



Research

Cite this article: Thompson B, Griffiths TL. 2021 Human biases limit cumulative innovation. *Proc. R. Soc. B* **288**: 20202752. <https://doi.org/10.1098/rspb.2020.2752>

Received: 3 November 2020

Accepted: 8 February 2021

Subject Category:

Behaviour

Subject Areas:

cognition, behaviour

Keywords:

cultural evolution, innovation, function
learning, inductive bias, Bayesian, optimization

Author for correspondence:

Bill Thompson

e-mail: wtd@princeton.edu

Human biases limit cumulative innovation

Bill Thompson and Thomas L. Griffiths

Departments of Psychology and Computer Science, Princeton University, Princeton, NJ 08544, USA

BT, 0000-0001-7348-3531

Is technological advancement constrained by biases in human cognition? People in all societies build on discoveries inherited from previous generations, leading to cumulative innovation. However, biases in human learning and memory may influence the process of knowledge transmission, potentially limiting this process. Here, we show that cumulative innovation in a continuous optimization problem is systematically constrained by human biases. In a large ($n = 1250$) behavioural study using a transmission chain design, participants searched for virtual technologies in one of four environments after inheriting a solution from previous generations. Participants converged on worse solutions in environments misaligned with their biases. These results substantiate a mathematical model of cumulative innovation in Bayesian agents, highlighting formal relationships between cultural evolution and distributed stochastic optimization. Our findings provide experimental evidence that human biases can limit the advancement of knowledge in a controlled laboratory setting, reinforcing concerns about bias in creative, scientific and educational contexts.

1. Introduction

All human societies depend on technology [1]—tools, concepts, skills and subsistence techniques that help people solve physical and computational problems. One of the most distinctive aspects of human technology is that many innovations we rely on every day would be ‘unlikely or impossible for us to have achieved by ourselves’ [2]. From kayaks to cryptographic hash functions, efficient solutions to environmental challenges depend on *cumulative innovation*: ‘in most times and places, individuals don’t invent tools; tools evolve gradually’ [3] as innovations are transmitted from one generation to the next. This process of transmission and accumulation of knowledge is called cumulative cultural evolution (CCE) [4].

Here, we examine whether biases in human learning and memory can influence the outcome of cumulative innovation in a large behavioural experiment. Mathematical models suggest that efficient accumulation of knowledge depends on high-fidelity transmission [5] of successful innovations [6] to new individuals who themselves may innovate further [7]—the so-called ratchet effect [8]. Aspects of this process can be studied experimentally. Behavioural studies (reviewed in [9]) show that cumulative innovation under controlled laboratory conditions allows experimental micro-societies to discover increasingly functional solutions to problem-solving and optimization tasks in both physical [10–14] and virtual [15–18] settings (e.g. paper aeroplane [10] or virtual arrowhead [16] design tasks). Performance depends on factors such as population structure [11,19–21] and transmission dynamics [22], as predicted by models [17,23,24].

However, theories of cultural evolution [6,25] also suggest that the outcome of cumulative innovation may be constrained by a more fundamental problem. Imagine trying to construct a canoe after travelling in one built by somebody else. In many contexts, the transmission of knowledge from person to person is an imperfect process relying on inference, estimation, and reconstruction of ideas [25]. This introduces a problem of induction—the need to generalize beyond what is observable—that can only be solved by inductive bias [26]. Bias is a requirement for inference whenever aspects of a learning problem are not fully determined by the data observed. For instance, the structure and parameters of tools, artefacts or subsistence techniques cannot be directly

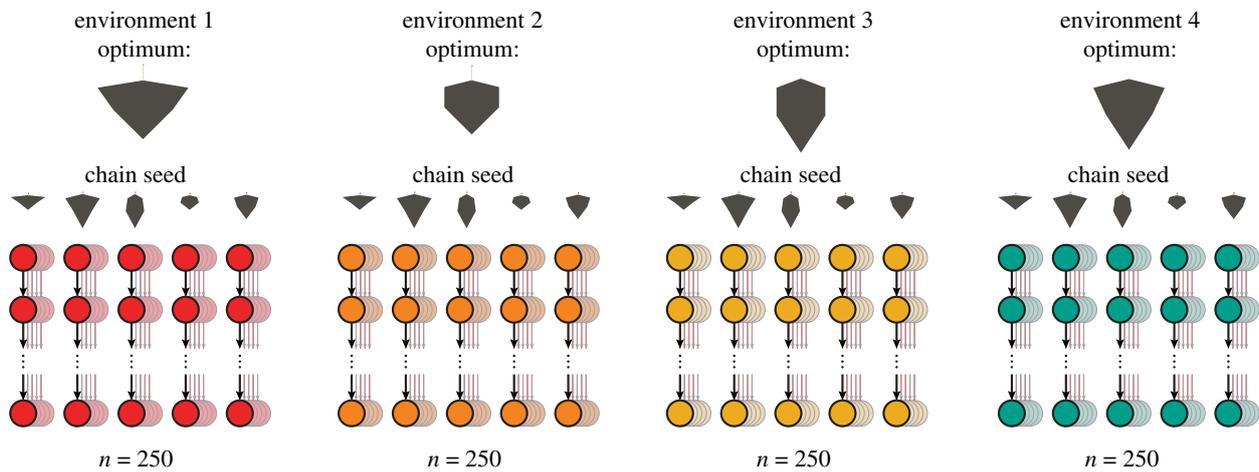


Figure 1. Transmission chain design. Participants were organized into transmission chains of 10 generations. A participant at generation t inherited the arrowhead designed by the participant at generation $t - 1$ of their chain. The principle experimental manipulation was the design of the highest-scoring virtual arrowhead (illustrated above), which varied between treatments (environments E1–E4). Each treatment was replicated in 25 independent chains. (Online version in colour.)

communicated from one mind to another: instead, these design features must often be inferred from limited experience, incomplete description, noisy perceptual information, or potentially inaccurate or partial memories.

Inductive biases can also play an important role in instances of ambiguous or conflicting social testimony. For example, Henrich *et al.* [27] showed that human biases in category-based induction may help to explain observed patterns of culturally evolved adaptive food taboos in a Fijian population. However, in specific cases where cultural information is indecisive, biases in categorization and similarity-based inferences can lead to non-adaptive taboos [27]. Cases like this exemplify the potential for inductive biases to influence decision-making and learning in functional settings when information transmission is limited. While teaching [28–31] and other pragmatic mechanisms for cooperative inference [32,33] can improve information transmission, inferential bottlenecks [3] create the opportunity for biases in human cognition to repeatedly influence the outcome of cultural transmission, raising the question of their long-term consequences.

Cognitive approaches to cultural evolution [25,34,35] have emphasized this possibility: ‘Human cognitive abilities act [...] as a filter on the representations capable or likely to be widely distributed in human populations’ [25]. As a result of the need for inductive generalization, artefact designs, tool-use techniques, behavioural repertoires, or technical concepts which people find difficult to remember or understand are systematically less likely to survive the process of cultural transmission. Instead, the process may favour decisions, concepts, or design choices that people find intuitive. This perspective is supported by formal models [34–36]. These dynamics may arise as a consequence of inductive biases that are themselves the result of development and learning [37], or from innate or adaptive features of human cognition [38]. Behavioural studies show that human biases can influence the transmission of conventional behaviours such as artificial languages [39], signalling systems [40], stereotypes [41], motor programs [42] and musical rhythms [43]. Cultural transmission by serial reproduction (SR) can even be viewed abstractly as an efficient algorithm for revealing cognitive biases [44].

However, the impact of human biases has been difficult to assess in functional settings. Understanding these effects poses

a challenge for theories of technology [3,45,46] and for experimental studies of cumulative innovation in structured problem-solving settings, which can be difficult to quantify in cognitive terms [4,46–49]. Studies of innovation in simpler optimization paradigms offer the potential for more continuity. People are excellent at solving optimization problems as individuals [50], networked groups [51] and cultural populations [24,52], often outperforming advanced algorithms [50,53]. Studies have focused on how social learning strategies (reviewed in [54]) influence the potential for accumulation in multi-modal, continuous optimization landscapes [22], and on causal understanding [24] for example. However, the potential for human inductive biases to influence these dynamics remains unknown, because studies often implement direct transfer of solutions between individuals. Without a quantitative theory relating psychological biases to inference and cumulative innovation in a specific experimental setting, these basic predictions have been difficult to test [4,55].

To help assess the impact of cognitive biases on CCE, we developed a probabilistic cognitive model of learning and innovation in chains of Bayesian learners. The model provides insight into functional and cognitive perspectives by revealing a formal connection between CCE and distributed optimization. Our analysis predicts that the biases of learners should limit cumulative innovation in continuous optimization problems. To test this prediction, we conducted a large behavioural experiment in which transmission chains of participants attempted to solve a multidimensional continuous optimization problem. Previous studies focused on complex functional landscapes [22,52] and learning from multiple models [16]. By contrast, our study used a transmission chain design (figure 1), and we constructed a simple hill-shaped landscape whose optimum should be easy to discover (figure 2). These simplifications allowed us to examine the biases of our participants that emerge directly as a consequence of the process of knowledge transmission.

Our key manipulation was to transpose the abstract optimization problem into four virtual environments, each defined by a different optimal solution to the same underlying problem. In this setting, the relevant biases are reflected in the kinds of solutions that people are likely to consider spontaneously or intuitively, before observing how other people solve the task or engaging in within-task

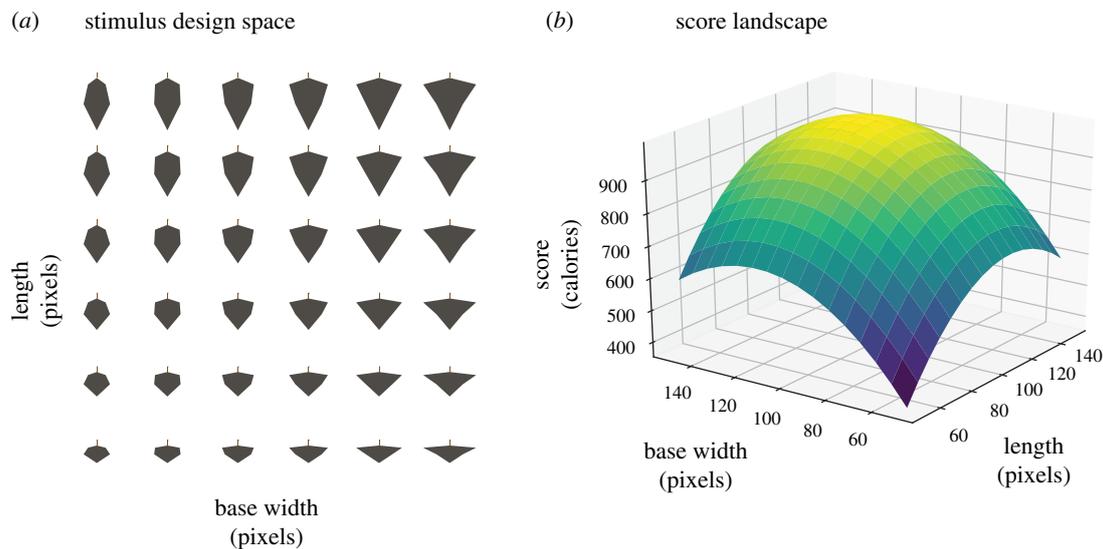


Figure 2. Optimization problem. (a) Illustration of the design space of possible arrowheads implied by the ability to manipulate base width and length (between 50 and 150 pixels). Arrowheads are regularly sampled points for illustration purposes in a continuous underlying space of possible designs. (b) The objective function (score landscape) $f(\theta)$. Yellow denotes higher scores, blue denotes lower scores. The score landscape was a simple, smooth, hill-shaped function (quadratic surface) with maximum at θ^* (here $\theta_{bw}^* = 115$, $\theta_l^* = 115$, illustrating environment E4). Across environments, the position of the optimum design (θ^*) varied but the shape of the landscape remained constant and as displayed here. (Online version in colour.)

learning. We estimated our participants' biases in the task using a SR paradigm, allowing us to independently measure the consistency between these biases and each of the optimal solutions in the four virtual environments. We then examined whether information accumulation was equally effective in all four environments, or whether information accumulation was constrained in environments less consistent with our participants' biases.

2. Model

One way to connect models of cultural transmission with cognitive theories of intelligence is to analyse populations of Bayesian learners [56]. Some of the earliest models of cultural transmission were based on this approach [6]. Bayesian models naturally express the cognitive structures important to theories of cultural transmission, such as technical knowledge [57], cognitive attractors [45], latent solutions [58], learning biases [36], adaptive instincts [38,59] and content biases [6]. These inductive biases are expressed by a prior distribution over a design space of possible solutions to a problem. Intuitive or learnable solutions have high prior probability. Counterintuitive solutions have lower prior probability. A prior distribution which assigns differential probabilities to solutions encodes inductive bias. For example, a simplicity bias [60] assigns more complex solutions lower probability. More structured or domain-specific priors can result from conceptual or semantic knowledge, such as intuitive physics [61], or inductive biases for particular structural forms [62].

Bayesian models have been instrumental in understanding how people's inductive biases can shape SR [63] and iterated learning [56]. These models have focused on the emergence of conventional behaviours [36,64,65], particularly human language [36,39,56,66]. We used this probabilistic paradigm to analyse CCE in populations of individuals attempting to solve an optimization problem and transmitting their solutions to new learners. Following influential models of CCE [6,7,34,67], we analysed transmission of a real-valued solution θ . This allowed us to view CCE in terms of two components:

inductive biases, expressed as a *prior* distribution $p(\theta)$; and the underlying optimization problem, expressed as a fitness landscape or objective function $f(\theta)$. We examined the dynamics of cultural transmission under the assumption that the relevant prior distribution is Gaussian with mean μ and variance δ^2 , that perception of transmitted designs is subject to Gaussian noise (with variance σ^2) and that innovation can be expressed as incremental steps along the gradient $\nabla_{\theta} f(\theta)$ of the fitness landscape (see Methods).

(a) Predictions

The mathematical framework that descends from combining these cognitive and functional assumptions has a surprisingly familiar structure. A transmission chain of individuals solving a continuous optimization problem through social learning by Bayesian inference and individual learning through bounded innovation has dynamics:

$$\theta_{t+1} = \theta_t - \alpha \nabla_{\theta} f(\theta_t) + \lambda(\mu - \theta_t) + \epsilon_{t+1}, \quad (2.1)$$

where ϵ_{t+1} is zero-mean Gaussian noise, α is an innovation rate parameter, and $\lambda = \sigma^2 / (\sigma^2 + \delta^2)$. Equation (2.1) describes a form of *stochastic gradient descent*—an optimization algorithm at the heart of modern artificial intelligence and machine learning [68]. This connection provides a formal basis for the view that CCE can be interpreted as a form of distributed computation [69], while also offering insight into the relationship between function and cognition—in equation (2.1), human inductive biases act as a form of *regularization*. The balance of transmission fidelity and bias strength plays the role of a regularization parameter, determining the impact of human biases on the process (see Methods).

We calculated the equilibrium behaviour of this process in hill-climbing problems (see Methods), and found that it converges towards a compromise between the mean μ of the prior distribution $p(\theta)$ and the optimum θ^* of the objective function $f(\theta)$. This compromise can be written as

$$\mathbb{E}_{t \rightarrow \infty}[\theta_t] = \mu + \beta(\theta^* - \mu). \quad (2.2)$$

Here, $0 \leq \beta \leq 1$ interpolates between processes dominated by function, converging towards optimal solutions ($\beta \rightarrow 1$), and processes dominated by inductive biases ($\beta \rightarrow 0$) converging towards intuitive solutions closer to μ . This interpolation parameter is determined by a principled relationship between bias strength, transmission fidelity, the rate of innovation, and the steepness of the fitness function (see Methods).

The general result in equation (2.2) shows that previous analyses of CCE [34,67] and SR [70] can be recapitulated as special cases. Equation (2.2) predicts that participants transmitting solutions to a continuous optimization problem will discover a compromise between the mean μ of the prior $p(\theta)$ and the optimum design θ^* . Task success should be higher in environments more consistent with participants' biases.

3. Results

To test these predictions, we conducted a pre-registered (<https://osf.io/hx9jd/>) behavioural study using a transmission chain design (see figure 1 and Methods). Participants completed an optimization task after briefly observing the solution produced by the previous participant and reconstructing it from memory. We adapted an influential paradigm developed by Mesoudi and colleagues [16,17,22,71,72]. Participants designed virtual *arrowheads* (figure 2) to earn as many calories of food as possible on a virtual hunt by modifying continuous properties of its design (the width of the arrowhead's base and its overall length, see figure 2a). Different combinations of arrowhead features (base width = θ_{bw} length = θ_l) lead to different scores determined by the landscape $f(\theta)$ in figure 2. Note that our goal is not to assess human biases in general for this task, which are likely to vary across populations [27]. Our aim is to estimate the biases of this specific population and predict their impact.

(a) Estimated biases

Participants ($n = 250$, 25 chains of 10 generations) assigned to the SR treatment completed a simplified version of the task (reconstruction of an inherited arrowhead design after brief observation, but no individual learning). Previous research [73] shows that SR results in samples from the prior $p(\theta)$. We used this connection to estimate people's biases (see Methods) about the stimulus design space (figure 2). The sample-mean arrowhead design can be taken as an estimate ($\hat{\mu}$) of the mean μ of the prior distribution—the prototype arrowhead design. Figure 3 illustrates the estimated prototype. In this figure, faint depictions of all 250 arrowheads produced by participants in the SR treatment are overlaid uniformly. The position of this prototype in the design space is also shown in figure 3 (inset). The prototype design is close to the centre of the design space, but is statistically significantly longer ($M = 107.5$, $Z = 20\,187$, $p < 0.001$, Wilcoxon signed-rank test) and wider ($M = 103.2$, $Z = 16\,874$, $p = 0.025$) than would be expected under random sampling.

(b) Task success

Participants ($n = 1000$) assigned to optimization treatments (E1–E4) first reconstructed the inherited design before completing four individual learning trials (practice *hunts*) in which they tested nearby variations on their arrowhead and gained

Estimated Prior Prototype Arrowhead Design

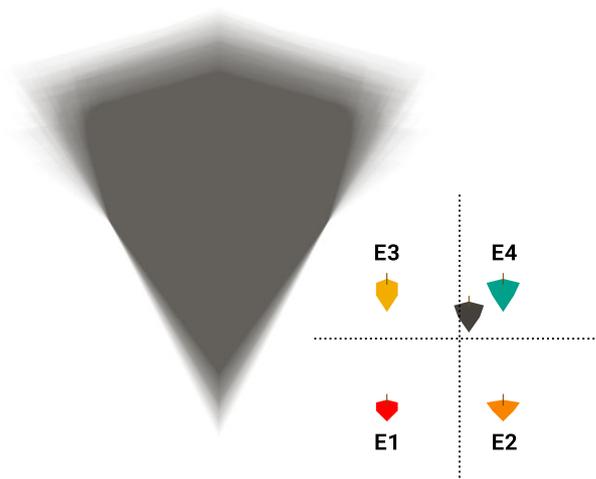


Figure 3. Experimental results (bias estimate). Visualization of the sample-based estimate of the mean $\hat{\mu}$ of the prior distribution $p(\theta)$ implicitly assumed by participants. Plot shows arrowheads produced by all $n = 250$ participants in the serial reproduction treatment faintly overlaid, illustrating $\hat{\mu}$. Inset shows prior mean arrowhead ($\hat{\mu}$, black) positioned in design space relative to optimum arrowheads in optimization treatments E1–E4. (Online version in colour.)

performance feedback. All four environments were determined by the landscape $f(\theta)$ shown in figure 2. However, we manipulated the *location* of the optimum arrowhead design (peak of the hill). Optimum designs are shown in figure 1. We calculated the consistency of each environment with human biases by measuring the absolute design space distance between the optimum design (θ^*) and the prototype ($\hat{\mu}$). Larger distance means the optimum is less consistent with human biases. Environments are ranked in order of least to most consistent: E1 is least consistent ($|\theta_{bw}^* - \hat{\mu}_{bw}| = 28.2$, $|\theta_l^* - \hat{\mu}_l| = 32.5$); E2 ($|\theta_{bw}^* - \hat{\mu}_{bw}| = 11.75$, $|\theta_l^* - \hat{\mu}_l| = 32.5$) and E3 ($|\theta_{bw}^* - \hat{\mu}_{bw}| = 28.22$, $|\theta_l^* - \hat{\mu}_l| = 7.5$) are more consistent, E4 is most consistent ($|\theta_{bw}^* - \hat{\mu}_{bw}| = 11.75$, $|\theta_l^* - \hat{\mu}_l| = 7.5$).

Figure 4 shows task success (produced arrowhead scores) across treatments. Participants assigned to E4 achieved the highest scores ($M = 950$, s.d. = 66), followed by E3 ($M = 934$, s.d. = 78) and E2 ($M = 883$, s.d. = 111). Participants assigned to the E1 treatment had least success ($M = 856$, s.d. = 151). A pre-registered one-way analysis of variance showed that the effect of environment on task success was significant ($F(3, 996) = 41.66$, $p < 0.0001$). However, a Levene test for homogeneity of variance showed that the assumption of equal variances was not met. A non-parametric alternative (the Kruskal–Wallis test by rank) confirmed that the effect of environment on task success was significant ($H(3) = 113.4$, $p < 0.001$). Pairwise comparisons using Wilcoxon's rank-sum test with Bonferroni adjustments showed that participants in the E4 treatment achieved significantly higher task success than participants in treatments E3 ($p = 0.006$), E2 ($p < 0.0001$) and E1 ($p < 0.0001$). Participants in the E3 treatment achieved significantly higher task success than participants in treatments E2 ($p < 0.001$) and E1 ($p < 0.0001$). There was no statistically significant difference in task success between participants in the E1 and E2 treatments ($p = 1$).

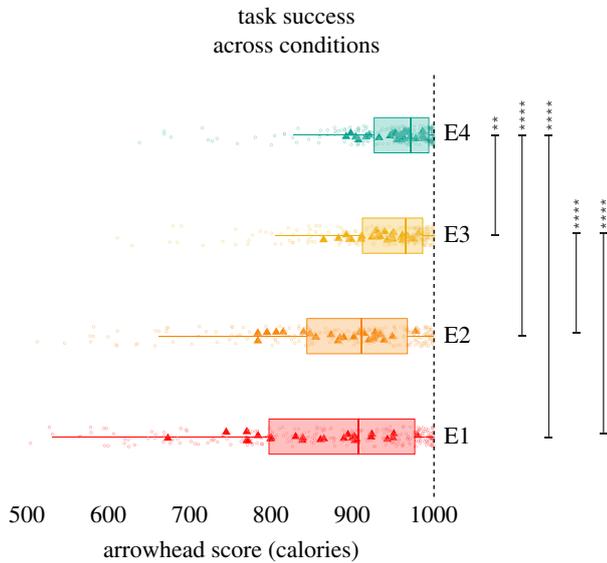


Figure 4. Experimental results (task success). Task success across optimization treatments. Triangles show mean score for individual chains. Faint dots show individual participant scores. Greater distance between the optimum and the prior mean resulted in worse performance. (Online version in colour.)

(c) Arrowheads discovered

The main prediction of our formal analysis is that people will converge on a linear compromise between the prototype $\hat{\mu}$ and the optimum θ^* . We performed a pre-registered regression analysis to test this prediction. We analysed base width and length separately. First, we performed a regression analysis with arrowhead base width as dependent variable and the difference between the prototype and the optimum ($\hat{\mu}_{\text{bw}} - \theta_{\text{bw}}^*$) as predictor. Our formal hypothesis was supported. Accounting for the mean of the prior ($\beta = 0.98$, $p < 0.0001$) and the difference between the prior and optimum ($\beta = 0.33$, $p < 0.0001$) accounted for 97% of the variance in arrowhead base width (adjusted $R^2 = 0.966$). Second, we performed the same analysis with arrowhead length as the dependent variable. Accounting for the mean of the prior ($\beta = 1.$, $p < 0.0001$) and the difference between prior and optimum ($\beta = 0.34$, $p < 0.0001$) accounted for 97% of the variance in arrowhead length (adjusted $R^2 = 0.968$).

(d) Exploratory analyses

We first established that optimization chains had converged to an equilibrium state. Regression analyses of the optimization treatments with environment, generation, and chain id as predictors showed generation was not a statistically significant predictor of task success ($\beta = 1.43$, $p = 0.45$), arrowhead base width ($\beta = -0.51$, $p = 0.12$), or length ($\beta = 0.16$, $p = 0.65$) after generation three, indicating convergence (figure 5b). We then examined the contributions of bias and individual learning to task success. Figure 5 shows expected task success under a random walk over the fitness landscape (dot-dash lines). We quantified this expectation by numerical integration over the fitness landscape within the range of the design space, which provides an approximation to the average task success that would be achieved by selecting an arrowhead design uniformly at random under the constraints of the task in each environment without transmission or innovation. This baseline helps to contextualize task success rates achieved through transmission and innovation.

As an additional reference point, we also examined task success rates under a model that includes innovation but not biased learning. Our mathematical analysis showed that unbiased learning via Bayesian inference ($\delta^2 \rightarrow \infty$) in the context of gradient-based innovation predicts discovery of the optimal solution (see equation (2.2) and Methods). Simulation of a simpler model of innovation supports this prediction. We simulated the experimental conditions (25 chains of 10 generations per treatment) under a model that assumes individuals sample four arrowhead designs uniformly at random within the bounded region of design space surrounding the design they inherited, and transmit the highest scoring sampled design. This process leads to convergence towards the optimum design in all treatments, and therefore does not provide an alternative explanation of our findings (Kruskal–Wallis test by rank showed no significant differences in task success between environments after generation 5 at $\alpha = 0.01$ in 912 of 1000 simulations of the experiment).

Figure 5 shows scores that would have been awarded to arrowheads designed by participants in the SR treatment (pale box plots). These hypothetical scores quantify the distribution of task success we would expect if participants relied exclusively on their biases as a source of information influencing their arrowhead designs. The comparison is useful because it establishes a baseline score distribution in the absence of cultural accumulation of information about the score landscape f . Examining arrowheads after generation three, we found that the hypothetical scores of participants in the SR treatment were significantly different from baselines in E4 ($Z = 14\,719$, $p < 0.001$, Wilcoxon signed-rank test), E3 ($Z = 13\,219$, $p < 0.001$) and E2 ($Z = 10\,577$, $p < 0.001$), but not in E1 ($Z = 8810$, $p = 0.098$). Participants' biases helped in some but not all environments. However, arrowheads designed by participants in the optimization treatments achieved significantly greater task success than arrowheads designed in the SR treatment in all environments (Mann–Whitney one-sided tests, E1: $W = 20\,255.5$, $p < 0.001$; E2: $W = 19\,985.5$, $p < 0.001$; E3: $W = 22\,830$, $p < 0.001$; E4: $W = 21\,035.5$, $p < 0.001$; two-sided tests concurred). Individual learning contributed to task success in all environments.

4. Discussion

Is technological advancement potentially constrained by human biases? This question is central to theories of human technological culture, but has been difficult to answer empirically. Functional perspectives highlight the capacity to discover efficient solutions to diverse environmental problems [1,5,6,16]. Cognitive perspectives highlight the transformative impact of human cognition on cultural transmission [25,34,55,74]. The integration of these perspectives into formal cognitive models and behavioural experiments is a challenge for modern approaches to human culture [4,46,48].

We developed a probabilistic model of CCE to quantify these effects. Our analysis combined a Bayesian formulation of cultural transmission with an optimization-based model of innovation through individual learning. The resulting mathematical framework predicts a trade-off between inductive biases and objective function in cumulative innovation.

Our analysis highlights formal connections between cultural evolution, stochastic optimization algorithms, and Bayesian inference, extending known connections between

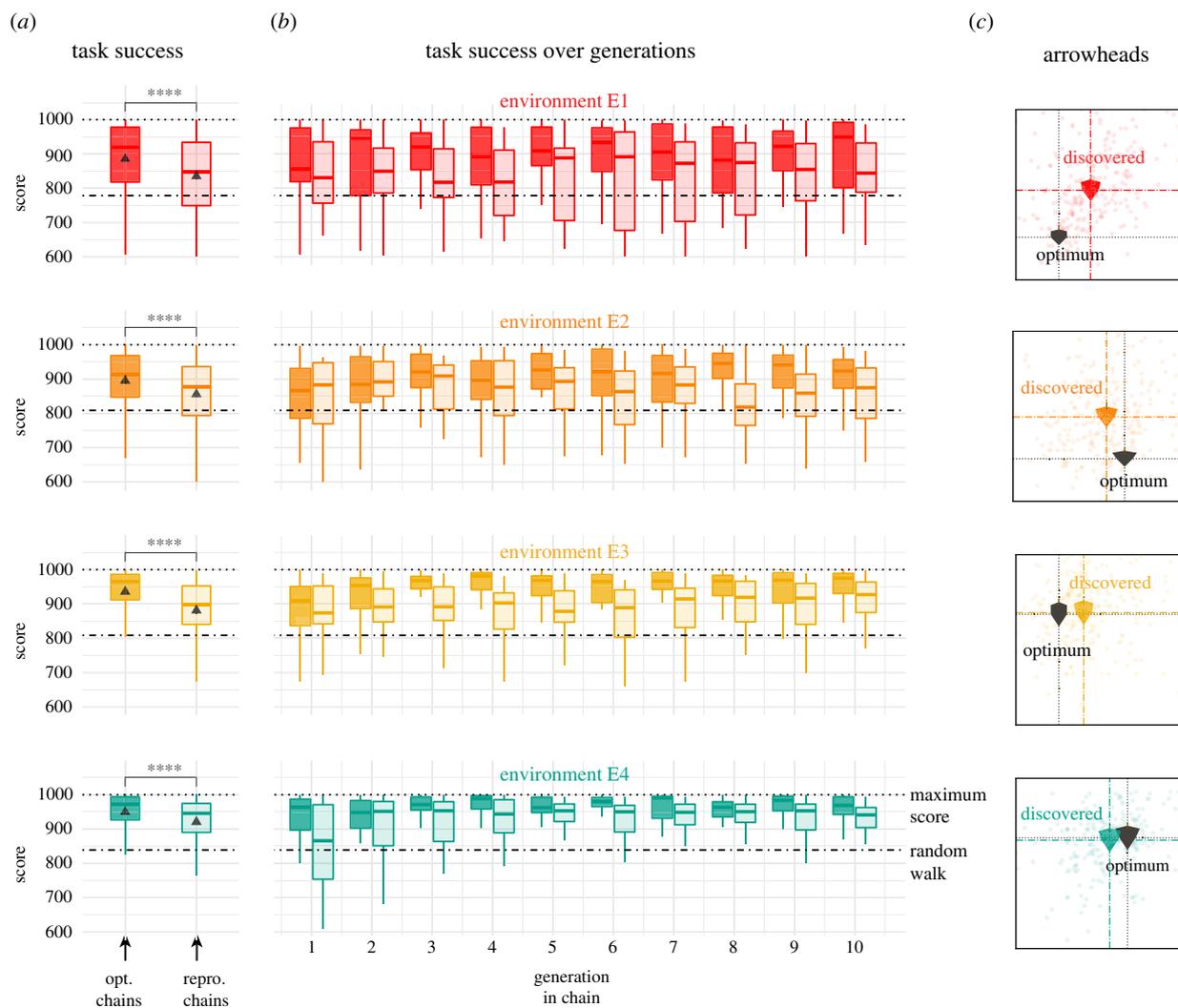


Figure 5. Experimental results (task success baselines and comparisons). Task success overall (a) and over generations (b). Solid box plots show arrowhead score within-treatment distributions. Faded box plots show scores that would have been awarded to arrowheads produced by participants in the serial reproduction treatment if they had been evaluated in against the scoring function. Triangles (a) show mean score. Dot-dashed lines show score expected under a random walk over the scoring function. Dotted lines show maximum score. (c) Optimum arrowhead and the mean design of produced arrowheads positioned within the stimulus design space. Faint dots show individual arrowhead designs in stimulus space. (Online version in colour.)

evolution, learning, and optimization [75–81] to cultural transmission. An optimization objective is an assumption of the model; however, the fact that the population dynamics has the form of a well-defined algorithm computing solutions to this optimization problem is not an assumption, but a finding that derives from the combination of simpler assumptions about bounded individual cognition. This connection to stochastic optimization algorithms establishes the relevance of a large mathematical literature on the properties of these processes that may help illuminate aspects of human technological cultures in the same way that evolutionary theory has offered quantitative insights and predictions. Here, for example, we found that human inductive biases can act as a form of regularization. Regularization is known to play an important role in the generalization profile of stochastic optimization processes [82], raising the question of whether analogous functional consequences of human learning biases may also be relevant to cultural accumulation, for instance.

More generally, our analysis illustrates a way of viewing cultural processes directly in terms of the computations performed by individuals and by the population, creating continuity between models of cognitive and cultural

processes. Computational approaches to cognition view learning in terms of the information processing carried out by individuals—the computations people perform and their dependencies on data and bias. Our results extend this perspective to the cultural process as a whole, by formalizing the computation performed by the *population*. This offers a quantitative formulation of the idea that populations can implement distributed computations over generations [69], and that cultural processes can lead to products that go beyond the information processing limits of individuals [1]—a core principle of CCE that has been notoriously difficult to quantify [4].

Our study contributes to an increasingly detailed empirical picture of how human cognition can influence cultural evolution [14,31,42,54], by quantifying biases that are introduced simply as a consequence of the process of cultural transmission. Our model makes predictions about the impact of inductive biases in general on the outcome of cultural transmission. To test these predictions, we selected a specific population of participants, measured their biases on a widely used task, and showed that the results aligned with our model predictions. Participants performed worse at solving a continuous optimization problem when the solution was less consistent

with the relevant prior distribution, confirming mathematical predictions and substantiating the assumptions of influential theories [3,25].

The current study focuses on a gradient optimization setting and is limited to transmission-chain population structures. These are significant simplifications. While these simplifications allowed us to identify a basic property of cumulative innovation, examining the impacts of inductive bias in more structured tasks and richer population dynamics is an important goal for future research. Extending our gradient-based model of innovation to more sophisticated algorithms that better approximate human search is also a priority. A more complete understanding requires analysis of important cases such as preferential selection among multiple social models [6], heterogeneous or dynamic priors, and varying environmental objectives. Whether these factors counteract or amplify the impacts of inductive bias remains unknown.

Another important assumption of our analyses is that the prototype arrowhead design measured using SR is a useful approximation to inductive biases that are at least partly shared among participants. SR has desirable properties as an estimation procedure: it provides a mathematically motivated, data-efficient approach to estimation of a prior distribution (see Methods and e.g. [83]) from participant responses in a social task, and previous research has shown that this approach accords with alternative methods [44,83]. However, this method and our model rely on the assumption that participant's inferences are approximately Bayesian with respect to a prior distribution that is shared, which is a clear simplification. Our model and experimental results indicate that quantifying a shared prototype design by estimating the mean of a prior distribution through SR provides an empirically useful approximation. However, the estimated prior prototype design should be interpreted as a useful summary statistic of a potentially complex distribution, rather than a platonic ideal design. Future research might profitably explore this simplification.

Our findings provide experimental evidence that human inductive biases can impact the discoveries people are able to make as a population in a controlled laboratory setting, as predicted by theories of cultural evolution [3,6,25,34,55]. Some theories have emphasized how technical knowledge [57] and other adaptive human priors [25,38] are well aligned with environmental problems. Other theories have emphasized the many problems people face with highly counterintuitive solutions [1,6,16]. Our framework is applicable to both contexts. Our results suggest that there are significant connections between evolutionary theories of cumulative culture and computational theories of inductive inference. Understanding these connections may help to integrate diverse approaches to the evolution of human technology and clarify conditions that facilitate effective accumulation from a cognitive perspective. These results reinforce the importance of mechanisms that help to alleviate bias in creative, educational and scientific contexts.

5. Methods

(a) Model

We made the simplifying assumption that dimensions θ_i of a multi-dimensional parametrization Θ can be treated independently. Under a Gaussian prior distribution $p(\theta_i) = \mathcal{N}(\theta_i; \mu_i, \delta_i^2)$ and Gaussian noise-corrupted observation of θ_i^t (with variance σ_i^2), the

sampling distribution $p(\hat{\theta}_i | \theta_i, \mu_i, \sigma_i, \delta_i)$ for the design induced ($\hat{\theta}_i$) is the posterior distribution, which is Gaussian with expectation

$$\mathbb{E}[\hat{\theta}_i] = \frac{\sigma_i^2}{s_i} \mu_i + \frac{\delta_i^2}{s_i} \theta_i, \quad (5.1)$$

where $s_i = \sigma_i^2 + \delta_i^2$. We quantified bounded innovation by assuming that people experiment with nearby variations on an inherited solution and use performance feedback to attempt a (guided) innovation. A generic way to capture the outcome of this process is *gradient-based learning* [84,85], casting innovation as a step along the gradient (slope) of f denoted by $\nabla_{\theta} f(\theta_i)$. The rate of innovation (the magnitude of the step $\alpha > 0$) is a model parameter. The chain dynamics are given by equation (2.1). The expected change per generation is

$$\begin{aligned} \mathbb{E}[\theta_i^{t+1} - \theta_i^t] &= \frac{\sigma_i^2}{s_i} \mu_i + \frac{\delta_i^2}{s_i} \theta_i^t - \alpha \nabla_{\theta} f(\theta_i^t) - \theta_i^t \\ &= \frac{\sigma_i^2}{s_i} (\mu_i - \theta_i^t) - \alpha \nabla_{\theta} f(\theta_i^t). \end{aligned} \quad (5.2)$$

which implies no further accumulation in expectation if $\nabla_{\theta} f(\theta_i)$ reaches the threshold

$$\nabla_i^* = \frac{\sigma_i^2 \mu_i - \theta_i^t}{s_i \alpha}. \quad (5.3)$$

If $f(\theta_i)$ is a quadratic surface with optimum θ_i^* then $f(\theta_i)$ has gradients $\nabla_{\theta} f(\theta_i) = a(\theta_i - \theta_i^*)$ for some constant a . Solving for ∇_i^* leads to equation (2.2) with

$$\beta_i = \frac{a\alpha s}{a\alpha s + \sigma^2} \quad (5.4)$$

(b) Prior estimation

Mathematical results show that cultural transmission by SR in a transmission chain can be analysed as a Markov chain in which the representation (the arrowhead design θ^t in the current context) induced by an individual at generation t depends only on the representation induced by the individual at generation $t - 1$ [56]. If the inferences made by the individuals in this chain can be approximately characterized as Bayesian inference under a prior distribution $p(\theta)$, then it can be shown that (under relatively general technical conditions: see [56]) the chain converges to a stationary equilibrium distribution in which the probability that any individual induces a design θ is the same as the probability of the design under the prior distribution $p(\theta)$. This result indicates that the arrowheads produced by individuals in SR chains can be seen as samples from the prior distribution.

One benefit of this finding is that it justifies analysing the designs produced along the chain, rather than for example being restricted to analysing the designs produced at the final generation of the process, providing a more stable estimate of the properties of the prior distribution (such as its expectation, which we analysed as $\hat{\mu}_i = (1/n) \sum_{k=1}^n \theta_{i,k}$ where k ranges over all $n = 250$ participants assigned to the SR treatment). Each of the 25 SR chains was initialized at one of five arrowhead designs (assigned using block randomization) chosen to offer an unbiased coverage of the design space (designs occupied four corners and the centre of a square positioned at the centre of design space; these designs are pictured in figure 1). Estimating prior distributions by SR has been evaluated by prior research and shown to have desirable characteristics and to accord with alternative methods in the domains of function learning [44] and category learning [73], for example.

(c) Participants

Participants were recruited on Amazon's Mechanical Turk over two sessions (US only, approval rating ≥ 95) and compensated

1 USD for participation plus performance-dependent compensation of up to 0.50 USD. Most participants completed in less than 5 min. Participants who failed attention checks (constructed from pilot data: minimum completion time; minimum reproduction accuracy; at least two unique arrowheads evaluated on learning trials) were excluded and a replacement was immediately re-recruited. No demographic data (e.g. age, gender identity) were collected.

(d) Stimuli

Scores were determined by a quadratic function $f(\theta_i) = (1/2)a(\theta_i - \theta_i^*)^2 + c$, where θ_i^* is the optimum value for feature θ_i . We required a parametrization which did not include negative scores and had an optimum (maximum) score in a semantically reasonable range. To meet these requirements, we set $a = -30$ and $c = 10\,000$ and divided the result by 100. The maximum available score was 1000 calories. To award a score, we computed $f(\theta_i)$ for both arrowhead dimensions and awarded the mean.

(e) Procedure

Participants were assigned to one of five treatments (E1 - E4, SR) using block randomization. The experiment had three trial types: an inheritance trial (IT); testing trials (TT \times 4) and a production trial (PT). In IT, participants observed an arrowhead produced by the participant at the previous generation (or one of five seeds at generation 1). The inherited design was presented in the centre of the screen on a white background for 3000 ms, preceded by a fixation cross (1000 ms). Participants were instructed to reconstruct the design as accurately as possible using HTML range-sliders which dynamically redrew the arrowhead in the centre of the screen. Range-slider initial position was

randomized. Participants could not proceed until both sliders had been moved. The arrowhead appeared upon modification of the sliders. Participants completed one task-familiarization trial (recreating a briefly observed random arrowhead design) before the inheritance trial.

During TT participants tested variations on their reconstruction of the inherited design. Range-sliders allowed testing of designs within 30 units (pixels) either side of the participant's reconstruction. Participants could click a 'Shoot' button at any time to test the current design. Clicking the shoot button triggered an animation in which the current design flew off screen right. Score was presented immediately after test, concluding TT. The participant's initial reconstruction and its score, along with any tested designs and their scores, were presented throughout at the top of the screen. In PT, participants were instructed that the AD produced on this trial will determine a bonus payment, at a rate of 0.10 USD per 200 calories. Previous designs and their scores remained visible. Score feedback was presented after PT, concluding the experiment.

Ethics. This research was approved by the Committee for Protection of Human Subjects at University of California, Berkeley.

Data accessibility. Experimental data, reproducible analyses, experiment implementation codebase and pre-registration are available at <https://osf.io/hx9jd/>.

Competing interests. We declare we have no competing interests.

Funding. This work was supported in part by NSF grant no. 1456709 and DARPA Cooperative agreement no. D17AC00004.

Acknowledgements. We thank Matthew Hardy for feedback. A preliminary version of this research was presented at the 41st Annual Meeting of the Cognitive Science Society and appears in its proceedings [86]. All experimental results presented here are new. We are grateful to the attendees of this meeting for valuable feedback.

References

- Henrich J 2015 *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton University Press.
- Caldwell CA, Atkinson M, Renner E. 2016 Experimental approaches to studying cumulative cultural evolution. *Curr. Dir. Psychol. Sci.* **25**, 191–195. (doi:10.1177/0963721416641049)
- Boyd R, Richerson PJ, Henrich J. 2013 The cultural evolution of technology. In *Cultural evolution* (eds PJ Richerson, MH Christiansen), pp. 119–142. New York, NY: The MIT Press.
- Mesoudi A, Thornton A. 2018 What is cumulative cultural evolution? *Proc. Biol. Sci.* **285**, 20180712.
- Lewis HM, Laland KN. 2012 Transmission fidelity is the key to the build-up of cumulative culture. *Phil. Trans. R. Soc. B* **367**, 2171–2180. (doi:10.1098/rstb.2012.0119)
- Boyd R, Richerson PJ 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Henrich J. 2004 Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *Am. Antiquity* **69**, 197–214. (doi:10.2307/4128416)
- Tomasello M, Kruger AC, Ratner HH. 1993 Cultural learning. *Behav. Brain Sci.* **16**, 495–511. (doi:10.1017/S0140525X0003123X)
- Caldwell CA, Atkinson M, Renner E. 2016 Experimental approaches to studying cumulative cultural evolution. *Curr. Dir. Psychol. Sci.* **25**, 191–195. (doi:10.1177/0963721416641049)
- Caldwell CA, Millen AE. 2008 Experimental models for testing hypotheses about cumulative cultural evolution. *Evol. Human Behav.* **29**, 165–171. (doi:10.1016/j.evolhumbehav.2007.12.001)
- Fay N, De Kleine N, Walker B, Caldwell CA. 2019 Increasing population size can inhibit cumulative cultural evolution. *Proc. Natl Acad. Sci. USA* **116**, 6726–6731. (doi:10.1073/pnas.1811413116)
- Muthukrishna M, Shulman BW, Vasilescu V, Henrich J. 2014 Sociality influences cultural complexity. *Proc. R. Soc. B* **281**, 20132511. (doi:10.1098/rspb.2013.2511)
- Scanlon LA, Lobb A, Tehrani JJ, Kendal JR. 2019 Unknottting the interactive effects of learning processes on cultural evolutionary dynamics. *Evol. Hum. Sci.* **1**, e17. (doi:10.1017/ehs.2019.17)
- Derex M, Bonnefon JF, Boyd R, Mesoudi A. 2019 Causal understanding is not necessary for the improvement of culturally evolving technology. *Nat. Hum. Behav.* **3**, 446–452.
- Derex M, Beugin MP, Godelle B, Raymond M. 2013 Experimental evidence for the influence of group size on cultural complexity. *Nature* **503**, 389–391. (doi:10.1038/nature12774)
- Mesoudi A. 2008 An experimental simulation of the 'copy-successful-individuals' cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evol. Hum. Behav.* **29**, 350–363. (doi:10.1016/j.evolhumbehav.2008.04.005)
- Mesoudi A, O'Brien MJ. 2008 The cultural transmission of great basin projectile-point technology I: an experimental simulation. *Am. Antiquity* **73**, 3–28. (doi:10.1017/S0002731600041263)
- Zwirner E, Thornton A. 2015 Cognitive requirements of cumulative culture: teaching is useful but not essential. *Sci. Rep.* **5**, 16781. (doi:10.1038/srep16781)
- Derex M, Boyd R. 2015 The foundations of the human cultural niche. *Nat. Commun.* **6**, 8398. (doi:10.1038/ncomms9398)
- Derex M, Boyd R. 2016 Partial connectivity increases cultural accumulation within groups. *Proc. Natl Acad. Sci. USA* **113**, 2982–2987. (doi:10.1073/pnas.1518798113)
- Mason W, Watts DJ. 2012 Collaborative learning in networks. *Proc. Natl Acad. Sci. USA* **109**, 764–769. (doi:10.1073/pnas.1110069108)
- Acerbi A, Tennie C, Mesoudi A. 2016 Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans.

- R. Soc. Open Sci.* **3**, 160215. (doi:10.1098/rsos.160215)
23. Creanza N, Feldman MW. 2014 Complexity in models of cultural niche construction with selection and homophily. *Proc. Natl Acad. Sci. USA* **111** (Suppl. 3), 10 830–10 837. (doi:10.1073/pnas.1400824111)
 24. Derex M, Perreault C, Boyd R. 2018 Divide and conquer: intermediate levels of population fragmentation maximize cultural accumulation. *Phil. Trans. R. Soc. B* **373**, 20170062. (doi:10.1098/rstb.2017.0062)
 25. Sperber D, Claidière N. 2008 Defining and explaining culture (comments on Richerson and Boyd, Not by genes alone). *Biol. Phil.* **23**, 283–292. (doi:10.1007/s10539-005-9012-8)
 26. Vapnik VN. 1999 An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **10**, 988–999. (doi:10.1109/72.788640)
 27. Henrich J, Heine SJ, Norenzayan A. 2010 The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83. (doi:10.1017/S0140525X0999152X)
 28. Caldwell CA, Renner E, Atkinson M. 2017 Human teaching and cumulative cultural evolution. *Rev. Phil. Psychol.* **9**, 1–20.
 29. Castro L, Toro MA. 2014 Cumulative cultural evolution: the role of teaching. *J. Theor. Biol.* **347**, 74–83. (doi:10.1016/j.jtbi.2014.01.006)
 30. Ho MK, MacGlashan J, Littman ML, Cushman F. 2017 Social is special: a normative framework for teaching with and learning from evaluative feedback. *Cognition* **167**, 91–106. (doi:10.1016/j.cognition.2017.03.006)
 31. Morgan TJH *et al.* 2015 Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun.* **6**, 6029. (doi:10.1038/ncomms7029)
 32. Yang SCH, Yu Y, Givchi A, Wang P, Vong WK, Shafto P. 2017 Optimal cooperative inference. See <http://arxiv.org/abs/1705.08971>.
 33. Wang J, Wang P, Shafto P. 2020 Sequential cooperative Bayesian inference. See <http://arxiv.org/abs/2002.05706>.
 34. Claidière N, Sperber D. 2007 The role of attraction in cultural evolution. *J. Cogn. Cult.* **7**, 89–111. (doi:10.1163/156853707X171829)
 35. Acerbi A, Charbonneau M, Miton H, Scott-Phillips T. 2019 Cultural stability without copying. See <https://osf.io/vjqc3>.
 36. Kirby S, Dowman M, Griffiths TL, Kirby S, Griffiths TL, Dowman M. 2007 Innateness and culture in the evolution of language. *Proc. Natl Acad. Sci. USA* **104**, 5241–5245. (doi:10.1073/pnas.0608222104)
 37. Kemp C, Perfors A, Tenenbaum JB. 2007 Learning overhypotheses with hierarchical Bayesian models. *Dev. Sci.* **10**, 307–321. (doi:10.1111/j.1467-7687.2007.00585.x)
 38. Boyer P. 1998 Cognitive tracks of cultural inheritance: how evolved intuitive ontology governs cultural transmission. *Am. Anthropol.* **100**, 876–889. (doi:10.1525/aa.1998.100.4.876)
 39. Kirby S, Tamariz M, Cornish H, Smith K. 2015 Compression and communication in the cultural evolution of linguistic structure. *Cognition* **141**, 87–102. (doi:10.1016/j.cognition.2015.03.016)
 40. Verhoef T, Kirby S, de Boer B. 2014 Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *J. Phon.* **43**, 57–68. (doi:10.1016/j.wocn.2014.02.005)
 41. Martin D, Hutchison J, Slessor G, Urquhart J, Cunningham SJ, Smith K. 2014 The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychol. Sci.* **25**, 1777–1786. (doi:10.1177/0956797614541129)
 42. Saldana C, Fagot J, Kirby S, Smith K, Claidière N. 2019 High-fidelity copying is not necessarily the key to cumulative cultural evolution: a study in monkeys and children. *Proc. R. Soc. B* **286**, 20190729. (doi:10.1098/rspb.2019.0729)
 43. Ravigani A, Delgado T, Kirby S. 2017 Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* **1**, 1–7. (doi:10.1038/s41562-016-0007)
 44. Kalish ML, Griffiths TL, Lewandowsky S. 2007 Iterated learning: intergenerational knowledge transmission reveals inductive biases. *Psychon. Bull. Rev.* **14**, 288–294. (doi:10.3758/BF03194066)
 45. Claidière N, Scott-Phillips TC, Sperber D. 2014 How Darwinian is cultural evolution? *Phil. Trans. R. Soc. B* **369**, 20130368. (doi:10.1098/rstb.2013.0368)
 46. Nettle D. 2020 Selection, adaptation, inheritance and design in human culture: the view from the Price equation. *Phil. Trans. R. Soc. B* **375**, 20190358. (doi:10.1098/rstb.2019.0358)
 47. Dietrich A. 2004 The cognitive neuroscience of creativity. *Psychon. Bull. Rev.* **11**, 1011–1026. (doi:10.3758/BF03196731)
 48. Heyes C. 2018 Enquire within: cultural evolution and cognitive science. *Phil. Trans. R. Soc. B* **373**, 20170051. (doi:10.1098/rstb.2017.0051)
 49. Fogarty L, Creanza N, Feldman MW. 2015 Cultural evolutionary perspectives on creativity and human innovation. *Trends Ecol. Evol.* **30**, 736–754. (doi:10.1016/j.tree.2015.10.004)
 50. Borji A, Itti L. 2013 Bayesian optimization explains human active search. See <https://dl.acm.org/doi/10.5555/2999611.2999618>.
 51. Mason WA, Jones A, Goldstone RL. 2008 Propagation of innovations in networked groups. *J. Exp. Psychol.: General* **137**, 422–433. (doi:10.1037/a0012798)
 52. Yehosseini KS, Moussaid M. 2020 Comparing groups of independent solvers and transmission chains as methods for collective problem-solving. *Sci. Rep.* **10**, 1–9. (doi:10.1038/s41598-020-59946-9)
 53. Wu Y, Ren M, Liao R, Grosse R. 2018 Understanding short-horizon bias in stochastic meta-optimization. See <http://arxiv.org/abs/1803.02021>.
 54. Kendal RL, Boogert NJ, Rendell L, Laland KN, Webster M, Jones PL. 2018 Social learning strategies: bridge-building between fields. *Trends Cogn. Sci.* **22**, 651–665. (doi:10.1016/j.tics.2018.04.003)
 55. Miton H, Charbonneau M. 2018 Cumulative culture in the laboratory: methodological and theoretical challenges. *Proc. R. Soc. B* **285**, 20180677. (doi:10.1098/rspb.2018.0677)
 56. Griffiths TL, Kalish ML. 2007 Language evolution by iterated learning with Bayesian agents. *Cogn. Sci.* **31**, 441–480. (doi:10.1080/15326900701326576)
 57. Osieurak F, Reynaud E. 2019 The elephant in the room: what matters cognitively in cumulative technological culture. *Behav. Brain Sci.* **43**, e156. (doi:10.1017/S0140525X19003236)
 58. Tennie C, Call J, Tomasello M. 2009 Ratcheting up the ratchet: on the evolution of cumulative culture. *Phil. Trans. R. Soc. B* **364**, 2405–2415. (doi:10.1098/rstb.2009.0052)
 59. Sperber D 1996 *Explaining culture: a naturalistic approach*. Oxford, UK: Blackwell.
 60. Chater N, Vitányi P, Vitanyi P. 2003 Simplicity: a unifying principle in cognitive science. *Trends Cogn. Sci.* **7**, 19–22. (doi:10.1016/S1364-6613(02)00005-0)
 61. Smith KA, Battaglia PW, Vul E. 2018 Different physical intuitions exist between tasks, not domains. *Comput. Brain Behav.* **1**, 101–118. (doi:10.1007/s42113-018-0007-3)
 62. Kemp C, Tenenbaum JB. 2008 The discovery of structural form. *Proc. Natl Acad. Sci. USA* **105**, 10 687–10 692. (doi:10.1073/pnas.0802631105)
 63. Canini KR, Griffiths TL, Vanpaemel W, Kalish ML. 2014 Revealing human inductive biases for category learning by simulating cultural transmission. *Psychon. Bull. Rev.* **21**, 785–793. (doi:10.3758/s13423-013-0556-3)
 64. Navarro DJ, Perfors A, Kary A, Brown SD, Donkin C. 2018 When extremists win: cultural transmission via iterated learning when populations are heterogeneous. *Cogn. Sci.* **42**, 2108–2149. (doi:10.1111/cogs.12667)
 65. Perfors A, Navarro DJ. 2014 Language evolution can be shaped by the structure of the world. *Cogn. Sci.* **38**, 775–793. (doi:10.1111/cogs.12102)
 66. Thompson B, Kirby S, Smith K. 2016 Culture shapes the evolution of cognition. *Proc. Natl Acad. Sci. USA* **113**, 4530–4535. (doi:10.1073/pnas.1523631113)
 67. Henrich J, Boyd R. 2002 On modeling cognition and culture: why cultural evolution does not require replication of representations. *J. Cogn. Cult.* **2**, 87–112. (doi:10.1163/156853702320281836)
 68. Kingma DP, Ba JL. 2015 Adam: a method for stochastic optimization. In *3rd Int. Conf. on Learning Representations, ICLR 2015: conference track proceedings*. See <http://arxiv.org/abs/1412.6980v9>.
 69. Smaldino PE, Richerson PJ. 2013 Human cumulative cultural evolution as a form of distributed computation. In *Handbook of human computation* (ed. P. Michelucci), pp. 979–992. New York, NY: Springer.
 70. Xu J, Griffiths TL. 2010 A rational analysis of the effects of memory biases on serial reproduction. *Cogn. Psychol.* **60**, 107–126. (doi:10.1016/j.cogpsych.2009.09.002)
 71. Mesoudi A. 2011 An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evol.*

- Hum. Behav.* **32**, 334–342. (doi:10.1016/j.evolhumbehav.2010.12.001)
72. Mesoudi A, Chang L, Murray K, Lu HJ. 2015 Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proc. R. Soc. B* **282**, 20142209. (doi:10.1098/rspb.2014.2209)
 73. Griffiths TL, Christian BR, Kalish ML. 2008 Using category structures to test iterated learning as a method for identifying inductive biases. *Cogn. Sci.* **32**, 68–107. (doi:10.1080/03640210701801974)
 74. Morin O. 2016 Reasons to be fussy about cultural evolution. *Biol. Phil.* **31**, 447–458. (doi:10.1007/s10539-016-9516-4)
 75. Watson RA, Szathmáry E. 2016 How can evolution learn? *Trends Ecol. Evol.* **31**, 147–157. (doi:10.1016/j.tree.2015.11.009)
 76. Kouvaris K, Clune J, Kounios L, Brede M, Watson RA. 2017 How evolution learns to generalise: using the principles of learning theory to understand the evolution of developmental organisation. *PLoS Comput. Biol.* **13**, e1005358. (doi:10.1371/journal.pcbi.1005358)
 77. Mandt S, Hoffman MD, Blei DM. 2017 Stochastic gradient descent as approximate Bayesian inference. See <http://arxiv.org/abs/1704.04289>.
 78. Harper M. 2009 The replicator equation as an inference dynamic. See <http://arxiv.org/abs/0911.1763>.
 79. Suchow JW, Bourgin DD, Griffiths TL. 2017 Evolution in mind: evolutionary dynamics, cognitive processes, and Bayesian inference. *Trends Cogn. Sci.* **21**, 522–530. (doi:10.1016/j.tics.2017.04.005)
 80. Maclaurin D, Duvenaud D, Adams RP. 2015 Early stopping is nonparametric variational inference. See <http://arxiv.org/abs/1504.01344>.
 81. Aitchison L. 2018 A unified theory of adaptive stochastic gradient descent as Bayesian filtering. See <http://arxiv.org/abs/1807.07540>.
 82. Krogh A, Hertz JA. 1992 A simple weight decay can improve generalization. See <https://papers.nips.cc/paper/1991/file/8eefcfd5990e441f0fb6f3fad709e21-Paper.pdf>.
 83. Griffiths TL, Kemp C, Tenenbaum JB. 2008 Bayesian models of cognition. In *Cambridge handbook of computational psychology* (ed. R Sun), pp. 59–100. Cambridge, UK: Cambridge University Press.
 84. Jordan C, Jordán K. 1965 *Calculus of finite differences*, vol. 33. Providence, RI: American Mathematical Society.
 85. Boyd S, Boyd SP, Vandenberghe L. 2004 *Convex optimization*. Cambridge, UK: Cambridge University Press.
 86. Thompson B, Griffiths TL. 2019 Inductive biases constrain cumulative cultural evolution. In *Proc. annual meeting of the Cognitive Science Society*, vol. **41**. See <https://cogsci.mindmodeling.org/2019/papers/0205/index.html>.