ &\quad + (p + \beta - 2p\beta) \log((\beta(1 - p) + (1 - \beta)p)c) - \log(C(m)). \end{aligned} \tag{Equation 33}$$

The optimal number of samples, m^* , is obtained by maximizing (Equation 33) over m . Here β will depend on the cumulative distribution function of the binomial distribution, and therefore m^* does not have an explicit solution.

The optimal number of samples versus cost. We demonstrate the trade-off between the growth rate of the population and the cost of additional samples. Figure S1A in Supplemental Information shows the growth rate for different numbers of samples when $q=0.8$ and $\eta=0.02$. The cost of sampling in this example is relatively small ($\eta=0.02$ implies that collecting 50 samples only cuts the fitness by half). As the number of samples increases, the growth rate increases initially, and achieves a maximum at 3.

Figure S1B in Supplemental Information shows the optimal number of samples as we change the magnitude of the cost. As the cost increases, the optimal number of samples decreases, and eventually goes to zero, indicating that it is not desirable to collect any sample.

Figure S1C in Supplemental Information shows that the optimal sample size is low when the amount of information in each sample is either low or high, given a specific cost, $\eta=0.02$. When the information in each sample is low (that is, when q is close to 0.5), there is little information to gain from each sample. As a result, it is not worth paying the cost for any sample. On the other hand, when the information in each sample is very high (that is, when q is close to 1), one sample is sufficient to infer the state reliably. As a result, it is not worth paying the cost beyond just one sample. More generally, Figure 2B in the main article illustrates the relationship between the optimal sample size and the information carried in each sample.

Idiosyncratic reproductive risks

Throughout this article, we have assumed that the number of offspring from actions a and b are given by the same two realizations of the random variables x_a and x_b , respectively, for all individuals. In other words, the risk in reproductive success is systematic across the population. In this section, we show that if the uncertainty in reproduction is instead idiosyncratic to each individual, the growth-optimal behavior always coincides with the individually optimal behavior with a linear utility. This distinction points to the central role that aggregate uncertainty plays in shaping the evolution of behavior and preferences (Brennan and Lo, 2011).

As before, we denote by $x_i^{f(S)}$ the random number of offspring produced by individual i :

$$x_i^{f(S)} = I_i^{f(S)} x_{a,i} + \left(1 - I_i^{f(S)}\right) x_{b,i}, \quad I_i^{f(S)} = \begin{cases} 1 & \text{with probability } f(S) \\ 0 & \text{with probability } 1 - f(S) \end{cases}.$$

We replace Assumption 2 with the following:

Assumption 6. Conditioned on state s_t , $(x_{a,i,t}, x_{b,i,t})$ are independently and identically distributed over time t and across individuals i in a given generation.

In contrast to the case with systematic risk as described in (Equation 6), $(x_{a,i,t}, x_{b,i,t})$ are now assumed to be independently and identically distributed across individuals as well as across time. This is a seemingly small change, but it has dramatic consequences for the evolutionary dynamics of the population and its growth-optimal behavior. In this case, the randomness in the number of offspring is strictly idiosyncratic in the sense that the correlation between the number of offspring for two individuals i and j is 0, even if both individuals choose the same course of action. Recall from (Equation 6) that in the systematic case, if two individuals choose the same action a , both will generate the same number of random offspring x_a , i.e., their reproductive success is perfectly correlated. The idiosyncratic risk implies that even if all individuals in a given population choose the same action, there will still be considerable cross-sectional variability in the number of offspring produced. This more diversified outcome has a very different set of implications for the growth-optimal behavior.

In particular, the total number of offspring $n_t^{f(\cdot)}$ across all type- $f(\cdot)$ individuals in generation t is:

$$n_t^{f(\cdot)} = \sum_{i=1}^{n_{t-1}^{f(\cdot)}} x_{i,t}^{f(S_t)} = \sum_{i=1}^{n_{t-1}^{f(\cdot)}} I_i^{f(S_t)} x_{a,i,t} + \sum_{i=1}^{n_{t-1}^{f(\cdot)}} \left(1 - I_i^{f(S_t)}\right) x_{b,i,t}$$

$$= n_{t-1}^{f(\cdot)} \left(\frac{\sum_{i=1}^{n_{t-1}^{f(\cdot)}} f_i(\mathcal{S}_t) x_{a,i,t}}{n_{t-1}^{f(\cdot)}} + \frac{\sum_{i=1}^{n_{t-1}^{f(\cdot)}} (1 - f_i(\mathcal{S}_t)) x_{b,i,t}}{n_{t-1}^{f(\cdot)}} \right).$$

Applying the Law of Large Numbers, we determine the population dynamics between two generations:

$$n_t^{f(\cdot) \text{ a.s.}} \doteq n_{t-1}^{f(\cdot)} (f(\mathcal{S}_t) \mathbb{E}[x_a | \mathcal{S}_t] + (1 - f(\mathcal{S}_t)) \mathbb{E}[x_b | \mathcal{S}_t]) \quad (\text{Equation 34})$$

where “ \doteq ” in (Equation 34) denotes almost sure convergence as $n_{t-1}^{f(\cdot)}$ increases without bound. The key difference between (Equation 7) and (Equation 34) is that here both the individual’s choice and the number of offspring are idiosyncratic, and therefore both are subject to the Law of Large Numbers.

Through backwards recursion, the population size of type- $f(\cdot)$ individuals in generation T is calculated to be

$$n_T^{f(\cdot) \text{ a.s.}} \doteq \prod_{t=1}^T (f(\mathcal{S}_t) \mathbb{E}[x_a | \mathcal{S}_t] + (1 - f(\mathcal{S}_t)) \mathbb{E}[x_b | \mathcal{S}_t]) \quad (\text{Equation 35})$$

$$\frac{1}{T} \log n_T^{f(\cdot) \text{ a.s.}} \doteq \frac{1}{T} \sum_{t=1}^T \log (f(\mathcal{S}_t) \mathbb{E}[x_a | \mathcal{S}_t] + (1 - f(\mathcal{S}_t)) \mathbb{E}[x_b | \mathcal{S}_t]) \quad (\text{Equation 36})$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}_{\mathcal{S}} [\log (f(\mathcal{S}) \mathbb{E}[x_a | \mathcal{S}] + (1 - f(\mathcal{S})) \mathbb{E}[x_b | \mathcal{S}])] \quad (\text{Equation 37})$$

where (Equation 37) differs from (Equation 12) in that x_a and x_b only matter through their mean conditional on \mathcal{S} , because the idiosyncratic nature of the cross-section of individuals has eliminated randomness in the population. This simple expression attains its optimum at the extremes of $f(\cdot)$. Consequently, we have:

Proposition 5 (Idiosyncratic risks). Under idiosyncratic environmental risks (Assumptions 1, 3, 4, and 6), the growth-optimal behavior $f^*(\cdot)$ is given by:

$$f^*(\mathcal{S}) = \underset{f(\cdot)}{\operatorname{argmax}} f(\mathcal{S}) \mathbb{E}[x_a | \mathcal{S}] + (1 - f(\mathcal{S})) \mathbb{E}[x_b | \mathcal{S}].$$

In particular, $f^*(\cdot)$ is explicitly given by:

$$f^*(\mathcal{S}) = \begin{cases} 1 & \text{if } \mathbb{E}[x_a | \mathcal{S}] > \mathbb{E}[x_b | \mathcal{S}] \\ 0 & \text{if } \mathbb{E}[x_a | \mathcal{S}] \leq \mathbb{E}[x_b | \mathcal{S}], \end{cases}$$

where the expected values is taken over the joint distribution $\Phi(X_a, X_b | \mathcal{S})$.

Proposition 5 shows that under idiosyncratic environmental risks, the growth-optimal behavior maximizes the expected linear utility, when the expected value is taken over the posterior distribution of the environmental states conditional on all available information \mathcal{S} .

Testable implications

To test our model in an experimental setting, one could create an artificial environment in which *Drosophila* (fruit fly) individuals would face a choice between two places to lay their eggs (media A and B) as a biological binary choice model. Different media would be associated with fruits with different odors, like orange and pineapple, as an olfactory signal to *Drosophila* individuals. Competing fitness functions would then be implemented as different rules for harvesting *Drosophila* eggs from the two media. (We thank Terence C. Burnham for suggesting this design). In principle, one could create any possible environmental distribution through different harvesting rules.

For instance, to test our sunny/rainy example, let one of the media represent the river, and the other the plateau. One could harvest 300 eggs from the river when the environment is “sunny,” and 300 eggs from the plateau otherwise. One could measure the percentage of eggs laid by *Drosophila* on each medium over many generations. By changing the degree of nonstationarity in the environment, this method could be used to infer the degree of memory that *Drosophila* individuals use to decide where to lay eggs (behavior $f(\cdot)$ in our model).

This setting provides a controlled experimental framework to study the evolution of memory in environments with different levels of nonstationarity. More generally, one could design similar experiments to empirically study the emergence of Bayesian inference and any form of intelligence.

QUANTIFICATION AND STATISTICAL ANALYSIS

This study does not include statistical analysis of empirical data.