scientific reports

OPEN

The evolution of discrimination under finite memory constraints

Andrew W. Lo[1,2,3,4], Ruixun Zhang[5,6,7,8,9] & Chaoyi Zhao[1,2✉]

We develop an evolutionary model for individual discriminatory behavior that emerges naturally in a mixed population as an adaptive strategy. Our findings show that, when individuals have finite memory and face uncertain environments, they may rely on prior biases and observable group traits to make decisions, changing their discriminatory practices. We also demonstrate that a finite memory is a consequence of natural selection because it leads to higher fitness in dynamic environments with mutations. This adaptability allows individuals with finite memory to better respond to environmental variability, offering a potential evolutionary advantage. Our study suggests that memory constraints and environmental changes are critical factors in sustaining biased behavior, suggesting insights into the persistence of discrimination in real-world settings and possible mitigation strategies across fields, including education, policymaking, and artificial intelligence.

Keywords Discrimination, Finite memory, Evolution, Mutation, Education, Policymaking

One of today's most hotly debated issues is a failure of collective intelligence—discrimination. Discrimination manifests itself in many different forms, including but not limited to racial, gender, and ethnic biases. As such, discrimination has widespread social, economic, and psychological implications, and remains a significant problem in the world's societies. The persistence of this form of collective ignorance across different cultures and environments highlights the profound impact that biased behavior can have on both individuals and groups, by shaping differential access to resources, opportunities, and social standing.

Discrimination has historically been examined through frameworks such as internal bias, prejudice, and ignorance[1]. Traditional economic theories often attribute discriminatory behavior to irrational beliefs[2–4] or incomplete information[5,6], where decision-makers rely on stereotypes or unfounded assumptions to guide their choices about individuals or groups. Other explanations in the economics-based literature include the emergence of biases and stereotyping via motivated reasoning[7–9] and the strategic benefits of distorted beliefs[10–15].

Discrimination, however, is not simply an individual act shaped by personal biases. It is also a collective phenomenon that emerges as a group-level strategy in response to environmental pressures. When the environment rewards non-discrimination, those who adapt to it will succeed. However, if the environment rewards discrimination, a portion of the population will continue to exhibit this behavior. In this context, discrimination is not merely a by-product of misinformed beliefs, but a practice that may evolve as an adaptive behavior that maximizes the survival and success of a group. For instance, when resources are limited and fitness—defined here as the capacity to survive and reproduce—relies on interactions with other groups, individuals may use observable traits such as race, gender, or ethnicity to optimize their decision-making process.

Both cognitive biases and evolutionary pressures may shape discriminatory behavior. On the one hand, stereotypes, which are shaped by prior beliefs and incomplete information, may guide decisions in uncertain environments. On the other hand, evolutionary pressures may encourage discrimination as a strategic behavior for long-term success by the discriminator, especially in resource-scarce conditions.

From a biological standpoint, one plausible hypothesis is that stereotypes may emerge as an unintended consequence of finite memory. When individuals cannot retain all relevant information about their interactions with others, they may resort to simplified cues, such as group traits, to make decisions under uncertainty. This reliance on incomplete information can foster the development of biases and stereotypes as imperfect but functional adaptations to cognitive constraints. Alternatively, such biases may arise not only from a lack of information, but also from the need to process overwhelming amounts of it—where heuristics serve to reduce

[1]MIT, Sloan School of Management, Cambridge, MA 02142, USA. [2]MIT, Laboratory for Financial Engineering, Cambridge, MA 02142, USA. [3]MIT, Computer Science and Artificial Intelligence Laboratory, Cambridge, MA 02139, USA. [4]Santa Fe Institute, Santa Fe, NM 87501, USA. [5]Peking University, School of Mathematical Sciences, Beijing 100871, China. [6]Peking University, Center for Statistical Science, Beijing 100871, China. [7]Peking University, National Engineering Laboratory for Big Data Analysis and Applications, Beijing 100871, China. [8]Peking University, Laboratory for Mathematical Economics and Quantitative Finance, Beijing 100871, China. [9]University of Oxford, Department of Statistics, Oxford OX1 3LB, UK. ✉email: cy_zhao@mit.edu

complexity rather than compensate for scarcity. In both cases, cognitive limitations—whether due to memory scarcity or informational overload—can play a central role in shaping discriminatory behavior over time.

This article models discrimination as a behavior influenced by evolutionary pressures in environments where individuals face uncertainty and possess finite memory. A finite memory limits the ability to recall past information, and in doing so, it shapes discriminatory behavior. Unlike scenarios where individuals possess an infinite memory, where biases might not persist over time, a finite memory leads individuals to overemphasize recent events, changing their discrimination behavior. This adaptation, while not always optimal, is still able to maximize long-term fitness in environments with uncertainty and variability.

To formalize this idea, we develop a mathematical framework built upon a binary choice model[16–18], where individuals from two distinct groups—referred to here as "Andorians" and "Tellarians"—must decide whether to discriminate against members of the other group. The decisions of individuals in one group are shaped by histories of adverse events and prior beliefs about future risks in the other group. The fitness, or reproductive success, of an individual depends on these decisions. The Andorians, representing the majority group, must decide whether to engage with or avoid the Tellarians, whose probability of adverse events is initially unknown to the Andorians and must be inferred over time. Here, we follow the literature, deliberately using fictitious species borrowed from science fiction to reduce the tension that accompanies a discussion of these highly emotionally charged issues, and also to illustrate the generality of our analysis. In particular, our framework can be applied to any marginalized group[18,19].

In our model, we compare the outcomes of infinite and finite memory in shaping discriminatory behavior. With an infinite memory, individuals can retain a comprehensive record of past interactions, which implies a more informed decision-making process and less reliance on biased shortcuts. This infinite-memory case is regarded as a theoretical benchmark and is compared with the more realistic finite-memory case. We find that a finite memory may lead to different patterns of discrimination, as individuals tend to overemphasize recent experiences. Despite this difference in capacity, however, we prove that a finite memory can be the optimal strategy in dynamic environments where mutations can occur in individual behavior. This is because the flexibility provided by a finite memory allows individuals to adapt more rapidly, making it a more effective choice for long-term survival and fitness. In other words, a finite memory, while seemingly inefficient in stable environments, is crucial for adapting to the variability introduced by mutations.

These findings present a challenge for education and policymaking by highlighting the role of memory and environmental uncertainty in fostering bias and discrimination. The key issue is how to balance teaching about the dangers of discrimination, while acknowledging the natural cognitive processes that encourage it, and emphasizing the need for strategies that enhance adaptability and critical thinking rather than focusing solely on increasing information and exposure to diverse viewpoints.

The contributions of this article are threefold. First, we model the decision to discriminate as a function of both memory and prior beliefs, revealing patterns in the emergent discriminatory strategies. Second, we demonstrate how finite memory, in contrast to infinite memory, alters the characteristics of discrimination, potentially leading to different discriminatory behavior. Third, we show that in environments with mutations in individual behavior, finite memory can be more advantageous than infinite memory by enabling more flexible responses to changing conditions. We also discuss the real world implications of our theoretical findings in areas such as education, policymaking, and artificial intelligence. Proofs of all theoretical results are provided in the Supplementary Material.

The framework

As an illustration of the basic intuition behind our evolutionary framework, we first present a toy example of discrimination before introducing the formal model.

A toy example

Consider a population of individuals who must decide whether to discriminate or not when interacting with others[16]. The environment is stochastic. Seventy percent of the time, the environment favors non-discrimination, meaning that individuals who choose *not* to discriminate gain an advantage such as more resources or opportunities, which leads to reproductive success (say, producing 3 offspring). In contrast, discriminators will suffer a penalty, yielding 0 offspring. In the remaining 30% of the time, the environment favors discrimination, rewarding those who discriminate with reproductive success (3 offspring), while non-discriminators yield 0 offspring.

At first glance, always choosing not to discriminate seems like the rational decision because it leads to success more frequently. However, if the entire population consistently refrains from discrimination, the 30% of the time when discrimination leads to success would result in the population having no offspring during those periods, as non-discriminators would not reproduce. Over time, this would lead to extinction, as the population would miss critical opportunities to reproduce and sustain itself. The reverse is true for always choosing to discriminate; this behavior would lead to failure in the 70% of cases where non-discrimination is rewarded. In such cases, discriminators would not reproduce, also leading to extinction. Thus, neither strategy—always discriminating or always not discriminating—is sustainable for the population in the long term.

The optimal behavior in this environment is for individuals to not discriminate 70% of the time and to discriminate 30% of the time, aligning their behavior with the environmental conditions. This is known as the probability matching strategy, and it leads to the highest reproductive success for the group as a whole, ensuring survival across different environmental contexts[16,20]. The mathematical intuition behind why probability matching is the growth-optimal behavior lies in the fact that the population growth rate, $3 \times p^{70\%} \times (1-p)^{30\%}$, is maximized when $p = 70\%$. Over time, individuals who adopt this adaptive behavior dominate the population.

This example illustrates how environmental conditions shape discriminatory behavior. When the environment rewards non-discrimination, those who adapt to it will succeed. However, if the environment still occasionally rewards discrimination, a portion of the population will continue to exhibit this behavior.

To fully appreciate how behavior can be shaped by other important factors in evolution—such as finite memory and mutation—we present the formal model in the next section, which generalizes the binary choice model in the literature[16–18].

The formal model

Consider a hypothetical world with a population composed of two groups, as described in previous literature[18]: a majority group, which we refer to as the "Andorians," and a minority group, which we refer to as the "Tellarians." Group membership is unambiguous, mutually exclusive (each individual is a member of one and only one group), immutable, and observable by all. There are two factors that determine each individual's fitness: λ_A and λ_T. They represent social interactions with Andorian and Tellarian individuals, respectively. An individual who interacts with Andorian individuals is subject to the Andorian factor, λ_A, whereas an individual who interacts with Tellarian individuals is subject to the Tellarian factor, λ_T. Both λ_A and λ_T are independent binary random variables distributed as follows:

$$\lambda_A = \left\{ \begin{array}{ll} \lambda_A^{\text{low}}, & \text{with probability } q, \\ \lambda_A^{\text{high}}, & \text{with probability } 1-q, \end{array} \right. \qquad \lambda_T = \left\{ \begin{array}{ll} \lambda_T^{\text{low}}, & \text{with probability } r, \\ \lambda_T^{\text{high}}, & \text{with probability } 1-r, \end{array} \right.$$

where $\lambda_A^{\text{high}} > \lambda_A^{\text{low}} > 0$, $\lambda_T^{\text{high}} > \lambda_T^{\text{low}} > 0$, and $q, r \in (0, 1)$. Here, for expositional simplicity and without loss of generality, we consider a two-factor model. A more general model with multiple factors can also be found in the literature[17].

Without loss of generality, we assume that each factor takes only one of two possible values: a low fitness of λ_A^{low} or λ_T^{low}, which happens in the context of an adverse event related to that group, and a high fitness of λ_A^{high} or λ_T^{high}, which represents the normal case. Parameters q and r denote the probability of adverse events for the Andorian and the Tellarian groups, respectively, which we refer to as "adverse probabilities" for simplicity. For example, a Tellarian individual may experience an adverse event with a (small) probability r, in which case anyone interacting with that individual will experience low fitness in that period. Examples of adverse events could include crime, disease, or economic hardship, among others.

Analogous to human societies, we will assume that the Tellarian community has been politically underrepresented, with less access to education and economic opportunities[18]. As a result, this inequality has led to a higher adverse probability for the Tellarian community compared to the population average. Note that the higher adverse probability is not innate, but the result of a complicated set of determinants, including limited historical access to resources. However, in this model, individuals observe only each other's group membership, which they use as a marker in the absence of any other information. The true underlying causes of the higher adverse probability, such as lack of educational opportunities, are assumed to be unobservable.

We now focus on the perspective of an Andorian, who faces a decision between one of two actions—whether or not to discriminate against a Tellarian—which determines their fitness[18]. We assume that an Andorian's number of offspring is given by x_{dis} if the individual chooses to discriminate, and x_{nodis} if the individual chooses not to discriminate:

$$x_{\text{dis}} = \beta_{\text{dis}}\lambda_T + (1 - \beta_{\text{dis}})\lambda_A, \qquad x_{\text{nodis}} = \beta_{\text{nodis}}\lambda_T + (1 - \beta_{\text{nodis}})\lambda_A.$$

Here, $0 \leq \beta_{\text{dis}} < \beta_{\text{nodis}} \leq 1$. The fitness of an Andorian depends on both λ_T and λ_A, and β_{dis} and β_{nodis} represent its degrees of interaction with Tellarians under discrimination and non-discrimination, respectively. For example, $\beta_{\text{dis}} = 0.2$ means that, when choosing to discriminate, an Andorian interacts with Tellarians only 20% of the time and with other Andorians 80% of the time. These parameters quantify the intensity of intergroup contact: the lower the value of β_{dis} or β_{nodis}, the stronger the avoidance behavior. In our framework, discrimination is modeled as a reduction in the degree of interaction with Tellarians, which is captured by the inequality $\beta_{\text{dis}} < \beta_{\text{nodis}}$.

Assume all Andorians choose to discriminate with probability $p \in [0, 1]$ and to not discriminate with probability $1 - p$, denoted by a Bernoulli random variable I^p. Hence, the number of offspring for an individual is given by the random variable:

$$x^p = I^p x_{\text{dis}} + (1 - I^p)x_{\text{nodis}},$$

where

$$I^p = \left\{ \begin{array}{ll} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1-p. \end{array} \right.$$

We henceforth refer to p as the probability of discrimination by Andorians. Note that p can be 0 or 1, which corresponds to always discriminating or always not discriminating. Generally, p can also be between 0 and 1, which corresponds to randomized behavior.

We normalize the initial number of Andorians to 1 without loss of generality, and denote the number of Andorians in generation T as n_T. Because n_T grows exponentially over time T, we consider the exponential growth rate of the population size, $T^{-1} \log n_T$. Assume that (λ_A, λ_T) is independent and identically distributed

(IID) over time and identical for all individuals in a given generation. Then, as proved in the literature, $T^{-1} \log n_T$ converges in probability to the log-geometric-average growth rate[16–18,21]:

$$\mu(p) = \mathbb{E}\left[\log\left(px_{\text{dis}} + (1-p)x_{\text{nodis}}\right)\right]. \tag{1}$$

This result aligns with the well-known principle of geometric mean fitness in evolutionary biology[22].

Equation (1) plays a central role in evolutionary dynamics, as it captures the long-run fitness of a strategy in fluctuating environments. As formally established in the literature[16–18,21], maximizing Eq. (1) leads to a "winner-take-all" outcome, as individuals who do not maximize Eq. (1) will be rapidly overrun by those who do. In other words, individuals who maximize Eq. (1) will, over time, dominate the population. For completeness and clarity, we include a formal statement and proof of this result in the Supplementary Material.

The evolutionary behavior of maximizing Eq. (1) could be driven by cultural transmission mechanisms. For example, vertical transmission allows individuals to adopt the strategies of their parents or other high-performing role models who achieve superior long-term growth. Alternatively, under natural selection dynamics, individuals employing suboptimal strategies may be gradually eliminated, as they are systematically outperformed by those who maximize long-run population fitness.

In this framework, the behavior of Andorians is completely characterized by the probability of discrimination, p, and degrees of interaction with Tellarians, β_{dis} and β_{nodis}. Although not entirely realistic from a biological perspective, this simplification clarifies the impact of evolution on behavioral dynamics, allowing us to derive the growth-optimal behavior explicitly.

Discrimination with infinite memory

In this section, we study the case in which individuals have infinite memory. Although such a setting is not realistic in practice, it serves as a theoretical benchmark that allows us to characterize the optimal behavior under perfect information. By comparing it with the finite-memory case, we can evaluate the extent to which cognitive limitations distort decision-making and lead to discriminatory outcomes.

When individuals have infinite memory, their decision to discriminate or not is based on the entire history of observed outcomes. Therefore, in this case we assume that the Andorians have already learned the behavior patterns of the Tellarians over time, and in particular, the probability of Tellarian adverse events, r, is known to Andorians. In this section, we derive the probability of discrimination for Andorians that maximizes their growth rate and analyze the patterns of this optimal strategy.

Optimal probability of discrimination

The following proposition gives the optimal probability of discrimination that maximizes the growth rate of the Andorian group.

Proposition 1 The optimal probability of discrimination, p, that maximizes the log-geometric-average growth rate defined by Eq. (1) is given by

$$p^*(q,r) = \begin{cases} 1, & \text{if } \mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] > 1 \text{ and } \mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] < 1, \\ \text{solution to Eq. (3)}, & \text{if } \mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] \geq 1 \text{ and } \mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] \geq 1, \\ 0, & \text{if } \mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] < 1 \text{ and } \mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] > 1, \end{cases} \tag{2}$$

where p^* is defined implicitly in the second case of Eq. (2) by

$$\mathbb{E}\left[\frac{x_{\text{dis}}}{p^* x_{\text{dis}} + (1-p^*)x_{\text{nodis}}}\right] = \mathbb{E}\left[\frac{x_{\text{nodis}}}{p^* x_{\text{dis}} + (1-p^*)x_{\text{nodis}}}\right]. \tag{3}$$

Proposition 1 demonstrates how the optimal probability of discrimination, p^*, is determined by the relationship between the number of offspring of an Andorian who chooses to discriminate (x_{dis}) and the number of offspring of an Andorian who chooses not to discriminate (x_{nodis}). The optimal probability, p^*, for choosing to discriminate is determined by the comparative fitness between these two strategies, as follows:

- $p^* = 1$ **(Always Discriminate):** When $\mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] > 1$ and $\mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] < 1$, discrimination yields strictly better fitness outcomes than non-discrimination. In this case, the optimal strategy is to always discriminate, as it maximizes the individual's reproductive success.
- $p^* = 0$ **(Never Discriminate):** When $\mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] < 1$ and $\mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] > 1$, the reverse is true—non-discrimination yields superior fitness. Hence, the optimal strategy is to never discriminate.
- $p^* \in (0,1)$ **(Partial Discrimination):** When $\mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] \geq 1$ and $\mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] \geq 1$, neither strategy strictly dominates. While this might appear contradictory, it reflects a situation where the relative advantage of discrimination versus non-discrimination fluctuates across different environments. For example, if there is some probability that discrimination yields significantly higher fitness than non-discrimination ($x_{\text{dis}} \gg x_{\text{nodis}}$), and also some probability that the opposite holds ($x_{\text{nodis}} \gg x_{\text{dis}}$), then it is possible for both expectations to exceed 1. This implies that no single strategy is uniformly optimal, and the best response involves randomizing between them. In this case, p^* lies strictly between 0 and 1 and is determined implicitly by Eq. (3), reflecting a balance between the two strategies.

Proposition 1 demonstrates that, from the perspective of the whole population, the optimal strategy that maximizes the growth rate of the population may not be full discrimination or no discrimination. Instead, if

$\mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] \geq 1$ and $\mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] \geq 1$, the optimal probability of discrimination lies between 0 and 1, implying a randomized discrimination strategy.

As illustrated in the toy example discussed previously, in fact, from an individual's perspective, the survival-maximizing behavior is to always choose the action with higher average fitness ($p = 0$ or 1). However, partial discrimination emerges because the group as a whole benefits from survival advantages beyond individual optimization. In our framework, these benefits arise purely from the stochastic nature of the environment[16–18,23–25]. This is also a generalization of the "adaptive coin-flipping" strategies[26], which are interpreted as a form of altruism, because individuals who engage in this behavior seem to be acting in the interest of the population at the expense of their own individual fitness.

Patterns of discrimination

To further illustrate how the optimal probability of discrimination, p^*, changes with respect to the adverse probabilities of Andorians and Tellarians, q and r, let us consider the following assumption:

Assumption 1 The fitness outcomes of Andorians and Tellarians are identical in adverse situations: $\lambda^{\text{low}} := \lambda_A^{\text{low}} = \lambda_T^{\text{low}}$, and in normal situations: $\lambda^{\text{high}} := \lambda_A^{\text{high}} = \lambda_T^{\text{high}}$.

Assumption 1 simplifies the model by removing any inherent differences in the potential benefits or costs associated with the two groups. As a result, our framework focuses on the probabilities of adverse events, q and r, rather than any intrinsic differences in group fitness.

Under Assumption 1, the optimal solution given by Proposition 1 can be characterized explicitly as follows.

Proposition 2 Under Assumption 1, the optimal probability of discrimination, p, that maximizes the log-geometric-average growth rate defined by Eq. (1) is given explicitly by

$$p^*(q,r) = \begin{cases} 1, & \text{if } r > r_{\text{upper}}, \\ p^{*\text{partial}}(q,r), & \text{if } r_{\text{lower}} \leq r \leq r_{\text{upper}}, \\ 0, & \text{if } r < r_{\text{lower}}, \end{cases} \tag{4}$$

$$= \text{Bound}_0^1\left(p^{*\text{partial}}(q,r)\right), \tag{5}$$

where $\text{Bound}_0^1(x) = \min\{\max\{x,0\},1\}$,

$$r_{\text{upper}} = \frac{[(1-\beta_{\text{dis}})\lambda^{\text{high}} + \beta_{\text{dis}}\lambda^{\text{low}}]q}{(1-2\beta_{\text{dis}})(\lambda^{\text{high}} - \lambda^{\text{low}})q + \beta_{\text{dis}}\lambda^{\text{high}} + (1-\beta_{\text{dis}})\lambda^{\text{low}}},$$

$$r_{\text{lower}} = \frac{[(1-\beta_{\text{nodis}})\lambda^{\text{high}} + \beta_{\text{nodis}}\lambda^{\text{low}}]q}{(1-2\beta_{\text{nodis}})(\lambda^{\text{high}} - \lambda^{\text{low}})q + \beta_{\text{nodis}}\lambda^{\text{high}} + (1-\beta_{\text{nodis}})\lambda^{\text{low}}},$$

and

$$p^{*\text{partial}}(q,r) = \frac{[\beta_{\text{nodis}}\lambda^{\text{high}} + (1-\beta_{\text{nodis}})\lambda^{\text{low}}](1-q)r - [\beta_{\text{nodis}}\lambda^{\text{low}} + (1-\beta_{\text{nodis}})\lambda^{\text{high}}]q(1-r)}{(\beta_{\text{nodis}} - \beta_{\text{dis}})(\lambda^{\text{high}} - \lambda^{\text{low}})(q + r - 2qr)}.$$

In addition, the optimal growth rate defined by Eq. (1) is given explicitly by

$$\begin{aligned} \mu(p^*(q,r)) = {} & qr\log\lambda^{\text{low}} + (1-q)(1-r)\log\lambda^{\text{high}} \\ & + q(1-r)\log\left[(\beta_{\text{dis}} - \beta_{\text{nodis}})(\lambda^{\text{high}} - \lambda^{\text{low}})p^*(q,r) + \beta_{\text{nodis}}\lambda^{\text{high}} + (1-\beta_{\text{nodis}})\lambda^{\text{low}}\right] \\ & + (1-q)r\log\left[(\beta_{\text{dis}} - \beta_{\text{nodis}})(\lambda^{\text{low}} - \lambda^{\text{high}})p^*(q,r) + \beta_{\text{nodis}}\lambda^{\text{low}} + (1-\beta_{\text{nodis}})\lambda^{\text{high}}\right]. \end{aligned}$$

This proposition provides an explicit solution for the optimal probability of discrimination given by Proposition 1. When the adverse probability of the Tellarian group, r, exceeds a threshold r_{upper}, the probability of adverse events is high enough that it becomes optimal to always discriminate, as avoiding interactions with Tellarians maximizes the fitness of Andorians. Conversely, when r is below a threshold r_{lower}, adverse events are infrequent, making it optimal to never discriminate. For intermediate values of r, where $r_{\text{lower}} \leq r \leq r_{\text{upper}}$, the optimal strategy involves partial discrimination with probability $p^{*\text{partial}}(q,r)$, reflecting a balance between risks and benefits. This corresponds to the case $\mathbb{E}[x_{\text{dis}}/x_{\text{nodis}}] \geq 1$ and $\mathbb{E}[x_{\text{nodis}}/x_{\text{dis}}] \geq 1$ in Eq. (2).

The following proposition shows how the optimal probability of discrimination changes with respect to the adverse probabilities for the two groups, q and r.

Proposition 3 The optimal probability of discrimination given by Eq. (4), $p^*(q,r)$, decreases with respect to q and increases with respect to r.

Proposition 3 demonstrates that the optimal probability of discrimination increases as the adverse probability, r, increases, reflecting a greater incentive to discriminate when the risk of adverse events on the part of Tellarians is higher. In contrast, the probability of discrimination decreases as q, the probability of adverse events for Andorians themselves, increases. This indicates that, as Andorians themselves face a higher risk of adverse events, the relative benefit of discriminating against Tellarians diminishes.

Figure 1 shows how the optimal probability of discrimination, $p^*(q, r)$, given by Proposition 2 varies with different values of q and r. For illustrative purposes, we set $\lambda^{\text{low}} = 1$ and $\lambda^{\text{high}} = 2$. In addition, we assume $\beta_{\text{dis}} = 0$, meaning that when an individual chooses to discriminate, they completely avoid interacting with Tellarians. On the other hand, if an individual chooses not to discriminate, they interact equally with both groups, and we let β_{nodis} represent the proportion of Tellarians in the population. Figure 1a,b show the optimal discrimination probabilities when $\beta_{\text{nodis}} = 0.5$ and $\beta_{\text{nodis}} = 0.2$, respectively.

From Fig. 1, we observe a strong polarization between non-discrimination and full discrimination. In the top-left corners (where q is large and r is small), Andorian individuals tend not to discriminate, as the risk to Andorians is higher and the risk from Tellarians is lower. Conversely, in the bottom-right corners (where q is small and r is large), Andorians fully discriminate, because the risk from Tellarians is higher and the risk to Andorians is smaller. The transitional region between these two extremes is relatively narrow, indicating that the shift between non-discrimination and full discrimination happens over a small range of q and r values. Figure 1 also illustrates that the probability of discrimination increases with r and decreases with q. This verifies the theoretical results of Proposition 3.

Comparing Fig. 1a,b, we see that the transitional region in the middle is narrower when $\beta_{\text{nodis}} = 0.2$ than when $\beta_{\text{nodis}} = 0.5$, implying a quicker shift between non-discriminatory and discriminatory behaviors. This means that when the proportion of Tellarians in the population, β_{nodis}, is smaller, Andorians are more likely to discriminate, as Andorian individuals see less benefit from interacting with the minority group of Tellarians. This also implies a stronger behavioral polarization—Andorians will either fully discriminate against Tellarians or will not discriminate at all, with little middle ground.
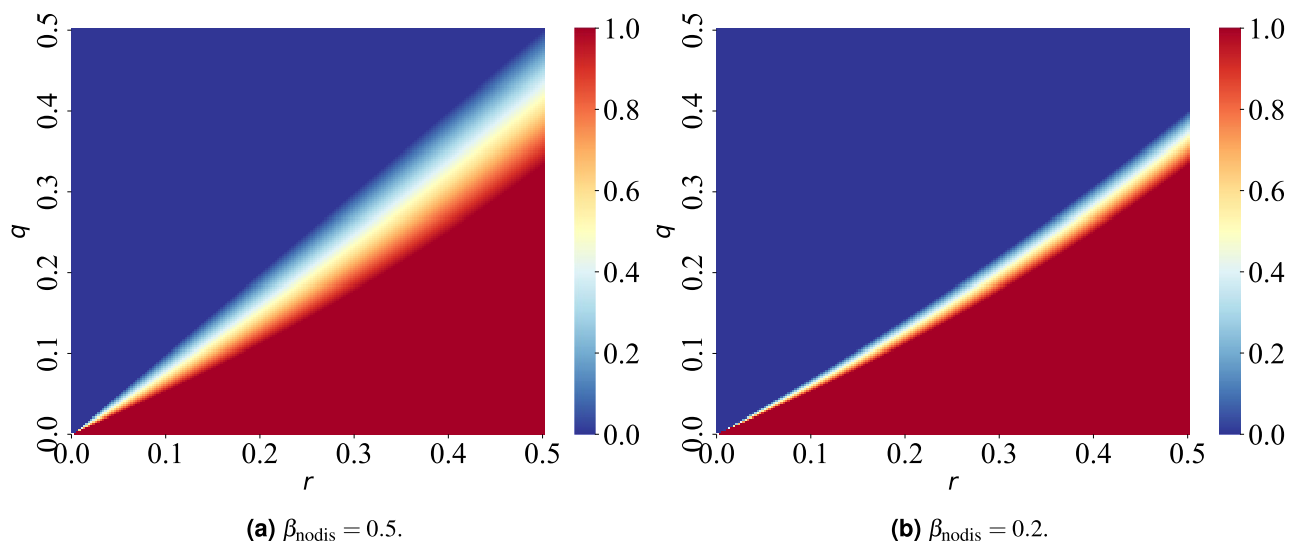
To summarize, we find that the discrimination strategy of Andorians changes with respect to the adverse probabilities of the two groups, and the polarization becomes stronger when the proportion of Tellarians in the population is smaller. These results are based on the assumption that the adverse probability of Tellarians, r, is known to Andorians. In the next section, we explore the discrimination patterns when individuals have finite memory, introducing new complexities to their decision-making.

Discrimination with finite memory

Unlike the case of infinite memory, individuals in practice have a finite memory due to natural cognitive constraints. Therefore, Andorians cannot perfectly estimate the true value of the probability of adverse events for Tellarians, r, based on the whole of their past history of interactions with Tellarians. In this section, we consider the scenario in which Andorians estimate the value of r based on their finite memory of Tellarians. This approach is inspired by the memory/prediction framework[27], which argues that individuals store memory patterns and use them to predict what will happen in the future.

We assume that Andorians use a Bayesian decision analysis framework to incorporate prior judgments and potential updates of their estimate of r into this framework. The basic idea of a Bayesian framework is that Andorians believe in advance that r will take certain values at some prior probability. These prior probabilities may be biased initially, but as Andorians interact more often with Tellarians, they will gradually update these probabilities to yield better estimates.

This Bayesian framework, although seemingly abstract, in fact has a biological motivation. It has been shown that Bayesian decisions can emerge naturally from evolution and adaptation[28]. In addition, human subjects tend to follow Bayesian strategies on average but not individually[29,30]. Human subjects have also been shown to make decisions based on a small number of samples instead of computing the fully Bayesian solution[31–33],

(a) $\beta_{\text{nodis}} = 0.5$.   **(b)** $\beta_{\text{nodis}} = 0.2$.

Figure 1. Optimal probability of discrimination given by Proposition 2. We set $\lambda^{\text{low}} = 1$, $\lambda^{\text{high}} = 2$, and $\beta_{\text{dis}} = 0$. (Hereinafter, we provide the detailed parameter setups of all figures in this article in the figure captions).

which aligns with our assumption of finite memory. The Bayesian paradigm has been widely used in research on cognitive science[34], such as animal learning[35], visual scene perception and concepts[36,37], motor control[38], semantic memory[39], symbolic reasoning[40], social cognition[41], and biological evolution[42–47].

Hereafter, we denote the *true* adverse probability of Tellarians by a constant r^*. Andorians do not know r^*. They apply a Bayesian decision framework, model the adverse probability of Tellarians as a random variable, r, and believe that r is either r_0 or r_1 with prior probabilities $\pi_0 \in (0, 1)$ and $\pi_1 = 1 - \pi_0$, respectively:

$$r = \begin{cases} r_0, & \text{with probability } \pi_0, \\ r_1, & \text{with probability } \pi_1. \end{cases} \tag{6}$$

Without loss of generality, let us assume that $0 < r_0 < r_1 < 1$.

The prior probabilities π_0 and π_1 represent Andorians' prior judgment about Tellarians. As Andorians interact with Tellarians, they may update their judgment based on their observations of Tellarians during their interactions. Assume that Andorians have observed N Tellarians with fitnesses $\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N$, respectively, where $\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N$ are IID random variables following the same distribution as λ_T. The parameter N can be interpreted as a proxy for the amount of information retained in memory—larger N implies more extensive exposure to past experiences, while smaller N captures a more limited view. In this sense, our use of finite N serves as a stylized representation of finite memory, which inherently limits the amount of usable information. While this is a simplification of how real memory functions, it provides a tractable way to model information constraints in belief updating.

Under this framework, the Bayesian estimation of r^* is explicitly given by the following proposition:

Proposition 4 The Bayesian estimation of r^*, $\hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$, is given explicitly by:

$$\hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N) = \mathbb{E}[r|\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N] = \frac{\pi_0 r_0^m (1-r_0)^{N-m} r_0 + \pi_1 r_1^m (1-r_1)^{N-m} r_1}{\pi_0 r_0^m (1-r_0)^{N-m} + \pi_1 r_1^m (1-r_1)^{N-m}}, \tag{7}$$

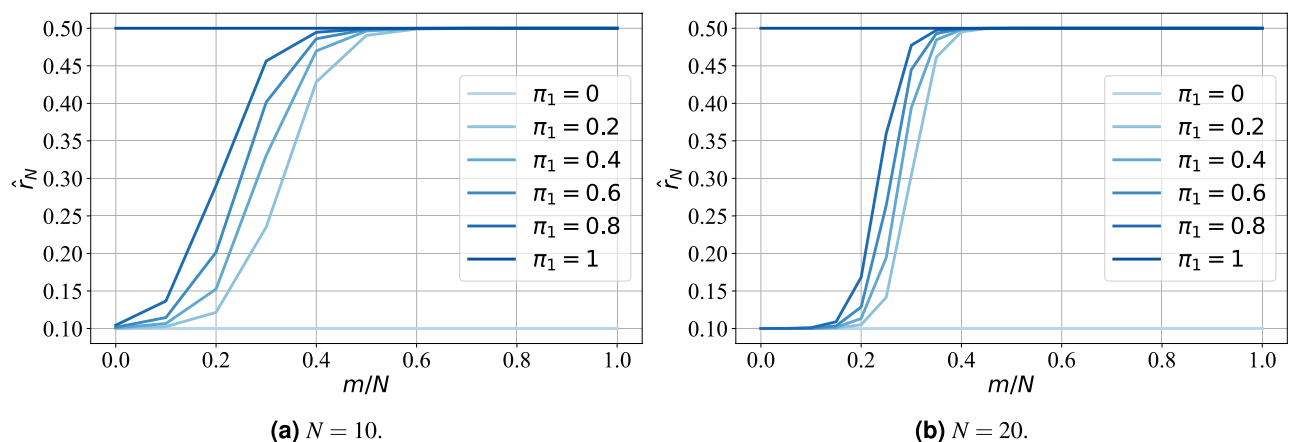where $m = \#\{i : \lambda_T^i = \lambda_T^{\text{low}}, 1 \le i \le N\}$ is the number of Tellarians observed with adverse events. Proposition 4 shows that the Bayesian estimation of r^*, $\hat{r}_N$, is a weighted average of r_0 and r_1, with weights depending on the prior probabilities, π_0 and π_1, as well as the observed outcomes of Tellarians. The estimated value, $\hat{r}_N$, falls between r_0 and r_1. In particular, in the extreme case where $\pi_0 = 1$ and $\pi_1 = 0$, Andorians are fully confident that $r = r_0$, so $\hat{r}_N = r_0$ regardless of the observations. Similarly, if $\pi_0 = 0$ and $\pi_1 = 1$, Andorians are certain that $r = r_1$, and $\hat{r}_N = r_1$.

If π_1 lies within 0 and 1, $\hat{r}_N$ will depend on the number of Tellarians that Andorians have observed, N, and the number of those individuals who experienced adverse events, m. As m increases, the estimation shifts towards the higher probability r_1, while a lower m makes $\hat{r}_N$ closer to r_0. Therefore, the estimator, $\hat{r}_N$, incorporates both prior beliefs and actual observations of Tellarians.

Figure 2 illustrates how $\hat{r}_N$ changes with respect to the observed proportion of Tellarians who experience adverse events, m/N, for different values of the prior probability, π_1. Figure 2a,b correspond to different total numbers of observations, $N = 10$ and $N = 20$, respectively.

Figure 2 shows that $\hat{r}_N$ increases as the proportion of Tellarians with adverse events (m/N) rises, reflecting the influence of observed outcomes on the estimation of r^*. When individuals experience more random adverse events from interactions with Tellarians, they tend to attribute it to the Tellarian species, because it is the most easily observable marker, leading to discrimination against Tellarians. This phenomenon is also referred to as statistical discrimination[2,3]—agents overweight the prevalence of a trait in a group when that trait appears to be highly representative of the group in question[48,49].

The estimation also relies on the prior probability, π_1. For a given observed proportion m/N, $\hat{r}_N$ increases as π_1 increases. This reflects the outcome that, as the weight given to the belief of $r = r_1$ grows, the estimate

Figure 2. The Bayesian estimation of r^* given by Proposition 4. We set $r_0 = 0.1$ and $r_1 = 0.5$.

of r^* will be higher—in the case of humans, this is consistent with the findings in the literature that they tend to anchor towards their original beliefs[50]. When $\pi_1 = 0$, meaning Andorians fully believe $r = r_0$ initially, the estimation remains to be r_0. Conversely, when $\pi_1 = 1$, corresponding to the belief that $r = r_1$, the estimation remains to be r_1.

Comparing Fig. 2a,b, we observe that a larger N leads to a sharper transition in $\hat{r}_N$ as a function of m/N. With more observations, the estimation becomes more sensitive to the proportion of adverse events, resulting in a steeper curve. This shows the impact of larger sample sizes in refining the estimate of r^*, making the estimation process more decisive as N grows.

The following proposition further shows that the estimation of r^* given by Proposition 4, $\hat{r}_N$, can be used to maximize the *conditional* expected growth rate of the population, given the observations of Tellarians.

Proposition 5 The optimal probability of discrimination, p, that maximizes the conditional expected growth rate,

$$\mu(p|\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N) = \mathbb{E}\left[\log\left(px_{\text{dis}} + (1-p)x_{\text{nodis}}\right)|\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N\right],$$

is

$$\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N) := p^*\left(q, \hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)\right), \tag{8}$$

where $p^*(\cdot, \cdot)$ is given by Eq. (2) and $\hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ is given by Eq. (7).

Proposition 5 demonstrates that the Bayesian decision is optimal when maximizing the conditional expected growth rate of the Andorian group if the true value of r^* is unknown and the memory is finite. In addition, Proposition 5 also provides a practical approach for determining the optimal probability of discrimination. With this result, we can first estimate r^* using the Bayesian estimation given by Proposition 4, $\hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$. Then, we substitute $\hat{r}_N$ into the expression for the optimal probability of discrimination $p^*(q, \hat{r}_N)$, as defined in Eq. (2).

Patterns of discrimination

In this section, we study the patterns of discrimination when Andorians have finite memory. Figures 3, 4, and 5 show how the optimal probability of discrimination, $\hat{p}_N^*$, given by Proposition 5 varies with respect to different parameters. In particular, Fig. 3 focuses on the impact of prior probability, π_1, and the prior belief of adverse probability of Tellarians, r_1; Fig. 4 focuses on the observed number of adverse events of Tellarians, m; and Fig. 5 further focuses on the number of observed Tellarians, N. As in Fig. 1, we observe that the optimal probability of discrimination increases as r^* increases and as q decreases.

In Fig. 3a–d, we observe that when r_1 increases from 0.3 to 0.5, the range where the discrimination probability equals 1 expands. This means that higher values of r_1 make fully discriminatory behavior more likely. In addition, as π_1 increases from 0.2 to 0.8, individuals are also more inclined to adopt full discrimination. Both indicate that prior beliefs play a crucial role in determining how easily discrimination is triggered.

Figure 4 further shows how the observed number of adverse events of Tellarians, m, influences the optimal probability of discrimination. As shown in Fig. 4a–d, the area with full discrimination expands from $m = 0.8Nr^*$ to $m = 1.2Nr^*$. This means that the more frequently adverse outcomes are observed among the Tellarians, the broader is the resulting range of discrimination. In other words, discrimination may also emerge due to biased observations.

Figure 5 illustrates how the optimal probability of discrimination responds to changes in the number of observed Tellarians, N, under the assumption that the observed number of adverse events is set to its expected value, $m = Nr^*$. In Fig. 5a,c, we fix $N = 5$, while in Fig. 5b,d, we increase the number of observed Tellarians to $N = 10$. We find that the difference between Fig. 5b,d—which is driven by variation in π_1—is noticeably smaller than the corresponding difference between Fig. 5a,c. This suggests that, as more evidence accumulates through increased observations, the influence of prior beliefs on discriminatory behavior gradually diminishes.
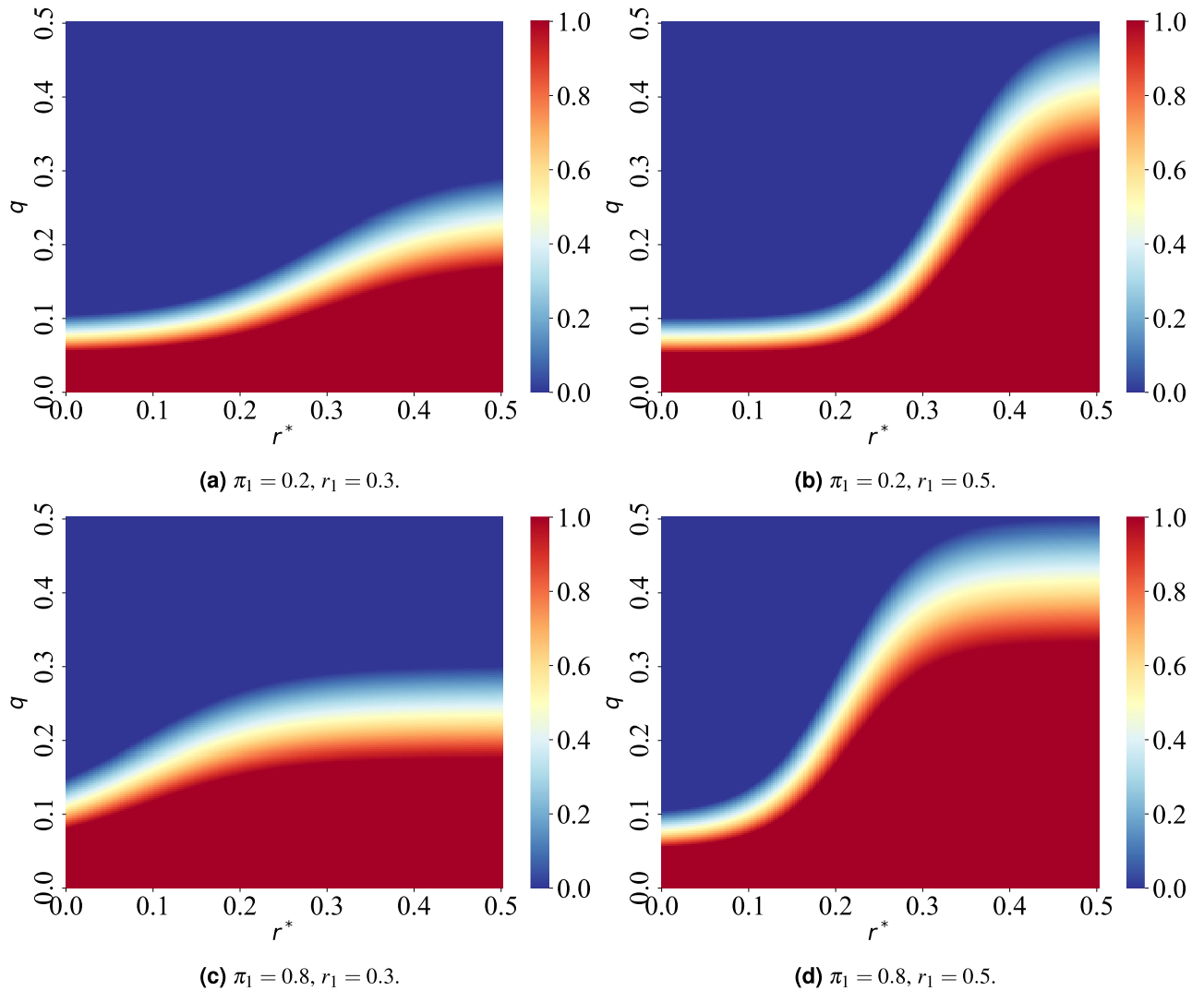
Finally, by comparing Figs. 3–5 and Fig. 1, we find that the patterns of discrimination are different if Andorians have finite or infinite memory. For example, from the figures, when the true adverse probability of Tellarians, r^*, is smaller than the prior value, $r_0 = 0.1$, discrimination tends to occur more frequently under finite memory. This is not simply due to the limited number of observations, but because the Bayesian estimate, $\hat{r}_N$, remains partially anchored to the prior belief, r_0, leading to an overestimation of risk and, consequently, greater discriminatory behavior. In the real world, because all individuals have only finite memory, prior beliefs and biased observations play significant roles in these patterns.

Types of errors

As we have shown, prior beliefs and biases in observations can significantly influence discriminatory behavior. This influence can lead to situations where Andorians either discriminate against Tellarians who pose a lower risk, or fail to discriminate against those who pose a higher risk. These two types of errors reflect the potential consequences of "incorrect" decisions based on imperfect information.

We formally define the two types of errors as follows. A Type I error occurs when Andorians discriminate against Tellarians who have a lower adverse probability (r_0), while a Type II error occurs when Andorians do not discriminate against Tellarians who have a higher adverse probability (r_1). In our context, we define the two types of errors as follows:

(a) $\pi_1 = 0.2, r_1 = 0.3.$

(b) $\pi_1 = 0.2, r_1 = 0.5.$

(c) $\pi_1 = 0.8, r_1 = 0.3.$

(d) $\pi_1 = 0.8, r_1 = 0.5.$

Figure 3. The optimal probability of discrimination under finite memory given by Proposition 5 with different prior probability, π_1, and prior belief of adverse probability of Tellarians, r_1. We set $N = 10$, $m = Nr^*$, $r_0 = 0.1$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0$, and $\beta_{\text{nodis}} = 0.5$.

$$\text{Type I error} = \mathbb{E}\left[\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \dots, \lambda_T^N)|r^* = r_0\right], \tag{9}$$

$$\text{Type II error} = \mathbb{E}\left[1 - \hat{p}_N^*(\lambda_T^1, \lambda_T^2, \dots, \lambda_T^N)|r^* = r_1\right], \tag{10}$$

where $\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \dots, \lambda_T^N)$ is given by Eq. (8).

By definition, if Andorians always choose to discriminate, the Type I error $= 1$ and the Type II error $= 0$; conversely, if they never discriminate, the Type II error $= 1$ and the Type I error $= 0$. In practice, Andorians may choose to discriminate with a certain probability, leading to a trade-off between the two types of errors.

The following proposition provides explicit formulas to compute the two types of errors.

Proposition 6 The two types of errors defined by Eqs. (9) and (10) can be computed by

$$\text{Type I error} = \sum_{m=0}^{N} \frac{N!}{m!(N-m)!} r_0^m (1-r_0)^{N-m} p^*(q, \hat{r}_N(m)),$$

$$\text{Type II error} = 1 - \sum_{m=0}^{N} \frac{N!}{m!(N-m)!} r_1^m (1-r_1)^{N-m} p^*(q, \hat{r}_N(m)),$$

where $p^*(\cdot, \cdot)$ is given by Eq. (2) and

(a) $\pi_1 = 0.2$, $m = 0.8Nr^*$.

(b) $\pi_1 = 0.2$, $m = 1.2Nr^*$.

(c) $\pi_1 = 0.8$, $m = 0.8Nr^*$.

(d) $\pi_1 = 0.8$, $m = 1.2Nr^*$.

Figure 4. The optimal probability of discrimination under finite memory given by Proposition 5 with different prior probability, π_1, and number of observed adverse events of Tellarians, m. We set $N = 10$, $r_0 = 0.1$, $r_1 = 0.3$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0$, and $\beta_{\text{nodis}} = 0.5$.

$$\hat{r}_N(m) = \frac{\pi_0 r_0^m (1-r_0)^{N-m} r_0 + \pi_1 r_1^m (1-r_1)^{N-m} r_1}{\pi_0 r_0^m (1-r_0)^{N-m} + \pi_1 r_1^m (1-r_1)^{N-m}}. \tag{11}$$

Figure 6 shows how the Type I and Type II errors given by Proposition 6 change with respect to the prior probability, π_1, and the adverse probability for Andorians, q. Figure 6a illustrates the Type I error, which increases as π_1 increases. This indicates that as the prior belief in a higher adverse probability r_1 becomes stronger, the probability of incorrectly choosing to discriminate when $r^* = r_0$ increases. Similarly, the Type I error also increases as the adverse probability of Andorians, q, decreases.
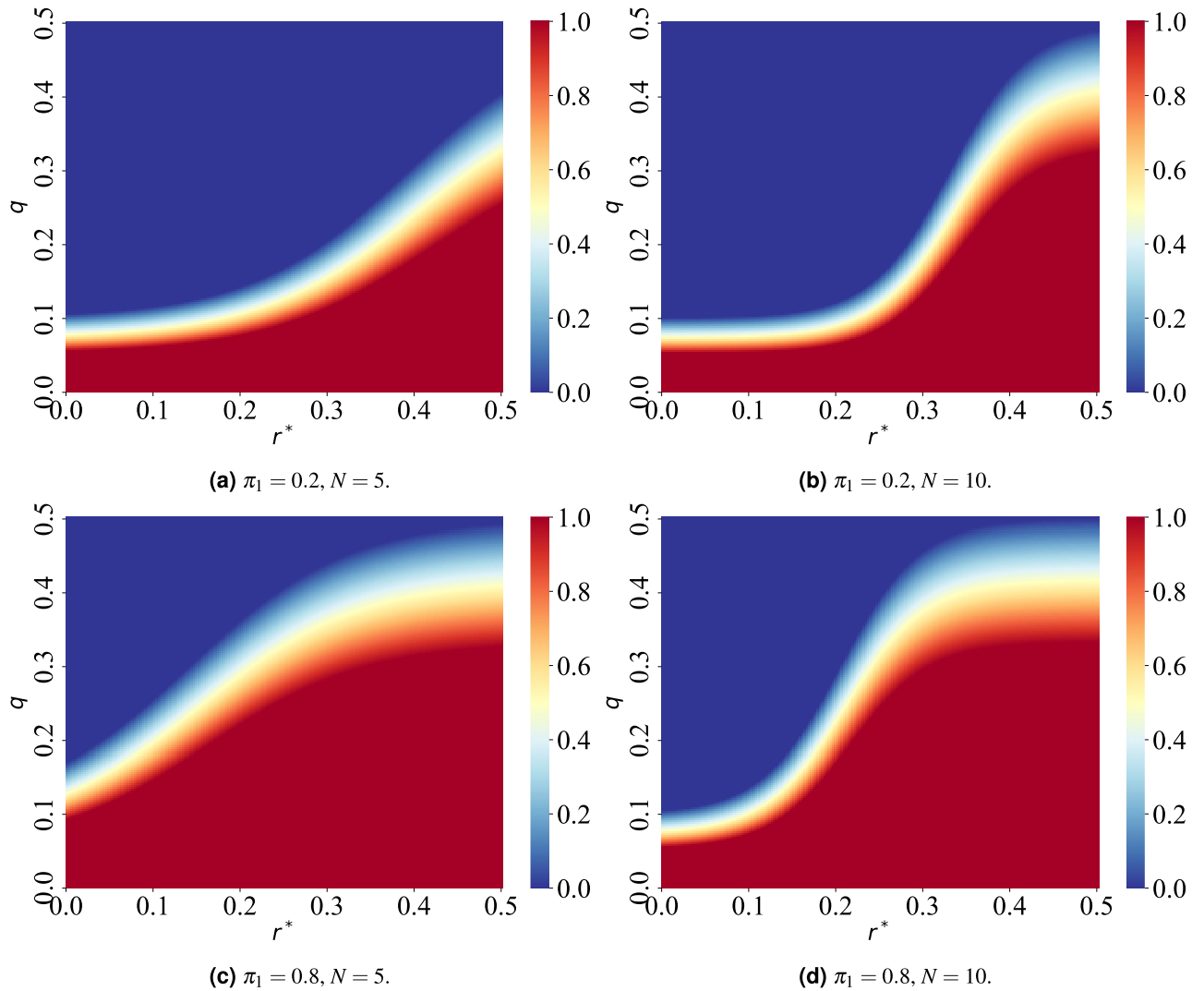
In contrast, Fig. 6b shows the Type II error, which decreases with π_1. This means that as the prior belief in r_1 strengthens, the probability of wrongly not discriminating when the true adverse probability is r_1 becomes lower. The Type II error also decreases as q decreases due to a higher probability of discrimination.

Overall, the figure demonstrates a classic trade-off between Type I and Type II errors. Increasing π_1 raises the chance of a Type I error while reducing the Type II error. This highlights the role of prior beliefs in shaping the balance between greater discrimination (a Type I error) and missing avoidable risks (a Type II error).

Asymptotic behavior
Our previous analysis shows that finite memory can significantly influence discriminatory behavior, as Andorians must rely on limited observations to make their discriminatory decisions. In this section, we examine the asymptotic behavior of discrimination as the number of observations increases without bound. This analysis allows us to explore the relationship between finite memory and infinite memory.

The following proposition shows the asymptotic behavior of $\hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ given by Eq. (7) as the number of observations of Tellarians, N, increases without bound.

(a) $\pi_1 = 0.2, N = 5$.

(b) $\pi_1 = 0.2, N = 10$.

(c) $\pi_1 = 0.8, N = 5$.

(d) $\pi_1 = 0.8, N = 10$.

Figure 5. The optimal probability of discrimination under finite memory given by Proposition 5 with different prior probability, π_1, and number of observed Tellarians, N. We set $r_0 = 0.1$, $r_1 = 0.3$, $m = 1.0Nr^*$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0$, and $\beta_{\text{nodis}} = 0.5$.

Proposition 7 Denote $\hat{r}_N = \hat{r}_N(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ given by Eq. (7). As N increases without bound, we have

$$\hat{r}_N \overset{a.s.}{\to} \begin{cases} r_0, & r^* < \tilde{r}, \\ r_1, & r^* > \tilde{r}, \end{cases} \tag{12}$$

where

$$\tilde{r} = \frac{\log[(1-r_0)/(1-r_1)]}{\log[(1-r_0)/(1-r_1)] + \log(r_1/r_0)}.$$
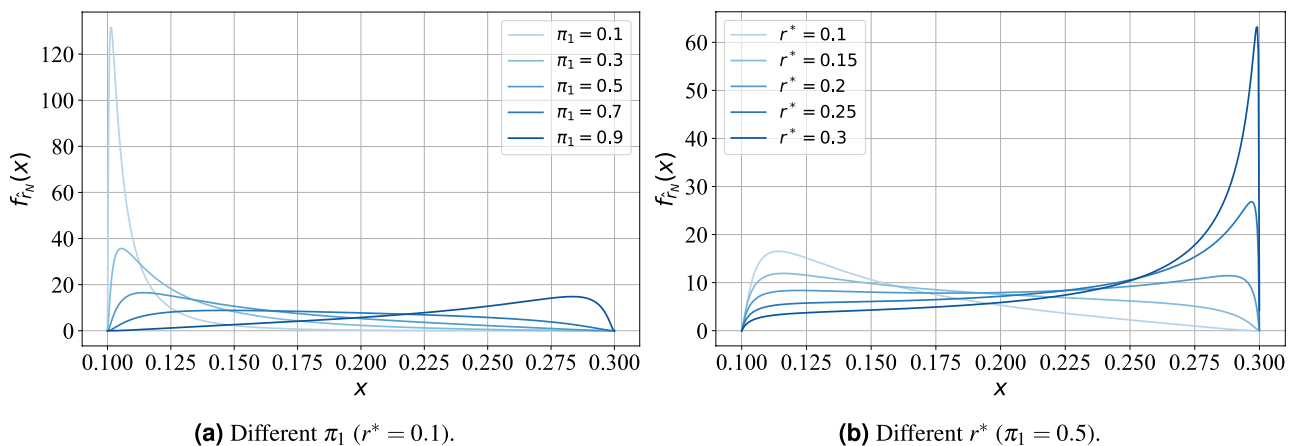
In addition, we have

$$\frac{\log \frac{\hat{r}_N - r_0}{r_1 - \hat{r}_N} + \log \frac{\pi_0}{\pi_1} - N\left[r^* \log \frac{r_1}{r_0} - (1-r^*)\log \frac{1-r_0}{1-r_1}\right]}{\sqrt{Nr^*(1-r^*)\left(\log \frac{r_1}{r_0} + \log \frac{1-r_0}{1-r_1}\right)^2}} \overset{d}{\to} \mathcal{N}(0,1), \tag{13}$$

which implies that for sufficiently large N, $\hat{r}_N$ has the following asymptotic density for $x \in (r_0, r_1)$:

$$f_{\hat{r}_N}(x) \approx \varphi\left(\frac{\log \frac{x - r_0}{r_1 - x} + \log \frac{\pi_0}{\pi_1} - N\left[r^* \log \frac{r_1}{r_0} - (1-r^*)\log \frac{1-r_0}{1-r_1}\right]}{\sqrt{Nr^*(1-r^*)\left(\log \frac{r_1}{r_0} + \log \frac{1-r_0}{1-r_1}\right)^2}}\right) \cdot \frac{r_1 - r_0}{(x - r_0)(r_1 - x)}, \tag{14}$$

(a) Type I error.　　　　　　　　　　　　　　　　　**(b)** Type II error.

Figure 6. Type I and Type II errors given by Proposition 6. We set $N = 10, r_0 = 0.1, r_1 = 0.3, \lambda_{\text{high}} = 2,$ $\lambda_{\text{low}} = 1, \beta_{\text{dis}} = 0,$ and $\beta_{\text{nodis}} = 0.5.$

(a) Different π_1 $(r^* = 0.1)$.　　　　　　　　　　**(b)** Different r^* $(\pi_1 = 0.5)$.

Figure 7. Asymptotic density of $\hat{r}_N$ given by Proposition 7. We set $r_0 = 0.1, r_1 = 0.3,$ and $N = 10.$

where $\varphi(\cdot)$ is the density of the standard normal distribution. Here, "$\overset{a.s.}{\to}$" and "$\overset{d}{\to}$" stand for almost surely convergence and convergence in distribution, respectively.

Equation (12) implies that $\hat{r}_N$ converges to r_0 almost surely if $r^* = r_0$, and converges to r_1 almost surely if $r^* = r_1$. This implies that the Bayesian estimator, $\hat{r}_N$, is a consistent estimator of r^*. In other words, as the number of observations of Tellarians, N, increases without bound, the estimation of r^* under finite memory will eventually converge to the case of infinite memory.

Equation (14) further gives the asymptotic density of $\hat{r}_N$ as N increases without bound. Figure 7 shows this asymptotic density under different values of π_1 and r^*. Figure 7a illustrates that, when π_1 is low, the density is concentrated around $r_0 = 0.1$, indicating a stronger belief that r^* is near r_0. As π_1 grows, the density shifts toward $r_1 = 0.3$, showing that the estimation increasingly aligns with the higher value of prior belief in Tellarians' adverse probability.

Figure 7b shows the asymptotic density for different values of the true adverse probability of Tellarians, r^*. As r^* increases from 0.1 to 0.3, the density shifts from being concentrated around $r_0 = 0.1$ to being closer to $r_1 = 0.3$. This illustrates that the Bayesian estimation, $\hat{r}_N$, gives a consistent estimation of the true adverse probability, r^*.

The following proposition gives the asymptotic behavior of the optimal probability of discrimination, $\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$, given by Eq. (8).

Proposition 8 Under Assumption 1, denote $\hat{p}_N^* = \hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ given by Eq. (8). As N increases without bound, if $\hat{r}_N \overset{a.s.}{\to} \hat{r}_\infty$, we have

$$\hat{p}_N^* \overset{a.s.}{\to} p^*(q, \hat{r}_\infty).\tag{15}$$

In addition, if $r_{\text{lower}} < \hat{r}_\infty < r_{\text{upper}}$, we have

$$\frac{\log \frac{(B-r_0 C)\hat{p}_N^* + A - r_0 D}{(r_1 C - B)\hat{p}_N^* + r_1 D - A} + \log \frac{\pi_0}{\pi_1} - N\left[r^* \log \frac{r_1}{r_0} - (1-r^*)\log \frac{1-r_0}{1-r_1}\right]}{\sqrt{Nr^*(1-r^*)\left(\log \frac{r_1}{r_0} + \log \frac{1-r_0}{1-r_1}\right)^2}} \overset{d}{\to} \mathcal{N}(0,1),\tag{16}$$

which implies that for sufficiently large N, $\hat{p}_N^*$ has the following asymptotic density for $y \in (0,1)$:

$$f_{\hat{p}_N^*}(y) \approx \varphi\left(\frac{\log \frac{(B-r_0 C)y + A - r_0 D}{(r_1 C - B)y + r_1 D - A} + \log \frac{\pi_0}{\pi_1} - N\left[r^* \log \frac{r_1}{r_0} - (1-r^*)\log \frac{1-r_0}{1-r_1}\right]}{\sqrt{Nr^*(1-r^*)\left(\log \frac{r_1}{r_0} + \log \frac{1-r_0}{1-r_1}\right)^2}}\right)$$
$$\cdot \frac{(r_1 - r_0)(BD - AC)}{[(B - r_0 C)y + A - r_0 D][(r_1 C - B)y + r_1 D - A]},\tag{17}$$

where

$$A = \left[\beta_{\text{nodis}}\lambda^{\text{low}} + (1 - \beta_{\text{nodis}})\lambda^{\text{high}}\right]q,$$
$$B = (\beta_{\text{nodis}} - \beta_{\text{dis}})(\lambda^{\text{high}} - \lambda^{\text{low}})q,$$
$$C = -(\beta_{\text{nodis}} - \beta_{\text{dis}})(\lambda^{\text{high}} - \lambda^{\text{low}})(1 - 2q),$$
$$D = (\lambda^{\text{high}} - \lambda^{\text{low}})(1 - 2\beta_{\text{nodis}})q + \beta_{\text{nodis}}\lambda^{\text{high}} + (1 - \beta_{\text{nodis}})\lambda^{\text{low}},$$
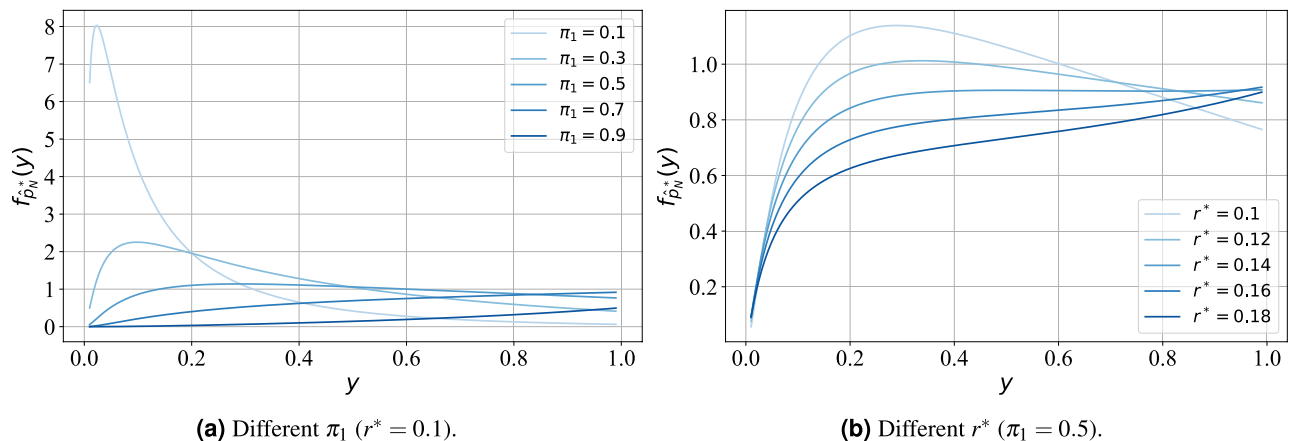
and $\varphi(\cdot)$ is the density of the standard normal distribution. Here, "$\overset{a.s.}{\to}$" and "$\overset{d}{\to}$" stand for almost surely convergence and convergence in distribution, respectively.

Equation (15) demonstrates that, as the number of observations of Tellarians, N, increases without bound, the optimal probability of discrimination under finite memory, $\hat{p}_N^*$, will converge to the result under the case of infinite memory. Equation (17) further gives the asymptotic density of $\hat{p}_N^*$ when the optimal strategy with infinite memory is partial discrimination ($r_{\text{lower}} < \hat{r}_\infty < r_{\text{upper}}$). Figure 8 shows this asymptotic density of $\hat{p}_N^*$ under different values of π_1 and r^*.

Figure 8a shows how the asymptotic density of $\hat{p}_N^*$ changes with varying prior probabilities π_1. As π_1 increases, the density of $\hat{p}_N^*$ concentrates more towards large values. This indicates that, with a higher prior belief in the higher adverse probability r_1, the optimal probability of discrimination tends to be higher.
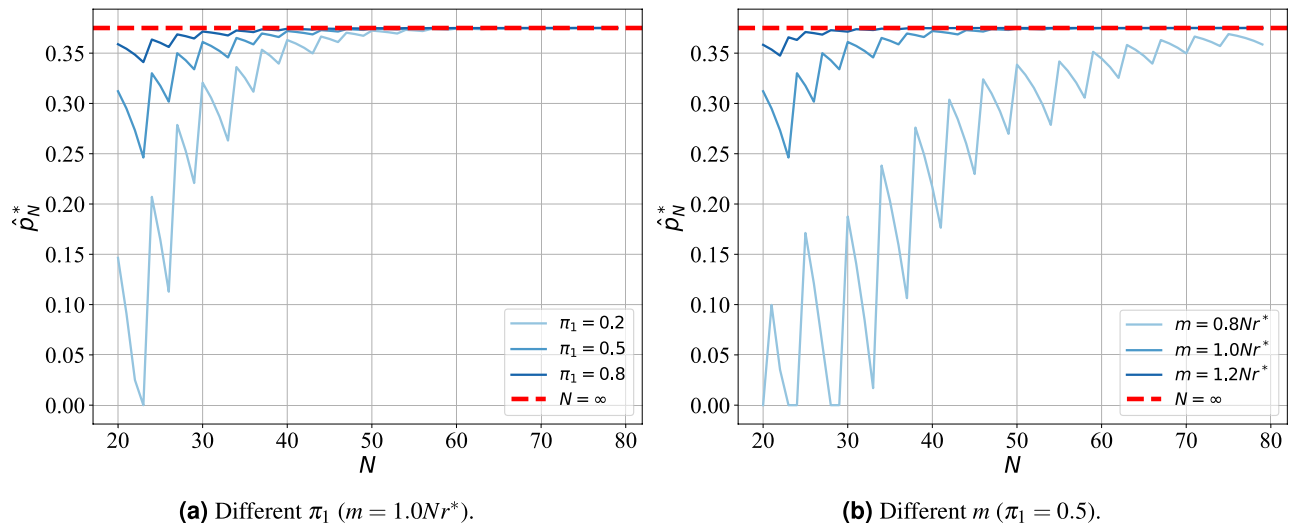
Figure 8b shows how the asymptotic density changes with different true values of r^*. As r^* increases, the density of $\hat{p}_N^*$ shifts towards larger values, indicating a stronger tendency toward discrimination, while when r^* is closer to r_0, the density of $\hat{p}_N^*$ is more concentrated near lower values. This implies that the optimal probability of discrimination depends on the true adverse probability of Tellarians, r^*. These observations meet our intuition and previous findings—finite memory will lead to Andorians making decisions based on their prior judgment and their finite observations of Tellarians, and if the prior information and observations are biased, unnecessary discrimination may emerge.

Figure 9 further illustrates the asymptotic behavior of the optimal probability of discrimination, $\hat{p}_N^*$, by showing how it evolves with the number of observed Tellarians, N, under finite memory. In Fig. 9a, we fix $m = Nr^*$ and vary the prior belief π_1, while in Fig. 9b, we fix $\pi_1 = 0.5$ and vary the number of observed adverse

(a) Different π_1 ($r^* = 0.1$).

(b) Different r^* ($\pi_1 = 0.5$).

Figure 8. Asymptotic density of $\hat{p}_N^*$ given by Proposition 8. We set $r_0 = 0.1$, $r_1 = 0.3$, $N = 10$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0$, and $\beta_{\text{nodis}} = 0.5$.

(a) Different π_1 ($m = 1.0Nr^*$). **(b)** Different m ($\pi_1 = 0.5$).

Figure 9. The optimal probability of discrimination given by Eq. (8), $\hat{p}_N^*$, versus the number of observed Tellarians, N. We set $r_0 = 0.1$, $r_1 = 0.3$, $r^* = 0.3$, $q = 0.25$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0$, and $\beta_{\text{nodis}} = 0.5$.

events of Tellarians, m. The red dashed lines represent the benchmark optimal discrimination probability under infinite memory, as given by Eq. (4). From Fig. 9a,b, we observe that, as N increases, the finite-memory solution, $\hat{p}_N^*$, gradually converges to the infinite-memory benchmark.

In addition, we observe that the speed of convergence depends crucially on both the prior belief and the observed outcomes. Specifically, in these simulations, we fix $r_0 = 0.1$, $r_1 = 0.3$, and set the true value of r^* to $r_1 = 0.3$. As shown in Fig. 9a, a higher prior weight on r_1 (i.e., larger π_1) leads to faster convergence, as the Andorians are already predisposed to believe in the true value of r^*. Figure 9b shows that, when more adverse outcomes are observed (i.e., higher m), the influence of the incorrect hypothesis, r_0, diminishes more rapidly, further facilitating convergence to the infinite-memory optimal discrimination probability. These findings highlight how both prior beliefs and observation quality jointly shape discrimination behavior.

One can also observe an interesting wave-like fluctuation in Fig. 9 as N increases. This arises because the estimate $\hat{r}_N$ is based on N binary outcomes, and the number of observed adverse events, m, can only take integer values. As a result, small changes in N can cause $\hat{r}_N$ to shift across different regions of the piecewise-defined strategy function in Eq. (4), resulting in discontinuous adjustments in the optimal discrimination probability. Similar patterns often appear in models involving binomial inputs, such as Edgeworth expansions for sums of Bernoulli random variables[51] and binomial tree models in option pricing[52].

Mutation and finite memory

We have thoroughly examined the optimal discrimination strategies under both infinite and finite memory scenarios. In particular, we observed how memory limitations affect decision-making and how the optimal probability of discrimination evolves as the number of observations grows. Theoretically, the case of infinite memory seems to provide a more comprehensive understanding of the environment, as it incorporates all available information to maximize population growth. This leads to a natural question: from the perspective of maximizing the long-term growth of a population, is infinite memory truly superior to finite memory?

Interestingly, we know that evolution has favored finite memory systems in practice. Populations do not rely on infinite historical data, but instead make decisions based on limited observations and experiences. Why has evolution favored finite memory? What advantages does it offer over infinite memory?

In this section, we demonstrate that, in the absence of mutations, infinite memory is indeed optimal for maximizing population growth. However, when mutations are introduced into the population—creating variability in behavior and outcomes—finite memory systems can outperform infinite memory systems in terms of maximizing population growth. We explore how finite memory allows populations to adapt more efficiently to changing environments, while infinite memory may lead to rigidity and suboptimal decisions in the presence of such mutations. By incorporating the effects of mutations into our analysis, we show that finite memory is not a limitation, but an evolutionary advantage in maximizing the adaptability and growth potential of populations.

The optimality of infinite memory without mutation

The following proposition establishes that, under our previous framework without mutation, the population growth rate achieved under infinite memory is always superior to that of finite memory.

Proposition 9 Assume $r^* \in \{r_0, r_1\}$. For the expected growth rate given by Eq. (1), $\mu(p)$, we have

$$\mathbb{E}\left[\mu\left(\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)\right)\right] \leq \mu\left(\hat{p}_\infty^*(\lambda_T^1, \lambda_T^2, \ldots)\right) = \mu\left(p^*(q, r^*)\right),$$

where $\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ is given by Eq. (8), $\hat{p}_\infty^*(\lambda_T^1, \lambda_T^2, \ldots) = \lim_{N \to \infty} \hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$, and $p^*(q, r)$ is given by Eq. (2).

The intuition behind Proposition 9 is that populations with infinite memory have access to all historical information, allowing them to make optimal decisions based on a complete understanding of the environment. Consequently, the expected growth rate under infinite memory will always be better than the growth rate under finite memory.

Figure 10 illustrates how the expected growth rate of the Andorian population changes as the number of observed Tellarians, N, increases. The red dashed line represents the growth rate achieved under infinite memory ($N = \infty$), and the blue curves correspond to varying prior probabilities π_1 for the higher adverse probability, r_1.

Figure 10 demonstrates that, the expected growth rate of Andorians under finite memory, $\mathbb{E}\left[\mu\left(\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)\right)\right]$, is always lower than the value achieved under infinite memory. In addition, the expected growth rate improves with more observations, as Andorians gain more information about the true value of r^*, allowing them to make more informed discrimination decisions.

Figure 10 also shows that, for higher values of π_1, the growth rate converges more slowly to the infinite memory result as the number of observations, N, increases. This is because a higher π_1 means that Andorians initially expect a higher risk from Tellarians, which leads to more discrimination even when r^* is low. Therefore, more observations are needed to correct this prior belief, resulting in a slower convergence. Conversely, for lower values of π_1, where the prior belief favors r_0, the expected growth rate converges faster to the infinite memory case, as fewer observations are required to adjust the discrimination strategy.

The optimality of finite memory with mutation

Up to now, we always assume that Tellarians have a fixed true probability of adverse events, r^*. However, in reality, mutations can occur. In this section, we incorporate the possibility that Tellarians can mutate[53]. Specifically, we assume that for each Tellarian, with a small probability p_m, its probability of adverse events mutates to $r_m \neq r^*$, and with a probability $1 - p_m$, its adverse event probability remains at r^*. Mutations occur independently across Tellarians. Andorians are unaware of the possibility of mutations and base their decisions on the assumption that r^* is constant.

The following proposition demonstrates that, when mutations are introduced, finite memory can outperform infinite memory in terms of maximizing the population growth rate.
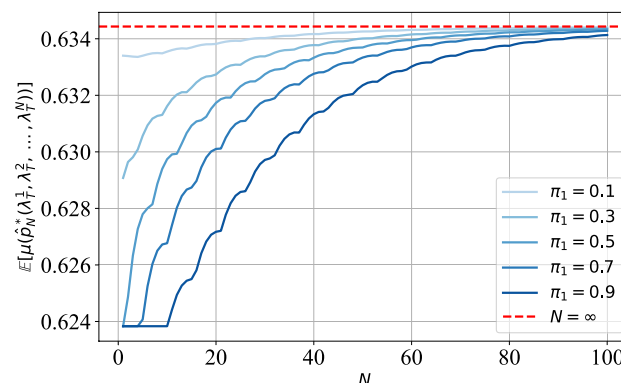
Proposition 10 Assume that the probability of mutation satisfies $p_m \in (0, 1)$ and the adverse probability after mutation $r_m \in (0, 1)$ satisfies $r_m \neq r^*$. Further assume that $r^* \in \{r_0, r_1\}$, and Assumption 1 holds. Then, given $p_m, r_m,$ and r^*, we can always find a set of parameters, $\pi_0, \pi_1, r_0, r_1, q, \beta_{\text{nodis}}, \beta_{\text{dis}}, \lambda^{\text{low}},$ and λ^{high}, such that there exists $N < \infty$ satisfying

$$\mathbb{E}\left[\mu\left(\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)\right)\right] > \mu\left(\hat{p}_\infty^*(\lambda_T^1, \lambda_T^2, \ldots)\right),$$
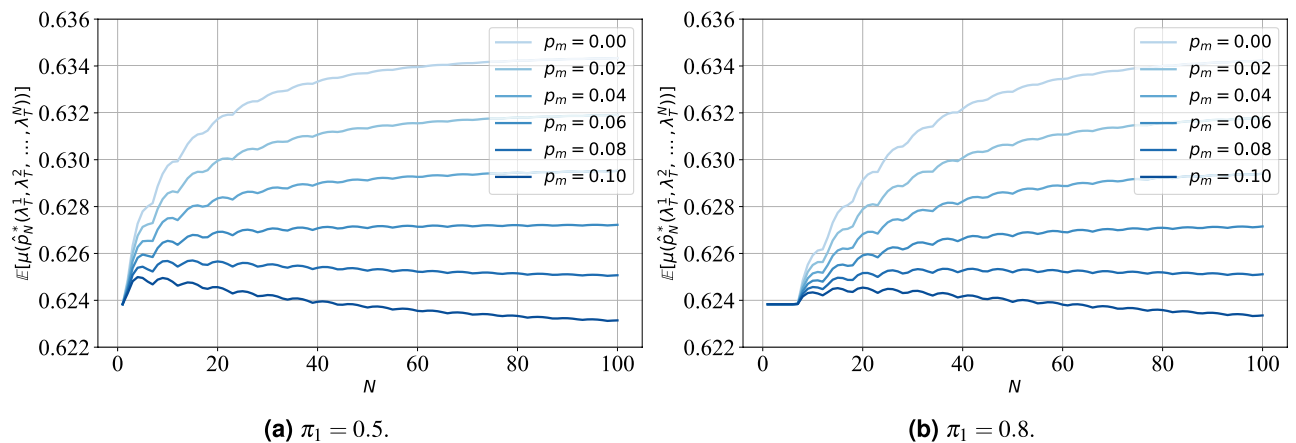
where $\hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$ is given by Eq. (8) and $\hat{p}_\infty^*(\lambda_T^1, \lambda_T^2, \ldots) = \lim_{N \to \infty} \hat{p}_N^*(\lambda_T^1, \lambda_T^2, \ldots, \lambda_T^N)$.

The intuition behind Proposition 10 lies in the adaptive flexibility that finite memory provides in dynamic environments. While infinite memory may appear advantageous in stable environments, it can hinder adaptation when the environment changes unexpectedly—for instance, due to mutations that alter the true adverse event probability. In such cases, individuals with infinite memory may become overly anchored to outdated prior beliefs, making their behavior less responsive to environmental changes.
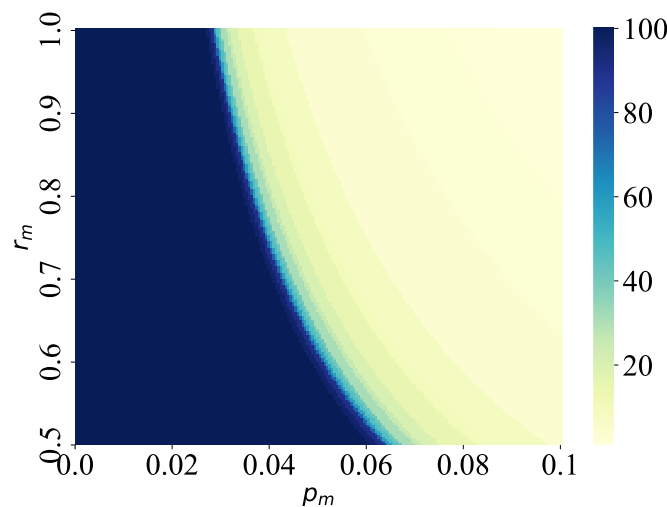
In contrast, finite memory allows individuals to update their beliefs based on a limited set of observations under the Bayesian framework, making a compromise between prior beliefs and observations. This constraint—rather than being a liability—can serve as a strength in dynamic settings. By avoiding overcommitment to prior

Figure 10. The expected growth rate of the Andorian population versus the number of observed Tellarians, N. We set $q = 0.1$, $r^* = 0.1$, $r_0 = 0.1$, $r_1 = 0.3$, $\beta_{\text{dis}} = 0.0$, $\beta_{\text{nodis}} = 0.5$, $\lambda_{\text{high}} = 2$, and $\lambda_{\text{low}} = 1$.

(a) $\pi_1 = 0.5$.   **(b)** $\pi_1 = 0.8$.

Figure 11. The expected growth rate of the Andorian population versus the number of observed Tellarians, N, under different mutation probabilities, p_m. We set $r^* = 0.1$, $r_0 = 0.1$, $r_1 = 0.3$, $\lambda_{\text{high}} = 2$, $\lambda_{\text{low}} = 1$, $\beta_{\text{dis}} = 0.0$, $\beta_{\text{nodis}} = 0.5$, and $r_m = 0.5$.

Figure 12. The optimal number of observations, N, that maximizes the expected growth rate of the Andorian population. We set $q = 0.1$, $r^* = 0.1$, $r_0 = 0.1$, $r_1 = 0.3$, $\pi_0 = \pi_1 = 0.5$, $\beta_{\text{dis}} = 0$, $\beta_{\text{nodis}} = 0.5$, $\lambda_{\text{high}} = 2$, and $\lambda_{\text{low}} = 1$.

beliefs, finite memory enables behavior to better reflect newly emerging patterns. This flexibility is what allows populations with finite memory to outperform those with infinite memory in environments with mutation.

Figure 11 illustrates how the expected growth rate of the Andorian population changes with the number of observed Tellarians, N, under different mutation probabilities, p_m. Figure 11a,b show the results for two different prior probabilities, $\pi_1 = 0.5$ and $\pi_1 = 0.8$, respectively. Different curves represent different values of mutation probability, p_m, with $p_m = 0$ corresponding to the no-mutation case.

From the figures, we observe that as the mutation probability, p_m, increases, the expected growth rate tends to decrease. This is because as p_m grows, the probability of adverse events changes more frequently due to mutation, making it harder for Andorians to accurately estimate r^* and adjust their discrimination strategy effectively. At higher mutation rates, the population becomes more prone to errors in estimating the true adverse probability, resulting in fewer optimal decisions and lower growth rates.

When comparing Fig. 11 with Fig. 10 (that is, without mutation), a key difference emerges: in the infinite memory scenario, the growth rate converges to the optimal value as N increases. In contrast, when mutations are introduced, finite memory populations experience a trade-off between the benefits of observing more Tellarians (increased N) and the challenge of adapting to an environment where Tellarians can mutate. As seen in Fig. 11, at higher mutation rates ($p_m > 0$), as N increases, the growth rate may first increase then decrease, indicating that finite memory may be better than infinite memory.

Finite memory can outperform infinite memory in the presence of mutations. However, what is the optimal number of observations, or the optimal "memory size," that will maximize population growth in such environments? Figure 12 numerically illustrates the optimal number of observations, N, that maximizes the

expected growth rate of Andorians. We compare the optimal growth rate with N at ranges from 1 to 100. The color indicates the optimal value of N, where darker colors represent higher optimal values of N, and lighter colors represent smaller optimal values of N.

The figure shows that, when p_m and r_m are small, the optimal N is closer to 100, indicating that Andorians benefit from gathering a large number of observations to make more informed decisions. However, as p_m and r_m increase, the optimal number of observations decreases, suggesting that a finite memory is more effective in environments with higher mutation probabilities or more drastic changes in the adverse probability. Fewer observations are needed to optimize growth, as relying on too many observations may result in outdated or misleading information. In such cases, a finite memory becomes advantageous, allowing Andorians to quickly adapt to the changing environment.

Overall, our theoretical and numerical results emphasize that, in environments with frequent or substantial mutations, a finite memory (i.e., a smaller N) is preferable, as it allows populations to adapt more quickly, maximizing their growth potential.

Discussion

This article presents a comprehensive model that explores the evolution of discrimination in populations under both finite and infinite memory constraints, highlighting how rational maximizing decision-making processes may lead to discriminatory behavior. Our results demonstrate that discrimination is not merely a byproduct of bias or prejudice, but can arise as a response to environmental uncertainty and limited information. The core contribution of our work shows how finite memory introduces complexities into decision-making that sustain discriminatory practices over time, even when infinite memory might theoretically eliminate the need for reliance on prior biases.

We model discrimination as a strategy where individuals must decide whether to engage with others based on their group membership and observable traits. When individuals have infinite memory, they accumulate extensive knowledge over time, reducing their reliance on potentially biased observable characteristics. However, under finite memory constraints, the limited ability to recall past experiences leads individuals to depend on fewer past observations. This induces a tendency toward discriminatory behavior, especially when group traits serve as imperfect signals of risk.

An important extension of our model incorporates mutations, where the probability of adverse events among Tellarians may change. We find that while infinite memory theoretically provides individuals with the capacity to make fully informed decisions, it becomes a liability in environments where mutations occur. In such dynamic settings, finite memory allows for greater adaptability and flexibility, enabling populations to adjust more quickly to environmental changes. This result challenges the traditional assumption that more information is always better for decision-making. Instead, we show that limited memory might be an evolutionary advantage in maximizing long-term fitness, especially in environments where the probability of adverse events is variable.

While our model emphasizes finite memory as a cognitive constraint, the resulting behavior can also be interpreted more broadly as a response to limited information. We represent memory limitations by assuming that individuals base their judgments on a fixed number N of past observations—capturing the idea that decisions are made based on a restricted information set. While our analysis focuses on memory, similar behavioral patterns could arise from other informational bottlenecks, such as perceptual noise or attentional filtering. These alternative interpretations do not conflict with our core results, but rather suggest broader implications for how various forms of cognitive constraints may shape discriminatory behavior.

The results of this study have important implications for education, especially in terms of understanding how biases and discrimination develop. Our findings suggest that bias is not only a product of ignorance or prejudice, but can also be an adaptive response to environmental uncertainty and limitations in memory. This idea challenges traditional educational approaches that focus solely on sharing information or encouraging exposure to diverse viewpoints[54]. It suggests that more focus should be placed on teaching individuals how to navigate changing environments and improve their critical and flexible thinking skills, which can reduce reliance on biased decision-making strategies.

However, this presents a significant challenge in educational settings: how do we effectively teach the dangers of discrimination while also acknowledging the natural cognitive processes that may encourage it? Our comparison between infinite and finite memory highlights that such biases are not necessarily rooted in malice or ignorance, but may arise as adaptive responses to cognitive constraints. Memory limitations can make it difficult for individuals to fully process complex or fluctuating information, leading them to adopt shortcuts in decision-making. Therefore, it becomes crucial to foster not only awareness of bias but also the ability to adapt to uncertainty without relying on stereotypes or simplified judgments.

This introduces a nuanced perspective on education's role in combating bias. It is not enough to simply provide more information or advocate for diversity. Instead, educational strategies should focus on enhancing cognitive flexibility, helping people recognize when they are relying on biased heuristics, and offering tools to better adapt to dynamic, unpredictable environments. This more balanced approach may lead to better outcomes in reducing discriminatory behavior in the long run.

For policymakers, understanding the role of finite memory in shaping bias offers new insights into why social biases persist despite efforts to promote equality. Policy efforts often assume that increasing knowledge or awareness will reduce bias, but our model suggests that even with more information, finite memory can lead to continued reliance on group-based traits when making decisions under uncertainty. This indicates that policies need to go beyond awareness campaigns, focusing instead on fostering adaptability and reducing the cognitive strain on decision-makers. Structural changes in institutions that guide behavior away from biased shortcuts may be needed to make lasting progress.

Our findings also have significant implications for the development of artificial intelligence (AI) systems, particularly in how they handle memory and adapt to uncertainty. AI systems, like human decision-makers, are limited by their ability to process and store information. Mechanisms such as Long Short-Term Memory (LSTM) networks[55] and attention mechanisms[56] have been introduced to help AI models retain and prioritize relevant information over time. These mechanisms are crucial in adapting to dynamic environments and minimizing biased decisions. Instead of merely increasing the volume of data, our results suggest the need for AI systems to balance memory constraints with adaptability, ensuring they can respond to changing conditions without reinforcing harmful biases. Designing algorithms that leverage memory mechanisms while prioritizing fairness will be essential for reducing discriminatory outcomes in AI decision-making.

While our model provides a tractable framework for analyzing the emergence of discrimination under cognitive constraints, it also has several limitations. First, our results rely on the log-geometric-average growth rate of the population, Eq. (1), as the evolutionary objective. Although this has been shown to produce long-run dominant strategies[16–18,21], it is not the only plausible evolutionary criterion. Alternative fitness concepts—such as maximizing the expected number of offspring, focusing on relative success, or including kin-selected strategies[23,57–59]—may lead to different behavioral dynamics.

Second, we represent finite memory as a fixed number N of IID observations. Although analytically convenient, this abstraction does not capture the full richness of real-world memory processes, such as the decay of older memories, the asymmetrical recall of negative versus positive events, or individual-level variation in the retention and weighting of past experiences.

Third, we assume that prior beliefs (π_0, π_1) are fixed and exogenously given. In practice, however, such priors may be shaped and reshaped by dynamic social influences—such as media exposure, education, or cultural narratives—which are not captured in our static Bayesian formulation.

Fourth, while we adopt a Bayesian framework, actual decision-making in animal and human minds may be a mix of different cognitive strategies. Alternatives include value-based heuristics such as prospect theory[60], cue-based decision rules[61], and reinforcement learning through trial and error[62].

Future work could extend the model along several directions. These include exploring alternative evolutionary objectives, incorporating more sophisticated memory dynamics, endogenizing belief formation, and comparing the fitness of Bayesian and non-Bayesian strategies under different conditions. Empirical validation—through behavioral experiments or observational data—would also help assess the extent to which our theoretical predictions align with real-world discrimination and belief formation.

This work contributes to the growing literature on evolutionary dynamics and decision-making under uncertainty. By demonstrating how memory constraints shape discriminatory behavior, we offer a new perspective on how discrimination can be a rational, albeit suboptimal, strategy in certain contexts. Overall, our study highlights the importance of considering both memory constraints and environmental variability when analyzing discrimination. These factors are critical in understanding the persistence of bias, inequality, and discrimination, and they provide a framework for developing more effective strategies to mitigate the harmful effects of discrimination in society.

Data availability

No datasets were generated or analyzed during the current study.

References

1. Becker, G. S. *The Economics of Discrimination* (University of Chicago Press, 1957).
2. Phelps, E. S. The statistical theory of racism and sexism. *Am. Econ. Rev.* **62**, 659–661 (1972).
3. Arrow, K. J. The theory of discrimination. In Ashenfelter, O. & Rees, A. (eds.) *Discrimination in Labor Markets*, 3–33 (Princeton University Press, Princeton, NJ, 1973).
4. Coate, S. & Loury, G. C. Will affirmative-action policies eliminate negative stereotypes? *Am. Econ. Rev.* 1220–1240 (1993).
5. Schneider, D. J. *The Psychology of Stereotyping* (Guilford Press, 2005).
6. Bordalo, P., Coffman, K., Gennaioli, N. & Shleifer, A. Stereotypes. *Q. J. Econ.* **131**, 1753–1794 (2016).
7. Fuligni, A. J. *Contesting Stereotypes and Creating Identities: Social Categories, Social Identities, and Educational Participation* (Russell Sage Foundation, 2007).
8. Gorski, P. The myth of the "culture of poverty". *Educ. Leadersh.* **65**, 32 (2008).
9. Bohren, J. A., Haggag, K., Imas, A. & Pope, D. G. Inaccurate statistical discrimination: An identification problem. *Rev. Econ. Stat.* 1–45 (2023).
10. Brocas, I. & Carrillo, J. D. The value of information when preferences are dynamically inconsistent. *Eur. Econ. Rev.* **44**, 1104–1115 (2000).
11. Carrillo, J. D. & Mariotti, T. Strategic ignorance as a self-disciplining device. *Rev. Econ. Stud.* **67**, 529–544 (2000).
12. Bénabou, R. & Tirole, J. Self-confidence and personal motivation. *Q. J. Econ.* **117**, 871–915 (2002).
13. Compte, O. & Postlewaite, A. Confidence-enhanced performance. *Am. Econ. Rev.* **94**, 1536–1557 (2004).
14. Charness, G., Rustichini, A. & Van de Ven, J. Self-confidence and strategic behavior. *Exp. Econ.* **21**, 72–98 (2018).
15. Heller, Y. & Winter, E. Biased-belief equilibrium. *Am. Econ. J. Microecon.* **12**, 1–40 (2020).
16. Brennan, T. J. & Lo, A. W. The origin of behavior. *Q. J. Finan.* **1**, 55–108 (2011).
17. Zhang, R., Brennan, T. J. & Lo, A. W. Group selection as behavioral adaptation to systematic risk. *PLoS ONE* **9**, e110848 (2014).
18. Lo, A. W. & Zhang, R. The wisdom of crowds versus the madness of mobs: An evolutionary model of bias, polarization, and other challenges to collective intelligence. *Collect. Intell.* **1**, 26339137221104784 (2022).
19. Lo, A. W. & Zhang, R. *The Adaptive Markets Hypothesis: An Evolutionary Approach to Understanding Financial System Dynamics* (Oxford University Press, 2024).
20. Lo, A. W., Marlowe, K. P. & Zhang, R. To maximize or randomize? An experimental study of probability matching in financial decision making. *PLoS ONE* **16**, e0252540 (2021).
21. Zhang, R., Brennan, T. J. & Lo, A. W. The origin of risk aversion. *Proc. Natl. Acad. Sci.* **111**, 17777–17782 (2014).

22. Seger, J. & Brockman, H. J. What is bet-hedging?. *Oxford Surverys in Evolutionary Biology* **4**, 182–211 (1987).
23. Frank, S. A. Natural selection. I. Variable environments and uncertain returns on investment. *J. Evol. Biol.* **24**, 2299–2309 (2011).
24. Frank, S. A. Natural selection. II. Developmental variability and evolutionary rate. *J. Evol. Biol.* **24**, 2310–2320 (2011).
25. Frank, S. A. Natural selection. III. Selection versus transmission and the levels of selection. *J. Evol. Biol.* **25**, 227–243 (2012).
26. Cooper, W. S. & Kaplan, R. H. Adaptive "coin-flipping": a decision-theoretic examination of natural selection for random individual variation. *J. Theor. Biol.* **94**, 135–151 (1982).
27. Hawkins, J. & Blakeslee, S. *On Intelligence* (Macmillan, 2004).
28. Lo, A. W. & Zhang, R. The evolutionary origin of Bayesian heuristics and finite memory. *iScience* **24** (2021).
29. Goodman, N. D., Tenenbaum, J. B., Feldman, J. & Griffiths, T. L. A rational analysis of rule-based concept learning. *Cogn. Sci.* **32**, 108–154 (2008).
30. Vul, E., Goodman, N., Griffiths, T. L. & Tenenbaum, J. B. One and done? optimal decisions from very few samples. *Cogn. Sci.* **38**, 599–637 (2014).
31. Griffiths, T. L. & Tenenbaum, J. B. Optimal predictions in everyday cognition. *Psychol. Sci.* **17**, 767–773 (2006).
32. Sanborn, A. & Griffiths, T. Markov chain Monte Carlo with people. In Platt, J., Koller, D., Singer, Y. & Roweis, S. (eds.) *Advances in Neural Information Processing Systems*, vol. 20 (Curran Associates, Inc., 2007).
33. Vul, E. & Pashler, H. Measuring the crowd within: Probabilistic representations within individuals. *Psychol. Sci.* **19**, 645–647 (2008).
34. Griffiths, T. L., Kemp, C. & Tenenbaum, J. B. *Bayesian Models of Cognition*, 59–100. Cambridge Handbooks in Psychology (Cambridge University Press, 2008).
35. Courville, A. C., Daw, N. D. & Touretzky, D. S. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* **10**, 294–300 (2006).
36. Yuille, A. & Kersten, D. Vision as Bayesian inference: Analysis by synthesis?. *Trends Cogn. Sci.* **10**, 301–308 (2006).
37. Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. Human-level concept learning through probabilistic program induction. *Science* **350**, 1332–1338 (2015).
38. Körding, K. P. & Wolpert, D. M. Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* **10**, 319–326 (2006).
39. Steyvers, M., Griffiths, T. L. & Dennis, S. Probabilistic inference in human semantic memory. *Trends Cogn. Sci.* **10**, 327–334 (2006).
40. Oaksford, M. & Chater, N. The probabilistic approach to human reasoning. *Trends Cogn. Sci.* **5**, 349–357 (2001).
41. Baker, C. L., Tenenbaum, J. B. & Saxe, R. R. Goal inference as inverse planning. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, no. 29 in 29 (2007).
42. Greaves, H. & Wallace, D. Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind* **115**, 607–632 (2006).
43. Leitgeb, H. & Pettigrew, R. An objective justification of Bayesianism I: Measuring inaccuracy. *Philos. Sci.* **77**, 201–235 (2010).
44. Leitgeb, H. & Pettigrew, R. An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philos. Sci.* **77**, 236–272 (2010).
45. Okasha, S. The evolution of Bayesian updating. *Philos. Sci.* **80**, 745–757 (2013).
46. Castellano, S. Bayes' rule and bias roles in the evolution of decision making. *Behav. Ecol.* **26**, 282–292 (2015).
47. Campbell, J. O. Universal Darwinism as a process of Bayesian inference. *Front. Syst. Neurosci.* **10**, 49 (2016).
48. Bordalo, P., Coffman, K., Gennaioli, N. & Shleifer, A. Stereotypes. *Q. J. Econ.* **131**, 1753–1794 (2016).
49. Tversky, A. & Kahneman, D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol. Rev.* **90**, 293 (1983).
50. Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science* **185**, 1124–1131 (1974).
51. Petrov, V. V. *Sums of Independent Random Variables*, vol. 82 (Springer Science & Business Media, 2012).
52. Boyle, P. P. & Lau, S. H. Bumping up against the barrier with the binomial method. *J. Deriv.* **1**, 6–14 (1994).
53. Brennan, T. J., Lo, A. W. & Zhang, R. Variety is the spice of life: Irrational behavior as adaptation to stochastic environments. *Q. J. Financ.* **8**, 1850009 (2018).
54. Anderson, E. *The Imperative of Integration* (Princeton University Press, 2010).
55. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
56. Vaswani, A. et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, vol. 30 (2017).
57. Frank, S. A. When to copy or avoid an opponent's strategy. *J. Theor. Biol.* **145**, 41–46 (1990).
58. Lo, A. W., Orr, H. A. & Zhang, R. The growth of relative wealth and the Kelly criterion. *J. Bioecon.* **20**, 49–67 (2018).
59. McNamara, J. M. Implicit frequency dependence and kin selection in fluctuating environments. *Evol. Ecol.* **9**, 185–203 (1995).
60. Kahneman, D. & Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **47**, 263–292 (1979).
61. Gigerenzer, G., Todd, P. M., ABC Research Group. *Simple Heuristics That Make Us Smart* (Oxford University Press, 2000).
62. Dayan, P. & Niv, Y. Reinforcement learning: The good, the bad and the ugly. *Curr. Opin. Neurobiol.* **18**, 185–196 (2008).

Author contributions
Andrew W. Lo and Ruixun Zhang conceived the idea, designed the study, and revised the manuscript. Chaoyi Zhao conducted the study, proposed and proved the theoretical results, and wrote the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-17089-9.

Correspondence and requests for materials should be addressed to C.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.