

RESEARCH ARTICLE

# Representational capacity shapes both the “whether” and “how” of social learning

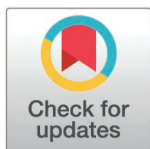
Max Taylor-Davies \*

School of Informatics, The University of Edinburgh, Edinburgh, United Kingdom,

\* [maxtaylordavies@gmail.com](mailto:maxtaylordavies@gmail.com)

## Abstract

Social learning is widely understood as offering a mechanism to mitigate the costs and risks of individual trial-and-error exploration. This cost-avoidance account implies a framing of social learning as a resource-rational adaptation, which should be most beneficial to populations with limited capacity to learn asocially. But in cases of peer-to-peer transmission, that same limited capacity may hinder the reliability of social information—rendering it less, not more, useful. So how do these conflicting intuitions resolve? Across a series of simulation experiments, we find evidence for an ‘inverse-U’ relationship where social information use emerges most strongly in populations with moderate capacity relative to the complexity of their environment. Furthermore, we demonstrate a co-evolutionary transition in social learning strategies: as population capacity rises, selection pressure shifts from favouring a success bias to a conformist approach. Our findings suggest that constraints on information-processing play an important role in the emergence of culture, determining not just if a species learns socially, but also how. All code for this project is available at <https://github.com/maxtaylordavies/capacity-and-social-learning>



## OPEN ACCESS

**Citation:** Taylor-Davies M (2026) Representational capacity shapes both the “whether” and “how” of social learning. PLoS Comput Biol 22(7): e1014504. <https://doi.org/10.1371/journal.pcbi.1014504>

**Editor:** Alireza Soltani, Dartmouth College Department of Psychological and Brain Sciences, UNITED STATES OF AMERICA

**Received:** March 22, 2026

**Accepted:** June 27, 2026

**Published:** July 15, 2026

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1014504>

**Copyright:** © 2026 Max Taylor-Davies. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

## Author summary

Many animals learn not only from their own trial-and-error exploration, but also by observing others. This “social learning” can be extremely useful, allowing young or inexperienced individuals to avoid the dangers of interacting with the environment themselves. This may be especially true for animals that are less able to learn by themselves due to limited attention or memory. But if these same limits affect those animals’ peers, the social information available will become less reliable, making social learning less, not more, useful. To reconcile these two opposing factors, we used computer simulations of agents learning which unfamiliar mushrooms were edible. We varied how much information agents could observe about each mushroom, and whether they learned alone or from others. We found that the benefits of social learning were strongest for neither

provided the original author and source are credited.

**Data availability statement:** Code for reproducing our results is available at <https://github.com/maxtaylordavies/capacity-and-social-learning>.

**Funding:** M.T.D is supported by the UKRI grant EP/S023208/1, EPSRC Centre for Doctoral Training in Robotics and Autonomous Systems (RAS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

the most nor least capable agents, but rather in between, where others were reliable enough to help but individual learning was still difficult. We also found that the same intervention influenced the best “strategy” for choosing whom to learn from: when agents could observe little information, it was best to copy the single most successful individual; when they could observe more information, it was best to follow the majority.

## Introduction

Why is social learning so taxonomically widespread [1–4]? What makes learning from others so beneficial, when most animals are perfectly capable of learning by themselves? The prevailing answer to this question goes back to Boyd & Richerson [5] and holds that asocial learning (via trial-and-error exploration) imposes certain costs; either through the baseline metabolic requirements of the exploration process itself, or more dramatically by exposing the learner to risks of serious harm. By leveraging the experience of their conspecifics, and thus reducing the amount of asocial exploration necessary to acquire adaptive behaviours, the social learner mitigates these costs to eke out a fitness advantage [6–8]. Take the example of mushroom foraging [9], where a novice must learn to consume safe species while avoiding poisonous ones. A trial-and-error approach, while not impossible, might kill an unlucky learner before they can reliably distinguish the two. More prudent would be to follow the cues given by an experienced forager, who (perhaps by learning from someone else in turn) already possess the relevant environmental knowledge [10–12].

Considering the general case where social information may be noisier or less reliable than direct first-hand experience, this cost-avoidance account bears a structural resemblance to the picture at the heart of resource-rational analysis—which hold that evolution shapes cognition towards strategies that trade off optimally between the rewards they produce and the costs involved in executing them [13–18]. This has led some to characterise social learning as being itself a resource-rational adaptation [19–21], with the implication that it should be most beneficial to agents with limited capacity to learn ‘by themselves’.

A large body of previous theoretical work has identified various factors that influence the adaptiveness of social learning. Much of this work has focused on the effect of environmental variability, finding for example that social learning is favoured when the rate of change is low enough that information does not become obsolete too quickly, but not so low that no learning is required [5,22–25]. Other studies have investigated the impact of task difficulty (typically via the size of the search space or the quality of available asocial evidence), demonstrating consistently that social learning is more strongly favoured in harder task environments [26,27].

Less attention has been paid to how properties of the learner themselves (rather than of the environment) affect the utility of learning from others, with consideration of this angle limited mostly to factors such as agents’ strategies for choosing whom to learn from or when to favour social over asocial learning, or the network

structure by which they are connected [6,28–30]. No prior work has yet given explicit consideration to the impact of limits on information-processing capacity. At first glance, it might be argued that this is just task difficulty under another name, and so does not merit separate study: as we decrease an agent's ability to process information and represent their environment, we increase the effective difficulty of the asocial learning problem they face—rendering social learning more attractive. But unlike most instances of resource-rational analysis, where it can make sense to consider the computational bounds of a single agent in isolation, the picture here is complicated by the fact that any realistic treatment of social learning must assume that the agents *providing* the social information are constrained in the same manner; any such capacity limit thus impacts the system in two distinct ways. This produces conflicting intuitions: decreasing capacity makes the asocial learning problem harder, which should favour social learning; at the same time, it reduces the quality of social information, which should favour *asocial* learning.

In this paper, we attempt to untangle the outcome of these two effects through a series of computational experiments. Using a contextual bandit problem modelled on the mushroom foraging setting, we investigate how the fitness advantage conferred by social learning varies as a function of agents' capacity to represent decision-relevant features, in environments of varying complexity. Our results suggest that the benefit of social learning neither increases nor decreases monotonically with capacity, but instead follows an 'inverse-U' relationship—where social information use emerges most strongly in populations of intermediate capacity relative to the complexity of their environment. Finally, we also investigate the relative performance of two social learning *strategies*, finding that increasing capacity produces a systematic transition from success bias to conformity.

## Methods

### Shared bandit problem

We model the foraging problem using a deterministic version of the *contextual bandit* paradigm [31], faced in parallel by each of a population of different agents. In a regular non-contextual bandit problem, an agent proceeds through a sequence of trials, at each of which they choose one of a number of 'arms' to pull, gaining some amount of reward as a result. The agent's task is to maximise cumulative payoff by estimating the reward distribution associated with each arm. In a contextual bandit, the only difference is that the agent additionally receives a piece of 'context'  $x_t$  at the start of each trial, the values of which correspond in some way to different settings of arm-reward probabilities. The agent must therefore learn a mapping from context value to optimal arm choice (e.g., "when  $x_t = 0$  pull arm 1, when  $x_t = 1$  pull arm 2").

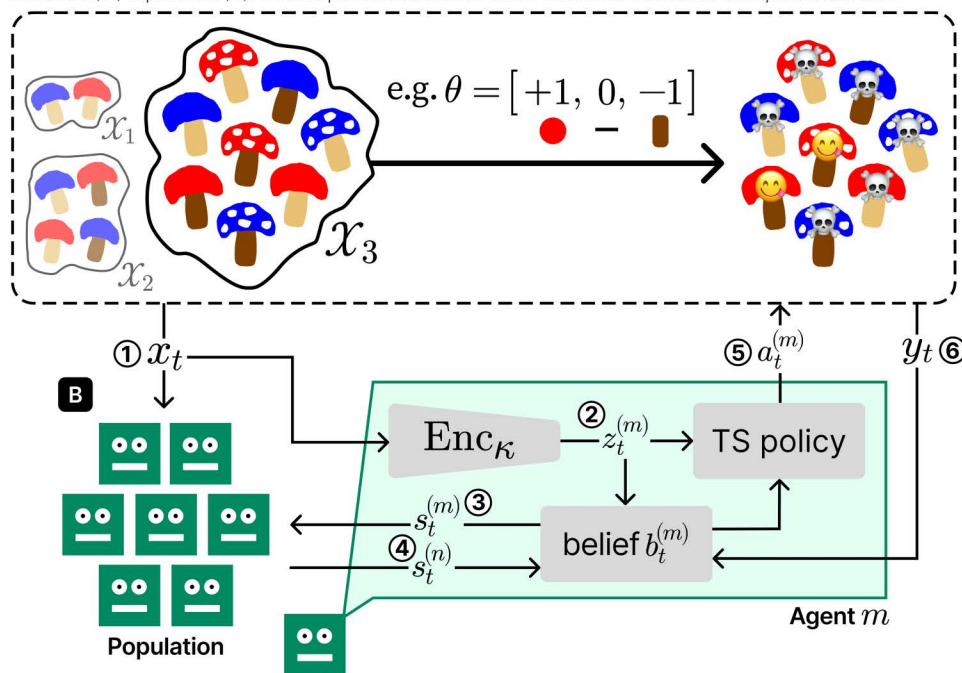
In our task, each context value  $x_t$  corresponds to the distinguishing features of a particular mushroom species. Specifically, we say that each species is characterised by a vector of  $n$  binary features  $x \in \mathcal{X}_n = \{-1, +1\}^n$  (i.e., an  $n$ -feature environment contains  $2^n$  different species). These features can be interpreted as referring to things like the cap colour of the mushroom, whether it has spots, its size, etc. A deterministic function  $f_\theta$  maps this feature space to a binary label of +1 (edible) or -1 (poisonous):

$$f_\theta(x) = \begin{cases} +1 & \text{if } x \cdot \theta = \|\theta\|_0 \\ -1 & \text{otherwise} \end{cases} \quad (1)$$

where  $\|\cdot\|_0$  returns the number of nonzero elements in its argument, and  $\theta \in \Theta_n = \{-1, 0, +1\}^n$  corresponds to a logical rule over some subset of context features. The allowance of zeros means that rules can 'ignore' certain features; for example,  $\theta = [+1 \ 0 \ -1]$  means that a species  $x$  must have  $x_1 = +1$ ,  $x_3 = -1$  and either value of  $x_2$  in order to be classified as edible (see Fig 1A for an illustration of this example). We note also that  $\theta$  is fixed for each instantiation of the task environment, so that a given population only ever experiences a single rule; all our results will be reported for simulation runs aggregated over many rules sampled uniformly from  $\Theta_n$ .

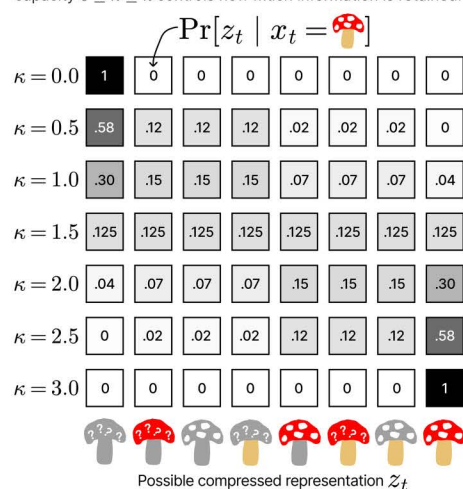
## A Environment

Mushroom feature vectors  $x$  are drawn from set  $\mathcal{X}_n$  where  $n$  is the number of binary features. A deterministic rule  $\theta$  classifies each  $x$  as edible (+1) or poisonous (-1). The example rule shown here is "edible mushrooms must have a red cap and a dark stem".



## C Feature encoder

Each agent forms a compressed representation  $z$  of feature vector  $x$  by discarding randomly selected features. The capacity  $0 \leq \kappa \leq n$  controls how much information is retained.



**Fig 1. A high-level diagram of the bandit task and agent model used in our simulations.** (A) An illustration of the environment for  $n=3$  and a single example rule. (B) A simplified depiction of a single trial from the perspective of a single agent  $m$  who is an asocial or belief-based social learner. The steps shown are: [1] population receives context  $x_t$  [2] agent  $m$  forms lossy representation  $z_t^{(m)}$  [3] agent  $m$  emits social cue  $s_t^{(m)}$  [4] if agent  $m$  is a social learner, they update their beliefs using cue  $s_t^{(n)}$  from agent  $n \neq m$  [5] agent  $m$  chooses arm  $a_t^{(m)}$  [6] iff  $a_t^{(m)} = \text{eat}$ , agent  $m$  updates beliefs with ground truth feedback  $y_t$ . If agent  $m$  were instead a *policy-based* social learner, they would simply copy agent  $n$ 's action directly. (C) The distributions over encoder outputs  $z_t$  for illustrative values of  $\kappa$  and a single example  $x_t$ .

<https://doi.org/10.1371/journal.pcbi.1014504.g001>

The task proceeds trial-by-trial: at trial  $t \in [0, T-1]$  each agent  $m$  in the population has fitness  $h_t^{(m)}$ , receives shared context  $x_t \in \mathcal{X}_n$  (Rather than sampling  $x_t$  uniformly from  $\mathcal{X}_n$  we apply a class-balanced correction, so that  $p(x_t \text{ edible} | \theta) = p(x_t \text{ poisonous} | \theta) = 0.5$ . This ensures that the survival challenge posed by a given environment is purely determined by the difficulty of the learning problem (a function of  $n$  and capacity), not by how hazardous it would be for a randomly acting agent.), independently selects arm  $a_t^{(m)} \in \{\text{eat}, \text{discard}\}$  and receives reward  $r_t^{(m)}$  according to

$$r(x, a; \theta) = \begin{cases} +1 & \text{if } a = \text{eat} \wedge f_\theta(x) = +1 \\ -c & \text{if } a = \text{eat} \wedge f_\theta(x) = -1 \\ 0 & \text{if } a = \text{discard} \end{cases} \quad (2)$$

i.e., there is a 'safe' arm (that always produces zero reward) and a 'risky' arm that produces rewards of +1 or  $-c$  for different contexts. The fitness of each agent in the population is then updated as

$$h_{t+1}^{(m)} = \min(h_t^{(m)} + r_t^{(m)} - h_{\text{const}}, h_{\text{max}}) \quad (3)$$



where  $h_{\max}$  is the maximum achievable fitness and  $h_{\text{const}}$  is the baseline rate of fitness loss. If  $h_{t+1}^{(m)} < 0$ , agent  $m$  ‘dies’ and is removed from the environment. Unless otherwise specified, we will use  $c = 1$ ,  $h_0 = 5$ ,  $h_{\max} = 10$ ,  $h_{\text{const}} = 0.01$  throughout all experiments; see [S1 Text](#) for a sensitivity analysis of these and other parameters.

### Asocial learning as Thompson sampling

Thompson Sampling (TS) is a principled and widely used approach for balancing exploration and exploitation in bandit problems [32,33], and has also been used to model human behaviour in tasks similar to ours [34,35]. A TS agent maintains a posterior belief  $b_t(\theta)$  over the parameters of their environment; at each trial, they sample a single point  $\tilde{\theta}_t$  from  $b_t$ , and then select the arm that maximises expected reward under  $\tilde{\theta}_t$ . The resulting environment feedback is then used to form updated belief  $b_{t+1}(\theta)$ . In our case, each agent in the population maintains an independent belief  $b_t^{(m)}(\theta)$  about the rule  $\theta$  that controls which species are edible via  $f_\theta$ :

$$\begin{aligned} \tilde{\theta}_t^{(m)} &\sim b_t^{(m)}(\theta) \\ a_t^{(m)} &= \begin{cases} \text{eat} & \text{if } f_{\tilde{\theta}_t^{(m)}}(x_t) = +1 \\ \text{discard} & \text{if } f_{\tilde{\theta}_t^{(m)}}(x_t) = -1 \end{cases} \end{aligned} \quad (4)$$

Since our goal is to study the impact of capacity constraints, we must somehow limit agents’ ability to process information. There are multiple ways we could approach this within the TS framework; we do so in terms of the *number of features* (in  $x_t$ ) that agents can perceive during each trial. This choice of implementation is motivated by a focus on *representational* capacity: that is, agents’ ability to represent the task-relevant features of their environment. This is distinct from, e.g., functional capacity—regardless of how much information is retained from  $x_t$ , all agents are equally capable of performing the computations necessary to evaluate rewards, update beliefs, select actions, etc.

Formally, given input  $x_t$ , we say that each agent forms an individual internal representation  $z_t^{(m)} \sim \text{Enc}_\kappa(x_t)$  where  $\text{Enc}_\kappa$  preserves on average  $\kappa$  bits of information per  $x_t$ . E.g. when  $\kappa = 0$ , agents are unable to process *any* input features; when  $\kappa \geq n$  agents can process all incoming information; when  $\kappa = \frac{n}{2}$  each agent perceives on average half of the features at each trial (see [Fig 1C](#)). Note that features are included or excluded independently with uniform probability, so an agent may perceive one feature at trial  $t$  and a different feature at trial  $t+1$ .

Our encoder-based implementation follows work in resource-rational analysis that uses the information bottleneck framework as a principled way to model capacity limits in human and animal cognition [18,36–39]. In [S1 Text](#), we report results for two related but alternative ways of degrading agents’ access to information; again, these variants constitute different representational bottlenecks rather than an exhaustive survey of possible capacity-related constructs.

Action selection then proceeds on the basis of  $z_t^{(m)}$  in place of  $x_t$ , with [Eq \(4\)](#) becoming (We reuse the  $f_\theta(\cdot)$  notation here for convenience, though technically the function that agents compute is slightly different to [Eq \(1\)](#), as it ignores elements of  $\theta$  that line up with zeroed-out entries in  $z_t^{(m)}$ .)

$$\begin{aligned} \tilde{\theta}_t^{(m)} &\sim b_t^{(m)}(\theta) \\ a_t^{(m)} &= \begin{cases} \text{eat} & \text{if } f_{\tilde{\theta}_t^{(m)}}(z_t^{(m)}) = +1 \\ \text{discard} & \text{if } f_{\tilde{\theta}_t^{(m)}}(z_t^{(m)}) = -1 \end{cases} \end{aligned} \quad (5)$$

If (and only if) an agent chooses  $a_t^{(m)} = \text{eat}$ , they obtain the ‘ground truth’ category label  $y_t = f_\theta(x_t)$  and update their beliefs by suppressing the probability of all rules inconsistent with datapoint  $(z_t^{(m)}, y_t)$ :

$$\begin{aligned} \log b_{t+1}^{(m)}(\theta) &= \text{update}(b_t^{(m)}(\theta), y_t, \epsilon) \\ &= \begin{cases} \log b_t^{(m)}(\theta) + \epsilon & \text{if } f_\theta(z_t^{(m)}) = y_t \\ \log b_t^{(m)}(\theta) + (1 - \epsilon) & \text{otherwise} \end{cases} \end{aligned} \quad (6)$$

where  $\epsilon \in [0, 0.5]$  controls how aggressively beliefs are updated (we will use  $\epsilon = 0.05$  throughout all experiments) (Note that the use of representation  $z_t^{(m)}$  in Eq (6) means that a given rule  $\theta$  might make an incorrect prediction for  $x_t$  but still not be downweighted, if the features responsible for the mismatch are not present in  $z_t^{(m)}$ ). It is worth emphasising here that the true (uncompressed) feature vector  $x_t$  plays no direct role in the agent's policy: *both* action selection *and* belief revision use only each agent's lossy representation  $z_t^{(m)}$ , ensuring that  $\kappa$  exerts a 'global' influence over agents' ability to learn and make decisions.

As an illustration of how complexity (number of features  $n$ ) and representational capacity ( $\kappa$ ) interact to determine the difficulty of the asocial learning problem, Fig 2 shows the fitness over time of different capacity-limited agents in the least ( $n=2$ ) and most ( $n=6$ ) complex environments. In both environments, we see a consistent positive relationship between capacity and fitness. We also see a significant difference between the two complexity levels: at  $n=6$ , even agents with unlimited capacity ( $\kappa = n$ ) are unable on average to reach the fitness ceiling  $h_{\max}$ .

### Adding social information use

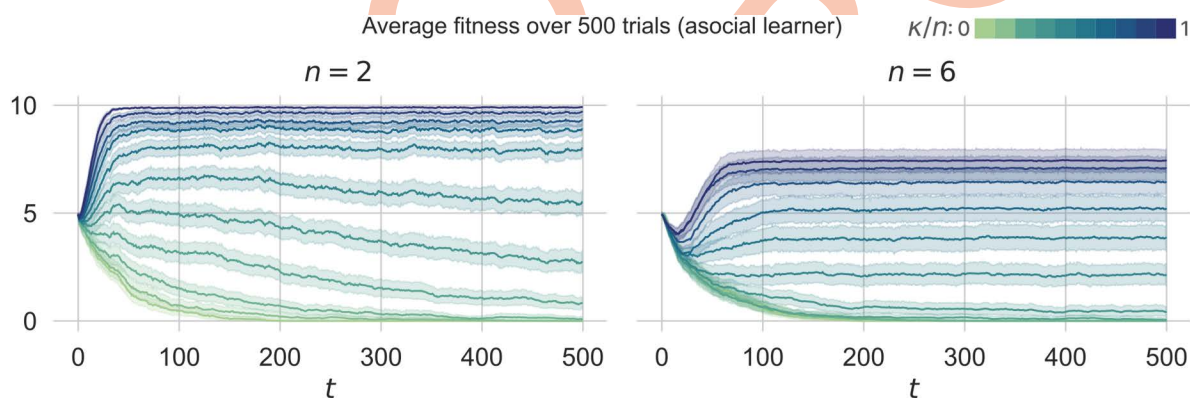
Eqs (5)–(6) define our *asocial* learning process. As with capacity, there are a number of different ways we could model social information use, each inviting different cognitive or biological interpretations. We implement two model variants that aim to capture distinct levels of the social learning hierarchy [6,40,41]. However, we do not claim that these two variants are exhaustive; generalising our experiments to additional forms of social learning would be a worthwhile direction for future work.

### Policy-based social learning

One of the simplest ways that an agent can learn socially is through direct imitation. In this variant, the social learner simply copies the action choice (eat or discard) from another agent chosen uniformly at random. This represents a maximalist version of social learning, in the sense that it entirely replaces the asocial learning and decision-making processes.

### Belief-based social learning

A more nuanced approach is for social information to influence learners' beliefs about the world (i.e.,  $\theta$ ), while action selection still proceeds according to Eq (5). We model this as follows. Immediately after forming representation  $z_t^{(m)}$  of



**Fig 2. The fitness over 500 trials of a purely asocial learner with different levels of representational capacity ( $\kappa$ ) in the simplest (left) and most complex (right) task environments.** Each line shows the mean over 50 rules sampled uniformly from  $\Theta_n$  (with shaded areas giving bootstrapped 95% CIs).

<https://doi.org/10.1371/journal.pcbi.1014504.g002>

context vector  $x_t$ , each agent in the population expresses a *social cue*  $s_t^{(m)}$  capturing their ‘best guess’ about whether  $x_t$  is edible or poisonous:

$$e_t^{(m)} = \mathbb{E}_{b_t^{(m)}}[y_t] = \sum_{\theta} f_{\theta}(z_t^{(m)}) b_t^{(m)}(\theta)$$

$$s_t^{(m)} = \begin{cases} +1 & \text{if } e_t^{(m)} > 0 \\ -1 & \text{otherwise} \end{cases} \quad (7)$$

If agent  $m$  is a belief-based social learner, they then select another agent  $n$  uniformly at random, and perform a pre-action belief update following Eq (6) as  $\log b_t^{(m)}(\theta) = \text{update}(b_t^{(m)}(\theta), s_t^{(n)}, \epsilon)$ —that is, social and asocial learning operate under this model via identical ‘mechanics’, with the only difference being that the social belief update uses cue  $s$  in place of ground truth  $y$ . Crucially, the ‘social belief update’ takes place before the learner acts, and is independent of which arm is subsequently chosen; it thus captures the idea that social learning provides a complementary source of information that sidesteps the need for direct agent-environment interaction. While we intentionally treat the cue  $s$  as an abstract signal of a particular agent’s beliefs, it could be interpreted either as something that is *inferred* by the social learner (on the basis of observed actions), or as the content of some explicit communication between agents (such as an utterance or gesture).

Note that for now we are modelling all social learners as using *unbiased transmission*; that is, whether following the policy-based or belief-based approach, they are equally likely to learn from any other member of the population. This approach is also sometimes referred to in the literature as ‘linear imitation’, since the probability of a social learner updating on cue value  $s$  or copying action  $a$  scales linearly with the frequency of  $s$  or  $a$  in the population. While decades of research have established that humans and other animals are in fact biased social learners, preferring information from some conspecifics over others [5,6,22,42–44], our focus on the *initial* emergence of social information use means that an indiscriminate ‘strategy’ is appropriate. We will relax this assumption in our final two experiments (Sections -).

## Evolutionary dynamics

Our evolutionary simulations are based on the Moran process [45,46], a simple model of selection in finite populations. In the Moran process, individuals are characterised by a set of fixed types; at each step, a single individual is chosen uniformly at random to die, replaced by a newborn whose parent is selected with probability proportional to their fitness, and who inherits their parent’s type (with some small probability of mutation  $p_{\text{mut}}$ )—thus maintaining a constant population size. We use the same “one out, one in” reproductive dynamics, but agents in our simulation die when their fitness falls to zero (with background probability  $p_{\text{death}}$  of dying randomly before then)—so any given timestep may see no deaths, one death, or multiple deaths. We will use  $p_{\text{mut}} = 0.01$ ,  $p_{\text{death}} = 0.001$  throughout our experiments—for a sensitivity analysis of these (and other) parameters, see S1 Text.

In our simulations, each agent’s type corresponds to and determines their learning behaviour: fully-asocial or unbiased-social in Experiment 2 (Section); the same plus conformist-social and copy-fittest-social in Experiments 3–4 (Sections -). In this way we are adopting a form of the phenotypic gambit [47,48], allowing a single mutation to transform a purely asocial learner into a social information user, or an unbiased social learner into a strategic one. Clearly, this abstracts away a huge amount of complexity; as a simplifying assumption, it is necessary to make practical the kinds of experiments we wish to run.

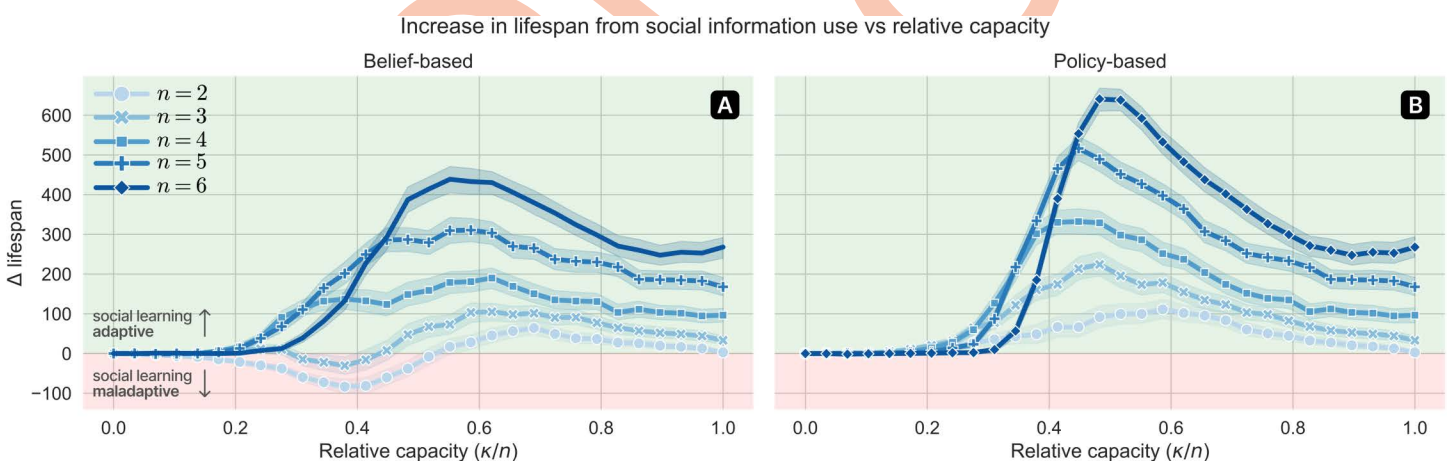
Since newborn agents inherit no actual *knowledge* from their parents, the overall process can in one sense be viewed as implementing a crude form of meta-learning: in the ‘inner loop’, each agent begins life as a blank slate and must learn the correct mushroom-eating behaviour in order to survive; in the ‘outer loop’, a selection pressure shaped by the combination of complexity and capacity pushes agents towards the most effective learning strategy.

## Results

### Capacity shapes the utility of social information

We first describe the results of a preliminary experiment in which we assess the impact on survival from social information use under different conditions of environment complexity ( $n$ ) and representational capacity ( $\kappa$ ). We simulate a background population of 100 *asocial-only* learners (i.e., their types are fixed, with mutation disabled) for  $T = 1500$  trials; after a warmup period of 500 trials, we ‘drop in’ two new agents—one that uses social information and one that doesn’t—and then record the difference in their subsequent lifespans. This is intended to capture the scenario of a single mutant social learner appearing within an established population of organisms that otherwise learn only via asocial interaction with their environment: how does the capacity of this population to process information determine the fitness advantage enjoyed by the mutant individual? To ensure that our results are robust, we run 500 separate simulations for each  $(n, \kappa)$  pair, over rules sampled uniformly from  $\Theta_n$ .

Earlier, we posed the question of whether the value of social information increases or decreases with a population’s representational capacity. We now offer initial evidence that, to some extent, *both* of these relationships hold true. Fig 3 shows the average excess lifespan for the social learner as a function of  $\kappa$ , for each value of  $n$ . Rather than either a monotonic increase or decrease, we instead observe for both social learning variants an ‘inverse-U’ relationship: at very low  $\kappa$  all agent types are ~equally bad at surviving; as we increase  $\kappa$  the social learner ekes out an advantage, which peaks before decreasing as  $\kappa$  approaches  $n$ . In the case of belief-based social learning (panel A) we also observe an interesting interaction between  $n$  and  $\kappa$ : in low-complexity environments ( $n = 2, 3$ ) there is a region of the curve ( $0 < \kappa \lesssim n/2$ ) where social information use is actually *maladaptive*, and the social learner is worse at surviving than their social-information-ignoring counterpart. This effect disappears in more complex environments ( $n \geq 4$ ), where at low capacity the asocial learning problem is so difficult that updating on unreliable cues makes little difference. Interestingly, an equivalent maladaptive region is not observed under policy-based social learning. This difference is explained by the fact that receiving an incorrect cue at trial  $t$  can “contaminate” the belief-based learner’s posterior, and thus their future behaviour—which can lead them to perform worse than they would have via asocial learning. By contrast, copying an incorrect action at time  $t$  impacts only the policy-based social learner’s performance on trial  $t$ , with no downstream consequences. Policy-based



**Fig 3. Results of Experiment 1.** The difference in lifespan (over 1000 timesteps) for an agent that does/doesn’t engage in policy-based (panel A) or belief-based (panel B) social learning, in a population otherwise composed entirely of asocial-only learners. Each point represents the mean over 500 simulation runs, with rules sampled uniformly from  $\Theta_n$  (shaded areas give bootstrapped 95% CIs).

<https://doi.org/10.1371/journal.pcbi.1014504.g003>



social learning may therefore be *ineffective* (i.e., provide zero advantage), but if actions are being copied solely from asocial learners, it cannot be actively maladaptive.

### Capacity shapes the emergence of social learning

To test whether the trend observed in Experiment 1 translates into an effect on the *emergence* of social information use, we now move beyond our analysis of the single-mutant case and allow all agents to occupy either type. Beginning with an initial population of 200 non-social-learners (where social learners can this time appear by mutation), each  $(n, \kappa, \theta)$  simulation is run for  $T = 10^4$  trials, after which we record the number of social learners in the population. We also run a comparable set of control simulations in which all social information use is disabled so that both agent types have identical learning behaviour. This allows us to estimate the expected type proportions that arise under neutral drift, solely from the effect of capacity on each population's turnover rate; subtracting these baseline proportions yields an estimate of the actual selection pressure.

As Fig 4 shows, we observe a similar qualitative pattern to that found in Experiment 1: selection for social learners varies roughly as an inverse-U-shaped function of capacity; and is stronger in more complex environments (higher  $n$ ). We also find again that, under certain combinations of  $n$  and  $\kappa$ , belief-based social learning is actually maladaptive. Interestingly—in contrast to Experiment 1—we now also observe a small amount of negative selection in the policy-based case. This is almost certainly the result of a population-level dynamic akin to Rogers' paradox [22], where a too-high frequency of imitators effectively parasitises the supply of new information. That this effect is weaker for policy-based than belief-based social learning reflects the fact that the latter can be rendered maladaptive by *both* individual-level (belief contamination) and population-level mechanisms. This combination of factors also explains why the maladaptive region for belief-based social learning is larger than in Experiment 1, persisting at higher values of  $n$ .

### Capacity shifts the balance between social learning strategies

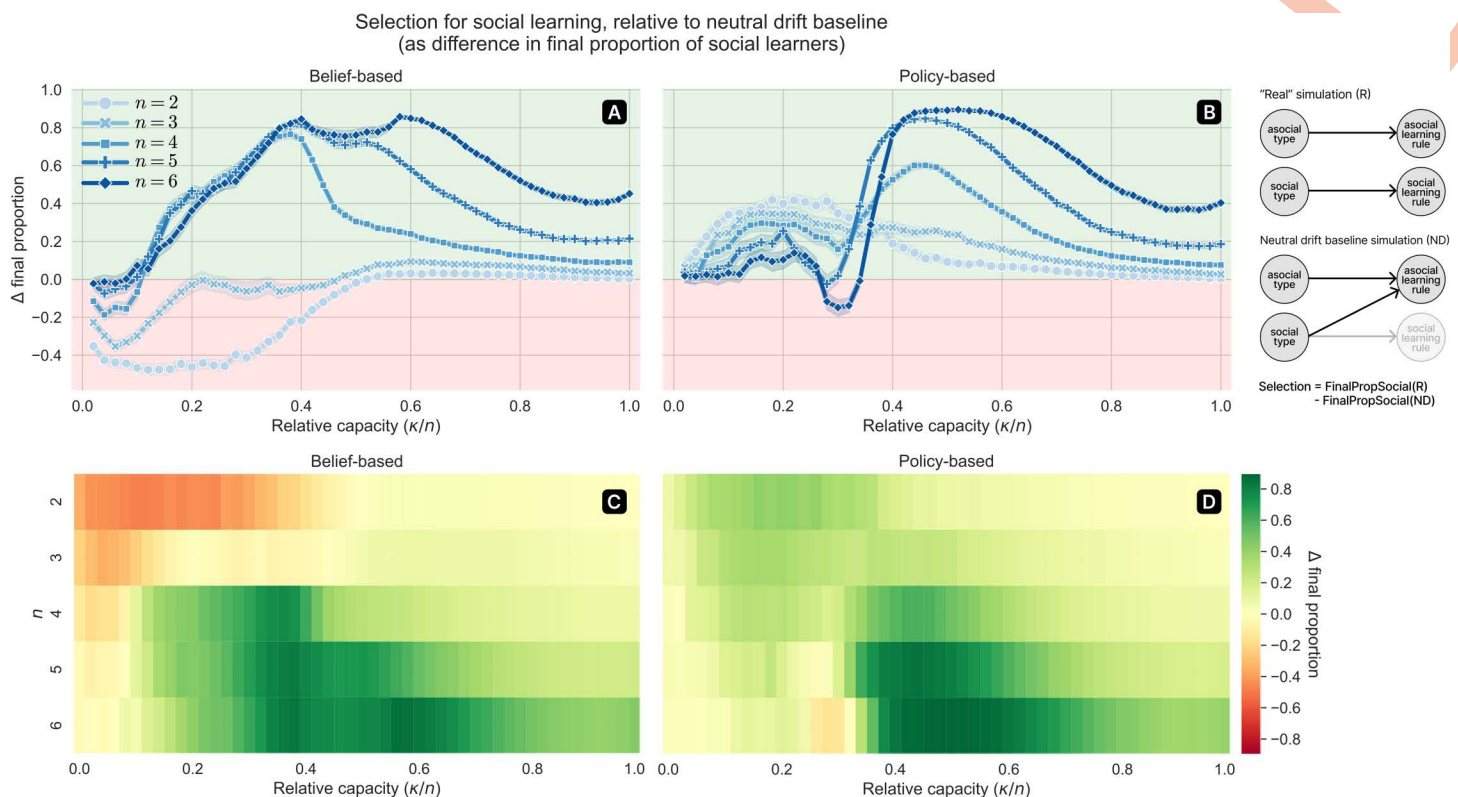
One limitation of our first and second experiments is their restriction of all social learners to simple unbiased transmission—which arguably describes only a very small minority of the social learning behaviour observed 'in the wild' [6,42,44,49]. It would be interesting to try and capture any interactions that capacity limits might have with the *strategies* that human and animal social learners rely on to target their learning towards particular conspecifics.

In our third experiment, we extend the model to accommodate two additional such strategies that are well-studied both empirically and theoretically [5,50–52]. We implement these strategies as different distributions over potential demonstrators, while otherwise holding fixed the content and mechanics of social learning. That is, all social learners still receive either a cue or action as described earlier; they differ in how they select the source of that cue or action. Formally, each strategy is implemented as a particular distribution over which 'other' agent  $n \neq m$  is selected by agent  $m$  as a source of social information:

$$p_{\text{unbiased}}(n) = \frac{1}{M-1} \quad (8)$$

$$p_{\text{conformist}}(n) \propto \begin{cases} 1 & \text{if } |i \neq m : s_t^{(i)} = s_t^{(n)}| > \frac{M-1}{2} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$p_{\text{fittest}}(n) \propto \begin{cases} 1 & \text{if } h_t^{(n)} = \max_{i \neq m} h_t^{(i)} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$



**Fig 4. Results of Experiment 2.** Empirical selection for social learning under different combinations of complexity ( $n$ ) and capacity ( $\kappa$ ). Selection is measured as the difference in final population proportion of social learners between simulations where the behavioural effect of the ‘social’ type was enabled (“real” simulation) or disabled (neutral drift baseline). Panels A and C show results for the belief-based social learning variant; panels B and D for the policy-based variant. Each point represents the mean over 500 independent simulation runs, with rules sampled uniformly from  $\Theta_n$ . All simulations were run for  $10^4$  timesteps.

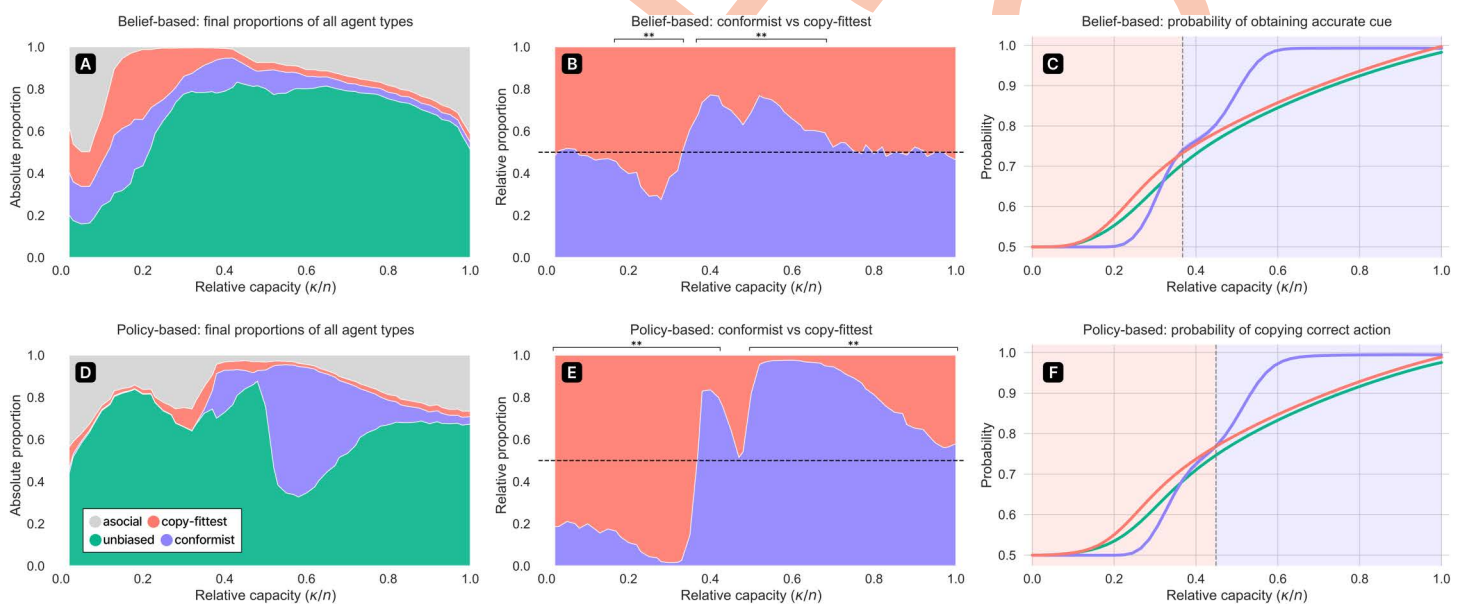
<https://doi.org/10.1371/journal.pcbi.1014504.g004>

where  $M$  is the population size and  $|\cdot|$  returns the number of elements in a set. Eqs (8)–(10) are given in terms of the belief-based model, but identical formulations are used for the policy-based variant, where all  $s$  terms are simply replaced with  $a$  terms. The unbiased learner (Eq (8)) samples uniformly from all its peers, as in Experiments 1–2. The conformist learner (Eq (9)) first identifies which cue or action is expressed by the majority of its peers, and then samples uniformly from agents expressing that value (Note that since there are only two possible values for both cues and actions, and we always use an even-numbered  $M$  (and therefore odd  $M - 1$ ), the conformist strategy doesn’t need to handle the possibility of there being no majority.). The copy-fittest learner (Eq (10)) samples uniformly from the currently fittest agent(s) among its peers. These hard-threshold rules should be read as idealised limiting cases of a broader space of strategies in which selection probability varies smoothly with frequency or fitness/success. While our motivation was to avoid introducing additional free parameters, we note that there is some precedent for these formulations [26,51,53].

Keeping all else unchanged, we re-ran the evolutionary simulations from Experiment 2, this time using an initial equal split of asocial and unbiased social learners, and focusing on the intermediate-complexity ( $n=4$ ) environment. This is intended to capture the scenario where ‘strategic’ (i.e., biased) social information use appears by mutation within a population in which unbiased social learning is moderately well established. As before, we report the final proportion of each type in the population following  $T=10^4$  trials.

Panels B and D of Fig 5 illustrate an interesting result: at lower capacity, social learners using the ‘copy-fittest’ strategy are favoured relative to those using the conformist strategy; at higher capacity the reverse is true. The intuition for this pattern is as follows. When capacity is low, large inter-agent variance in knowledge (and thus fitness) can arise simply from a few individuals getting ‘lucky’—variance which is successfully exploited by the copy-fittest strategy. At higher capacity, where most agents are making predictions above chance, these differences in knowledge are less pronounced, and are outweighed by the insulation that conformity provides against individual noise. To validate this intuition, we estimated the probabilities that strategic belief-based social learners obtain accurate social cues (i.e.,  $\Pr[s_t^{(n)} = y_t]$ ; see S1 Text for details on how this is estimated). Similarly, for policy-based social learners, we estimated the probability of each strategy copying the correct action. As suggested by panels B and D, panels C and F confirm that at low capacity these probabilities are highest for copy-fittest, but that there is a crossover point after which it is always highest for conformity. Strikingly, this analysis further reveals that in the low-capacity regime, simple unbiased transmission is *also* likelier to obtain both accurate cues and correct actions than conformity—suggesting that under some conditions, an apparently naïve approach to social learning can be preferable to a more strategic one.

As in previous experiments, it is worth noting that we again see a difference between the belief-based and policy-based variants; with relative selection for both strategies, and absolute selection for conformism, considerably stronger in the policy-based case. One plausible explanation for this is simply that the difference between conformist and copy-fittest strategies is sharper in the policy-based model, where the impact of social information on behaviour is much more *direct*. The more roundabout nature of the belief-based approach might mean that the benefits of either strategy are softened—especially at low capacity.



**Fig 5. Results of Experiment 3.** (A) the final proportion of each agent type in the population after  $10^4$  timesteps of evolutionary simulation under different values of capacity ( $\kappa$ ), in the  $n=4$  environment and using the belief-based social learning model. Proportions were aggregated over 500 independent simulation runs, with rules sampled uniformly from  $\Theta_4$ . (B) from the same data, the relative proportions of conformist and copy-fittest strategies. Brackets indicate  $p$ -values  $< 0.01$ , using a Wilcoxon signed-rank test with a null hypothesis of equal proportions. (C) the estimated probability that the cues obtained by each social learning strategy are accurate (i.e.,  $s_t = y_t$ ), based on a simulation where no social learning actually occurs. Shaded regions denote timesteps at which copy-fittest learners outnumber conformists and vice versa. Panels D-F are equivalent to A-C but for the policy-based social learning model.

<https://doi.org/10.1371/journal.pcbi.1014504.g005>

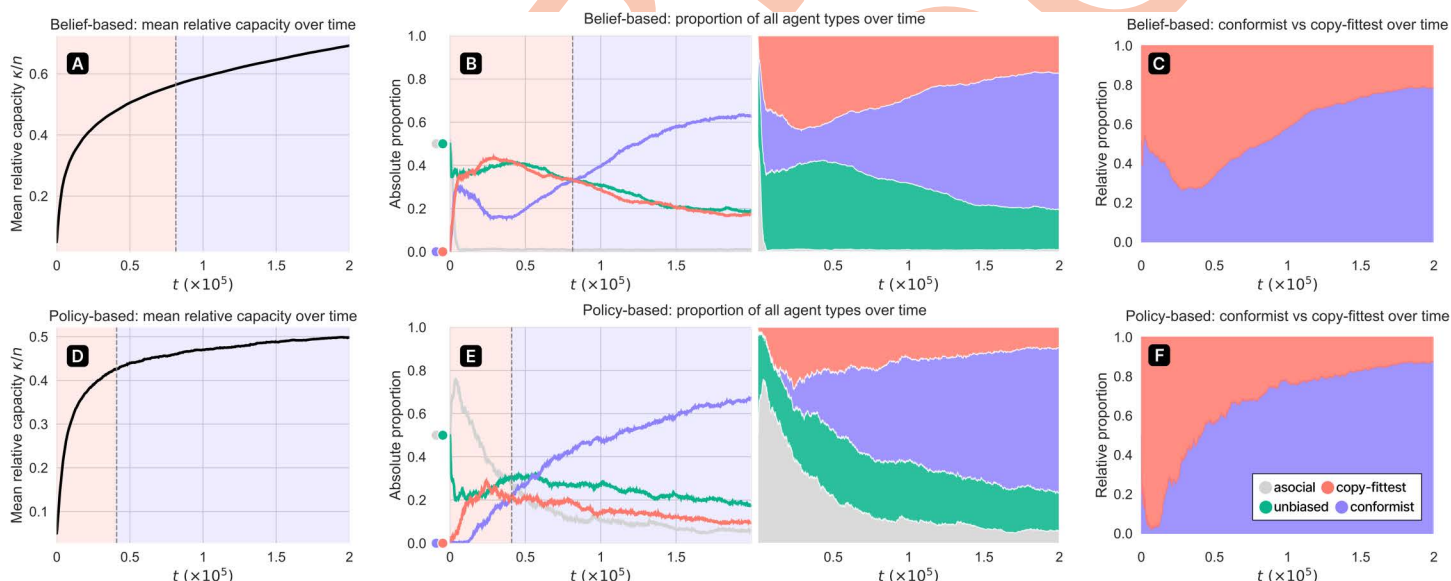
## Learning strategies co-evolve with capacity

All of the evolutionary simulations reported thus far have held each population's representational capacity constant, while allowing only agents' learning types to evolve. The results of Experiment 3 suggest that, within a population whose capacity evolves upwards over time, and which initially contains no 'strategic' social learners, we should expect to see two separate transitions; where copy-fittest social learners first obtain a majority before themselves being overtaken by conformists. For our final experiment, we test this prediction by jointly evolving both capacity and agent types within a single population. To capture the intuition that capacity, as a more 'fundamental' attribute of cognition, should evolve more slowly than specific learning strategies, we allow  $\kappa$  to mutate only every 20 timesteps:

$$\kappa_0 \sim \mathcal{N}(0.2, 0.05)$$

$$\kappa_{\text{offspring}} = \begin{cases} \kappa_{\text{parent}} + \mathcal{N}(0, 0.05) & \text{if } 20 \mid t \\ \kappa_{\text{parent}} & \text{otherwise} \end{cases} \quad (11)$$

The results of Experiment 4 are shown in Fig 6. As we would expect, mean capacity (panels A,D) increases monotonically over time, while panels B,C,E,F confirm that we do indeed observe the two predicted transitions. Initially, copy-fittest agents are more effective at invading the population. These dynamics are then steadily reversed, with conformists soon gaining a plurality that is retained for the remainder of the simulation time. Interestingly, the frequency of asocial learners crashes to  $\approx 0$  almost immediately in the belief-based model (panel B), but follows a much steadier decline in the policy-based model (panel E). This is almost certainly a consequence of the same Rogers'-paradox-like dynamics that contributed to negative selection in Experiment 2. The difference arises because belief-based social learners are not pure free-riders; while they use information from others to update their beliefs, they still choose actions and emit their own cues independently. By contrast, policy-based social learners are pure free-riders, contributing no independent information



**Fig 6. Timeseries data from agents undergoing simultaneous co-evolution of representational capacity and learning type, in the  $n=4$  environment.** (A) The change in average capacity over time. (B) The proportion of each agent type in the population over time. (C) The relative proportions of conformist and copy-fittest social learners over time. All plots report mean values over 200 independent simulation runs, with rules sampled uniformly from  $\Theta_4$ . Shaded regions of panels A-B denote timesteps at which copy-fittest learners outnumber conformists and vice versa.

<https://doi.org/10.1371/journal.pcbi.1014504.g006>



back to the common pool. Asocial learners are thus more valuable to the population under policy-based than belief-based social learning; leading us to observe a much stronger frequency-based effect in the former case.

## Discussion

Many animals, either in infancy or as a response to migration or environmental change, must at some point learn about an unfamiliar world while avoiding the various dangers that it poses. Faced with this dilemma, social learning offers a valuable way to acquire adaptive behaviours while mitigating the cost and risks of trial-and-error exploration. In understanding the factors that strengthen or weaken this effect, previous work has studied in depth the role of environment properties such as temporal variability [5,22–25] and task difficulty [26,27]. Where consideration has been given instead to the impact of *agent*-centric properties, it has mostly been in terms of different social learning strategies, balances between social and asocial learning, or network structures [28, 6, 29, 30]. Less attention has been given to the role of more ‘fundamental’ or domain-general properties of cognition—which, by virtue of being slower to evolve, ought to exert influence over the emergence of narrower adaptations such as whether or how learners integrate social information.

In this paper, as a step towards a broader understanding of how agent and environment properties interact to determine which learning strategies are most successful, we have explored how the evolutionary benefits of social learning depend on agents’ ability to represent task-relevant information. Our focus on capacity is motivated by the resemblance of the cost-avoidance argument to ideas in resource-rational analysis [15]. While this resemblance has led some to characterise social information use as a resource-rational solution to the problem of learning in complex and dangerous environments [19–21], the relationship between information-processing capacity and the utility of social learning had not previously been tested.

Such an exploration may initially seem redundant: decreasing an agent’s ability to process or represent information is surely equivalent to increasing the difficulty of the asocial learning problem, and work such as [26,27] already tells us what to expect as a result. But this is only true if we assume that the learning agent’s *sources* of social information are not similarly constrained. Since any realistic treatment must reject this assumption, the picture becomes less straightforward: on one hand, decreasing capacity makes the asocial learning problem harder, which should favour social learning; on the other hand, it reduces the quality of social information, which should favour asocial learning. Through a series of simulation experiments, we have shown that the combined result of these two effects is that the adaptiveness of social learning follows an ‘inverse-U’ relationship with capacity, decreasing towards both extremes of the spectrum.

How should this result be interpreted? Whether social learning operates as an auxiliary information source complementary to the primary channel of asocial feedback (as in our belief-based model) or as a surrogate policy that replaces individual exploration (as in our policy-based model), its value will depend on the balance of two factors: the amount of reliable information conveyed, and the extent of the learner’s marginal need for such information. The first of these factors increases with population capacity: higher  $\kappa$  means that demonstrators express more accurate cues and take more adaptive actions (and in the belief-based model, cues can furthermore precipitate larger belief updates). Meanwhile, the learner’s informational need decreases with capacity—as  $\kappa$  approaches  $n$ , the asocial learning problem becomes more straightforward, reducing the space available for improvements in fitness from access to extra information. The combined effect is that social learning is most useful when capacity is *high* enough that social information is moderately informative but *low* enough that the learning problem is still difficult (see Table 1).

While our main qualitative findings held across both belief-based and policy-based social learning, each of our experiments surfaced differences between the two variants—highlighting some interesting consequences of how social information is integrated into downstream behaviour. Perhaps the most interesting of these is that belief-based social learning was more susceptible to “individual-level” maladaptivity (via contamination of future choices) whereas policy-based social learning was more susceptible to frequency-dependent effects in the vein of Rogers’ paradox. Future work may explore connections to recent findings from Roberts-Gaal et al. [54], who show that ecological stability influences the format of

**Table 1. A summary of the effect of representational capacity on the two factors that determine the value of social information use.**

| $\kappa/n$ | social information quality | learner need |
|------------|----------------------------|--------------|
| low        | low                        | high         |
| moderate   | moderate                   | moderate     |
| high       | high                       | low          |

<https://doi.org/10.1371/journal.pcbi.1014504.t001>

cultural learning by shaping a tradeoff between efficient but inflexible procedural information and more flexible but costly goal- or causal-structure information. This difference in “failure modes” between our two social learning variants highlights a related tradeoff; namely that more flexible, less localised integration of social information can also increase the downstream consequences of exposure to unreliable demonstrators.

Also of relevance to our results is the work of Fogarty et al. [55], who studied the impact of task difficulty on the evolution of explicit costly teaching, as opposed to ‘inadvertent’ social learning. Notably, the authors also found an “n-shaped function” where teaching was not favoured for traits that were either too easy to learn (where learners can readily acquire them through other means) or too difficult (where teachers are unlikely to possess the relevant information). One way to view our findings is that they demonstrate how Fogarty et al.’s results for explicit teaching can hold also for inadvertent social learning, given a construction of task difficulty that arises from constraints on individuals’ representational capacity.

### Limitations and directions for future work

While our current models are focused on the relative adaptiveness of social vs asocial learning (and between different social learning biases), it is important to recognise that resource constraints also influence the strategies that agents deploy during asocial learning itself. For instance, constraints on executive function have long been known to shift human learners away from computationally expensive model-based planning and toward simpler, model-free heuristics [56]. More recent work has shown, similarly, that human learners will limit (costly) uncertainty-directed exploration in favour of (cheaper) random exploration under increases in working memory load [57], time pressure [58] or search space complexity [59]. This line of work not only offers a potential alternative formulation of capacity in terms of ability to explore, but also suggests an additional route by which generic capacity limits might favour social learning (as an alternative to costly directed-exploration strategies). Explicitly connecting our findings to richer and more structured forms of asocial exploration thus represents another promising direction for future research.

A more significant limitation of the current work is that our models do not consider any cognitive costs specific to social learning itself—capacity limits manifest only in agents’ ability to represent their environment *in general*, and have no impact on how successfully they obtain the actions or cues of others. Doing things this way allows us to isolate how representational capacity changes the supply of and demand for social information. It is, however, an oversimplification, and future work should investigate the effect of introducing social-learning-specific costs. We suspect that doing so would not fundamentally change the inverse-U dynamics found in our experiments, but would likely narrow the adaptive region by rendering social learning less effective at lower capacity values. Such considerations would be particularly salient under an interpretation of our belief-based social learning variant in which belief cues are not directly observable but are instead inferred from actions or similar evidence by a theory-of-mind-like process [40,60]. The same could be said for a more explicitly pedagogical setting where learner and demonstrator alike must make inferences about one another’s knowledge and intentions [61,62].

Our comparison between conformist and success-biased social learning also leaves aside any costs of implementing these strategies. In our simulations, learners can always identify the majority or currently fittest individual without error. Success-biased learning in the wild, however, requires estimating which individuals are competent on the basis of noisy,

delayed, and/or domain-specific cues. Similarly, conformist learning involves sampling enough demonstrators to estimate the majority, and so may be costly when social access is limited or behavioural frequencies are difficult to observe. Taking such costs into account may shift the two strategies' relative advantage.

Finally, our foraging environment encapsulate a specific class of learning problem—inferring the discrete (deterministic) rule that explains a sequence of sampled experiences. An alternative approach would have agents generalise from prior encounters based on some measure of perceived similarity between different mushrooms (feature vectors). Rule-based and similarity-based generalisation involve different tradeoffs: rules allow for rapid transfer to novel stimuli, while similarity-based representations can capture a wider space of functional relationships [63]. Future work might test how adopting the latter in place of our symbolic rule inference model affects selection for social learning. For example, it is possible that agents' asocial learning performance would degrade more gracefully under limited feature access, potentially narrowing the region over which social learning is adaptive, or allowing the maladaptive effect to persist at higher environment complexities.

## Conclusion

In summary, when it comes to the usefulness of cultural learning, not all agents are created equal. Information-poor populations are unable to support reliable social learning, while information-rich populations don't really need it; only when agents lie within a goldilocks zone of relative capacity does social learning reliably take root. A possible biological implication of our results, which future work should seek to test, is that species most strongly reliant on social learning should be found disproportionately at an intermediate level of representational capacity relative to their ecological demands. Crucially, we suggest that species below this threshold may fail to evolve social learning not because they lack either the need for cheaper, less risky alternatives to trial-and-error exploration, or the 'machinery' for integrating social information—but because their conspecifics do not generate cues reliable enough to be worth making use of. This is analogous to observations that cumulative culture emerges only once distinct thresholds of individual problem-solving are breached [64,65].

But what about humans? We are extremely cognitively sophisticated, and yet each of us typically learns a huge fraction of our skills and knowledge from parents, teachers, mentors and peers: does this not contradict our central message? Not necessarily. While it's true that humans enjoy highly developed brains, we also continuously modify our environment to create conditions of increasing ecological and cultural complexity. Once social information use emerges, it can play into a feedback loop of cultural niche construction [66,67], where the cultural spread of information produces more complex environments, which in turn keep humans within the adaptive regime of *relative* capacity and reinforce the demand for more social learning.

Building upon our main finding (the “whether”), we also demonstrated through an additional set of simulations evidence for an effect on the relative adaptiveness of different social learning *strategies* (the “how”). Specifically, we found that learners who selectively use information from only the fittest agents performed well in lower-capacity populations, whereas conformist learners that simply followed the majority opinion were most successful in higher-capacity populations. We offer the following intuition for this result. In a capacity-poor population, a few individuals experiencing a ‘lucky run’ of trials can produce significant variance in knowledge and therefore fitness—in other words, the expected difference in social information quality between the fittest and median-fitness agents is relatively high. In a capacity-rich population, where the majority of agents are above chance, these differences are dampened, and are outweighed by the protection against individual noise that conformity provides.

An analysis of each strategy's probability of obtaining correct social cues or optimal actions revealed something else surprising: though always worse than copy-fittest, at sufficiently low capacity the naïve approach of selecting a source at random (unbiased transmission/linear imitation) yields better odds than conformist social learning. This stands in contrast with predictions from prior work that conformist social learning should always outperform unbiased transmission [26], and future work should seek to explore this result in greater detail.

## Supporting information

### S1 Text. Supplementary simulation results for alternative capacity variants, robustness analyses, and additional methodological details.

(PDF)

## Acknowledgments

This section is intended only for general acknowledgements and thanks. Any information related to funding, data availability, author contributions, etc. should be entered directly into their dedicated fields in the PLOS Editorial Manager submission system, which will then be incorporated into the appropriate section in your article during the production process.

## Author contributions

**Conceptualization:** Max Taylor-Davies.

**Investigation:** Max Taylor-Davies.

**Methodology:** Max Taylor-Davies.

**Visualization:** Max Taylor-Davies.

**Writing – original draft:** Max Taylor-Davies.

**Writing – review & editing:** Max Taylor-Davies.

## References

1. Galef Jr BG. Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. Social learning: Psychological and biological perspectives. Lawrence Erlbaum Associates, Inc. 1988. p. 3–28.
2. Heyes CM. Social learning in animals: categories and mechanisms. Biological Reviews. 1994;69(2):207–31.
3. Laland KN, Atton N, Webster MM. From fish to fashion: experimental and theoretical insights into the evolution of culture. Philos Trans R Soc Lond B Biol Sci. 2011;366(1567):958–68. <https://doi.org/10.1098/rstb.2010.0328> PMID: 21357218
4. Bridges AD, Royka A, Wilson T, Lockwood C, Richter J, Juusola M, et al. Bumblebees socially learn behaviour too complex to innovate alone. Nature. 2024;627(8004):572–8. <https://doi.org/10.1038/s41586-024-07126-4> PMID: 38448580
5. Boyd R, Richerson PJ. Culture and the evolutionary process. Chicago: University of Chicago Press. 1985.
6. Rendell L, Fogarty L, Hoppitt WJE, Morgan TJH, Webster MM, Laland KN. Cognitive culture: theoretical and empirical insights into social learning strategies. Trends Cogn Sci. 2011;15(2):68–76. <https://doi.org/10.1016/j.tics.2010.12.002> PMID: 21215677
7. Turner CR, Mann SF, Spike M, Magrath RD, Sterelny K. Joint evolution of traits for social learning. Behav Ecol Sociobiol. 2023;77(4). <https://doi.org/10.1007/s00265-023-03314-w>
8. Gelpí RA, Buchsbaum D. Children as cultural explorers: How imitation, pedagogy, and selective trust prepare children for learning in the cultural niche. Oxford Handbook of Cultural Evolution. Oxford University Press. 2025.
9. Kameda T, Nakanishi D. Does social/cultural learning increase human adaptability?: Rogers's question revisited. Evolution and Human Behavior. 2003;24(4):242–60.
10. Kaaronen RO. Mycological rationality: Heuristics, perception and decision-making in mushroom foraging. Judgm decis mak. 2020;15(5):630–47. <https://doi.org/10.1017/s1930297500007841>
11. Kaaronen RO, Manninen MA, Eronen JT. Rules of thumb, from Holocene to Anthropocene. Anthropocene Review. 2023.
12. Papa A. Cultural Transmission and Evolution of Mushroom Knowledge: Insights From Mycophobic Norway. Top Cogn Sci. 2025;:10.1111/tops.70027. <https://doi.org/10.1111/tops.70027> PMID: 40971659
13. Lieder F, Griffiths TL. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behav Brain Sci. 2019;43:e1. <https://doi.org/10.1017/S0140525X1900061X> PMID: 30714890
14. Bhui R, Lai L, Gershman SJ. Resource-rational decision making. Current Opinion in Behavioral Sciences. 2021;41:15–21. <https://doi.org/10.1016/j.cobeha.2021.02.015>
15. Icard TF. Resource rationality. 2023.
16. Arumugam D, Ho MK, Goodman ND, Van Roy B. Bayesian Reinforcement Learning With Limited Cognitive Load. Open Mind (Camb). 2024;8:395–438. [https://doi.org/10.1162/opmi\\_a\\_00132](https://doi.org/10.1162/opmi_a_00132) PMID: 38665544



17. Lai L, Gershman SJ. Human decision making balances reward maximization and policy compression. *PLoS Comput Biol*. 2024;20(4):e1012057. <https://doi.org/10.1371/journal.pcbi.1012057> PMID: 38669280
18. Quillien T, Taylor-Davies M. Factive mindreading reflects the optimal use of limited cognitive resources. *Proc Biol Sci*. 2026;293(2070):20251852. <https://doi.org/10.1098/rspb.2025.1852> PMID: 42086243
19. Taylor-Davies M, Bramley N, Lucas CG. A rational framework for group-based selective social learning. *Open Mind*. 2025;9:677–708.
20. Witt A, Toyokawa W, Lala KN, Gaissmaier W, Wu CM. Humans flexibly integrate social information despite interindividual differences in reward. *Proc Natl Acad Sci U S A*. 2024;121(39):e2404928121. <https://doi.org/10.1073/pnas.2404928121> PMID: 39302964
21. Fränken J-P, Valentin S, Lucas CG, Bramley NR. Naïve information aggregation in human social learning. *Cognition*. 2024;242:105633. <https://doi.org/10.1016/j.cognition.2023.105633> PMID: 37897881
22. Rogers AR. Does biology constrain culture. *American Anthropologist*. 1988;90(4):819–31.
23. Feldman MW, Aoki K, Kumm J. Individual Versus Social Learning: Evolutionary Analysis in a Fluctuating Environment. *AS*. 1996;104(3):209–31. <https://doi.org/10.1537/ase.104.209>
24. Aoki K, Wakano JY, Feldman MW. The Emergence of Social Learning in a Temporally Changing Environment: A Theoretical Model. *Current Anthropology*. 2005;46(2):334–40. <https://doi.org/10.1086/428791>
25. Toelch U, van Delft MJ, Bruce MJ, Donders R, Meeus MTH, Reader SM. Decreased environmental variability induces a bias for social information use in humans. *Evolution and Human Behavior*. 2009;30(1):32–40. <https://doi.org/10.1016/j.evolhumbehav.2008.07.003>
26. McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, et al. Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior*. 2005;26(6):483–508. <https://doi.org/10.1016/j.evolhumbehav.2005.04.003>
27. Acerbi A, Jacquet PO, Tennie C. Behavioral constraints and the evolution of faithful social learning. *Curr Zool*. 2012;58(2):307–18. <https://doi.org/10.1093/czoolo/58.2.307>
28. Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, et al. Why copy others? Insights from the social learning strategies tournament. *Science*. 2010;328(5975):208–13. <https://doi.org/10.1126/science.1184719> PMID: 20378813
29. Yaman A, Bredeche N, Çaylak O, Leibo JZ, Lee SW. Meta-control of social learning strategies. *PLoS Comput Biol*. 2022;18(2):e1009882. <https://doi.org/10.1371/journal.pcbi.1009882> PMID: 35226667
30. Barkoczi D, Galesic M. Social learning strategies modify the effect of network structure on group performance. *Nat Commun*. 2016;7:13109. <https://doi.org/10.1038/ncomms13109> PMID: 27713417
31. Schulz E, Konstantinidis E, Speekenbrink M. Putting bandits into context: How function learning supports decision making. *J Exp Psychol Learn Mem Cogn*. 2018;44(6):927–43. <https://doi.org/10.1037/xlm0000463> PMID: 29130693
32. Russo DJ, Van Roy B, Kazerouni A, Osband I, Wen Z. A Tutorial on Thompson Sampling. *Foundations and Trends® in Machine Learning*. 2018;11(1):1–96. <https://doi.org/10.1561/22000000070>
33. Agrawal S, Goyal N. Analysis of Thompson Sampling for the Multi-armed Bandit Problem. In: *Proceedings of the 25th Annual Conference on Learning Theory*, 2012. 39.1–39.26.
34. Speekenbrink M, Konstantinidis E. Uncertainty and exploration in a restless bandit problem. *Top Cogn Sci*. 2015;7(2):351–67. <https://doi.org/10.1111/tops.12145> PMID: 25899069
35. Gershman SJ. Deconstructing the human algorithms for exploration. *Cognition*. 2018;173:34–42. <https://doi.org/10.1016/j.cognition.2017.12.014> PMID: 29289795
36. Futrell R. Information-theoretic principles in incremental language production. *Proc Natl Acad Sci U S A*. 2023;120(39):e2220593120. <https://doi.org/10.1073/pnas.2220593120> PMID: 37725652
37. Zaslavsky N, Hu J, Levy RP. A rate-distortion view of human pragmatic reasoning. In: 2020. <https://arxiv.org/abs/2005.06641>
38. Lai L, Gershman SJ. Human decision making balances reward maximization and policy compression. *PLoS Comput Biol*. 2024;20(4):e1012057. <https://doi.org/10.1371/journal.pcbi.1012057> PMID: 38669280
39. Kinney DB, Lombrozo T. Building Compressed Causal Models of the World. *Cognitive Psychology*. 2023.
40. Wu CM, Vélez N, Cushman FA. Representational exchange in human social learning: Balancing efficiency and flexibility. *The drive for knowledge: The science of human information seeking*. Cambridge: Cambridge University Press. 2022. p. 169–92. <https://doi.org/10.1017/9781009026949.009>
41. Charpentier CJ, Iigaya K, O'Doherty JP. A Neuro-computational Account of Arbitration between Choice Imitation and Goal Emulation during Human Observational Learning. *Neuron*. 2020;106(4):687–699.e7. <https://doi.org/10.1016/j.neuron.2020.02.028> PMID: 32187528
42. Rendell L, Fogarty L, Laland KN. Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution*. 2010;64(2):534–48. <https://doi.org/10.1111/j.1558-5646.2009.00817.x> PMID: 19674093
43. Mesoudi A, O'Brien MJ. The Cultural Transmission of Great Basin Projectile-Point Technology II: An Agent-Based Computer Simulation. *Am antiq*. 2008;73(4):627–44. <https://doi.org/10.1017/s0002731600047338>
44. Laland KN. Social learning strategies. *Learn Behav*. 2004;32(1):4–14. <https://doi.org/10.3758/bf03196002> PMID: 15161136
45. Moran PAP. Random processes in genetics. *Math Proc Camb Phil Soc*. 1958;54(1):60–71. <https://doi.org/10.1017/s0305004100033193>
46. Ewens WJ. *Mathematical population genetics: theoretical introduction*. 2nd ed. New York: Springer-Verlag. 2004.

47. Grafen A. Natural Selection, Kin Selection and Group Selection. *Behavioural Ecology: An Evolutionary Approach*. 2nd ed. Oxford: Blackwell Scientific Publications. 1984. p. 62–84.
48. West SA, Gardner A. Adaptation and Inclusive Fitness. *Current Biology*. 2013;23(13):R577–84. <https://doi.org/10.1016/j.cub.2013.05.031>
49. Kendal RL, Boogert NJ, Rendell L, Laland KN, Webster M, Jones PL. Social Learning Strategies: Bridge-Building between Fields. *Trends Cogn Sci*. 2018;22(7):651–65. <https://doi.org/10.1016/j.tics.2018.04.003> PMID: 29759889
50. Mesoudi A. An experimental simulation of the “copy-successful-individuals” cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*. 2008;29(5):350–63. <https://doi.org/10.1016/j.evolhumbehav.2008.04.005>
51. Mesoudi A. An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*. 2011;32(5):334–42. <https://doi.org/10.1016/j.evolhumbehav.2010.12.001>
52. Morgan TJH, Rendell LE, Ehn M, Hoppitt W, Laland KN. The evolutionary basis of human social learning. *Proc Biol Sci*. 2012;279(1729):653–62. <https://doi.org/10.1098/rspb.2011.1172> PMID: 21795267
53. Schlag KH. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*. 1998;78(1):130–56. <https://doi.org/10.1006/jeth.1997.2347>
54. Roberts-Gaal X, Bolic M, Cushman FA. Environmental variability shapes the representational format of cultural learning. *Proc Natl Acad Sci U S A*. 2025;122(28):e2505283122. <https://doi.org/10.1073/pnas.2505283122> PMID: 40623192
55. Fogarty L, Strimling P, Laland KN. The evolution of teaching. *Evolution*. 2011;65(10):2760–70.
56. Otto AR, Gershman SJ, Markman AB, Daw ND. The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol Sci*. 2013;24(5):751–61. <https://doi.org/10.1177/0956797612463080> PMID: 23558545
57. Cogliati Dezza I, Cleeremans A, Alexander W. Should we control? The interplay between cognitive control and information integration in the resolution of the exploration-exploitation dilemma. *J Exp Psychol Gen*. 2019;148(6):977–93. <https://doi.org/10.1037/xge0000546> PMID: 30667262
58. Wu CM, Schulz E, Pleskac TJ, Speekenbrink M. Time pressure changes how people explore and respond to uncertainty. *Sci Rep*. 2022;12(1):4122. <https://doi.org/10.1038/s41598-022-07901-1> PMID: 35260717
59. Brown VM, Hallquist MN, Frank MJ, Dombrovski AY. Humans adaptively resolve the explore-exploit dilemma under cognitive constraints: Evidence from a multi-armed bandit task. *Cognition*. 2022;229:105233. <https://doi.org/10.1016/j.cognition.2022.105233> PMID: 35917612
60. Apperly IA, Butterfill SA. Do humans have two systems to track beliefs and belief-like states?. *Psychological Review*. 2009;116(4):953–70. <https://doi.org/10.1037/a0016923>
61. Shafto P, Goodman ND, Griffiths TL. A rational account of pedagogical reasoning: teaching by, and learning from, examples. *Cogn Psychol*. 2014;71:55–89. <https://doi.org/10.1016/j.cogpsych.2013.12.004> PMID: 24607849
62. Gweon H. Inferential social learning: cognitive foundations of human social learning and teaching. *Trends Cogn Sci*. 2021;25(10):896–910. <https://doi.org/10.1016/j.tics.2021.07.008> PMID: 34417094
63. Wu CM, Meder B, Schulz E. Unifying principles of generalization: Past, present, and future. *Annual Review of Psychology*. 2025;76(1):275–302. <https://doi.org/10.1146/annurev-psych-021524-110810>
64. Tennie C, Call J, Tomasello M. Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos Trans R Soc Lond B Biol Sci*. 2009;364(1528):2405–15. <https://doi.org/10.1098/rstb.2009.0052> PMID: 19620111
65. Pradhan GR, Tennie C, van Schaik CP. Social organization and the evolution of cumulative technology in apes and hominins. *J Hum Evol*. 2012;63(1):180–90. <https://doi.org/10.1016/j.jhevol.2012.04.008> PMID: 22658335
66. Boyd R, Richerson PJ, Henrich J. The cultural niche: why social learning is essential for human adaptation. *Proc Natl Acad Sci U S A*. 2011;108 Suppl 2(Suppl 2):10918–25. <https://doi.org/10.1073/pnas.1100290108> PMID: 21690340
67. Laland K, Matthews B, Feldman MW. An introduction to niche construction theory. *Evol Ecol*. 2016;30:191–202. <https://doi.org/10.1007/s10682-016-9821-z> PMID: 27429507