bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Outsourcing Memory Through Niche Construction

Edward D. Lee[a], **Jessica C. Flack**[b], **and David C. Krakauer**[b]

[a]Complexity Science Hub Vienna, Josefstædter Strasse 39, Vienna, Austria; [b]Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501

This manuscript was compiled on January 7, 2023

Adaptation to changing environments is a universal feature of life and can involve the organism modifying itself in response to the environment as well as actively modifying the environment to control selection pressures. The latter case couples the organism to environment. Then, how quickly should the organism change in response to the environment? We formulate this question in terms of how memory duration scales with environmental rate of change when there are trade-offs in remembering vs. forgetting. We derive a universal scaling law for optimal memory duration, taking into account memory precision as well as two components of environmental volatility, bias and stability. We find sublinear scaling with any amount of environmental volatility. We use a memory complexity measure to explore the strategic conditions (game dynamics) favoring actively reducing environmental volatility—outsourcing memory through niche construction—over investing in neural tissue. We predict stabilizing niche construction will evolve when neural tissue is costly, the environment is variable, and it is beneficial to be able to encode a rich repertoire of environmental states.

adaptation | learning | stigmergy | niche construction | scaling

What is the optimal timescale of adaptation—how long should memory of the environment persist when the environment is changing? And when should the organism invest in changing the rate of environmental change? Research in a wide range of fields suggests that bidirectional organism-environment feedback through niche construction and symbiosis is common and plays a significant role in shaping evolutionary dynamics. Slowly evolving genes co-evolve with quickly evolving culture (1), as illustrated by the evolution of dairy-farming facilitating selection of alleles for adult lactase persistence (2). Quickly evolving organisms modify their otherwise slowly changing niches and alter selection pressures (3–5), illustrated by yeast modifying fruit environments to attract Drosophilid flies that enhance yeast propagation (6). Institutions feed back to influence individual decisions by changing cost landscapes and enhancing cultural transmission (7, 8) (e.g. legislation in support of same-sex marriage that increases the willingness to voice support in the face of risk (9)). To gain information about noisy, hidden variables and reduce social uncertainty, error-prone individual pigtailed macaques collectively compute a social power structure that reduces uncertainty about the cost of social interaction, making accessible new forms of conflict management (reviewed in references 10, 11). Bacteria quorum sense, controlling group behavior in dynamically complex, changing environments (reviewed in reference 12). Individuals, institutions, and firms all adapt to audit targets (Goodhart's Law), creating new feedbacks as they attempt to game the system (13–16). In order to undermine competitors, agents can destabilize a system like in the recent Reddit-Gamestop event in which powerful hedge funds are thought to have introduced volatility to markets by manipulating Reddit users to short squeeze yet other hedge funds (17). Motivated by these examples, we develop a synthetic framework that combines information theory, game dynamics, and scaling theory, in order to determine how adaptation scales in a range of plausible strategic settings including niche construction.

We start by reformulating adaptation as rate of discounting of the past, building a conceptual and mathematical bridge to work on memory (18–21). We take into account four factors: *bias* as preference in the environment for a particular state, *stability* as the rate at which environment fluctuates (22, 23), *precision* as the capacity agents have to resolve environmental signal, and *feedback* as the rate of agent modification of the environment. In Table 1, we provide examples of studies addressing the interaction of bias, stability, and precision. We also drop the separation of timescales assumption commonly made in modeling papers and explicitly consider feedback. We allow modification of the environment to be either passive or active, such that active modification can be destabilizing (increasing entropy) as well as stabilizing (reducing entropy). The Reddit-Gamestop event is one example of this "destabilizing" niche construction. Another is guerrilla warfare in which a weaker party randomly determines which battles to neglect by allocating zero resources (24). In contrast, active agents can stabilize the environment by buffering against variation (5) or slowing its rate of change to reduce uncertainty about the future (10, 25). A relatively simple example is stigmergy in which trails or routes are consolidated through repeated use (26). More complicated examples include the collective computation of slowly changing power structures in macaque groups (27) and foraging subgroup size distributions by spider monkeys (28) in which social structures are computed through communication and decision networks. Finally, we take into account how the precision (29) of an agent's or organism's

Significance Statement

All organisms must adapt to changing environments, but adaptation can modify the environment itself. We solve a version of this problem in terms of how long organisms remember. Shorter memory should be better for variable environments and longer for slow changing ones, but environmental variability depends on feedback. Surprisingly, we find the same mathematical law in both cases, revealing how much shorter memory should be relative to the environmental timescale. We consider how this depends on memory complexity and metabolic costs in populations, allowing us to predict a general set of conditions for when organism will outsource memory to the environment: when maintaining a brain is costly, the environment fluctuates quickly, and organisms inhabit a complex environment.

All authors helped develop the initial idea. E.D.L. did the analysis, modeling, and wrote the code. The authors drafted the manuscript jointly.

The authors declare no competing interests.

[2]To whom correspondence should be addressed. E-mail: edlee@csh.ac.at

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

I Bias-Stability	II Stability-Precision	III Bias-Precision	IV Integrated
Taxis of larval invertebrates (30)	Seed dormancy/germ banking (31)	Bandit problems (32)	Volatile bandits (33)
Stochastic voting models (34)	Particle swarms (35)	Microbial chemotaxis (36)	Learning changing data sources (37)
Learning changing distributions (38)	Cognitive aging (39)	Speed-accuracy trade-offs (40–42)	Consensus with link failure (43)
Loss/Change aversion (44)		Optimal foraging (45)	Retinal sensitivity rescaling (19)
		Page Rank consensus (46)	

Table 1. Classification of studies in terms of bias, stability, and precision. We group studies according to the pairs of factors they combine (I) bias-stability, (II) bias-precision, and (III) stability-precision. Studies that implicitly combine all of these factors are noted under "integrated" (IV). Studies in category I focus on rules that apply in variable environments (bias) where environmental distributions are prone to rapid change (stability). Studies in category II focus on rules that apply when environments are likely to change (stability) and where making the correct decision depends on sensitivity to signals (precision or also "accuracy" in some literature). Studies in category III focus on rules that apply in variable environments (bias) and where making the correct decision depends on power of sensors to detect signal (precision). Studies in category IV include elements of I–III and apply to variable environments prone to rapid change, where sensory precision varies. Feedback with the environment, where agent inference modifies the input statistics and timescales are formally coupled, remains little considered despite being a central premise of research on niche construction and stigmergy.

estimates of environmental state influences its ability to fit the environment at a given degree of volatility.

In "Result 1," we explore the conditions under which long memory is beneficial. In "Result 2," we derive the scaling relationship for optimal memory duration and environmental change. In "Result 3," we derive by way of a back-of-the-envelope calculation the costs of memory using the literature on metabolic scaling. In "Result 4," we introduce game dynamics introducing a complexity cost of memory to explore the evolution of active modification and outsourcing of memory to the environment.

Model structure & assumptions

We summarize the structure of our model in Figure 1, which combines the essential components of adaptive agents. As a result, it connects passive agents that learn the statistics of a fluctuating environment with those that modify the environment itself. We summarize notation in Appendix Table S2.

The environment E at time t is described by a probability distribution $p_E(s,t)$ over configurations s, a vector of descriptive properties. The environment has a bias for preferred states that changes on a relatively slow timescale. Here, we represent the state of the environment with a single bit $s \in \{-1, 1\}$, analogous to the location of a resource as a choice between left and right (41, 47–49). In one configuration, the distribution of resources p_E is biased to the left at a given time t, or $p_E(s=-1,t) > p_E(s=1,t)$, such that an agent with matching preference would do better on average than an agent with misaligned preference. In the mirrored configuration, the environment shows a bias of equal magnitude to the right $p_E(s=-1,t) < p_E(s=1,t)$. Such probabilistic bias can be represented as an evolving "field" $h_E(t)$,

$$p_E(s,t) = \frac{1}{2} + \frac{s}{2}\tanh h_E(t), \qquad [1]$$

such that reversal in bias corresponds to flip of sign $h_E(t) \to -h_E(t)$ that naturally embodies a symmetry between left and right. At every time point, the environment has clearly defined bias in one direction or another, determined by setting the external field to either $h_E(t) = -h_0$ or $h_E(t) = h_0$. With probability $1/\tau_E$ per unit time, the bias in the environment reverses such that over time τ_E the environment remains correlated with its past. When τ_E is large, we have long correlation times and a slow environment or a "slow variable." This formulation yields a stochastic environment whose uncertainty depends on both fluctuation rate, such that low rate implies high stability, and the strength of bias for a particular state, such that a strong bias yields a clear environmental signal.

Passive agents sample from the environment and choose a binary action. In principle, the precision of the choice is dependent on the number of sensory cells contributing to the estimate of environmental state, the sensitivity of those cells, and the number of samples each cell collects while the contribution of each factor to the estimate can differ. In our model, all the alternatives are captured by τ_c. When τ_c is high (either because the sensory cells sampled from the environment for a long time, many sensory cells contributed estimates, or each sensory cell is very sensitive) agents obtain exact measurements of the environment. A small τ_c corresponds to noisy estimates. The resulting estimate of environmental state $\hat{p}$ thus incurs an error ϵ_{τ_c},

$$\hat{p}(s,t) = p_E(s,t) + \epsilon_{\tau_c}(t). \qquad [2]$$

From this noisy signal, sensory cells obtain an estimate of bias $\hat{h}(t)$, which is related to environmental bias $h_E(t)$ plus measurement noise $\eta_{\tau_c}(t)$,

$$\hat{h}(t) = h_E(t) + \eta_{\tau_c}(t). \qquad [3]$$

In the limit of large precision τ_c and given that the noise in the estimated probabilities $\epsilon_{\tau_c}(t)$ from Eq 2 is binomial distributed, the corresponding error in field $\eta_{\tau_c}(t)$ converges to a Gaussian distribution (see Materials and Methods). Then, at each time step the agent's measurement of the environment includes finite-sample noise which is inversely related to precision.

An aggregation algorithm determines how much to prioritize the current measurement over historical ones. This gives the duration of memory by recording the agent's estimate of the state of the environment at the current moment in time $h(t)$ and feeding it to sensory cells at time $t+1$ with some linear weighting $0 \le \beta \le 1$ (50),

$$h(t+1) = (1-\beta)\hat{h}(t+1) + \beta h(t). \qquad [4]$$

This estimate is stored in an "aggregator" A_t, and we define $h(0) = 0$. The weight β determines how quickly the previous state of the system is forgotten such that when $\beta = 0$ the agent is constantly learning the new input and has no memory and when $\beta = 1$ the agent ceases to learn preserving its initial

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

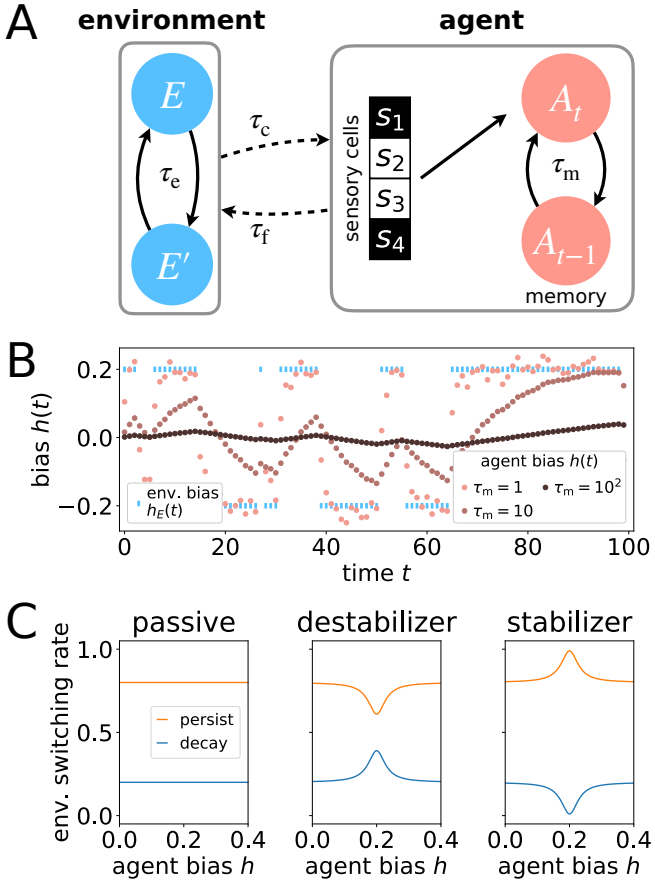

Fig. 1. (A) Overview of framework. Environment E switches configuration on timescale τ_E. The agent measures the current environment through sensory cells with precision τ_c, here worth 4 bits. To obtain an estimate of environment statistics at time t, the agent A_t combines present sensory estimates with memory of previous estimates recorded in an aggregator A_{t-1} (Eq 4) such that memory decays over time τ_m (Eq 5). Coupling with the environment speeds up or slows down environmental change on timescale τ_f (Eq 6). (B) Example trajectories of agents adapting to environmental state $h_E(t)$ with short, medium, and long memory. (C) Rate of environment switching per time step as a function of agent bias h relative to environmental bias $h_E = 0.2$. For passive agents, switching rate does not depend on agent bias. For destabilizers $\alpha = 0.95$, for stabilizers $\alpha = -0.95$. For both, $v = 0.1$ from Eq 6 and environmental timescale $\tau_E = 5$.

state. In between, agent memory decays exponentially with lifetime

$$\tau_m \equiv -1/\log\beta. \qquad [5]$$

We think of the weight β that the aggregation algorithm places on the current estimate relative to the stored value as the timescale of adaptation τ_m, or agent memory duration.

The output of this computation is the agent's behavior, $p(s,t)$. We measure the effectiveness of adaptation, or fit to the environment, with the divergence between a probability vector describing an agent and that of the environment. Measures of divergence, like Kullback-Leibler (KL) divergence, and, more generally, mutual information, have been shown to be natural measures of goodness of fit in evolutionary and learning dynamics from reinforcement learning through to Bayesian inference (51, 52).

Here we extend the model to include feedback by allowing agents to alter environmental stability, which is operationalized as the probability of switching. We add to the switching rate $1/\tau_E$, the *active* construction rate,

$$\frac{1}{\tau_f(t)} \equiv \frac{v^2/\tau_E}{[h(t)-h_E(t)]^2+v^2}, \qquad [6]$$

such that the probability q that the environment changes at the next point in time is

$$q[h_E(t+1) \neq h_E(t)] = 1/\tau_E + \alpha/\tau_f(t). \qquad [7]$$

Eq 6 is written so that it remains normalized for arbitrary v and that the rate gets smaller as the squared distance between agent bias and environmental bias $[h(t)-h_E(t)]^2$ goes to zero. The probability q of the environment switching to the opposite configuration includes weight $\alpha \in (0,1]$ to tune the strength of destabilizers, or $\alpha \in [-1,0)$ for stabilizers. This means that for positive α, the rate of switching increases as the agent matches the environment more closely and the opposite for negative α, whereas the parameter v controls how closely the agent must match the environment to have an effect (i.e. the width of the peak as plotted in Figure 1C). The two types of active agents capture two ways adaptive behavior can feedforward to influence the timescale environmental change.* We note that when $1/\tau_f = 0$, we obtain *passive* agents that do not modify their environment, thus connecting passive and active agents to one another along a continuum scale.

Putting these elements of adaptation together, as shown in Figure 1A, we obtain a toy learning agent that infers the statistics of a time-varying and stochastic environment.

Result 1: Long memory and adaptation favored when sensory cells are imprecise & environments are slow

The timescale of adaptation represents a balance between the trade-offs of preserving an internal state for too long or losing it too fast. We explore this trade-off by calculating an agent's fit to a changing environment. The fit can be quantified with the KL divergence between environment $p_E(s,t)$ with bias $h_E(t)$ and agent $p(s,t)$,

$$D_{KL}[p_E||p](t) = \sum_{s\in\{-1,1\}} p_E(s,t)\log_2\left(\frac{p_E(s,t)}{p(s,t)}\right). \qquad [8]$$

When the KL divergence is $D_{KL} = 0$, the agents use optimal bet-hedging, known as "proportional betting," which is important for population growth dynamics (53, 54). Eq 8 is also minimized for Bayesian learners under optimal encoding (55). Assuming agents are playing a set of games in which they must guess the state of the environment at each time step, Eq 8 is the information penalty paid by imperfect compared to perfect agents.

After averaging over many environmental bias switches, we obtain the agent's typical *divergence*,

$$\bar{D} \equiv \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} D_{KL}[p_E||p](t), \qquad [9]$$

The bar notation signals an average over time. Thus, fit improves as $\bar{D}$ decreases.

*Note that in this binary example, the new environmental configuration when switching is unique, enforcing a deterministic switch, but in general there may be a large number of $K \gg 1$ options such that the agent cannot easily guess at the results of environmental fluctuations.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

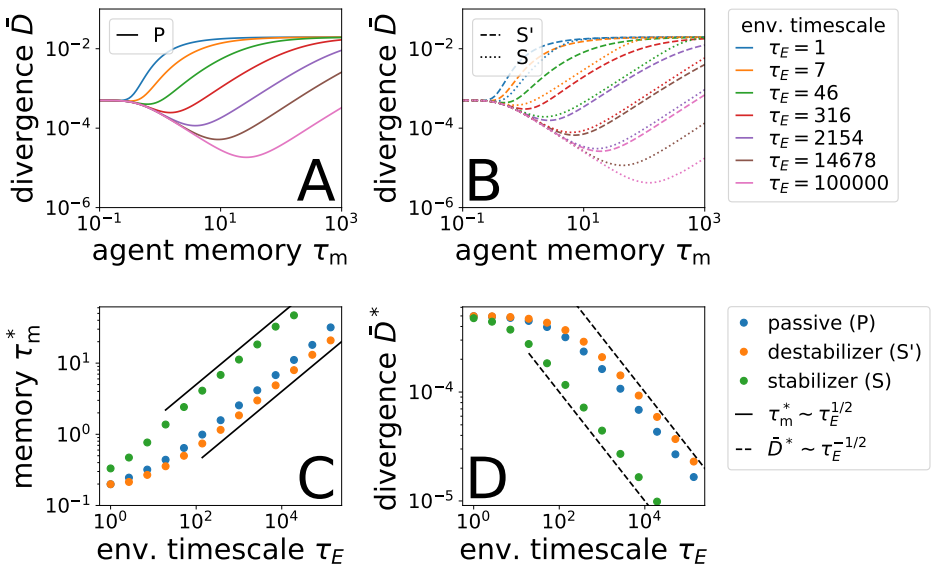

Fig. 2. Divergence $\bar{D}$ as a function of agent memory τ_m and environmental timescale τ_E for (A) passive and (B) active agents including destabilizers S' and stabilizers S. For longer τ_E, agents with longer memory always do better, a pattern emphasized for stabilizers and diminished for destabilizers. (C) Scaling of optimal memory duration τ_m^* with environmental timescale τ_E, corresponding to minima from panels A and B. (D) Divergence at optimal memory duration $\bar{D}^* \equiv \bar{D}(\tau_\mathrm{m}^*)$. Environmental bias $h_0 = 0.2$.

In Figures 2A and B, we show divergence $\bar{D}(\tau_\mathrm{m}, \tau_E)$ as a function of the agent's memory τ_m given environmental timescale τ_E. In the limiting cases in which an agent has either no memory and is constantly adapting or has infinite memory and adaptation is absent, the timescale on which environmental bias changes ultimately has no effect—we observe convergence across all degrees of bias and stability. When an agent has no memory, or $\tau_\mathrm{m} = 0$, an agent's ability to match the environment is solely determined by its sensory cells. Low precision τ_c leads to large errors on measured environmental bias $h_E(t)$ and large divergence $\bar{D}(\tau_\mathrm{m} = 0)$. On the other hand, high precision τ_c increases performance and depresses the intercept (Eq 23). At the right hand side of Figure 2A, for large $\tau_\mathrm{m} \gg 1$, behavior does not budge from its initial state. Assuming that we start with an unbiased agent such that the transition probability is centered as $q(h) = \delta(h)$, the Dirac delta function, the agent's field is forever fixed at $h = 0$. Then, divergence $\bar{D}(\tau_\mathrm{m} = \infty)$ reduces to a fixed value that only depends on environmental bias (Eq 24). In between the two limits of zero and infinite agent memory, the model produces a minimum divergence $\bar{D}(\tau_\mathrm{m} = \tau_\mathrm{m}^*)$. This indicates the optimal duration of memory τ_m^* for a given degree of environmental bias and stability.

The benefits of memory are more substantial for agents with imprecise sensory cells. This benefit is the difference $\bar{D}(\tau_\mathrm{m} = 0) - \bar{D}(\tau_\mathrm{m} = \tau_\mathrm{m}^*)$ as shown in Figure 3A. As one might expect, integrating over longer periods of time provides more of a benefit when the present estimate $\hat{p}$ is noisy, ${\tau_\mathrm{c}}^{-1}$ is large, and sensory cells are not particularly precise, a deficiency in precision that memory counters by allowing organisms to accumulate information over time. This intuition, however, only applies in the limit of large environmental bias h_0 where the contours of optimal memory flatten and become orthogonal to precision ${\tau_\mathrm{c}}^{-1}$. When the bias in the environment is weak, the curved contours show that the benefits of memory come to depend strongly on nontrivial interaction of precision and environmental bias. The complementary plot is the benefit from forgetting, $\bar{D}(\tau_\mathrm{m} = \infty) - \bar{D}(\tau_\mathrm{m} = \tau_\mathrm{m}^*)$ in Figure 3 B, which is largely determined by bias h_0. When bias is strong, the costs of estimating the environment inaccurately are large, and it becomes important to forget if sensory cells are imprecise. Thus, our model encapsulates the trade-off between remembering and forgetting both in terms of their absolute benefits as well as the emergence of simple dependence of the respective benefits in the limits of high environmental bias and high sensory precision. An agent has optimally tuned its timescale of adaptation $\tau_\mathrm{m} = \tau_\mathrm{m}^*$ when it has balanced the implicit costs of tuning to fluctuations against the benefits of fitting bias correctly.

Result 2: Adaptation and environmental change scale sublinearly

For sufficiently slow environments, or sufficiently large τ_E, we find that optimal memory duration τ_m^* scales with the environmental timescale τ_E sublinearly as in Figure 2C. To derive the scaling between optimal memory and environmental timescale, we consider the limit when agent memory persistence is small relative to the environmental persistence $\tau_\mathrm{m} \ll \tau_E$. Under this condition, optimal memory represents a trade-off between a poor fit lasting time τ_m and a good fit for time $\tau_E - \tau_\mathrm{m}$. During the poor fit, the agent pays a typical cost at every single time step such that the cost grows linearly with its duration, $\mathcal{C}\tau_\mathrm{m}$, for constant $\mathcal{C}$. When the environment is stable, agent precision is enhanced by a factor of τ_m because it effectively averages over many random samples, or a gain of $\mathcal{G} \log \tau_\mathrm{m}$ for constant $\mathcal{G}$. When we weight each term by the fraction of time spent in either transient or stable phases, τ_m/τ_E and $(\tau_E - \tau_\mathrm{m})/\tau_E$ respectively, we obtain the trade-off

$$\mathcal{C}\frac{\tau_\mathrm{m}^2}{\tau_E} - \mathcal{G}\frac{\tau_E - \tau_\mathrm{m}}{\tau_E} \log \tau_\mathrm{m}. \qquad [10]$$

At optimal memory τ_m^*, Eq 10 will have zero derivative. Keeping only the dominant terms and balancing the resulting equa-

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

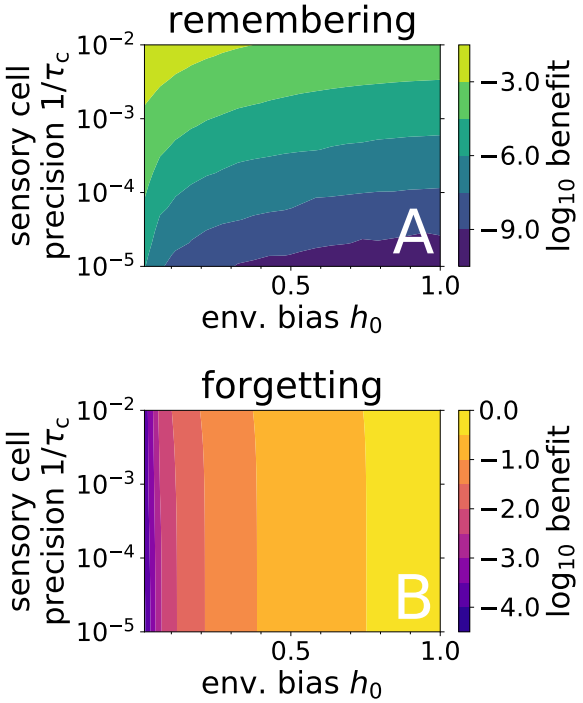

Fig. 3. Benefit from (A) remembering and from (B) forgetting defined as the reduction in divergence at optimal memory duration relative to no memory, $\bar{D}(\tau_m = 0) - \bar{D}(\tau_m^*)$, and optimal memory duration to infinite memory, $\bar{D}(\tau_m = \infty) - \bar{D}(\tau_m^*)$, respectively. We show passive agents given environmental timescale $\tau_E = 10$. All contours must converge (A) when $h_0 = 0$ and (B) when $\tau_c = 0$. Agent-based simulation parameters specified in accompanying code.

tion, we find

$$\tau_m^* \sim \tau_E^{1/2}. \qquad [11]$$

This scaling argument aligns with numerical calculation as shown in Figure 2C.

Similarly, we calculate how optimal divergence $\bar{D}^*$ scales with environmental timescale. Assuming that the agent has a good estimate of the environment such that the error in average configuration $\epsilon_{\tau_c}(t)$ is small, agent behavior is $p_E(s,t) + \epsilon_{\tau_c}(t)$ and $\epsilon_{\tau_c}(t)$ is normally distributed. Then, we expand the divergence about $p_E(s,t)$ in Taylor series of error $\epsilon_{\tau_c}(t)$ (Materials & Methods). Over a timescale of τ_m^*, the precision of this estimate is further narrowed by a factor of τ_m^* such that

$$\bar{D}^* \sim 1/\tau_m^* \sim \tau_E^{-1/2}. \qquad [12]$$

Although we do not account for the transient phase, we expect the relation in Eq 12 to dominate in the limit of large τ_E, and our numerical calculations indeed approach the predicted scaling in Figure 2C. In contrast, when environment does not fluctuate, or bias $h_0 = 0$, agents pay no cost for failing to adapt to new environments and infinite memory is optimal. Overall, the sublinear scaling between memory duration and rate of environmental change indicates an economy of scale. Agents require proportionally less expenditure on adaptation in slow environments than would be true under a linear relationship. Hence a slow environment is in this sense highly favorable to an adaptive agent when considering the costs of poor adaptation.

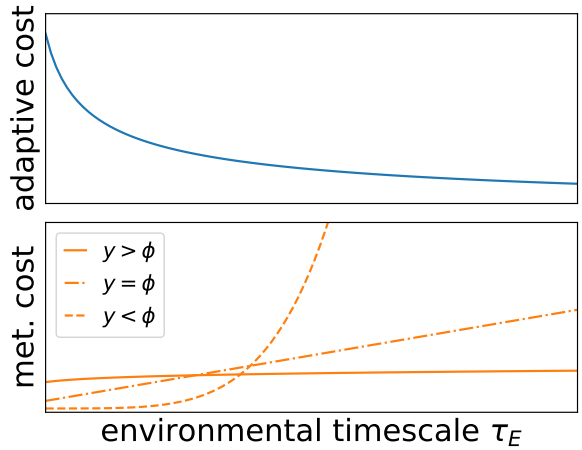

Fig. 4. Scaling of adaptive and metabolic costs with environmental timescale τ_E. (A) Adaptive cost $\bar{D}$ is largest at small τ_E, but (B) metabolic costs are largest for longer-lived organisms with scaling dependent on exponents y and ϕ such as for "elephants" that experience slower environments (Eq 14).

Result 3: Metabolic cost of memory can become prohibitive in slow environments

Here we ask how memory might become limited by the metabolic costs of neural tissue.

We start with the well-documented observation that physical constraints on circulatory networks responsible for energy distribution influence organismal traits including lifespan and size across the animal kingdom from microorganisms to blue whales (56, 57). Metabolic costs matter for brain mass M_{br}, which scales with body mass M_{bo} sublinearly, $M_{br} = AM_{bo}^a$, where $a = 3/4$ across taxa (within individual taxa it spans the range 0.24 to 0.81 (58)). To account for memory cost, we make the simple assumption that the quantity of brain mass required for memory is proportional to the number and duration of environmental states (the "environmental burden") the organism encounters,

$$M_{br} \propto N\tau_E. \qquad [13]$$

After all, we say, "an elephant never forgets" and not the same of a mouse.

Now, we use predictions of allometric scaling theory to relate metabolic rate B to mass, $B \propto M^{1/4}$ (59), and lifespan to body mass, $T \propto M_{bo}^b$ for metabolic exponent $b = 1/3$ (60). From Eq 13, we obtain a relationship between metabolic rate and memory burden, $B \propto N^\phi \tau_E^\phi$, where $\phi \equiv a/4b$.[†] Note that this scaling is sublinear for biological organisms, $\phi < 1$. Although the adaptive cost decays with τ_E in Figure 4A, metabolism grows as τ_E^ϕ as shown in Figure 4B. The competing scalings suggest that for small organisms the cost of adaptation will make a disproportionate contribution to the lifetime energy budget of an organism. This is consistent with observations on developmental neural growth in butterflies (61).[‡]

[†] When we use $a = 3/4$, we obtain the range $\phi = [5/8, 15/16]$, the endpoints depending on whether $b = 0.3$ or $b = 0.2$, respectively, while accounting for taxa-specific variation in a leads to much wider range of $\phi \in [0.2, 1.01]$. Thus, we hypothesize that longer environmental timescales lead to increased brain mass and metabolic expenditure with sublinear scaling.

[‡] As noted in the cited study and its citations, experience leads to larger brain size, indicating that learning from such experience is sufficiently valuable to warrant concomitant constitutive and induced costs.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

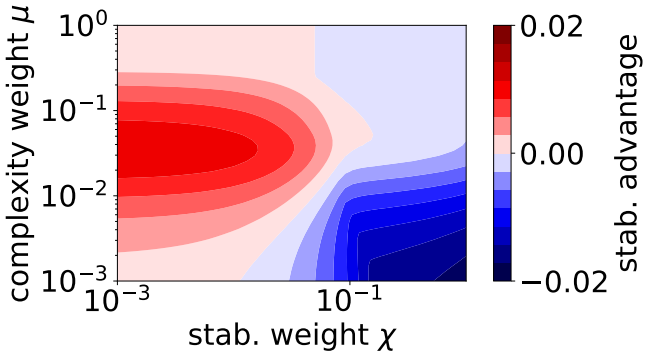

Fig. 5. Comparison of total divergence for stabilizers $\mathcal{D}_S$ and destabilizers $\mathcal{D}_{S'}$, or $\mathcal{D}_{S'} - \mathcal{D}_S$, in a fixed environment and common sensory precision and costs. The difference is between agents poised at optimal memory duration given μ, χ, and β. Small stabilization weight χ favors stabilizers, whereas high monopolization cost μ favors destabilizers. Simulation parameters are specified in accompanying code.

To generalize the previous argument, we assume larger organisms experience longer environmental timescales. Then, $\tau_E \propto T^y$, where $y \in [0, 1]$ to ensure that τ_E and N increase together since $\tau_E^{1/y-1} \propto N$. We now find the relationship between metabolic rate and environmental timescale

$$B \propto \tau_E^{\phi/y} \propto N^{\phi/(1-y)}, \qquad [14]$$

which reduces to the previous case when $y = 1$ (and N is a constant). Such dependence implies that the metabolic cost of memory will explode with environmental timescale (and organism lifetime) as y approaches zero and grow slowly and sublinearly when $y = 1$. Both possibilities are shown in Figure 4B. More generally, lifespan is expected to influence the relative contributions of adaptive versus metabolic costs (62, 63).

Result 4: Niche construction, memory complexity, & the outsourcing principle

In Result 3, we explored the metabolic cost of memory versus adaptation, emphasizing the metabolic constraints on long memories. In this section we focus on the information costs of adaptation when allowing for active modification of the environment. We explore how outsourcing memory to the environment by slowing it down is beneficial when costs of poor adaptation are dominant (10).

A slow environmental timescale increases the advantages of persistent memory, but it also reduces the amount of new information an organism requires by reducing uncertainty about the state of the environment. In this sense, slow environmental variables reflect a form of niche construction. Whether ant pheromone trails, food caching, collectively computing power structures, writing, or map-making, niche construction that promotes the stability or predictability of the local environment (5, 64) reduces the number of environmental configurations that an organism needs to encode. Stabilizing niche construction, however, also creates a public good that by reducing environmental uncertainty, provides a benefit to all agents, and can be exploited by free riders. This can lead to a tragedy of the commons (65).

We explore the conditions under which active modification of the environment can evolve given the free-rider problem, and how this overcomes the costs of adaptation. We introduce stabilizing mutants into a population of passive agents. Assuming other organisms are poorly adapted to regularities in the environment, we expect stabilizing mutants to gain a competitive advantage but only over the short term. In the longer term, an established stabilizer population is susceptible to invasion by free-riders exploiting outsourced memory; said another way, stabilizers slow environmental timescales and reduce divergence for all individuals sharing the environment, but they uniquely pay for stabilization. Thus, as in the classical example of niche construction, the usual "tragedy of the commons" argument makes it an evolutionary dead end (65).

It follows that stabilization is only a competitive strategy if individuals can monopolize extraction of resources from the stabilized environment. In the natural world, this could occur through physical encryption (e.g. undetectable pheromones (66)), the erasure of signal (e.g. food caching (67)), or the restriction of social information (e.g. concealment (68)). To model competition between monopolistic stabilizers and other strategies, we account for the costs of memory, stabilization, and precision. We introduce a new memory cost of encoding complex environments as

$$H(\tau_{\mathrm{m}}) = \log_2(1 + 1/\tau_{\mathrm{m}}). \qquad [15]$$

Eq 15 can be thought of as a cost of exploring more configurations over a short period time versus agents that are temporally confined. This is different from costs associated with the environmental burden in Result 3, which emphasizes the costs of persistence, not variability.

We define the cost stabilizers pay for niche construction as the extent of change to the environmental switching rate, or the KL divergence between the natural environmental rate $1/\tau_E$ and the time-averaged, modified rate $\langle 1/\tilde{\tau}_E \rangle$,

$$G(1/\tau_E, \langle 1/\tilde{\tau}_E \rangle) = \frac{1}{\tau_E} \log_2 \left(\frac{1/\tau_E}{\langle 1/\tilde{\tau}_E \rangle} \right) + \left(1 - \frac{1}{\tau_E}\right) \log_2 \left(\frac{1 - 1/\tau_E}{1 - \langle 1/\tilde{\tau}_E \rangle} \right). \qquad [16]$$

The quantity G depends implicitly on stabilization strength α because smaller α slows the environment further. For passive agents and destabilizers, $G = 0$ by definition because non-stabilizers fit to τ_E and only stabilizers benefit from the slower timescale with monopolization.

We finally consider the cost of precision, which we assume to be given by the information obtained by the agent from sampling the environment,

$$C(\tau_{\mathrm{c}}) = \log_2 \tau_{\mathrm{c}}. \qquad [17]$$

Sensory complexity means that higher precision implies higher expenditure to obtain such precision, given by the KL divergence between environment configuration and agent behavior, $C \sim -\log_2(\sigma^2)$ leaving out constants. This depends on the variance of agent measurement noise $\sigma^2 = p_E(s,t)[1 - p_E(s,t)]/\tau_{\mathrm{c}}$. Infinitely precise sensory cells lead to diverging cost, whereas imprecise cells are cheap.

Putting these costs together with divergence $\bar{D}$, we obtain the total divergence

$$\mathcal{D} = \bar{D} + \mu H + \chi G + \beta C. \qquad [18]$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Weights $\mu \geq 0$, $\chi \geq 0$, $\beta \geq 0$ represent the relative contribution of these costs. As a result, we can distinguish dominant strategies by comparing total divergence such as between the pair of destabilizer and stabilizer strategies shown in Figure 5. Large μ, or high complexity cost, means that a pure population of stabilizers would be stable to invasion from destabilizers. Whereas for large χ, or heavy stabilization cost, the opposite is true. The generalized measure of adaptive cost in Eq 18, given the weights, carves out regions of agent morphospace along axes of computational cost. This is a morphospace that captures the relative advantage of internal versus external memory that can be thought of as a space of evolutionary outsourcing.

As has often been remarked in relation to evolution, survival is not the same as arrival. We now determine when stabilizer strategies can emerge in this landscape. We start with a pure population of passive agents with stabilization strength $\alpha = 0$ and poised about optimal memory duration $\tau_{\mathrm{m}} = \tau_{\mathrm{m}}^*$ determined by minimizing both divergence $\bar{D}$ and complexity μH. Whether or not stabilizers emerge under mutation and selection can be determined through adaptive dynamics (69–71), that is by inspecting the gradient of the total divergence along the parameters $(\partial_{\tau_{\mathrm{m}}}\mathcal{D}, \partial_{\alpha}\mathcal{D}, \partial_{\tau_{\mathrm{c}}}\mathcal{D})$, or memory complexity, stabilizer strength, and precision. As we show in SI Appendix C and Eq S16, the gradient terms can be calculated under a set of perturbative approximations. Using local convexity about optimal memory τ_{m}^*, we show that the term $\partial_{\alpha}\mathcal{D}$ drives passive agents to smaller α and slower timescales; it originates from combining the scaling law from Eq 12 and complexity of memory. The term $\partial_{\tau_{\mathrm{c}}}\mathcal{D}$ shows that precision tends to decrease when the cost gradient $\partial_{\tau_{\mathrm{c}}}(\beta C)$ dominates over $\partial_{\tau_{\mathrm{c}}}\bar{D}$. In this case, the general conditions $\partial_{\alpha}\mathcal{D} < 0$ and $\partial_{\tau_{\mathrm{c}}}\mathcal{D} < 0$ funnel a passive population towards stabilization and reduced precision.

Discussion

Life is adaptive, but optimal adaptation would seem to depend on a multitude of properties of both organism and environment, which have been studied in a wide literature (Table 1). To the contrary, we predict that it does not. This becomes clear once we organize crucial aspects of adaptation into a unified framework in terms of timescales including niche construction that speeds up or slows down the environment (Figure 1). We find that memory duration, under a wide range of assumptions and conditions, scales sublinearly with environmental rates of change (Figure 2). This essentially derives from the competition between using current but noisy information and the reliance on outdated but precise information, leading to a universal, optimal timescale for adaptation. Importantly, sublinear scaling implies that persistent features of the environment can be more efficiently encoded the longer-lasting they become; there is an economy of scale.

Yet, memory remains costly as it requires investing in neural tissue. To estimate this cost and how it might affect adaptation, we use metabolic scaling theory to estimate how much neural tissue an organism must allocate to memory for a given rate of environmental change. We find that the metabolic costs of memory can increase super-linearly with the persistence time of environmental statistics. Thus, while memory need not grow in proportion to environmental stability, costs of memory could increase disproportionately (Figure 4). Because adaptive costs peak at short timescales, this suggests that adaptive costs are most important for organisms with short lifespans such as insects.

When the costs of adaptation are greater than the metabolic costs of memory, active modification of the environment such as stabilizing niche construction can be favored. In this case the organism intervenes on the environmental timescale to decrease volatility. Although outsourcing of memory to the environment reduces the organism's need to adapt, it introduces two new problems. First, active modification is itself not free. Second, slow environmental variables created by active modification are public goods that can be exploited by free riders.

To address the costs of active modification and free riding, we introduce game dynamics considering the information costs of adaptation including the complexity of memory. Unlike memory duration, memory complexity quantifies the effective number of states that agents occupy. Starting with passive agents, we find that the spontaneous emergence of adaptive dynamics stabilizes the environment, lengthening the optimal memory duration τ_{m}^* and thereby making weak stabilizers less competitive. This moves a population as a whole towards slower timescales. In other words, stabilizing niche construction, because of the economy of scale with respect to memory, requires proportionally less neural tissue for memory relative to the size of the whole brain as given by metabolic scaling theory. This is effectively outsourcing memory from neural tissue to the environment. As a possible consequence, organisms could reduce absolute brain size or invest in a larger behavioral repertoire, increasing competitiveness by monopolizing a larger number of environmental states. Do learning agents in volatile environments "given a choice" to invest in additional memory or to directly change the environment favor the latter?

This hypothesis is consistent with related work on institutions and social structure as a form of collectively encoded memory (72–75) or as devised constraints (e.g. reference 76) that slow down the need to acquire functional information. In pigtailed macaque society (reviewed in reference 10), individuals collectively compute a social-power distribution from status signaling interactions. The distribution of power as a coarse-grained representation of underlying fight dynamics changes relatively slowly and consequently provides a predictable social background against which individuals can adapt. By reducing uncertainty and costs, the power distribution facilitates the emergence of novel forms of impartial conflict management. Conflict management, in turn, further reduces volatility, allowing individuals to build more diverse and cohesive local social niches and engage in a greater variety of socially positive interactions (77). In other words, outsourcing memory, in this case, of fight outcomes, to a stable social structure in the power distribution allows for a significant increase in social complexity. More generally, we anticipate that one of the features of slowing environmental timescales, including social environments fostered by institutions, might the emergence of new functions (78).

Without feedforward and feedback loops between environment and agent such as in the case of the passive agent, our framework is akin to the classical problem of learning. This has been a major problem of interest in foraging (79), neural circuits that adapt to changing input distributions (19, 80, 81) and modes of prediction in order to best adapt to multiple clustered sets of statistics (20, 80). We introduce here a min-

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

imal modeling framework for connecting learners to active agents that modify the environment through the act of adaptation. Our framework provides a first-order approximation to this extended space, which could itself be extended in several directions to include how agents physically modify their environments, connecting to the physics of behavior with the physics of information (82).

Materials and Methods

The code used to generate these results will be made available on GitHub at https://github.com/eltrompetero/adaptation.

Numerical solution to model. Given Eqs 1–4 defining the binary agent, we calculate agent behavior in two ways. The first method is with agent-based simulation (ABS). We generate a long time series either letting the environment fluctuate independently and training the agent at each moment in time or coupling environmental fluctuations at each time step with the state of the agent. By sampling over many such iterations, we compute the distribution over agent bias given environmental bias, $q(h|h_E)$, which converges to a stationary form.

This principle of stationarity motivates our second solution of the model using an eigenfunction method. If the distribution is stationary, then we expect that under time evolution that the conditional agent distribution map onto itself

$$q(h|h_E) = \mathcal{T}[q(h|h_E)]. \tag{19}$$

If the time-evolution operator $\mathcal{T}$ evolves the distribution over a single time step, the external field can either stay the same with probability $1 - 1/\tau_E$ or reverse with probability $1/\tau_E$.

For either for these two possible alternatives over a single time step, we must convolve the distribution with the distribution of noise for the field η_{τ_c}. The distribution of noise derives from agent perceptual errors ϵ_{τ_c} on the estimated probabilistic bias of the environment (Eq 2). Hence, the corresponding error distribution for the bias η_{τ_c} originates from the binomial distribution through a transformation of variables. We can simplify this because in the limit of large sensory cell sample size τ_c the binomial distribution converges to a Gaussian and a concise representation of the distribution of η_{τ_c} becomes accurate. Using Eq 1, we find that the distribution of perceptual errors in the bias yields

$$\begin{aligned}\rho(\eta_{\tau_c}, t) = (8\pi\sigma^2)^{-1/2} \exp\Big\{ -[\tanh h_E(t) - \\ \tanh(h_E(t) + \eta_{\tau_c})]^2 / 8\sigma^2 \Big\} \operatorname{sech}^2(h_E(t) + \eta_{\tau_c}).\end{aligned} \tag{20}$$

Here, the agent's perceptual estimate of the environment includes finite-sample noise determined by the sensory cell precision $1/\tau_c$. At finite τ_c, there is the possibility that the agent measure a sample from the environment of all identical states. In our formulation, the fields then diverge as do the fields averaged over many separate measurements. We do not permit such a "zero-temperature" agent that freezes in a single configuration in our simulation just as thermodynamic noise imposes a fundamental limit on invariability in nature. Our agents inhabit an *in silico* world, where the corresponding limit is fixed by the numerical precision of the computer substrate, so we limit the average of the bits sampled from the environment to be within the interval $[-1 + 10^{-15}, 1 - 10^{-15}]$. This is one amongst variations of this idea that inference is constrained by regularization, Bayesian priors, Laplace counting (in the frequentist setting), etc. Regardless of the particular approach with which finite bounds might be established, they are only important in the small τ_c limit. See SI Appendix A.

Given the Gaussian approximation to precision error, we propagate the conditional distribution over a single time step, defining a self-consistent equation that can be solved by iterated application. To make this calculation more efficient, we only solve for abscissa of the Chebyshev basis in the domain $\beta \in [0, 1]$, fixing both the endpoints of the interval including the exact value for $\beta = 1$ from Eq 24 (83) (more details in SI Appendices A and B). In Figure S7, we show that our two methods align for a wide range of agent memory τ_m. Importantly, the eigenfunction approach is much faster than ABS for large τ_c because the latter can require a large number of time steps to converge. On the other hand, ABS is relatively fast for small τ_c. Thus, these two approaches present complementary methods for checking our calculation of agent adaptation.

Divergence curves. To measure how well agent behavior is aligned with the environment, we compare environment $p_E(s,t)$ and agent $p(s,t)$ with the KL divergence at each time step to obtain the agent's typical loss in Eq 9. Equivalently, we can average over the stationary distribution of fields conditional on environment

$$\bar{D} = \frac{1}{N_E} \sum_E \int_{-\infty}^{\infty} dh\, q(h|h_E) D_{\mathrm{KL}}[p_E(h_E)||p(h)], \tag{21}$$

where we sum over all possible environments E and weight them inversely with the number of total environments N_E. For the binary case, $N_E = 2$. We furthermore simplify this for the binary case as

$$\bar{D} = \int_{-\infty}^{\infty} dh\, q(h|h_E = h_0) D_{\mathrm{KL}}[p_E(h_E = h_0)||p(h)]. \tag{22}$$

In Eq 22, we have combined the two equal terms that arise from both positive $h_E = h_0$ and negative $h_E = -h_0$ biases of the environment.

In Figure 2A and B, we show divergence as a function of agent memory over a variety of environments of varying correlation time $\bar{D}(\tau_m, \tau_E)$. When the agent has no memory, its behavior is given solely by the properties of the sensory cells as is determined by the integration time τ_c. Then, we only need account for the probability that the environment is in either of the two symmetric configurations and how well the memoryless agent does in both situations. Since the configurations are symmetric, the divergence at zero memory is

$$\begin{aligned}\bar{D}(\tau_m = 0) = \int_{-\infty}^{\infty} d\eta_{\tau_c}\, \rho(\eta_{\tau_c}|h_E = h_0) \times \\ \sum_{s\in\{-1,1\}} p_E(s|h_E = h_0) \log_2 \left(\frac{p_E(s|h_E = h_0)}{p(s)} \right),\end{aligned} \tag{23}$$

where the biased distribution of environmental state p_E and the error distribution ρ from Eq 20 are calculated with environmental bias set to $h_E = h_0$. Note that this is simply Eq 22 explicitly written out for this case.

At the limit of infinite agent memory, as in the right hand side of Figure 2A, passive agents have perfect memory and behavior does not budge from its initial state. Assuming that we start with an unbiased agent such that $q(h) = \delta(h)$, the Dirac delta function, the agent's field is forever fixed at $h = 0$. Then, divergence reduces to

$$\bar{D}(\tau_m = \infty) = 1 - S[p_E], \tag{24}$$

where the conditional entropy $S[p_E] = -p_E(s|h = h_0) \log_2 p_E(s|h = h_0) - [1 - p_E(s|h = h_0)] \log_2 [1 - p_E(s|h = h_0)]$.

Scaling argument for optimal memory. As is summarized by Eq 10, the value of optimal memory can be thought of as a trade-off between the costs of mismatch with environment during the transient adaptation phase and gain from remembering the past during stable episodes. In order to apply this argument to the scaling of divergence, we consider the limit where the environment decay time τ_E is very long and agent memory τ_m is long though not as long as the environment's. In other words, we are interested in the double limit $\tau_m \to \infty$ and $\tau_m/\tau_E \to 0$. Then, it is appropriate to expand divergence in terms of the error in estimating the bias

$$\begin{aligned}\bar{D} = \Big\langle \sum_{s\in\{-1,1\}} p_E(s,t) \log p_E(s,t) - \\ p_E(s,t) \log[p_E(s,t) + \epsilon_{\tau_c}(t)] \Big\rangle,\end{aligned} \tag{25}$$

where the average is taken over time. Considering only the second term and simplifying notation by replacing $\epsilon_{\tau_c}(t)$ with ϵ,

$$\begin{aligned}&\langle p_E(s,t) \log p_E(s,t) + \log[1 + \epsilon/p_E(s,t)] \rangle \\ &\approx \left\langle p_E(s,t) \log p_E(s,t) + \frac{\epsilon}{p_E(s,t)} - \frac{1}{2} \frac{\epsilon^2}{p_E(s,t)^2} \right\rangle,\end{aligned} \tag{26}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where the average error $\langle \epsilon \rangle = 0$ and assuming that the next non-trivial correlation of fourth order $\mathcal{O}\left(\left\langle \epsilon^4 \right\rangle\right)$ is negligible. Plugging this back into Eq 25,

$$\bar{D} \approx \sum_{s \in \{-1,1\}} \frac{\tau_E - \tau_\mathrm{m}}{\tau_E} \frac{\left\langle \epsilon^2 \right\rangle}{p_E(s)^2} + \frac{\tau_\mathrm{m}}{\tau_E} \left\langle \frac{\epsilon^2}{p_E(s,t)^2} \right\rangle. \qquad [27]$$

The first term in Eq 27 relies on the fact that when environmental timescales are much longer than agent memory, the errors become independent of the state of the environment. Thus, we can average over the errors separately, and the environment configuration average can be treated independently of time $p_E(s,t) \to p_E(s)$. The second term, however, encases the transient dynamics that follow immediately after a switch in environmental bias while the agent remembers the previous bias. It is in the limit $\tau_\mathrm{m}/\tau_E \to 0$ that we can completely ignore this term and the scaling for optimal memory $\tau_\mathrm{m}^* \sim \tau_E^{1/2}$ from Eq 11 is the relevant limit that we consider here.

Since the errors with which the agent's matching of environmental bias is given by a Gaussian distribution of errors, the precision increases with the number of samples taken of the environment: it should increase with both sensory cell measurement time τ_c as well as the typical number of time steps in the past considered, $\tau_\mathrm{m} = -1/\log\beta$. Thus, we expect the scaling of divergence at optimal memory to be

$$\bar{D}^* \sim \frac{1}{\tau_\mathrm{m}^* \tau_\mathrm{c}}, \qquad [28]$$

which with Eq 11 leads to the scaling of optimal memory with environment decay time Eq 12. Though the scaling with precision timescale τ_c in Eq 28 is at $\tau_\mathrm{m} = \tau_\mathrm{m}^*$, it is clear that a similar scaling with τ_c holds at $\tau_\mathrm{m} = 0$, where only precision determines divergence. However, such a scaling does not generally hold for any fixed τ_m, the trivial case being at $\tau_\mathrm{m} = \infty$, where divergence must go to a constant determined by environmental bias.

ACKNOWLEDGMENTS. E.D.L. was supported by the Omega Miller Program at the Santa Fe Institute. D.C.K. and J.F. are grateful for support from the James S. McDonnell Foundation 21st Century Science Initiative-Understanding Dynamic and Multi-scale Systems.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A. Agent-based simulation

To complement the eigenfunction solution described in Appendix B, we present the agent-based simulation.

After having specified the environmental bias $h_E(t)$, we generate a sample of τ_c binary digits from the distribution $p_E(s,t)$. From this sample, we calculate the mean of the environment $\langle s \rangle$ which is bounded in the interval $[-1+10^{-15}, 1-10^{-15}]$. These bounds are necessary to prevent the measured field $\hat{h}(t)$ from diverging and reflects the fact that *in silico* agents have a finite bound in the values they can represent, mirroring finite cognitive resources for biological or social agents as discussed in Materials and Methods. We combine this estimated field $\hat{h}(t)$ with the one from the aggregator having set the initial value condition $H(0) = 0$. Given the estimate of the field $h(t)$, we compute the Kullback-Leibler (KL) divergence between the agent distribution $p(s)$ and the environment $p_E(s)$.

When we calculate the divergence landscape across a range of different agent memories, we randomly generate the environment using the same seed for the random number generator. Though this introduces bias in the pseudorandom variation between divergence for agents of different types, it makes clearer the form of the divergence landscape by eliminating different offsets between the points. Our comparison of this approach with the eigenfunction solution in Appendix B provides evidence that such bias is small with sufficiently long simulations. For the examples shown in the main text, we find that total time $T = 10^7$ or $T = 10^8$ are sufficient for convergence to the stationary distribution after ignoring the first $t = 10^4$ time steps.

B. Eigenfunction solution

We present more details on top of those in Materials and Methods on the iterative, eigenfunction solution to the divergence of an agent relying on the fact that the distribution of agent bias $q(h)$ becomes stationary at long times.

Let us first consider the case of the passive agent. After sufficiently long time, the distribution of agent behavior $q(h)$ and the distribution conditioned on the two states of the environment $q(h|h_E = h_0)$ and $q(h|h_E = -h_0)$ converge to stationary forms. Assuming that the distributions have converged, we evolve the distribution a single time step. If the external field $h_E(t) = h_0$, then it either stays fixed with probability $1 - 1/\tau_E$ or it switches to the mirrored configuration with probability $1/\tau_E$.

Considering now the evolution of the conditional probability $q(h|h_E = h_0)$, we note that the state of the agent will be either be convolved by the distribution of sampling error at the next time step or lose probability density from a switching field. Since we are considering a symmetric configuration, however, the mirrored conditional density will reflect the same probability density back such as in Eq S1. Thus, Eq S1 is satisfied by the conditional density of agent bias that is solved by the eigenfunction for $q(h|h_E)$ with eigenvalue 1. By the Perron-Frobenius theorem when considering normalized eigenvectors, this is the unique and largest eigenvalue that returns the stationary solution.

To extend this formulation to active agents, we must also account for the dependence of the rate of switching on the distance between agent and environmental bias. This additional complication only requires a modification of Eq S1 to include such dependence in the rate coefficients. Thus, all types of agents can be captured by this eigenfunction solution and solved by iteration til convergence.

Eq S1 is only independent of time when agent memory $\tau_m = 0$. When there is finite memory, or $\beta > 0$, the distribution $q(h,t)$ "remembers" the previous state of the environment such that we must iterate Eq S1 again. Over many iterations, we will converge to the solution, but the convergence slows with agent memory which introduces ever slower decaying eigenfunctions. An additional difficult arises because the narrowing in the peak of the agent's estimate of the environment, like the peaks shown in Figure S7, require increased numerical precision. As a result, increasing memory and computational costs make it infeasible to calculate the eigenfunction with high precision for β close to 1.

Instead of calculating the full functional form directly below but not at the limit $\beta \to 1$, we use the output of the iterative eigenfunction procedure as input for an interpolation procedure using Chebyshev polynomials. We iterate Eq S1 for β equal to the Gauss-Lobatto abscissa of the Chebyshev polynomial of degree d, mapping the interval $\beta \in [0,1]$ to the domain $x \in [-1,1]$ for the set of Chebyshev polynomials (83). The Gauss-Lobatto points include the endpoints $\beta = 0$ and $\beta = 1$, the first of which is trivial numerically and the latter for which we have an exact solution given in Eq 24. Then, we exclude calculated values for large β that show large iteration error $\epsilon > 10^{-4}$. This threshold, however, leaves the coefficients of the Chebyshev polynomial undetermined. We instead interpolate these remaining $N - k$ points by by fitting a Chebyshev polynomial of degree $N - k - 1$ with least-squares on the logarithm of the divergence. A similar procedure can be run for the stabilization cost from Eq 16 to obtain Figure S6B. We find that typically $N = 30$ or $N = 40$ starting abscissa with a maximum of 10^3 iterations are sufficient to obtain close agreement with the agent-based simulation (ABS) from Appendix A (Figure S8). This interpolation procedure does not work well with ABS because small stochastic errors can lead to high-frequency modes in interpolation (and thus large oscillations), errors that can be essentially driven to zero exponentially fast for the eigenfunction method.

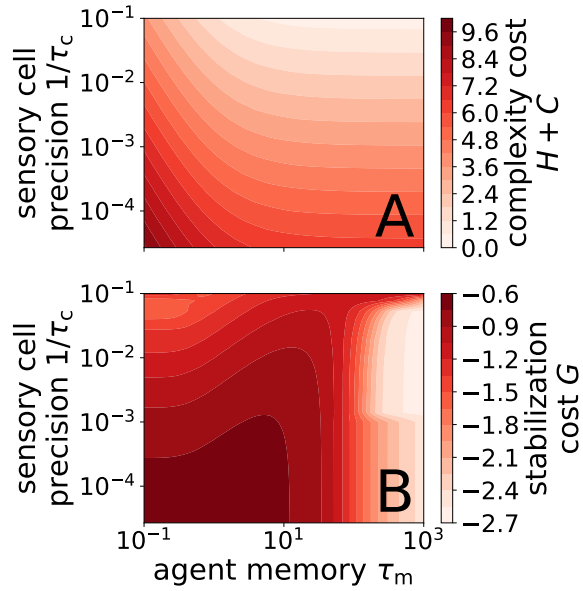

Fig. S6. Landscape of the costs we consider as (A) a combined agent complexity and (B) a stabilization cost. (A) Isocontours defined as sum of memory complexity and sensory precision costs. The values have been offset to ensure that the costs are positive over the shown landscape calculated from memory (Eq 15) and sensory cost (Eq 17). (B) As memory $\tau_m \to \infty$, stabilization cost converges to a finite value that can be calculated exactly from noting that agent behavior has probability density fixed at its starting point, $q(h) = \delta(h)$. A kink in the contours at $1/\tau_c = 10^{-3}$ arises from numerical precision errors where we matched up ABS and eigenfunction methods.

C. Evolution of reduced complexity

We consider a population of passive agents, or an agent with stabilization parameter $\alpha = 0$, precision timescale τ_c, and optimal memory τ_m^*, the variables that determine agent fitness. Assuming that the canonical equation for evolution applies (i.e. mutations only change phenotype and fitness slightly, the population dynamics move much faster than the evolutionary landscape such that we can assume a single phenotype dominates), the rate at which the population evolves across the phenotypic landscape is proportional to the fitness gradient. In addition to this assumption, we will assume that the population is always poised at optimal memory, an assumption that will be made clear below.

We recall that the total divergence consists of the time-averaged divergence $\bar{D}$, statistical complexity cost H, stabilization cost G, and precision cost C

$$\mathcal{D} = \bar{D} + \mu H(\tau_m) + \chi G(\tau_E, \tilde{\tau}_E) + \beta C(\tau_c) \qquad \text{[S2]}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$q(h,t|h_E=h_0) = \left(1-\frac{1}{\tau_E}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho(\eta_{\tau_c}|h_E=h_0)q(h,t-1|h_E=h_0)\delta(h-h_0-\eta_{\tau_c})\,d\eta_{\tau_c}\,dh+ \\ \frac{1}{\tau_E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho(\eta_{\tau_c}|h_E=-h_0)q(h,t-1|h_E=-h_0)\delta(h+h_0-\eta_{\tau_c})\,d\eta_{\tau_c}\,dh. \quad [S1]$$

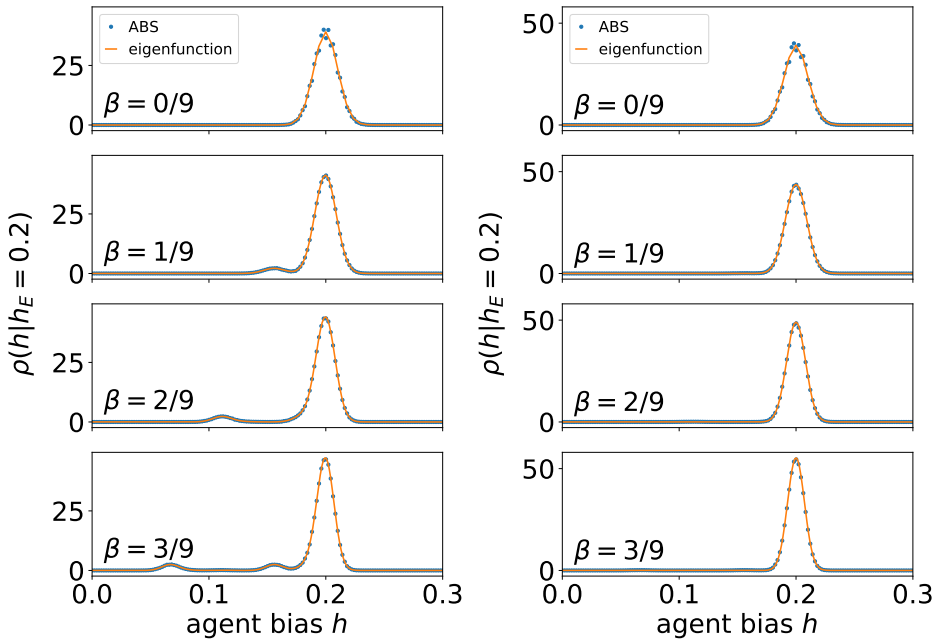

Fig. S7. Comparison of agent-based simulation (ABS) and eigenfunction solution for the conditional probability distribution of agent bias $q(h|h_0)$ for (left) a passive agent and (right) stabilizer. Agent-based simulation returns a normalized histogram that aligns closely with the eigenfunction solution. Environment timescale $\tau_E = 20$ and bias $h_0 = 0.2$. Spacing of discrete domain in eigenfunction solution determined in proportion with typical width of the peak around $h = h_0$, which scales as in Eq 20 and inversely with the square root of agent memory τ_m.

with semi-positive weights μ, χ, and β. In Figure S9, we show each the divergence of a stabilizer without such costs in blue, each of these costs separately in black, and their sum in orange to generate the total divergence in Eq 18. For the evolutionary dynamics, we must calculate the gradient $(\partial_{\tau_\mathrm{m}}\mathcal{D}, \partial_\alpha\mathcal{D}, \partial_{\tau_c}\mathcal{D})$ determining the evolution in the properties of the agent. We calculate these term by term and then put them together at the end.

We assume that agent memory τ_m is at the minimum of the combination of time-averaged divergence $\bar{D}$ and statistical complexity cost μH (stabilization is zero for passive agents). Since divergence has a unique minimum and complexity monotonically approaches $H(\tau_\mathrm{m}=\infty)=0$, the addition of complexity only shifts optimal memory to a larger value. Without the complexity cost, we have that small deviations about optimal memory can be represented by a quadratic function for some positive constant a,

$$\bar{D} = D^* + a(\tau_\mathrm{m}-\tau_\mathrm{m}^*)^2, \quad [S3]$$

where we write

$$D^* = D_0(\tau_\mathrm{m}^*)^{1/2} \quad [S4]$$

for some positive constant D_0. Once we have accounted for a perturbative addition from memory complexity, however, we have a shifted optimal memory

$$\tau_\mathrm{m}^{**} = \tau_\mathrm{m}^* + \frac{\mu}{2(\log 2)a\tau_\mathrm{m}^*(\tau_\mathrm{m}^*+1)} + \mathcal{O}(\mu^2) \quad [S5]$$

obtained from $\partial_{\tau_\mathrm{m}}[\bar{D}+\mu H]=0$ and using the approximation that μ is small. Eq S18 shows us that memory complexity, the term proportional to μ, drives optimal memory τ_m^{**} up.

Taking the approximation in Eq S18 the shifted optimal divergence, denoted by an apostrophe, becomes

$$\bar{D}'(\tau_\mathrm{m}^{**}) = D^* + a\frac{\mu^2}{4(\log 2)^2(\tau_\mathrm{m}^*)^2(\tau_\mathrm{m}^*+1)^2} + \mathcal{O}(\mu^3). \quad [S6]$$

Again, perturbations about the local optimum lead to

$$\bar{D}'(\tau_\mathrm{m}) \approx D^* + a\frac{\mu^2}{4(\log 2)^2(\tau_\mathrm{m}^*)^2(\tau_\mathrm{m}^*+1)^2} + \\ b(\tau_\mathrm{m}-\tau_\mathrm{m}^{**})^2 \quad [S7]$$

for some positive constant b, which implicitly depends on the complexity cost. Eq S20 expresses local convexity about shifted optimal memory τ_m^{**} according to the corresponding shifted divergence $\bar{D}'$. This indicates how the population is poised along the ridge of optimal memory given a perturbative cost of memory complexity.

Then, time-averaged divergence will grow because optimal memory changes. Assuming that the population is at optimal memory, we obtain for the partial derivative with respect to α

$$\partial_\alpha\bar{D}' = \left[\frac{D_0}{2}(\tau_\mathrm{m}^*)^{-\frac{3}{2}} + a\frac{\mu^2(2\tau_\mathrm{m}^*+1)}{2(\log 2)(\tau_\mathrm{m}^*)^3(\tau_\mathrm{m}^*+1)^3}\right]\left|\frac{\partial\tau_\mathrm{m}^*}{\partial\alpha}\right|, \quad [S8]$$

where we have used the fact that optimal memory must increase with stronger stabilizer, or that $\partial_\alpha\tau_\mathrm{m}^*<0$, to explicitly pull out a negative sign. Given that we are in the scalin regime, this confirms that in Eq S8 divergence at optimal memory decreases as α approaches -1 from above as expected.

Niche-constructing stabilization changes the environmental timescale through feedback. We start by considering over a long

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

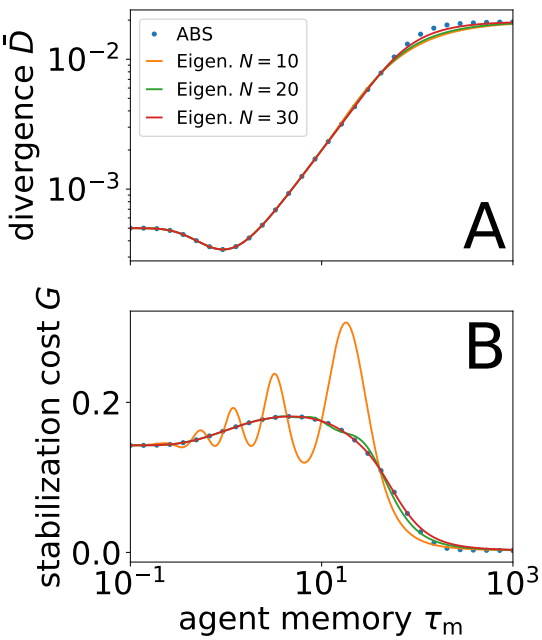

Fig. S8. Example of convergence of least-squares fit of Chebyshev polynomial with increasing number of abscissa N with the eigenfunction solution. (top) For comparison, divergence D as calculated from the agent-based simulation (ABS). The eigenfunction solution is close even with a relatively small number of points fit to a 9th-degree Chebyshev polynomial. Both methods are especially effective when the environmental timescale is small as is here, where $\tau_E = 10$. The bias $h_0 = 0.2$. (bottom) Stabilization cost is similarly interpolated, but it is slower to converge with visible oscillations disappearing by $N = 30$. For $N = 20$ and $N = 30$, not all the points fell within the convergence criterion and only 19 and 28 points were fit, respectively. For both plots, the Chebyshev polynomial approximation is slowest to converge near the sharp bends at large τ_m. ABS is run for 10^7 time steps.

period of time the average over many environmental switches

$$\begin{aligned}\langle 1/\tilde{\tau}_E\rangle &= 1/\tau_E + \alpha\left\langle \frac{v^2}{v^2+(h-h_E)^2}\right\rangle, \\ &= 1/\tau_E + \alpha f(\tau_m).\end{aligned} \quad [S9]$$

Since we do not know the exact form of the second term on the right hand side, we represent it as some function f that represents an average over time. For notational simplicity, we only make explicit f's dependence on τ_m, but it depends on agent properties and environmental timescale. Now, a change in α also indirectly affects τ_m^* because the environmental timescale will change, reducing or increasing the agents ability to track the new environment. For example, with the passive agent, an increase in α introduces environmental stabilization, driving the effective environmental timescale slower and moving the optimal memory timescale up. Accounting for these derivatives means that

$$d_\alpha \langle 1/\tilde{\tau}_E\rangle = f(\tau_m) + \alpha\partial_{\tau_m} f(\tau_m)\partial_\alpha \tau_m. \quad [S10]$$

Now, we will again make use of the assumption that τ_m is close τ_m^* such that we can make the linear approximation $f(\tau_m) \approx f(\tau_m^*) + (\tau_m - \tau_m^*)f'(\tau_m^*)$. Putting this in, we find

$$\begin{aligned} d_\alpha \langle 1/\tilde{\tau}_E\rangle(\tau_m) = f(\tau_m^*) + (\tau_m - \tau_m^*)f'(\tau_m^*) + \\ \alpha\partial_{\tau_m}[f(\tau_m^*) + (\tau_m - \tau_m^*)f'(\tau_m^*)]\partial_\alpha\tau_m^* \end{aligned} \quad [S11]$$

For a passive agent, this simplifies because $\alpha = 0$. Furthermore, we know that $f'(\tau_m^*) = 0$ because we have assumed that the agent is at optimal memory so any deviation from optimal memory must generally increase the typical distance between environmental and agent bias $(h - h_E)^2$. Then,

$$d_\alpha \langle 1/\tilde{\tau}_E\rangle(\tau_m) = f(\tau_m^*). \quad [S12]$$

Eq S12 is already clear from Eq S9 given the assumptions we have made, but these steps take us through the general problem (when not situated exactly at optimal memory and when $\alpha \neq 0$ are more complicated). In other words, decreasing α for the weak stabilizer will reduce the probability that the environment switches by the term in Eq S12 because $f > 0$ and f' — the change in probability is not just dependent on the rate effect f but also its derivative.

Under such a change, the new environmental timescale will deviate from τ_E and so the stabilization cost can be expanded as

$$\begin{aligned} G(\tau_E, \tilde{\tau}_E) &= \frac{1}{\tau_E}\log\left(\frac{1/\tau_E}{\langle 1/\tilde{\tau}_E\rangle}\right) + \left(\frac{1}{\tau_E}\right)\log\left(\frac{1-1/\tau_E}{1-\langle 1/\tilde{\tau}_E\rangle}\right) \\ &\approx \frac{1}{2\tau_E}\left[\langle 1/\tilde{\tau}_E\rangle - 1/\tau_E\right]^2 + \\ &\quad \frac{1}{2}\left(1-\frac{1}{\tau_E}\right)\left[\langle 1/\tilde{\tau}_E\rangle - 1/\tau_E\right]^2 \\ &= \frac{1}{2}\left[\langle 1/\tilde{\tau}_E\rangle - 1/\tau_E\right]^2, \end{aligned} \quad [S13]$$

a cost that increases quadratically with the change in the averaged switch probability $\langle 1/\tilde{\tau}_E\rangle$ away from $1/\tau_E$. For a passive agent, this direction is 0 unless we allow for α to vary, which leads to the relation

$$G(\tau_E, \tilde{\tau}_E) = \frac{\alpha^2}{2} f(\tau_m^*)^2. \quad [S14]$$

Eq S14 tells us that if we vary α, we must pay a stabilization cost that, at least locally, grows quadratically with the strength of stabilization with zero gradient.

The simplest contribution is with respect to the change in the precision timescale τ_c. Divergence, as derived in Materials & Methods, is proportional to $1/\tau_c$. On the other hand, precision cost is $C = \log\tau_c$. Since optimal memory timescale does not depend on τ_c, the change of the total divergence is

$$\partial_{\tau_c}\left[D_1/\tau_c + \beta\log_2\tau_c\right] = -D_1/{\tau_c}^2 + \beta/\tau_c, \quad [S15]$$

where we take $D^* = D_1/\tau_c$ to encapsulate the terms in the divergence apart from the scaling with precision timescale. If this has a minimum at positive τ_c, the value of τ_c at which the minimum is reached is ${\tau_c}^* = D_1/\beta$.

Putting all of these together, we have the terms in the gradient

$$\begin{aligned} \partial_{\tau_m}\mathcal{D} &= 2a(\tau_m - \tau_m^*) \\ \partial_\alpha\mathcal{D} &= \left[\frac{D_0}{2}(\tau_m^*)^{-\frac{3}{2}} + a\frac{\mu^2(2\tau_m^*+1)}{2\log(2)(\tau_m^*)^3(\tau_m^*+1)^3}\right]\left|\frac{\partial\tau_m^*}{\partial\alpha}\right| \\ \partial_{\tau_c}\mathcal{D} &= \beta/\tau_c - D_1/{\tau_c}^2 \end{aligned} \quad [S16]$$

When the cost gradient $\partial_\alpha\mathcal{D} < 0$, a population of passive agents is driven towards niche construction and when $\partial_{\tau_c}\mathcal{D} < 0$ towards precision reduction. Thus, the conditions that lead to reduction in agent complexity by increasing memory, enhancing stabilization, and lowering precision are captured by these gradients.

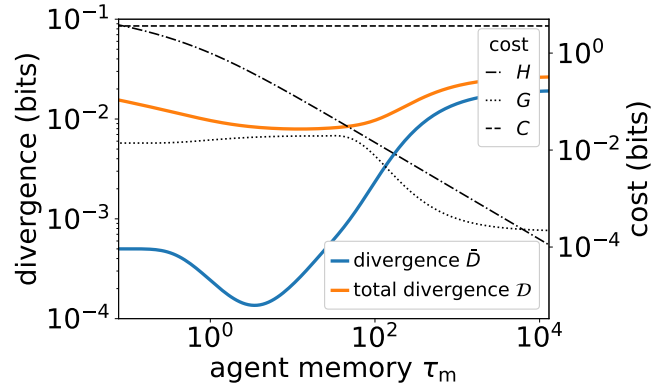

Fig. S9. Example of cost functions for stabilizers with varying memory but fixed sensory precision. (blue) Without costs, divergence profile shows only a single global minimum. (orange) With costs, we obtain degenerate minima at memory values around $\tau_m = 0$ and $\tau_m = 20$. Eigenfunction solution parameters specified in Materials and Methods code.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A similar derivation can be made for the evolution of a starting population of destabilizers, or agents with $\alpha > 0$, instead a pure population of passive agents. However, this requires us to deal with all the terms in Eq S11 and to account for a term from the gradient of stabilization cost in Eq S16 instead of assuming $\alpha = 0$. The change in the environmental timescale is more complicated to calculate because we must then consider the way that destabilization determines the modified environmental timescale in Eq S9, but it is clear that the qualitative results will be the same because of the adaptive gain from slower environmental timescales, i.e. decreasing α, but the exact rate at which α changes will depend on the curvature of the stabilization cost.

D. Metabolic costs of neural tissue for memory

In the total divergence in Eq 18 and as discussed in Appendix C, we consider information costs separately from energetic, metabolic costs of neural tissue. An important consideration for comparing the costs directly with one another is that that the right units for comparison are not clear, an issue that we avoid by only considering the scaling exponents presented in Result 3. Furthermore, while the scaling argument makes clear that the metabolic costs will dominate at sufficiently long lifetimes, the differences in how information and energetic costs affect reproductive fitness make a direct comparison in a combined "total divergence" equation problematic.

Nonetheless, if we do entertain the inclusion of metabolic costs into the total divergence, we will find that rising metabolic costs with environmental timescale will lead to a upper cutoff, i.e. truncating memory at some point beyond which the benefits of increasing stabilization are counteracted by the monotonically increasing costs of supporting neural tissue for memory.

To show this more formally, we redo the calculations in the previous section with an additional metabolic cost of memory,

$$\mathcal{D} = \bar{D} + \mu H(\tau_{\mathrm{m}}) + \chi G(\tau_E, \tilde{\tau}_E) + \beta C(\tau_{\mathrm{c}}) + \gamma F(\tau_{\mathrm{m}}), \qquad \text{[S17]}$$

with the new term $F(\tau_{\mathrm{m}}) = \tau_{\mathrm{m}}^{2\phi}$ defining the metabolic cost from Result 3. Then, we have for the shifted optimal memory

$$\tau_{\mathrm{m}}^{**} = \tau_{\mathrm{m}}^{*} + \frac{\mu}{2(\log 2) a \tau_{\mathrm{m}}^{*}(\tau_{\mathrm{m}}^{*}+1)} - \frac{\phi(\tau_{\mathrm{m}}^{*})^{2\phi}}{a(\tau_{\mathrm{m}}^{*}+1)} + \mathcal{O}(\mu^2) + \mathcal{O}(\gamma^2) + \mathcal{O}(\mu\gamma), \qquad \text{[S18]}$$

obtained from $\partial_{\tau_{\mathrm{m}}}[\bar{D} + \mu H + \gamma F] = 0$ and using the approximation that both μ and γ are small. The perturbative assumption is not necessary to take, but then there is no closed analytical solution for the shifted optimal memory τ_{m}^{**} that we can write down. Eq S18 shows us that memory complexity, the term proportional to μ, tends to drive optimal memory τ_{m}^{**} up but the metabolic cost, the term proportional to γ, tends to drive it down, the balance of which determine the exact change in optimal memory.

Taking the approximation in Eq S18 the shifted optimal divergence, denoted by an apostrophe, becomes

$$\bar{D}'(\tau_{\mathrm{m}}^{**}) = D^* + a\frac{\mu^2}{4(\log 2)^2(\tau_{\mathrm{m}}^{*})^2(\tau_{\mathrm{m}}^{*}+1)^2} + \frac{\gamma^2\phi^2(\tau_{\mathrm{m}}^{*})^{4\phi}}{a^2(\tau_{\mathrm{m}}^{*}+1)^2} - \frac{a\phi\mu\gamma(\tau_{\mathrm{m}}^{*})^{2\phi}}{2(\log 2) a\tau_{\mathrm{m}}^{*}(\tau_{\mathrm{m}}^{*}+1)^2} + \mathcal{O}(\mu^3) + \mathcal{O}(\gamma^3) + \mathcal{O}(\mu\gamma^2) + \mathcal{O}(\mu^2\gamma). \qquad \text{[S19]}$$

Again, perturbations about the local optimum lead to

$$\bar{D}'(\tau_{\mathrm{m}}) \approx D^* + a\frac{\mu^2}{4(\log 2)^2(\tau_{\mathrm{m}}^{*})^2(\tau_{\mathrm{m}}^{*}+1)^2} + \frac{\phi^2(\tau_{\mathrm{m}}^{*})^{4\phi}}{a^2(\tau_{\mathrm{m}}^{*}+1)^2} - \frac{a\phi\mu(\tau_{\mathrm{m}}^{*})^{2\phi}}{2(\log 2) a\tau_{\mathrm{m}}^{*}(\tau_{\mathrm{m}}^{*}+1)^2} + b(\tau_{\mathrm{m}} - \tau_{\mathrm{m}}^{**})^2 \qquad \text{[S20]}$$

for some positive constant b, which implicitly depends on the complexity cost. Assuming that the population is at optimal memory, we obtain for the partial derivative with respect to α

$$\partial_\alpha \bar{D}' = -\left[\frac{D_0}{2}(\tau_{\mathrm{m}}^{*})^{-3/2} + a\frac{\mu^2(2\tau_{\mathrm{m}}^{*}+1)}{2(\log 2)(\tau_{\mathrm{m}}^{*})^3(\tau_{\mathrm{m}}^{*}+1)^3} + \frac{2(\log 2)\phi^2 a^{-2}(\tau_{\mathrm{m}}^{*})^{4\phi} + \mu\phi(\tau_{\mathrm{m}}^{*})^{2\phi-1}}{(\log 2)(\tau_{\mathrm{m}}^{*}+1)^3} - \frac{4(\log 2)\phi^3(\tau_{\mathrm{m}}^{*})^{4\phi-1} + 2^{-1}\mu\phi(2\phi-1)(\tau_{\mathrm{m}}^{*})^{2\phi-2}}{(\log 2)(\tau_{\mathrm{m}}^{*}+1)^2}\right]\frac{\partial\tau_{\mathrm{m}}^{*}}{\partial\alpha}. \qquad \text{[S21]}$$

Unlike the previous outcome in Eq S8, it is not necessarily the case that a stronger stabilizer will decrease divergence because sufficiently large metabolic costs will counteract the adaptive benefits of a slower environment.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Variables used in main text organized by section in which they are first introduced or used.

Model structure & assumptions

Parameter	Description
A_t	discrete agent state at time t, e.g. $\{-1, 1\}$
E_t	discrete environmental state at time t, e.g. $\{-1, 1\}$
h_0	parameter for strength of environmental bias
h	agent bias
$\hat{h}$	agent's estimate of environmental bias
h_E	environmental bias
p	agent's probability distribution over possible states of A_t after time integration
$\hat{p}$	agent's estimate of environment probability distribution at time t based on present samples
p_E	environmental probability distribution over possible states of E_t
q	probability of change in environmental bias at a single time step
s	state of environment taking values of -1 or 1
t	time
v	construction rate curvature
α	construction rate weight, $\alpha < 0$ for stabilizers and $\alpha > 0$ for destabilizers
β	learning weight in Eq 4; coefficient of precision cost in Eq 18
ϵ_{τ_c}	perceptual error
η_{τ_c}	estimated bias error
τ_{c}	sampling duration, inverse precision
τ_E	environment duration
τ_{f}	niche construction duration
τ_{m}	agent memory duration

Result 1

Parameter	Description
$\bar{D}$	time-averaged Kullback-Leibler (KL) divergence
$\bar{D}^*$	time-averaged KL divergence at optimal memory duration
D_{KL}	KL divergence
τ_{m}^*	optimal memory duration

Result 3

Parameter	Description
B	metabolic rate
M_{br}	brain mass
N	number of episodes of environmental change
T	lifespan of organism
y	exponent relating environment duration and organism lifetime
ϕ	exponent relating metabolic rate and memory duration, $\phi = a/4b$ for energetic exponents a and b

Result 4

Parameter	Description
C	sensory precision cost
$\mathcal{D}$	total divergence
G	stabilization cost
H	complexity of memory cost
β	coefficient for sensory cost
μ	coefficient for memory complexity cost
$\tilde{\tau}_E$	modified environment duration
χ	coefficient for stabilization cost

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Feldman MW, Laland KN (1996) Gene-culture coevolutionary theory. *Trends in Ecology & Evolution* 11(11):453–457.
2. Gerbault P, et al. (2011) Evolution of lactase persistence: an example of human niche construction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1566):863–877.
3. Odling-Smee FJ, Laland KN, Feldman MW (1996) Niche Construction. *The American Naturalist* 147(4):641–648. Publisher: The University of Chicago Press.
4. Laland KN, O'Brien MJ (2011) Cultural Niche Construction: An Introduction. *Biological Theory* 6(3):191–202.
5. Clark AD, Deffner D, Laland K, Odling-Smee J, Endler J (2020) Niche Construction Affects the Variability and Strength of Natural Selection. *The American Naturalist* 195(1):16–30.
6. Buser CC, Newcomb RD, Gaskett AC, Goddard MR (2014) Niche construction initiates the evolution of mutualistic interactions. *Ecology Letters* 17(10):1257–1264. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ele.12331.
7. Bowles S (2006) *Microeconomics*.
8. Poon P, Flack JC, Krakauer DC (2022) Institutional dynamics and learning networks. *PLOS ONE* 17(5):e0267688. Publisher: Public Library of Science.
9. Ofosu EK, Chambers MK, Chen JM, Hehman E (2019) Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *Proceedings of the National Academy of Sciences* 116(18):8846–8851. Publisher: Proceedings of the National Academy of Sciences.
10. Flack JC (2017) Coarse-graining as a downward causation mechanism. *Phil. Trans. R. Soc. A* 375(2109):20160338.
11. Flack J (2017) Life's information hierarchy in *From Matter to Life: Information and Causality*, eds. Ellis GFR, Davies PCW, Walker SI. (Cambridge University Press, Cambridge), pp. 283–302.
12. Mukherjee S, Bassler BL (2019) Bacterial quorum sensing in complex and dynamically changing environments. *Nature Reviews Microbiology* 17(6):371–382. Number: 6 Publisher: Nature Publishing Group.
13. Merton RK (1948) The Self-Fulfilling Prophecy. *The Antioch Review* 8(2):193–210. Publisher: Antioch Review, Inc.
14. Strathern M (1997) Improving ratings: audit in the British University system. *European Review* 5(3). Publisher: Cambridge University Press.
15. Soros G (2013) Fallibility, reflexivity, and the human uncertainty principle. *Journal of Economic Methodology* 20(4):309–329. Publisher: Routledge _eprint: https://doi.org/10.1080/1350178X.2013.859415.
16. Manheim D, Garrabrant S (2019) Categorizing Variants of Goodhart's Law. *arXiv:1803.04585 [cs, q-fin, stat]*. arXiv: 1803.04585.
17. Jakab S (2022) *The Revolution That Wasn't: GameStop, Reddit, and the Fleecing of Small Investors*. (Portfolio, New York, NY).
18. Kalman RE (1960) A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* pp. 35–45.
19. Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive Rescaling Maximizes Information Transmission. *Neuron* 26(3):695–702.
20. Gershman SJ, Radulescu A, Norman KA, Niv Y (2014) Statistical Computations Underlying the Dynamics of Memory Updating. *PLoS Comput Biol* 10(11):e1003939.
21. Davis RL, Zhong Y (2017) The Biology of Forgetting—A Perspective. *Neuron* 95(3):490–503.
22. Kussell E, Leibler S (2005) Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. *Science* 309(5743):2075–2078.
23. Rivoire O, Leibler S (2011) The Value of Information for Populations in Varying Environments. *J Stat Phys* 142(6):1124–1166.
24. Chowdhury SM, Kovenock D, Sheremeta RM (2013) An experimental investigation of Colonel Blotto games. *Economic Theory* 52(3):833–861.
25. Krakauer D, Bertschinger N, Olbrich E, Flack JC, Ay N (2020) The information theory of individuality. *Theory in Biosciences* 139(2):209–223.
26. Theraulaz G, Bonabeau E (1999) A Brief History of Stigmergy. *Artificial Life* 5(2):97–116.
27. Brush ER, Krakauer DC, Flack JC (2018) Conflicts of interest improve collective computation of adaptive social structures. *Science Advances* 4(1):e1603311. Publisher: American Association for the Advancement of Science.
28. Ramos-Fernandez G, Smith Aguilar SE, Krakauer DC, Flack JC (2020) Collective Computation in Animal Fission-Fusion Dynamics. *Frontiers in Robotics and AI* 7.
29. Ratcliff R (1978) A Theory of Memory Retrieval. *Psychological Review* 85(2):59–108.
30. Koehl MAR, Cooper T (2015) Swimming in an Unsteady World. *Integr. Comp. Biol.* 55(4):683–697.
31. Evans MEK, Dennehy JJ (2005) Germ Banking: Bet-Hedging and Variable Release from Egg and Seed Dormancy. *The Quarterly Review of Biology* 80(4):431–451.
32. Slivkins A (2019) *Introduction to Multi-Armed Bandits*. (Now Publishers).
33. Kaspi H, Mandelbaum A (1995) Levy Bandits: Multi-Armed Bandits Driven by Levy Processes. *Ann. Appl. Probab.* 5(2):541–565.
34. Schofield N (2008) Divergence in the spatial stochastic model of voting in *Power, Freedom, and Voting*, eds. Braham M, Steffen F. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 259–287.
35. Yuan DL, Chen Q (2010) Particle swarm optimisation algorithm with forgetting character. *International Journal of Bio-Inspired Computation* 2(1):59.
36. Tindall MJ, Gaffney EA, Maini PK, Armitage JP (2012) Theoretical insights into bacterial chemotaxis: Theoretical bacterial chemotaxis. *WIREs Syst Biol Med* 4(3):247–259.
37. Bifet A, Gavaldà R (2007) Learning from Time-Changing Data with Adaptive Windowing in *Proceedings of the 2007 SIAM International Conference on Data Mining*. (Society for Industrial and Applied Mathematics), pp. 443–448.
38. Kosina P, Gama J (2012) Handling time changing data with adaptive very fast decision rules in *Machine Learning and Knowledge Discovery in Databases*, eds. Flach PA, De Bie T, Cristianini N. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 827–842.
39. Rolls ET, Deco G (2015) Stochastic cortical neurodynamics underlying the memory and cognitive changes in aging. *Neurobiology of Learning and Memory* 118:150–161.
40. Ratcliff R, Rouder JN (1998) Modeling Response Times for Two-Choice Decisions. *Psychol Sci* 9(5):347–356.
41. Brunton BW, Botvinick MM, Brody CD (2013) Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science* 340(6128):95–98.
42. Miletic S, Boag R, Mathiopoulou V, Forstmann B (2019) Speed and accuracy in learning: A combined Q-learning diffusion decision model analysis in *2019 Conference on Cognitive Computational Neuroscience*. (Cognitive Computational Neuroscience, Berlin, Germany).
43. Kar S, Moura J (2009) Distributed Consensus Algorithms in Sensor Networks With Imperfect Communication: Link Failures and Channel Noise. *IEEE Trans. Signal Process.* 57(1):355–369.
44. Schweitzer ME, Cachon GP (2000) Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. *Management Science* 46(3):404–420.
45. Tregenza T (1995) Building on the Ideal Free Distribution in *Advances in Ecological Research*. (Elsevier) Vol. 26, pp. 253–307.
46. Musa HH, Noureldien A (2018) Comparing the ranking performance of page rank algorithm and weighted page rank algorithm. *Advanced Science Letters* 24(1):750–753.
47. Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. *Nature* 433(7025):513–516.
48. Franks NR, et al. (2007) Reconnaissance and latent learning in ants. *Proc. R. Soc. B.* 274(1617):1505–1509.
49. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113(4):700–765.
50. McNamara JM, Houston AI (1987) Memory and the efficient use of information. *Journal of Theoretical Biology* 125(4):385–395.
51. Donaldson-Matasci MC, Bergstrom CT, Lachmann M (2010) The fitness value of information. *Oikos* 119(2):219–230.
52. Krakauer DC (2011) Darwinian demons, evolutionary complexity, and information maximization. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 21(3):037110.
53. Kelly JL (1956) A New Interpretation of Information Rate. *the bell system technical journal* p. 10.
54. Cover TM, Thomas JA (2006) *Elements of Information Theory*. (John Wiley & Sons, Hoboken), Second edition.
55. Conti D, Mora T (2020) Non-equilibrium dynamics of adaptation in sensory systems. *arXiv:2011.09958 [nlin, q-bio]*.
56. West GB (1999) The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms. *Science* 284(5420):1677–1679.
57. White CR, Seymour RS (2003) Mammalian basal metabolic rate is proportional to body mass2/3. *Proceedings of the National Academy of Sciences* 100(7):4046–4049.
58. Burger JR, George MA, Leadbetter C, Shaikh F (2019) The allometry of brain size in mammals. *Journal of Mammalogy* 100(2):276–283.
59. West GB, Woodruff WH, Brown JH (2002) Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proceedings of the National Academy of Sciences* 99(Supplement 1):2473–2478.
60. Savage VM, Deeds EJ, Fontana W (2008) Sizing Up Allometric Scaling Theory. *PLoS Comput. Biol.* 4(9):e1000171.
61. Snell-Rood EC, Papaj DR, Gronenberg W (2009) Brain Size: A Global or Induced Cost of Learning? *Brain Behav Evol* 73(2):111–128.
62. Liefting M, Rohmann JL, Le Lann C, Ellers J (2019) What are the costs of learning? Modest trade-offs and constitutive costs do not set the price of fast associative learning ability in a parasitoid wasp. *Anim Cogn* 22(5):851–861.
63. Woude E, Groothuis J, Smid HM (2019) No gains for bigger brains: Functional and neuroanatomical consequences of relative brain size in a parasitic wasp. *J Evol Biol* p. jeb.13450.
64. Klyubin AS, Polani D, Nehaniv CL (2004) Tracking Information Flow through the Environment: Simple Cases of Stigmergy in *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, ed. Pollack J. (MIT Press).
65. Krakauer DC, Page KM, Erwin DH (2009) Diversity, Dilemmas, and Monopolies of Niche Construction. *The American Naturalist* 173(1):26–40.
66. Wen XL, Wen P, Dahlsjö CAL, Sillam-Dussès D, Šobotník J (2017) Breaking the cipher: Ant eavesdropping on the variational trail pheromone of its termite prey. *Proc. R. Soc. B.* 284(1853):20170121.
67. Smith CC, Reichman OJ (1984) The Evolution of Food Caching by Birds and Mammals. *Ann. Rev. Ecol. Syst.* 15:329–351.
68. Hall K, et al. (2017) Chimpanzee uses manipulative gaze cues to conceal and reveal information to foraging competitor. *Am J Primatol* 79(3):e22622.
69. Brännström Å, Johansson J, von Festenberg N (2013) The Hitchhiker's Guide to Adaptive Dynamics. *Games* 4(3):304–328.
70. Dieckmann U (year?) Ph.D. thesis.
71. Dieckmann U, Law R (1996) The dynamical theory of coevolution: A derivation from stochastic ecological processes. *J. Math. Biology* 34(5-6):579–612.
72. Poon P, Flack JC, Krakauer DC (2022) Institutional dynamics and learning networks. *PLoS ONE* 17(5):e0267688.
73. Brush ER, Krakauer DC, Flack JC (2018) Conflicts of interest improve collective computation of adaptive social structures. *Sci. Adv.* 4(1):e1603311.
74. Ramos-Fernandez G, Smith Aguilar SE, Krakauer DC, Flack JC (2020) Collective Computation in Animal Fission-Fusion Dynamics. *Front. Robot. AI* 7:90.
75. Lee ED, Daniels BC, Krakauer DC, Flack JC (2017) Collective memory in primate conflict implied by temporal scaling collapse. *J. R. Soc. Interface* 14(134):20170223.
76. North DC (2005) *Understanding the Process of Economic Change*. (Princeton University Press).
77. Flack JC, Girvan M, de Waal FBM, Krakauer DC (2006) Policing stabilizes construction of social niches in primates. *Nature* 439(7075):426–429.
78. Flack J (2017) Life's Information Hierarchy in *From Matter to Life: Information and Causality*, eds. Walker S, Davies PCW, Ellis GFR. (Cambridge University Press), pp. 283–302.
79. McNamara JM, Houston AI (1985) Optimal foraging and learning. *Journal of Theoretical*

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.01.506204; this version posted January 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Biology 117(2):231–249.

80. Gershman SJ, Wilson RC (2010) The Neural Costs of Optimal Control in *Advances in Neural Information Processing Systems.* (Curran Associates, Inc.), Vol. 23, pp. 712–720.
81. Fox E, Sudderth EB, Jordan MI, Willsky AS (2011) Bayesian Nonparametric Inference of Switching Dynamic Linear Models. *IEEE Trans. Signal Process.* 59(4):1569–1585.
82. Lee ED, Chen X, Daniels BC (2022) Discovering sparse control strategies in neural activity. *PLoS Comput Biol* 18(5):e1010072.
83. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) *Numerical Recipes: The Art of Scientific Computing.* (Cambridge University Press, New York), 3rd edition.