INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Vidiella B, Carrignon S, Bentley RA, O'Brien MJ, Valverde S. 2022 A cultural evolutionary theory that explains both gradual and punctuated change. *J. R. Soc. Interface* **19**: 20220570.
https://doi.org/10.1098/rsif.2022.0570

Received: 6 August 2022
Accepted: 24 October 2022

Subject Category:
Life Sciences–Earth Science interface

Subject Areas:
evolution, environmental science

Keywords:
cultural evolution, punctuated evolution, social learning, transparency, popularity bias

Author for correspondence:
Sergi Valverde
e-mail: sergi.valverde@ibe.upf-csic.es

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6277167.

THE ROYAL SOCIETY PUBLISHING

A cultural evolutionary theory that explains both gradual and punctuated change

Blai Vidiella[1], Simon Carrignon[2], R. Alexander Bentley[3], Michael J. O'Brien[4,5] and Sergi Valverde[1,6]

[1]Evolution of Networks Lab, Institute of Evolutionary Biology (UPF-CSIC), Passeig Marítim de la Barceloneta 37, 08003 Barcelona, Spain
[2]McDonald Institute for Archaeological Research, Downing Street, Cambridge CB2 3ER, UK
[3]Department of Anthropology, University of Tennessee, Knoxville, TN 37996, USA
[4]Department of Communication, History, and Philosophy and Department of Life Sciences, Texas A&M University–San Antonio, Texas 78224, USA
[5]Department of Anthropology, University of Missouri-Columbia, Missouri 65201, USA
[6]European Centre for Living Technology (ECLT), Ca' Bottacin, 3911 Dorsoduro Calle Crosera, 30123 Venezia, Italy

BV, 0000-0002-4819-7047; SC, 0000-0002-4416-1389; RAB, 0000-0001-9086-2197; SV, 0000-0002-2150-9610

Cumulative cultural evolution (CCE) occurs among humans who may be presented with many similar options from which to choose, as well as many social influences and diverse environments. It is unknown what general principles underlie the wide range of CCE dynamics and whether they can all be explained by the same unified paradigm. Here, we present a scalable evolutionary model of discrete choice with social learning, based on a few behavioural science assumptions. This paradigm connects the degree of transparency in social learning to the human tendency to imitate others. Computer simulations and quantitative analysis show the interaction of three primary factors—information transparency, popularity bias and population size—drives the pace of CCE. The model predicts a stable rate of evolutionary change for modest degrees of popularity bias. As popularity bias grows, the transition from gradual to punctuated change occurs, with maladaptive subpopulations arising on their own. When the popularity bias gets too severe, CCE stops. This provides a consistent framework for explaining the rich and complex adaptive dynamics taking place in the real world, such as modern digital media.

1. Introduction

Cumulative cultural evolution (CCE) [1–6], in which innovations accumulate over time through social learning, has been integral to human evolution [7–9] and inter-generational cultural adaptations of small traditional societies [10–15]. When expertise and/or performance are transparent, the rate of CCE correlates with the number of interacting individuals [16–19], in the complexity of forager assemblages [20–22] and in controlled social psychology experiments [23–27] in which small groups can outperform the most skilled/knowledgeable group member on short-term tasks [28–32].

It is not clear what implications CCE studies in small groups should have for larger populations, such as the urban environments humans have lived in for millennia. Extrapolating the hypothesized correlation between CCE and population size, it would seem that larger social learning networks would surface the best technologies [16,33,34], productive organizations [35], government institutions [36,37] and technical knowledge [38–40]. As innovation rate scales with population density [18,41], however, the number of similar options can

increase by orders of magnitude, and social learners need to update more frequently to keep up. While copying recent success is an adaptive strategy in a highly variable environment [26,42–44] at some point this capacity may be overwhelmed. When intrinsic pay-offs are no longer transparent, copying recent *popularity* can become a substitute for copying recent success—an understandable shortcut by human psychology evolved for a few hundred social relationships [9,45–48]. Long-run persistence of old information may limit the adoption of adaptations in this scenario [49].

Previous models must be expanded in order to disentangle the interplay between individual and collective levels. Game theory models of binary options, which are effective at characterizing social conformity in animals and small-scale human societies [44,50], cannot adequately reflect modern environments in which numerous 'games' are being played concurrently among thousands, if not millions, of agents competing for popularity of their views. The bewildering array of choices is not reducible to a single binary decision, and the transparency of information and social learning criteria vary widely, ranging from zero to a vast range [51,52].

In order to bridge this 'population gap' in CCE theory, here we propose a non-equilibrium model that can accommodate any number of similar options as well as varying degrees of information transparency and degree of popularity bias in a population. To recreate a wide range of cultural behaviours, we aim here for the simplest model that relies on the fewest assumptions and parameters. When simulated through heterogeneous interactions, the transparency–popularity link produces a spectrum of collective dynamics spanning from gradual [53] to punctuated change [17,34,54,55] that is essentially independent of specific system properties. Here, discontinuous events do not require highly skilled individuals making 'great leaps' [17]. Instead, a frequency-dependent balance of transparency and popularity bias limits the pace of cultural evolution, resulting in stasis associated with spontaneously arising 'barriers' that only infrequent events may overcome. In the future, this framework might provide a broad foundation for CCE, allowing it to be adjusted to specific real-world contexts.

2. Evolutionary model of discrete choice with social learning

Our conceptual model [56] consists of, firstly, the continuum between individual and social learning (horizontal axis in figure 1), and secondly, transparency of information or learning criteria (vertical axis). This model, grounded in discrete theory under social influence [57], unifies a range of approaches, from those emphasizing intrinsic pay-offs of the choices [58], learning from experts or successful individuals [22,42,59], all the way to models that assume copying is done with zero transparency of pay-offs or expertise [52,60–66].

This continuous parameter space includes certain well-known reference points. Rational decisions, per the standard social science model, are individual and transparent (upper left in figure 1), whereas learning from experts, per much of cultural evolutionary theory, is transparent and social (upper right). The bottom half of the continuous space includes random copying (lower right) and

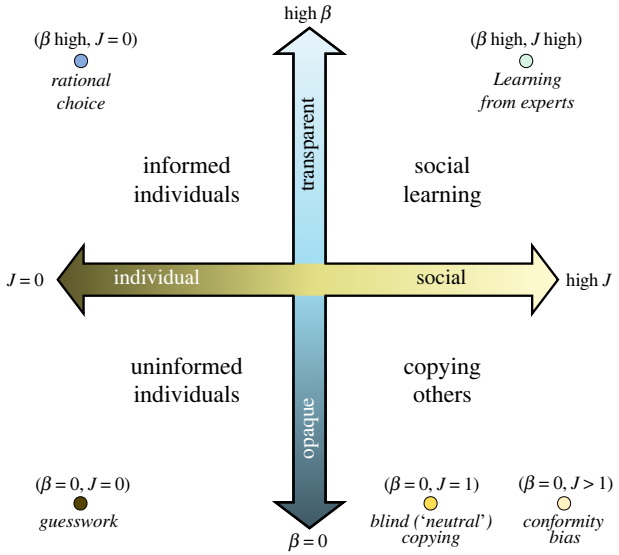

Figure 1. Conceptual 'dimensions' of information transmission, with points indicating established models in the continuous space.

guesswork as opaque individual learning (lower left). As these disparate models are reference points within a continuous space of possibility (figure 1), the framework could serve as a conceptual bridge from CCE in small-scale experiments and traditional societies to rapid change under massive, globalized communication.

Although there are many strategies for social learning [59], here we start pragmatically with popularity bias as the horizontal dimension, which we parametrize as J [56,67,68]. High J means doing as others do. The second key parameter, the vertical dimension, is transparency of learning, which ranges from informed to uninformed [56,69–71]. We parametrize information transparency as β, where high β enables selection for the best option. As the parameters are continuous, unlimited intermediate scenarios are possible.

Transparency represents an individual's sensitivity to differences in choice, acting on the intrinsic utility difference between options. Effectively the weight of individual learning, transparency amounts to the extent to which one's behavioural choice is influenced by the objective pay-off related to that behaviour. In the absence of popularity bias ($J = 0$), the larger the transparency is, the smaller is the variance in decision-making across the alternatives. When transparency is near zero in the absence of popularity bias, choice is random over the choice set, and each option is chosen with the same probability. When transparency is large, the relative values of pay-offs of each choice are high, such that the choice with the highest pay-off is reliably identified.

In each time step of the model, a new set of N choices are made. This could represent a new 'generation' of N agents that replace the previous generation, or it could represent the N agents making a new choice to replace the choice held in the previous time step. As both are mathematically equivalent, the model can represent successive generations over long time periods of cultural evolution, or successive time intervals within a time span as agents make Bayesian updates to their decisions. It could also represent successive samples, of size N of a large interconnected population,

whose choices are evolving through time. Popular choices have their own 'inertia' by virtue of stochastic change under popularity bias and/or the persistence of choices with high intrinsic utility under significant transparency.

The model proceeds in each time interval with each of the N agents identifying another agent to potentially copy. First, an option, i, is observed with probability $Pr(i, t+1)$ among all the $k(t)$ different alternatives in the most recent time interval. Following cultural evolutionary theory [72,73], we set this probability to its frequency, $p_i(t)$, modified by exponent J

$$Pr(i, t+1) \propto p_i(t)^J, \tag{2.1}$$

where the parameter J can bias this frequency-dependent selection. This popularity bias parameter ranges continuously from $J = 0$ for zero frequency effect, to $J = 1$ for probability in strict proportion to frequency (random copying) to $J > 1$ for conformist bias.

Subsequently, agent evaluates the intrinsic utility, U_i of its chosen option i and keeps it with probability, $Pr(i, t+1)$, which is determined logistically by βU_i, where β is the transparency of choice [22,57,74–76]

$$Pr(i, t+1) \propto e^{\beta U_i + \psi}, \tag{2.2}$$

where ψ represents the random Gaussian error in the pay-off estimation [77,78]. Equation (2.2) is typical in formulations of discrete choice theory or quantal response theory [57,79,80]. For parsimony, we leave aside the matter of different intrinsic preferences [81], which would necessitate an additional error term on utilities U_i.

Combining the steps, equations (2.1) and (2.2), yields the probability, $Pr(i, t+1)$, that the agent selects variant i at time $t+1$

$$Pr(i, t+1) = \frac{1}{Y_t} p_i(t)^J e^{\beta U_i + \psi}, \tag{2.3}$$

where Y_t is a normalizing term across all $k(t)$ variants

$$Y_t = \sum_{j=1}^{k} p_j(t)^J e^{\beta U_j + \psi}. \tag{2.4}$$

Alternatively, we can define the propensity $\Pi(U_i, p_i)$ of choosing the trait with utility U_i and popularity p_i as follows:

$$\Pi(U_i, p_i) = Pr(i, t+1) Y_t = p_i(t)^J e^{\beta U_i} e^{\psi}. \tag{2.5}$$

Equations (2.3)–(2.5) span a decision space. The formulation also informs our basic expectations at certain reference points. In the high-transparency realm without social learning ($J = 0$), we recover (2.2), which is a form of bounded rationality. We will see in our results what happens as J is increased from this reference point. At the other end of the spectrum, along the zero-transparency extreme, if $\beta = 0$ and $J = 0$, there is random selection among the $k(t)$ alternatives, $Pr(i, t+1) = 1/k(t) + \psi$. If $\beta = 0$ and $J = 1$, then it becomes a random copying, or Yule, model, where $Pr(i, t+1) \propto p_i(t) e^{\psi}$, i.e. proportional to approximated frequency.

Lastly, a small fraction of agents, μ, invents something new by modifying an existing variant i, such that its pay-off becomes $U_i + \epsilon$, where the random perturbation ϵ is drawn from a normal distribution with mean zero. In simulations, we vary μ from 0.005 to 0.1, consistent with ranges proposed for human invention [82–85]. With new index $k(t) + 1$, this new variant becomes part of the pool from which agents may choose in the next time step. Iterating (2.3) generates different probability distribution functions from the same starting point, under conditions of high transparency β or high popularity bias J.

3. Results

The model is multifarious, with an endless number of alternative outcomes. Here, we examine how the major modes of behaviour are affected by β and J, using a range of population sizes N and low mutation rates in our simulations (more comprehensive findings are reported in the electronic supplementary material throughout a grid of β–J pairings). We also provide analytical predictions for how the rate of utility gain (CCE) varies not just with population size but also with the J/β ratio.

Figure 2 shows the evolution of utility values chosen by 1000 agents over 200 time steps, showing a rich diversity of behaviours from this simple model. We identify general evolutionary regimes in the β–J space, including steady evolution, punctuated evolution, stochastic drift and random noise (clockwise from upper left in figure 2). Under zero transparency (bottom half of figure 2), there is little to no regular increase in utility. Without transparency or popularity bias ($\beta = 0$, $J = 0$), there is noise (figure 2, lower left). With popularity bias and no transparency ($\beta = 0$, $J = 1.5$), as in the lower right of figure 2, there are two notable effects. First, a majority of agents never escape their choices (horizontal red bar). Second, among the remaining minority of agents, a drifting 'consensus' (greenish band)—a minority of agents overall, but a majority of those not in the red band—is held together by popularity bias. If we were to draw a cross-section at a time step in the $\beta = 0$, $J = 1.5$ case, the frequency distribution would have multiple modes: a sharp peak at the majority utility value (red bar), but also a wider, secondary mode (greenish band) and other smaller modes, drifting through time.

At the top of figure 2, transparency β facilitates a cumulative increase in median utility, as we would expect. For positive transparency ($\beta = 0.25$) in the upper half of figure 2, the utility values undergo steady increase for $J = 0$ (upper left) and punctuated increases for $J = 1.5$ (upper right). Comparing utility increase by the end of these two simulations under the same value of $\beta = 0.25$ indicates that CCE can be faster with popularity bias ($J = 1.5$) than without ($J = 0$). The progress at ($\beta = 0.25$, $J = 1.5$) is not optimal, however—for example, the group drifting into negative utility values in the first 50 time steps and other suboptimal 'tendrils' are visible throughout that simulation (figure 2, upper right). This resembles homophily, via sorting into subpopulations around drifting modal utility values, some of which decline. This sorting of the population through popularity bias simply emerges, without having imposed any network or group structure on the model, or any intrinsic preferences instilled among the agents.

The optimal rate of CCE should require moderate values of both transparency and popularity bias. Locating the exact optimum β–J coordinate is not trivial, however, as it depends on the other parameters (μ, N) and requires a large number of simulations [68]. More coarsely, we can determine that CCE (rate of utility increase) is typically optimized with J at or slightly above 1, as long as transparency β is positive.

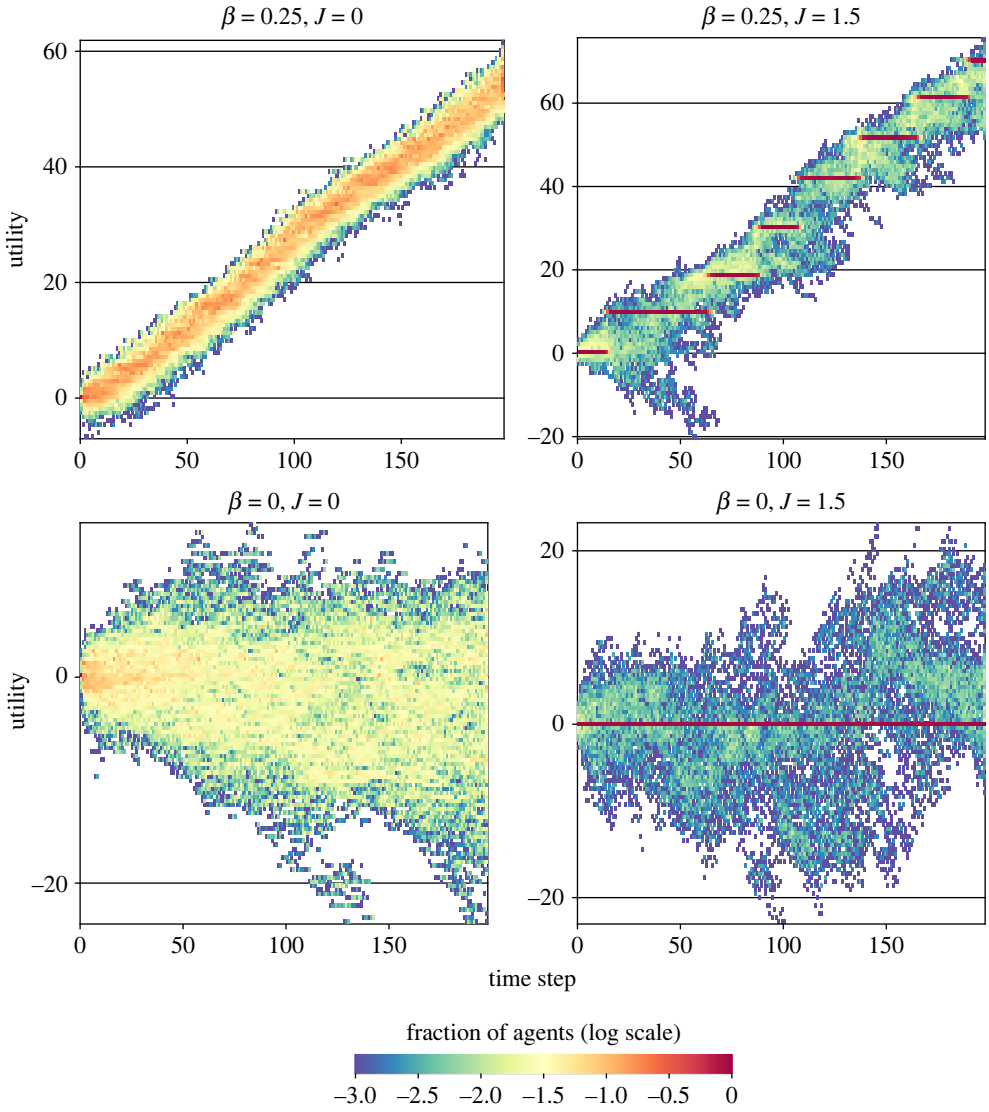

Figure 2. Trait utilities (*y*-axes) through time (*x*-axes) in simulations with different values of transparency (β) and popularity bias (J). Zero utility is the initial reference value. Colour indicates the relative fraction of agents in logarithmic scale (red = 100%). Note different *y*-axis range for each plot. In each simulation for 200 time steps, $N = 1000$ agents, $\mu = 0.1$, $\psi = 1$. Despite the system's stochastic nature, the reported behaviours are typical (statistically). The electronic supplementary material shows more replicas as well as a strong statistical signature in the rank frequency distributions associated with each behaviour.

Computational simulations show CCE decreasing as J is moved away from 1 (figure 3). For large populations and low transparency values, this optimal CCE occurs towards $J = 1$ (discussed below).

Note the increases in utility on the top row of figure 2 are smooth with $J = 0$ and punctuated with $J = 1.5$. In between, we find the transition from gradual to punctuated evolution lies at $J = 1$, which is simple frequency-proportional copying. Increasing J above 1, while maintaining sufficient β (see below), induces stagnation where plateaus of utility emerge for extended periods, with sporadic jumps to higher plateaus. The population bifurcates upon each new jump, as some agents continue copying the same choice while other agents increase their pay-offs with better (higher utility) choices. We can zoom in to look for early warning signals in time series activity before each jump, such as increased variance before a critical transition [86]. Indeed, for $J > 1$, we do see such signals. In a case with $\beta = 0.1$ and $J = 1.75$, figure 4*a* shows the utility values of each agent choice over the first 1000 time steps, which increase cumulatively through punctuated leaps. Each red bar indicates a large concentration of individual pay-offs around one utility value; each red point indicates nearly all the agents have copied the same variant (see §3 of the electronic supplementary material for a three-dimensional rendering of the same dataset). Above and below each red plateau in utility are those agents copying variants that differ from the majority. Just before each jump, the variance in utility values, as well as the entropy, increases as a spike (figure 4*b*). Aggregating results from 20 steps before and after each jump, figure 4*c* reveals the abrupt rise and fall of variance and entropy in utility values before and after each jump. This is a recurrent (and statistically repeatable) pattern (figure 4*c*).

Another approach to explain the punctuated mode is to consider whether uncommon utility gains are likely to be adopted by others. It has been argued that social systems exist in a condition of stasis for extended periods of time, punctuated by rapid shifts resulting in radical transformations, which are frequently associated with major inventions [34]. Here, we investigate whether punctuated changes are an emergent characteristic of an inherent conflict between transparency and popularity bias. Such 'utility barriers' to improvement (figure 5*a*) resemble valleys on a fitness landscape [87], but they are a fundamentally

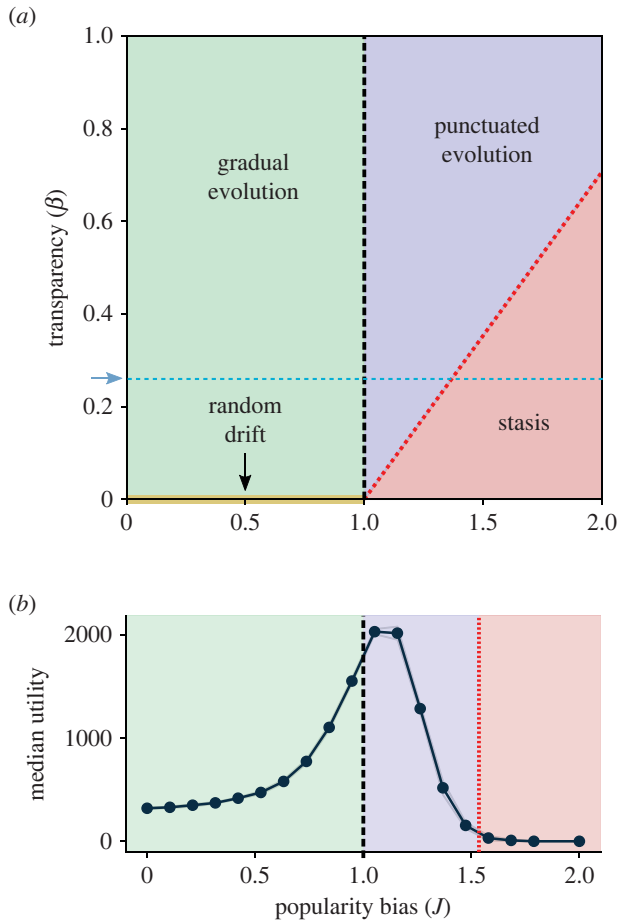

Figure 3. Theoretical boundaries for cumulative cultural evolution defined by equations (3.1) and (5.2). The $J = 1$ transition divides this space between gradual ($J < 1$) and punctuated evolution ($J > 1$). Random drift without selection occurs at $\beta = 0$ (yellow line), where $\Delta_{Uc} \sim -\infty$. In the red area in the lower right, there is no CCE as utility barriers are no longer reached. Bottom panel shows average median utility at $\beta = 0.25$ (blue arrow on top), after T time steps in 10 replicates (see Methods). Median utility is maximized at $J \approx 1.1$ (the so-called 'soft spot' of social learning, see text), then declines as J is increased further. Errors bars are within the line width.

dynamical phenomena caused by cultural evolutionary mechanisms rather than the underlying landscape features. In the mode of punctuated evolution, the population alternates between periods of apparent stasis and abrupt shifts to the next utility $U + \Delta_U$ (light grey region in figure 6). When J is larger than 1, utility barriers emerge, such that small utility increases below a threshold ($\Delta_U < \Delta_{Uc}$) are not selected. Depending on Δ_{Uc}, the system may be unable to select for new variants of higher utility, due to their low initial frequency (red bar in figure 2, lower right), such that some agents remain stuck with their initial utility value, U_0.

Figure 5 illustrates the effect on CCE of the probability, $P(U_{t+1} > U_t)$, that an agent obtains a higher utility in the next time step $t + 1$ (see Methods). For each new invention (produced at rate μ), the utility change, (Δ_U), from the agent's existing variant is drawn from a Gaussian distribution. The Gaussian distribution is centred at zero, so (Δ_U) can be either beneficial (+) or deleterious (−). Figure 5*a* shows three different scenarios of increasing conformity from bottom to top. Below a utility barrier (Δ_{Uc}), the inventions will probably not be copied (figure 5*a*). Increasing popularity bias J raises the value of the utility barrier (figure 5*a*, bottom to top), such that popularity bias that is too strong ($J > 1$) will slow down CCE (figure 5*b*,*c*). Conversely, if popularity bias is too weak or absent, CCE is not maximized either (figure 5*b*,*c*). The optimal value is very close to $J = 1$, where maladaptations are forbidden and every possible random change can be selected by the population.

The suggested CCE 'soft spot' reflects the utility barrier being minimized at $J = 1$. Overcoming these utility barriers requires the accumulation of sufficient diversity in utility values within the population, as generated by the Gaussian error, ψ, in each agent's pay-off estimation. While population size determines the number of inventions, μN, the stronger the popularity bias, J, the more trials (inventions) the population needs to make to be likely to overcome the utility barrier $\Delta_{Uc} > 0$. Since β and J determine the utility barrier, the relationship between population size and CCE (utility increase rate) will also depend on these parameters. As derived in the Methods section, we can describe the minimal utility improvement, Δ_U, that can reliably be copied by at least one agent in one time step, as

$$\Delta_U \geq \Delta_{Uc} = \frac{J-1}{\beta} \ln(N-1). \tag{3.1}$$

Note $\Delta_{Uc} = 0$ when $J = 1$. As equation (3.1) shows, utility barriers are lowered by transparency β and increased by popularity bias J and the logarithm of population size N.

The larger the popularity bias, the larger transparency β needs to be for utility gains to be discovered. Figure 6 further highlights the interplay between these variables, the utility barrier and CCE. In the Δ_U versus J space of figure 6*b*, the prediction of (3.1) is the blue diagonal—changes can occur above this line. Whether changes are positive is delineated by the horizontal at $\Delta_U = 0$. The space in figure 6*b* is thus divided into four distinct regions: adaptive behaviour (light blue), maladaptive learning (red), forgettable (white) and 'missed opportunities' (grey). The adaptive (blue) region yields gradual, steady evolution without significant plateaus. In the red triangular (maladaptive) region, the population may adopt negative utility displacements, which can send subpopulations into decline (see down-sloping tendrils in figure 4*a*). The grey area is called missed opportunities because this is where new inventions with positive utility differences are not adopted because they lie below the utility barrier. These spaces in figure 6*b* change with the ratio $\Delta_{Uc} \approx \ln N/\beta$: as this ratio increases, the boundary rotates counterclockwise, expanding the region of maladaptive adoptions (red) and missed opportunities (grey), that is, the maladaptive parameter space increases and adaptive space decreases.

When conformity is significant ($J > 1$), utility improvements (Δ_U) below the critical utility barrier Δ_{Uc}, or utility barrier, cannot be widely adopted by the population. The duration of the utility plateaus increases with J until eventually, at some conformity limit, J_C, the utility barrier is never overcome and CCE ceases. As derived in the Methods (5.2), we can predict how J_C increases with population size, N

$$J < J_C = 1 + \beta \sqrt{\frac{2\log(N/2)}{\log(N-1)}} \tag{3.2}$$

which, if we assume that population size is large enough (then $N - 1 \approx N$), indicates that CCE can occur as long as $J \leq 1 + \sqrt{2}\beta$. In other words, the crucial ratio is $(J-1)/\beta$. If this ratio exceeds the variance of the innovation distribution

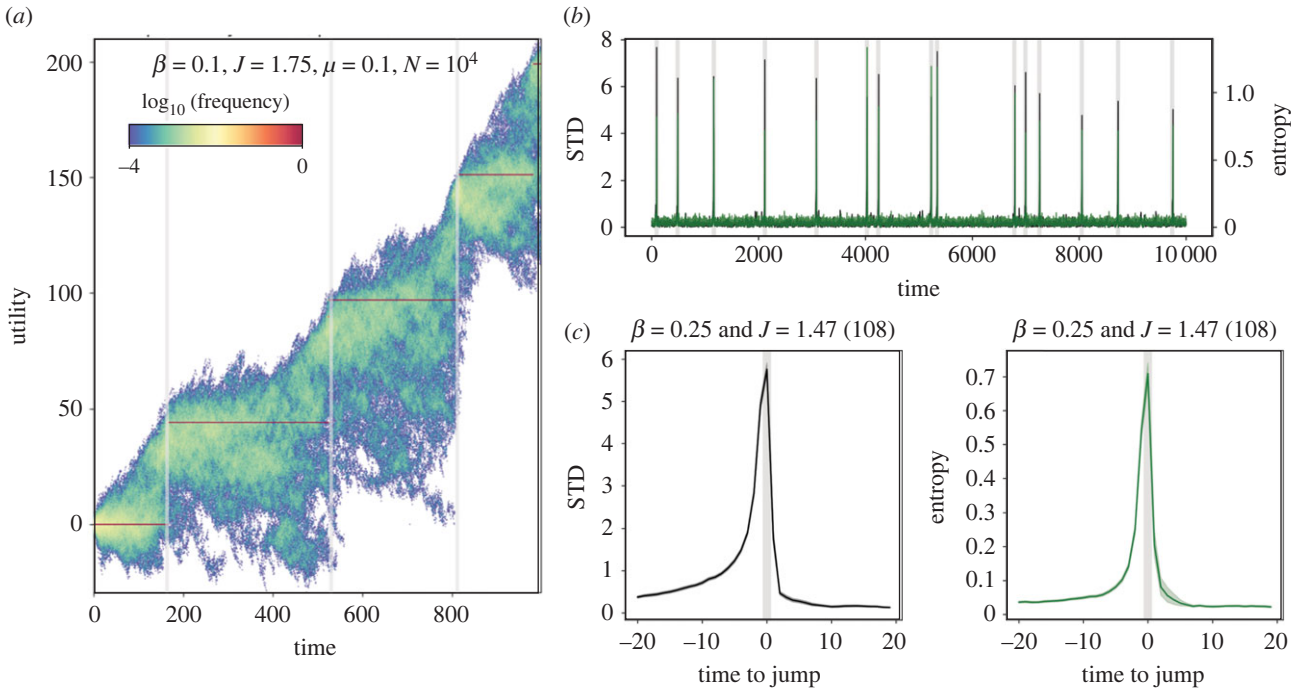

Figure 4. Detecting abrupt jumps in evolutionary cultural dynamics. (*a*) Punctuated pay-off increases take place under low transparency and high conformity. Colour indicates their population frequency (in logarithmic scale, see colour bar). (*b*) Trait diversity peaks very close to sharp jumps in median utility, which take place when the population overcomes the underlying utility barrier (see text). Here, we measure trait diversity using entropy (green curve) and standard deviation (black curve). (*c*) Aggregated statistics collapsing 108 jumps. Simulation parameters: $N = 10^4$ individuals, $\beta = 0.1$, $J = 1.75$, $\mu = 0.1$, $m = 1$, $\psi = 1$, $T_{max} = 10^4$ timesteps.

(approx. 1.4 for the Gaussian), CCE ceases. As long as popularity bias is less than J_C, CCE can proceed. When the bias exceeds J_C, most selections remain stuck at their initial value, U_0. Figure 7 shows how the conformity limit, J_C, depends on population size. Increasing population from small to medium sizes pushes the boundary up to higher J values, but this effect is asymptotic as the boundary converges to a finite quantity for very large populations.

4. Discussion

Unifying distinct aspects of different behavioural disciplines [88], we established a cultural evolutionary model based on the interaction between social learning and transparency of information. As the model is non-deterministic and every simulation is unique, we focus on general effects, their implications for cultural evolution, and future potential modifications. The model is relatively simple, but there are numerous variants, aspects and components of individual and social learning that were left out; we will go through some of them below. This parsimony was essential in order to examine model behaviours, which were already quite rich with just the few parameters used.

Our study demonstrates multiple phases in the space of CCE, ranging from stasis to stochastic drift, gradual change and punctuated change. By integrating these behaviours through few parameters, the model suggests fundamental insights. Firstly, it begins to untangle the much-discussed dependence of CCE on population size [20,21,24,69]. In the model, this relationship depends on a interaction between information transparency β and popularity bias J. Secondly, while CCE is facilitated by the combination of β and J, it is usually optimized close to $J = 1$, or simple frequency-dependence, which we suggest was significant in human evolution. Third, when popularity bias increases to conformist bias $J > 1$, changes in median utility gains become punctuated, with spikes in variance before each leap that may serve as early warning signals. It is worth noting that we do not incorporate any extra (e.g. cognitive) mechanisms to account for the emergence of punctuated changes [17,89]. That is, the model shows how gradual and punctuated change are two sides of the same basic set of evolutionary rules. Conformist bias also induces homophily, where suboptimal groups lag behind the population's best utility. When the conformist bias becomes too strong, CCE ceases totally, which might be a dramatic transition.

In terms of cultural evolution, the parameters of information transparency and popularity bias are fundamental [56]. Consistent with expectations, popularity bias and information transparency combine to maximize CCE. While the best combination depends on other parameters, generally innovation is optimized near $J = 1$. When popularity bias is increased from frequency dependence ($J = 1$) into conformist bias ($J > 1$), it begins to hinder innovation until eventually ($J \gg 1$) collapses into a single choice. Conformist bias ($J > 1$) also underlies a punctuated mode of change. As J is increased beyond 1, stasis periods emerge and then become longer as J increases. The pauses in CCE reflect clustering of agents that become stuck at suboptimal utility values through conformist bias toward popular traits. In a process resembling homophily, these maladaptive groups emerge without any network structure [23,30] or changes in invention process [17] imposed upon the model.

In the model, increasing J modestly above 1 can compensate for a decrease in information transparency β. Conversely, the effects of excessive conformity ($J \gg 1$) can be countered by increasing information transparency β, but the larger population size N, the more transparency is needed to overcome the utility barriers induced by conformist bias. Increasing J too far, without increasing β to compensate, stalls CCE. The relationship between CCE and population size depends on

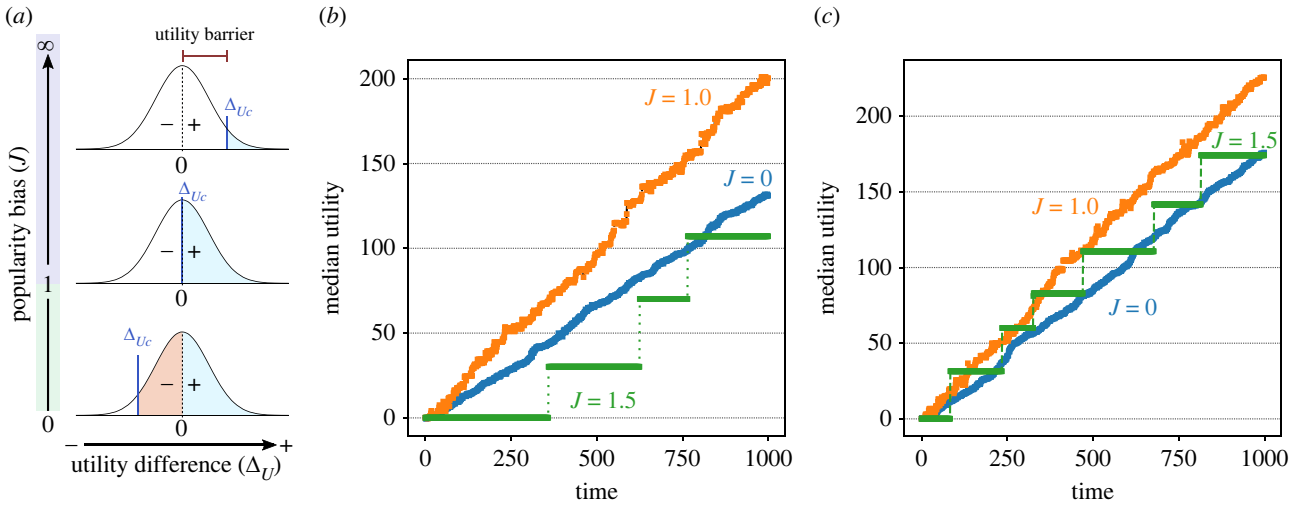

Figure 5. Statistical distribution of innovations and the pace of CCE. (*a*) Gaussian distributions of utility difference $\mathcal{N}(\Delta_U)$ under $J = 0$, $J = 1$ and $J > 1$. Each blue line indicates the utility barrier, Δ_{Uc}. Shading indicates whether random changes are rejected (white), accepted despite negative Δ_U (red) or accepted with positive Δ_U (blue). Panels at right show change in median utility during simulations ($\psi = 0$, $N = 1000$, $\beta = 0.05$) for the three different values of $J \in \{0, 1, 1.5\}$ with (*b*) $\mu = 0.05$ and (*c*) $\mu = 0.1$.

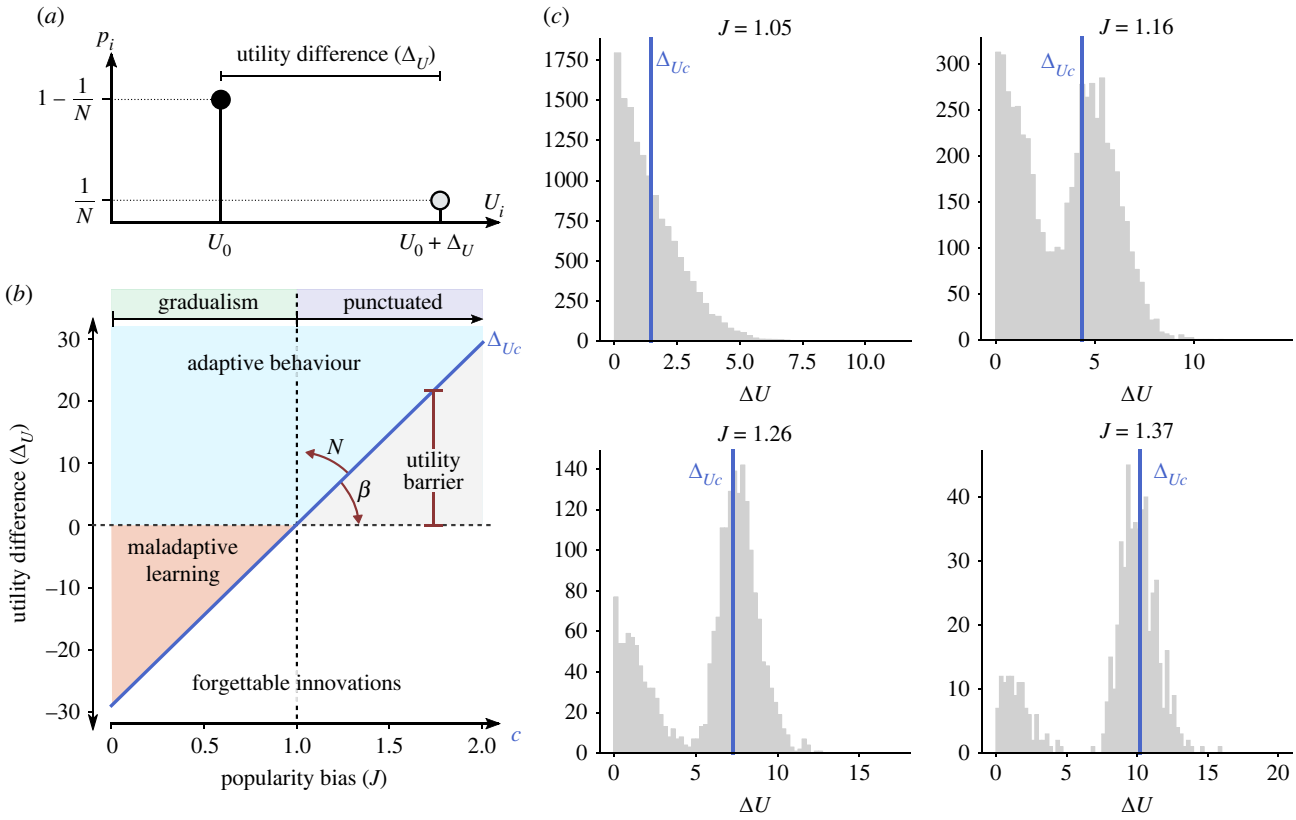

Figure 6. Comparison of theoretical and empirical utility barriers. Our model predicts the emergence of utility thresholds that must be crossed in order for novel variants to be transmitted. (*a*) Utility thresholds (or barriers) can be computed in a basic set-up where the population is centred at utility U_0 (black circle) and there is a single individual at a better variant with utility $U_0 + \Delta_U$ (grey circle). (*b*) Predicted utility difference (Δ_U) versus conformity (J). The blue diagonal described by equation (3.1) indicates where adaptations can be maintained (above line) or forgotten (below line). The slope of the diagonal depends on the population size (N) and information transparency (β). Negative utility barriers are linked with low popularity bias $J < 1$, permitting the fixation of maladaptations; $J = 1$ establishes the boundary between gradual and punctuated change; and positive utility barriers emerge when $J > 1$, limiting the fixation of small adaptations (grey area). (*c*) For various conformity levels, the frequency distribution of the empirical utility barrier is shown (grey bars). Numerical simulations support the theoretical prediction for Δ_{U_c} (solid blue line) (see text). Simulation parameters: $\beta = 0.25$, $N = 1000$, $T_{\max} = 10000$, 10 replicas, $\psi = 0$, $\mu = 0.005$.

the ratio of $J - 1$ to β. This helps contextualize the discussion of CCE and whether improvements can be discovered in large populations [19,31,69,90,91].

Among our goals has been to establish a foundation for understanding the future of cultural evolution using fundamental principles that applied in the deep past. In terms of the latter, we find it significant that CCE is optimized near $J = 1$, equivalent to simple frequency-proportional copying. This result was unexpected, as we did not model any cost–benefit function [92] to facilitate it.

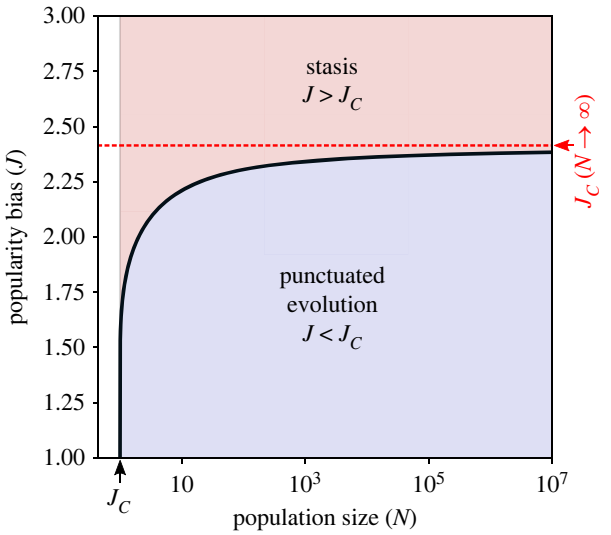

Figure 7. Prediction for the critical value of popularity bias, (J_C), separating punctuated evolution from founder effect, using $\beta = 1$ in (3.2).

As this was a robust outcome of the model, we suspect that it has wider evolutionary significance. We hypothesize that a cognitively frugal copy-popular strategy, was a natural starting point for the further evolution of the uniquely human Social Brain [93,94]. This would be similar to the copy-recent strategy that allows a majority to follow the well-informed minority [95–101], analogous to high β and $J \approx 1$. As hominin groups became larger in the last two million years [102], they would have benefited from conformist bias, $J > 1$, made possible by cognitive evolution towards numeracy and working memory [103,104]. In human groups, conformity can enhance group cooperation and learning through the collective awareness of shared attention [9,47,105,106].

Subsequent modifications of this model could incorporate a parameter for memory, as reputation—memory of an individual's interaction history [59]—affects levels of reciprocity among humans and non-human primates [48,107–109] with different functional network components of the brain activated for direct and indirect reciprocity [110–112]. Other parameters could include decay of intrinsic utility with age [113], emotions [114], boom-bust population dynamics [115], variable invention rates [17,82,116] or intrinsic preferences [81]. Parameters should be added incrementally, however, as each will multiplicatively increase the complexity of outcomes. Another question is how our results would change under skewed error distributions, such as the Gumbel distribution [22,69,90], such that utility barriers can be exceeded more often. Also, our parameters were applied to the entire simulated population; future work could model subpopulations with different β and J values.

As the model accords with other approaches to CCE featuring popularity bias as a key parameter [117,118], the additional parameter of transparency is relevant to contemporary contexts [119–121]. Popularity bias is exacerbated by algorithms that prioritize recent and popular information [119,122], such that low-cost, low-utility information is copied across massive online networks [123–125]. In the modern era, information is often not transparent, and popularity is no longer the best proxy for quality. Contemporary social media, for instance, do not necessarily surface the best ideas. Instead, online homophily, or the tendency for similar people in social groups to be connected together, has frequently polarized ideas and beliefs [126–129]. For example, clear pay-offs often lose out to the spread of misinformation [119,124,130–136]. In addition, such a large number of competing, similar options exist that informed, social learning may not be the best null hypothesis—thousands of daily social media influences amid billions of users, tens of thousands of discernible topics [125] and thousands of daily decisions [137].

For understanding online misinformation, a key factor should be how β and J act upon the sharing of information that are explicitly labelled by popularity (likes) as well as varying micro-levels of intrinsic utility, iterated over thousands/millions of online actions [61,124]. The interaction between β and J indicates a reason why culture with massive numbers of users (large J and N, low β) arguably change continually without necessarily getting 'better' [61,66,91,138–140]. Additionally, the homophily under conformity in the model resembles the well-known sorting and polarization in social media and politics [129,134,135,141].

For these reasons, the model of CCE we have presented could help explore how popularity bias and transparency of information have been pivotal from the evolution of social learning to contemporary culture. In the future, the model could be used as a foundation for connecting the evolution of the Social Brain, in small groups of transparent social learning, towards anticipating cumulative 'cultural evolution in the digital age' [119] and its limitless options of varying quality and degree of social conformity.

5. Methods

5.1. Emergence of utility barriers

In our model, reproduction rate of each variant i depends on both its utility U_i and trait frequency p_i (2.3). If there is no transparency ($\beta = 0$) and popularity bias is small ($J \approx 0$), however, then their is neither population bias nor selection based on utility, and all traits are equally probable $Pr(U_i) \approx 1/k$, where k is the number of different variants (see equation (2.4)). If $\beta = 0$ and $J \approx 0$ and also there are no new innovations introduced, $\mu \approx 0$, then eventually one variant will become fixed, i.e. the homogeneous choice, with probability $1/k$ [72].

Alternatively, assume a homogeneous population with $J > 1$, $\beta > 0$ and innovation rate $\mu > 0$, in which all but one of the N agents, fraction $1 - 1/N$, have a variant with utility U_0 or less. The single agent, fraction $1/N$, has chosen (or invented) a better variant, with utility $U_0 + \Delta_U$, where $\Delta_U > 0$ (figure 6*a*). Here, the propensity (2.5) of choosing the variant with utility U_0 is

$$\Pi(U_0, 1 - 1/N) = \left(1 - \frac{1}{N}\right)^J e^{\beta U_0 + \psi} \equiv \Pi_0,$$

while the propensity of choosing the better variant is increased by a factor of e^{Δ_U}

$$\Pi(U_0 + \Delta_U, 1/N) = \left(\frac{1}{N}\right)^J e^{\beta U_0 + \beta \Delta_U + \psi} \equiv \Pi_\Delta.$$

We hypothesize that there is a critical utility improvement, $\Delta_U > \Delta_{U_c}$, such that we expect the variant to be chosen at least once in the next interval, among the N selections of the N agents. As a mean-field approximation, this occurs when the variant can be expected to be selected by at least one agent (a conservative assumption, as this allows for it to be selected

multiple times as well)

$$N\left(\frac{\Pi_\Delta}{Y_t}\right) \geq 1,$$

where the normalization term equals the sum of the propensities, i.e. $Y_t = \Pi_0 + \Pi_\Delta$. This condition may be rewritten as

$$\Pi_\Delta(N-1) \geq \Pi_0.$$

By taking natural logarithms of both sides of the previous equation, we obtain

$$\ln(N-1) + J\ln\left(\frac{1}{N}\right) + \beta\Delta_U + \beta U_0 \geq J\ln\left(1-\frac{1}{N}\right) + \beta U_0.$$

Hence, the critical utility barrier (Δ_{U_c}) relates to the model parameters as follows:

$$\Delta_U \geq \Delta_{U_c} = \frac{J-1}{\beta}\ln(N-1).$$

This inequality predicts the critical utility barrier grows with the logarithm of population size (N) and popularity bias (J). Note we have assumed all alternative variants have utility U_0 *or less*, so this is a conservative threshold.

Figure 6*c* compares the theoretical estimate for Δ_{U_c} (blue line) with numerical simulations for various parameter values and multiple choices. The simulations show our theoretical prediction applies even to a non-homogeneous population. Multiple new variants could exceed the utility barrier; due to stochastic factors, it is impossible to anticipate which one of these inventions will finally be transmitted; all that can be predicted is that at least one of those located beyond the critical utility barrier will survive to the next generation.

5.2. The pace of cultural evolution

The pace of CCE is driven by the rate of invention μ and social adoption (the interaction between popularity bias J and transparency β). Stochastic effects are important here because populations would become stationary and incapable of adapting to environmental changes if there is no source of variation. In this context, populations can improve existing utilities (U_t) by performing many small random changes to them, i.e. $U_t + \Delta \rightarrow U_{t+1}$. What is the probability $P(U_{t+1} > U_t)$ of making this transition? How probable it is that innovations drive populations towards greater utility?

Let us assume that random changes (Δ) follow a normalized Gaussian distribution $\mathcal{N}(\Delta)$. Given the symmetry in the normal distribution, changes can be either beneficial (Δ_+) or deleterious (Δ_-) (figure 5*a*). That is, we will observe utility increases when there are beneficial but not deleterious change events

$$P(U_{t+1} > U_t) = P(\Delta_+ \cap \neg\Delta_-) = P(\Delta_+) - P(\Delta_+ \cap \Delta_-),$$

where the first term on the right side defines the probability of beneficial change

$$P(\Delta_+) = \int_0^\infty \mathcal{N}(\Delta)\,\mathrm{d}\Delta = \frac{1}{2}.$$

and the second term is the probability of maladaptative and neutral changes. As previously stated, variations of size Δ must overcome the utility barrier to be inherited in the next time step, i.e. $\Delta > |\Delta_{U_c}|$. Taking into account the utility barrier, this probability can be expressed as follows:

$$P(\Delta_+ \cap \Delta_-) = \int_0^{|\Delta_{U_c}|} \mathcal{N}(\Delta)\,\mathrm{d}\Delta.$$

When we put these definitions into the first equation, we get the following expression for the probability of increasing population utility:

$$P(U_{t+1} > U_t) = \frac{1}{2} - \int_0^{|\Delta_{U_c}|} \mathcal{N}(\Delta)\,\mathrm{d}\Delta. \tag{5.1}$$

This expression implies that the maximum rate of improvement (1/2) is restored at $J = 1$, i.e. when there is no utility barrier ($\Delta_{U_c} = 0$) and maladaptive and neutral changes are excluded.

5.3. Approaching the optimal popularity bias

The popularity bias $J > 1$, which amplifies the pace of cultural evolution as a function of degree of transparency β and population size N, is approximated in this section. According to computer simulations, removing maladaptive and neutral changes enhances the rate of improvement dramatically. Theoretical arguments suggests this takes place in the vicinity of the $J = 1$ transition (see above section and the peak in figure 3*b*). On the other hand, evolutionary stasis is prolonged when popularity bias is too strong, only amplifying the current dominant trait. What is the optimal value of $J = J_s$ that discards irrelevant changes without slowing down evolution?

We can use a geometrical argument to estimate the popularity bias's value. The utility barrier $\Delta_{Uc}(J_s)$ should be as near to the Gaussian distribution's centre as allowed, so that most neutral changes (however minor) are ignored (figure 5*a*). Formally, we need to solve the following equation:

$$\int_{\Delta_{Uc}(J_s)}^\infty \mathcal{N}(\Delta)\,\mathrm{d}\Delta = \frac{1}{2}\int_0^\infty \mathcal{N}(\Delta)\,\mathrm{d}\Delta,$$

where the left- and right-hand sides represent the probability of adaptive changes (5.1) and the probability of discarded positive changes, respectively. Using the exponential approximation for the complementary error function

$$\int_{\Delta_{Uc}(J_s)}^\infty \mathcal{N}(\Delta)\,\mathrm{d}\Delta = \frac{\mathrm{e}^{-(\Delta_{Uc}(J_s)/2)^2}}{2},$$

we can obtain the equivalent expression

$$\mathrm{e}^{-(\Delta_{Uc}(J_s)/2)^2} = \frac{1}{2}.$$

At this point, we can substitute the definition of the utility barrier $\Delta_{Uc}(J_s)$ (following equation (3.1)) to find that popularity bias J_s satisfies this condition

$$\frac{J_s - 1}{\beta}\log(N-1) = \pm 2\sqrt{\log(2)}.$$

Taking the positive boundary,

$$J_s = 1 + \frac{\beta}{\log(N-1)}2\sqrt{\log(2)}.$$

This equation suggests that the value of the popularity bias J_s optimizing the pace of cultural evolution is a balance between transparency and population size. Increasing the transparency of information $\beta > 0$ pushes the popularity bias J_s away from 1 (and increasing the likelihood of neutral changes). On the other hand, popularity bias $J_s \approx 1$ for large population sizes ($N \rightarrow \infty$).

5.4. Effect of population size

In this work, we hypothesized that utility barriers occur when there is a large popularity bias $J > 1$. The rate of cultural change is determined by the location of the utility barrier, which is influenced by both transparency and population size (see previous section). The likelihood of evolutionary slowdown increases as the utility barrier rises. It is also possible that the effect of

frequency-dependence copying is so powerful for large population sizes that it halts the evolutionary process.

To compute the size-dependent bounds of popularity bias $J_C > J$, we estimate the probability that no one in the population finds a variant crossing the utility barrier $\Delta_{Uc}(J_C)$. As before, the probability of adaptive changes is

$$\int_{\Delta_{Uc}(J_C)}^{\infty} \mathcal{N}(\Delta)\, d\Delta = \frac{e^{-(\Delta_{Uc}(J_C)/2)^2}}{2}.$$

Using a mean-field approach, we assume that the maximum popularity bias corresponds to an average number of individuals discovering an innovation being less than one

$$\frac{N}{2}\exp\left[-\left(\frac{J_C-1}{2\beta}\log(N-1)\right)^2\right] < 1.$$

After some algebra, we arrive at the expression

$$\log\left(\frac{2}{N}\right) > \frac{(J_C-1)^2}{2\beta^2}\log(N-1).$$

At this point, we can extract a limit of popularity bias $J_C > J$ that depends on the population size N

$$J_C = 1 + \beta\sqrt{\frac{2\log(N/2)}{\log(N-1)}}.$$

For large population N with $J > 1$, we obtain the following precondition for CCE:

$$J_C(N \to \infty) = \lim_{N\to\infty} 1 + \beta\sqrt{\frac{2\log(N/2)}{\log(N-1)}} \sim 1 + \beta\sqrt{2},$$

which can be alternatively described as

$$\frac{J-1}{\beta} \leq \sqrt{2}. \tag{5.2}$$

This equation suggests that in order to observe cumulative cultural behaviour, the ratio between conformity and transparency must be bounded by the corrected variance of the innovation distribution. Above this threshold, the probability of reaching higher utility values decreases exponentially.

Data accessibility. The data are provided in electronic supplementary material [142].

Authors' contributions. B.V.: conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; S.C.: investigation, software, writing—original draft; R.A.B.: conceptualization, formal analysis, investigation, methodology, supervision, validation, writing—original draft, writing—review and editing; M.J.O.: conceptualization, investigation, methodology, supervision, writing—original draft, writing—review and editing; S.V.: conceptualization, formal analysis, funding acquisition, investigation, methodology, supervision, validation, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. B.V. is supported through the 2020–2021 Biodiversa and Water JPI joint call under the BiodivRestore ERA-NET Cofund (grant no. 101003777) project MPA4Sustainability with funding organizations: Innovation Fund Denmark (IFD), Agence Nationale de la Recherche (ANR), Fundação para a Ciência e a Tecnologia (FCT), Swedish Environmental Protection Agency (SEPA) and grant no. PCI2022-132936 funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. S.V. is supported by the Spanish Ministry of Science and Innovation through the State Research Agency (AEI), grant no. PID2020-117822GB-I00/AEI/10.13039/501100011033. We acknowledge the support of the network PIE-202120E047- Conexiones-Life.

Acknowledgements. We thank Niles Eldredge for discussions about punctuated equilibria in natural and artificial evolution. B.V. and S.V. want to thank Salva Duran-Nebreda for the useful comments on the manuscript. B.V. thanks Josep Sardanyés and Lluís Alsedà for the good time together at Centre de Recerca Matemática.

References

1. Cavalli-Sforza LL, Feldman MW. 1973 Models for cultural inheritance. I. Group mean and within group variation. *Theor. Popul. Biol.* **4**, 42–44. (doi:10.1016/0040-5809(73)90005-1)
2. Cavalli-Sforza LL, Feldman MW. 1981 *Cultural transmission and evolution: a quantitative approach*. Princeton, NJ: Princeton University Press.
3. Boyd R, Richerson PJ. 2005 *Not by genes alone: how culture transformed human evolution*. Chicago, IL: University of Chicago Press.
4. Mesoudi A. 2011 *Cultural evolution*. Chicago, IL: University of Chicago Press.
5. Creanza N, Kolodny O, Feldman MW. 2017 Cultural evolutionary theory: how culture evolves and why it matters. *Proc. Natl Acad. Sci. USA* **114**, 7782–7789. (doi:10.1073/pnas.1620732114)
6. Mesoudi A, Thornton A. 2018 What is cumulative cultural evolution? *Proc. R. Soc. B* **285**, 20180712. (doi:10.1098/rspb.2018.0712)
7. 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)
8. Tomasello M, Kruger AC, Ratner HH. 1993 Cultural learning. *Behav. Brain Sci.* **16**, 495–511. (doi:10.1017/S0140525X0003123X)
9. Tomasello M, Carpenter M, Call J, Behne T, Moll H. 2005 Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–727. (doi:10.1017/S0140525X05000129)
10. 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.
11. Askarisichani O, Lane JN, Bullo F, Friedkin NE, Singh AK, Uzzi B. 2019 Structural balance emerges and explains performance in risky decision-making. *Nat. Commun.* **10**, 2648. (doi:10.1038/s41467-019-10548-8)
12. Henrich J, Broesch J. 2011 On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Phil. Trans. R. Soc. B* **366**, 1139–1148. (doi:10.1098/rstb.2010.0323)
13. Atkisson C, O'Brien MJ, Mesoudi A. 2012 Adult learners in a novel environment use prestige-biased social learning. *Evol. Psychol.* **10**, 519–537. (doi:10.1177/147470491201000309)
14. Boyd R, Richerson PJ, Henrich J. 2011 The cultural niche: why social learning is essential for human adaptation. *Proc. Natl Acad. Sci. USA* **108**(suppl. 2), 10 918–10 925. (doi:10.1073/pnas.1100290108)
15. Kraft TS, Venkataraman VV, Wallace IJ, Crittenden AN, Holowka NB, Stieglitz J, Harris J, Raichlen DA, Wood B, Gurven M, Pontzer H. 2021 The energetics of uniquely human subsistence strategies. *Science* **374**, 0130. (doi:10.1126/science.abf0130)
16. Valverde S. 2016 Major transitions in information technology. *Phil. Trans. R. Soc. B* **371**, 20150450. (doi:10.1098/rstb.2015.0450)
17. Kolodny O, Creanza N, Feldman MW. 2015 Evolution in leaps: the punctuated accumulation and loss of cultural innovations. *Proc. Natl Acad. Sci. USA* **112**, E6762–E6769. (doi:10.1073/pnas.1520492112)
18. Pan W, Ghoshal G, Krumme C, Cebrian M, Pentland A. 2013 Urban characteristics attributable to density-driven tie formation. *Nat. Commun.* **4**, 1961. (doi:10.1038/ncomms2961)

19. Kline MA, Boyd R. 2010 Population size predicts technological complexity in Oceania. *Proc. R. Soc. B* **277**, 2559–2564. (doi:10.1098/rspb.2010.0452)
20. Collard M, Buchanan B, O'Brien MJ. 2013 Population size as an explanation for patterns in the Paleolithic archaeological record: more caution is needed. *Curr. Anthropol.* **54**, S388–S396. (doi:10.1086/673881)
21. Vaesen K, Collard M, Cosgrove R, Roebroeks W. 2016 Population size does not explain past changes in cultural complexity. *Proc. Natl Acad. Sci. USA* **113**, E2241–E2247. (doi:10.1073/pnas.1520288113)
22. Henrich J. 2004 Demography and cultural evolution: why adaptive cultural processes produced maladaptive losses in Tasmania. *Am. Antiq.* **69**, 197–214. (doi:10.2307/4128416)
23. Centola D, Baronchelli A. 2015 The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proc. Natl Acad. Sci. USA* **112**, 1989–1994. (doi:10.1073/pnas.1418838112)
24. 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)
25. Mesoudi A, Lei C, Keelin M, Jing LH. 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)
26. Mesoudi A. 2008 An experimental simulation of the 'copy-successful-individuals' cultural learning strategy. *Evol. Hum. Behav.* **29**, 350–363. (doi:10.1016/j.evolhumbehav.2008.04.005)
27. 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)
28. Woolley AW, Aggarwal I, Malone TW. 2015 Collective intelligence and group performance. *Curr. Dir. Psychol.* **24**, 420–424. (doi:10.1177/0963721415599543)
29. Clément RJG *et al.* 2013 Collective cognition in humans: groups outperform their best members in a sentence reconstruction task. *PLoS ONE* **8**, e77943. (doi:101371/journal.pone.0077943)
30. Derex M, Beugin MP, Godelle B, Raymond M. 2013 Experimental evidence for the influence of group size on cultural complexity. *Nature* **503**, 389. (doi:10.1038/nature12774)
31. Derex M, Mesoudi A. 2020 Cumulative cultural evolution within evolving population structures. *Trends Cog. Sci.* **24**, 654–667. (doi:10.1016/j.tics.2020.04.005)
32. Dean LG, Vale GL, Laland KN, Flynn E, Kendal RL. 2014 Human cumulative culture: a comparative perspective. *Biol. Rev. Camb. Phil. Soc.* **89**, 284–301. (doi:10.1111/brv.12053)
33. Hidalgo CA. 2021 Economic complexity theory and applications. *Nat. Rev. Phys.* **3**, 92–113. (doi:10.1038/s42254-020-00275-1)
34. Valverde S, Solé RV. 2015 Punctuated equilibrium in the large-scale evolution of programming languages. *J. R. Soc. Interface* **12**, 20150249. (doi:10.1098/rsif.2015.0249)
35. Brahm F, Poblete J. 2021 The evolution of productive organizations. *Nat. Hum. Behav.* **5**, 39–48. (doi:10.1038/s41562-020-00957-x)
36. Ruck DJ, Bentley RA, Lawson DJ. 2020 Cultural prerequisites of socioeconomic development. *R. Soc. Open Sci.* **7**, 7190725. (doi:10.1098/rsos.190725)
37. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, Pinior B, Thurner S, Klimek P. 2020 Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat. Hum. Behav.* **4**, 1303–1312. (doi:10.1038/s41562-020-01009-0)
38. Yang Y, Wu Y, Uzzi B. 2020 Estimating the deep replicability of scientific findings using human and artificial intelligence. *Proc. Natl Acad. Sci. USA* **117**, 10 762–10 768. (doi:10.1073/pnas.1909046117)
39. Weis JW, Jacobson JM. 2021 Learning on knowledge graph dynamics provides an early warning of impactful research. *Nat. Biotech.* **39**, 1300–1307. (doi:10.1038/s41587-021-00907-6)
40. Zeng A, Fan Y, Di Z, Wang Y, Havlin S. 2021 Fresh teams are associated with original and multidisciplinary research. *Nat. Hum. Behav.* **5**, 1314–1322. (doi:10.1038/s41562-021-01084-x)
41. Bettencourt LW, Lobo J, Helbing D, Kühnert C, West GB. 2007 Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306. (doi:10.1073/pnas.0610172104)
42. Rendell L *et al.* 2010 Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213. (doi:10.1126/science.1184719)
43. Laland KN. 2004 Social learning strategies. *Learn. Behav.* **32**, 4–14. (doi:10.3758/BF03196002)
44. Henrich J, Boyd R. 1998 The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* **19**, 215–241. (doi:10.1016/S1090-5138(98)00018-X)
45. Dunbar RIM, Shultz S. 2007 Evolution in the social brain. *Science* **317**, 1344–1347. (doi:10.1126/science.1145463)
46. Dunbar RIM, Shultz S. 2017 Why are there so many explanations for primate brain evolution? *Phil. Trans. R. Soc. B* **372**, 20160244. (doi:10.1098/rstb.2016.0244)
47. Shteynberg G, Hirsh JB, Bentley RA, Garthoff J. 2020 Shared worlds and shared minds: a theory of collective learning and a psychology of common knowledge. *Psychol. Rev.* **127**, 918–931. (doi:10.1037/rev0000200)
48. Hill KR *et al.* 2011 Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science* **331**, 1286–1289. (doi:10.1126/science.1199071)
49. Duran-Nebreda S, O'Brien MJ, Bentley RA, Valverde S. 2022 Dilution of expertise in the rise and fall of collective innovation. *Humanit. Soc. Sci. Commun.* **9**, 365. (doi:10.1057/s41599-022-01380-5)
50. Aplin LM, Sheldon BC, McElreath R. 2017 Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. *Proc. Natl Acad. Sci. USA* **114**, 7830–7837. (doi:10.1073/pnas.1621067114)
51. Ehrenberg ASC. 1959 The pattern of consumer purchases. *J. R. Stat. Soc. C* **8**, 26–41. (doi:10.2307/2985810)
52. Farmer JD, Patelli P, Zovko II. 2005 The predictive power of zero intelligence in financial markets. *Proc. Natl Acad. Sci. USA* **102**, 2254–2259. (doi:10.1073/pnas.0409157102)
53. Nia HT, Jain AD, Liu Y, Alam M-R, Barnas R, Makris NC. 2015 The evolution of air resonance power efficiency in the violin and its ancestors. *Proc. R. Soc A* **471**, 20140905. (doi:10.1098/rspa.2014.0905)
54. O'Brien MJ, Lyman RL. 2000 *Applying evolutionary archaeology: a systematic approach*. New York, NY: Springer Science & Business Media.
55. Miu E, Gulley N, Laland KN, Rendell L. 2018 Innovation and cumulative culture through tweaks and leaps in online programming contests. *Nat. Commun.* **9**, 2321. (doi:10.1038/s41467-018-04494-0)
56. Bentley RA, O'Brien MJ, Brock WA. 2014 Mapping collective behavior in the big-data era. *Behav. Brain Sci.* **37**, 63–119. (doi:10.1017/S0140525X13000289)
57. Brock WA, Durlauf SN. 2001 Discrete choice with social interactions. *Rev. Econ. Stud.* **68**, 229–272. (doi:10.1111/1467-937X.00168)
58. Motes-Rodrigo A, Mundry R, Call J, Tennie C. 2021 Evaluating the influence of action- and subject-specific factors on chimpanzee action copying. *R. Soc. Open Sci.* **8**, 200228. (doi:10.1098/rsos.200228)
59. 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)
60. Acerbi A, Bentley RA. 2014 Biases in cultural transmission shape the turnover of popular traits. *Evol. Hum. Behav.* **35**, 228–236. (doi:10.1016/j.evolhumbehav.2014.02.003)
61. Carrignon S, Bentley RA, Ruck DJ. 2019 Modelling rapid online cultural transmission. *Palgrave Commun.* **5**, article 83. (doi:10.1057/s41599-019-0295-9)
62. Leroi AM, Lambert B, Rosindell J, Zhang X. 2020 Neutral syndrome. *Nat. Hum. Behav.* **4**, 780–790. (doi:10.1038/s41562-020-0844-7)
63. Mesoudi A, Lycett SJ. 2009 Random copying, frequency-dependent copying and culture change. *Evol. Hum. Behav.* **30**, 41–48. (doi:10.1016/j.evolhumbehav.2008.07.005)
64. Reali F, Griffiths TL. 2010 Words as alleles: connecting language evolution with Bayesian learners to models of genetic drift. *Proc. R. Soc. B* **277**, 429–436. (doi:10.1098/rspb.2009.1513)
65. Ruck DJ, Bentley RA, Acerbi A, Garnett P, Hruschka DJ. 2017 Role of neutral evolution in word turnover during centuries of English word popularity. *Adv.*

Complex Syst. **20**, 1750012. (doi:10.1142/S0219525917500126)

66. Bentley RA, Carrignon S, Ruck DJ, Valverde S, O'Brien MJ. 2021 Neutral models are a tool, not a syndrome. *Nat. Hum. Behav.* **5**, 807–808. (doi:10.1038/s41562-021-01149-x)
67. Brock WA, Bentley RA, O'Brien MJ, Caiado CCS. 2014 Estimating a path through a map of decision making. *PLoS ONE* **9**, e111022. (doi:10.1371/journal.pone.0111022)
68. Caiado CCS, Brock WA, Bentley RA, O'Brien MJ. 2016 Fitness landscapes among many options under social influence. *J. Theor. Biol.* **405**, 5–16. (doi:10.1016/j.jtbi.2015.12.013)
69. Vaesen K. 2012 Cumulative cultural evolution and demography. *PLoS ONE* **7**, e40989. (doi:10.1371/journal.pone.0040989)
70. Bentley RA, O'Brien MJ, Earls M. 2011 *I'll have what she's having: mapping social behavior*. Cambridge, MA: MIT Press.
71. O'Brien MJ, Bentley RA, Brock WA. 2019 *The importance of small decisions*. Cambridge, MA: MIT Press.
72. Nowak MA. 2006 *Evolutionary dynamics: exploring the equations of life*. Cambridge, MA: Harvard University Press.
73. Morgan TJH, Thompson B. 2020 Biased transformation erases traditions sustained by conformist transmission. *Biol. Lett.* **16**, 20200660. (doi:10.1098/rsbl.2020.0660)
74. Skaperdas S. 1996 Contest success functions. *Econ. Theor.* **7**, 283–290. (doi:10.1007/BF01213906)
75. Lee D. 2006 Best to go with what you know? *Nature* **441**, 822–823. (doi:10.1038/441822a)
76. Daw N, O'Doherty J, Dayan P. 2006 Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879. (doi:10.1038/nature04766)
77. Anderson S, de Palma A, Thisse JF. 1992 *Discrete choice theory of product differentiation*. Cambridge, MA: MIT Press.
78. Manski C. 1977 The structure of random utility models. *Theor. Decis.* **8**, 229–254. (doi:10.1007/BF00133443)
79. McKelvey RD, Palfrey TR. 1995 Quantal response equilibria for normal form games. *Games Econ. Behav.* **10**, 6–38. (doi:10.1006/game.1995.1023)
80. Arcidiacono P, Ellickson PB. 2011 Practical methods for estimation of dynamic discrete choice models. *Annu. Rev. Econ.* **3**, 363–394. (doi:10.1146/annurev-economics-111809-125038)
81. Acerbi A, Ghirlanda S, Enquist M. 2012 The logic of fashion cycles. *PLoS ONE* **7**, e32541. (doi:10.1371/journal.pone.0032541)
82. Eerkens JW, Lipo CP. 2005 Cultural transmission, copying errors, and the generation of variation in material culture in the archaeological record. *J. Anthropol. Archaeol.* **24**, 316–334. (doi:10.1016/j.jaa.2005.08.001)
83. Diederen P, van Meijl H, Wolters A. 2003 Modernisation in agriculture: what makes a farmer adopt an innovation? *Int. J. Agric. Resour. Gov. Ecol.* **2**, 328–342.
84. Srinivasan V, Mason CH. 1986 Nonlinear least squares estimation of new product diffusion models. *Mark Sci.* **5**, 169–178. (doi:10.1287/mksc.5.2.169)
85. Rogers EM. 1995 *Diffusion of innovations*. New York, NY: Free Press.
86. Scheffer M *et al.* 2012 Anticipating critical transitions. *Science* **338**, 344–348. (doi:10.1126/science.1225244)
87. van Nimwegen E, Crutchfield JP. 2000 Metastable evolutionary dynamics: crossing fitness barriers or escaping via neutral paths?. *Bull. Math. Biol.* **62**, 799–848. (doi:10.1006/bulm.2000.0180)
88. Gintis H. 2007 A framework for the unification of the behavioral sciences. *Behav. Brain Sci.* **30**, 1–61. (doi:10.1017/S0140525X07000581)
89. Mesoudi A. 2021 Blind and incremental or directed and disruptive? On the nature of novel variation in human cultural evolution. *Am. Phil. Q.* **58**, 7–20. (doi:10.2307/48600682)
90. Bentley RA, O'Brien MJ. 2011 The selectivity of social learning and the tempo of cultural evolution. *J. Evol. Psych.* **9**, 125–141. (doi:10.1556/JEP.9.2011.18.1)
91. Vaesen K, Houkes W. 2021 Is human culture cumulative? *Curr. Anthropol.* **62**, 218–238. (doi:10.1086/714032)
92. Mesoudi A. 2011 Variable acquisition costs constrain cumulative cultural evolution. *PLoS ONE* **6**, e18239. (doi:10.1371/journal.pone.0018239)
93. Henrich J, Gil-White FJ. 2001 The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* **22**, 165–196. (doi:10.1016/S1090-5138(00)00071-4)
94. Mesoudi A, Whiten A, Dunbar R. 2006 A bias for social information in human cultural transmission. *Brit. J. Psychol.* **97**, 405–423. (doi:10.1348/000712605X85871)
95. Couzin ID, Krause J, Franks NR, Levin SA. 2005 Effective leadership and decision-making in animal groups on the move. *Nature* **433**, 513–516. (doi:10.1038/nature03236)
96. Dyer JRG, Johansson A, Helbing D, Couzin ID, Krause J. 2009 Leadership, consensus decision making and collective behaviour in human crowds. *Phil. Trans. R. Soc. B* **364**, 781–789. (doi:10.1098/rstb.2008.0233)
97. Herbert-Read JE, Krause S, Morrell LJ, Schaerf TM, Krause J, Ward AJW. 2013 The role of individuality in collective group movement. *Proc. R. Soc. B* **280**, 20122564. (doi:10.1098/rspb.2012.2564)
98. Wolf MR, Kurvers RHJM, Ward AJW, Krause S, Krause J. 2013 Accurate decisions in an uncertain world: collective cognition increases true positives while decreasing false positives. *Proc. R. Soc. B* **280**, 20122777. (doi:10.1098/rspb.2012.2777)
99. Kurvers RHJM, Wolf M, Krause J. 2014 Humans use social information to adjust their quorum thresholds adaptively in a simulated predator detection experiment. *Behav. Ecol. Sociobiol.* **68**, 449–456. (doi:10.1007/s00265-013-1659-6)
100. Couzin ID *et al.* 2011 Uninformed individuals promote democratic consensus in animal groups. *Science* **334**, 1578–1580. (doi:10.1126/science.1210280)
101. Strandburg-Peshkin A, Farine DR, Couzin ID, Crofoot MC. 2015 Shared decision-making drives collective movement in wild baboons. *Science* **348**, 1358–1361. (doi:10.1126/science.aaa5099)
102. Ponce de León MS *et al.* 2021 The primitive brain of early *Homo*. *Science* **372**, 165–171.
103. Wynn T, Coolidge FL. 2010 Beyond symbolism and language: an introduction to Supplement 1, Working Memory. *Curr. Anthropol.* **51**, S5–S16. (doi:10.1086/650526)
104. d'Errico F, Doyon L, Colagé I, Queffelec A, Le Vraux E, Giacobini G, Vandermeersch B, Maureille B. 2018 From number sense to number symbols: an archaeological perspective. *Phil. Trans. R. Soc. B* **373**, 20160518. (doi:10.1098/rstb.2016.0518)
105. Meltzoff AN, Kuhl PK, Movellan J, Sejnowski TJ. 2009 Foundations for a new science of learning. *Science* **325**, 284–288. (doi:10.1126/science.1175626)
106. Tomasello M, Farrar MJ. 1986 Joint attention and early language. *Child Dev.* **57**, 1454–1463. (doi:10.2307/1130423)
107. Apicella CL, Marlowe FW, Fowler JH, Christakis NA. 2012 Social networks and cooperation in hunter-gatherers. *Nature* **481**, 497–501. (doi:10.1038/nature10736)
108. Sliwa J. 2021 Towards a collective animal neuroscience. *Science* **374**, 397–398. (doi:10.1126/science.abm3060)
109. Báez-Mendoza R, Mastrobattista EP, Wang AJ, Williams ZM. 2021 Social agent identity cells in the prefrontal cortex of interacting groups of primates. *Science* **374**, eabb4149.
110. Falk EB, Bassett DS. 2017 Brain and social networks: fundamental building blocks of human experience. *Trends Cog. Sci.* **21**, 674–690. (doi:10.1016/j.tics.2017.06.009)
111. Hein G, Morishima Y, Leiberg S, Sul S, Fehr E. 2016 The brain's functional network architecture reveals human motives. *Science* **351**, 1074–1078. (doi:10.1126/science.aac7992)
112. Parkinson C, Kleinbaum AM, Wheatley T. 2018 Similar neural responses predict friendship. *Nat. Commun.* **9**, article 332. (doi:10.1038/s41467-017-02722-7)
113. Lewandowsky S, Griffiths TL, Kalish ML. 2009 The wisdom of individuals: exploring people's knowledge about everyday events using iterated learning. *Cogn. Sci.* **33**, 969–998. (doi:10.1111/j.1551-6709.2009.01045.x)
114. Rocklage MD, Rucker DD, Nordgren LF. 2021 Mass-scale emotionality reveals human behaviour and marketplace success. *Nat. Hum. Behav.* **5**, 1323–1329. (doi:10.1038/s41562-021-01098-5)
115. Shennan S, Downey SS, Timpson A, Edinborough K, Colledge S, Kerig T, Manning K, Thomas MG. 2013 Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat. Commun.* **4**, 2486. (doi:10.1038/ncomms3486)

116. Findling C, Chopin N, Koechlin E. 2021 Imprecise neural computations as a source of adaptive behaviour in volatile environments. *Nat. Hum. Behav.* **5**, 99–112. (doi:10.1038/s41562-020-00971-z)
117. Van Cleve J. 2016 Cooperation, conformity, and the coevolutionary problem of trait associations. *J. Theor. Biol.* **396**, 13–24. (doi:10.1016/j.jtbi.2016.02.012)
118. Denton KK, Ram Y, Liberman U, Feldman MW. 2020 Cultural evolution of conformity and anticonformity. *Proc. Natl Acad. Sci. USA* **117**, 13 603–13 614. (doi:10.1073/pnas.2004102117)
119. Acerbi A. 2019 Cognitive attraction and online misinformation. *Palgrave Commun.* **5**, article 15. (doi:10.1057/s41599-019-0224-y)
120. Bentley RA, O'Brien MJ. 2017 *The acceleration of cultural change: from ancestors to algorithms*. New York, NY: MIT Press.
121. Bak-Coleman JB *et al.* 2021 Stewardship of global collective behavior. *Proc. Natl Acad. Sci. USA* **118**, e2025764118. (doi:10.1073/pnas.2025764118)
122. Turkle S. 2015 *Reclaiming conversation: the power of talk in a digital age*. Harmondsworth, UK: Penguin Press.
123. Barfuss W, Donges JF, Vasconcelos VV, Kurths J, Levin SA. 2020 Caring for the future can turn tragedy into comedy for long-term collective action under risk of collapse. *Proc. Natl Acad. Sci. USA* **117**, 12 915–12 922. (doi:10.1073/pnas.1916545117)
124. Vosoughi S, Roy D, Aral S. 2018 The spread of true and false news online. *Science* **359**, 1146–1151. (doi:10.1126/science.aap9559)
125. Obradovich N *et al.* 2020 Expanding the measurement of culture with a sample of two billion humans. (doi:10.2139/ssrn.3689334).
126. Aral S, Muchnik L, Sundararajan A. 2009 Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl Acad. Sci. USA* **106**, 21 544–21 549. (doi:10.1073/pnas.0908800106)
127. Druckman JN, Klar S, Krupnikov Y, Levendusky M, Ryan JB. 2021 Affective polarization, local contexts and public opinion in America. *Nat. Hum. Behav.* **5**, 28–38. (doi:10.1038/s41562-020-01012-5)
128. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. 2019 Fake news on Twitter during the 2016 U.S. presidential election. *Science* **363**, 374–378. (doi:10.1126/science.aau2706)
129. Johnson NF, Leahy R, Restrepo NJ, Velásquez N, Zheng M, Manrique P, Devkota P, Wuchty S. 2019 Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature* **573**, 261–265. (doi:10.1038/s41586-019-1494-7)
130. Johnson NF *et al.* 2020 The online competition between pro- and anti-vaccination views. *Nature* **582**, 230–233. (doi:10.1038/s41586-020-2281-1)
131. Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, Fowler JH. 2012 A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298. (doi:10.1038/nature11421)
132. Garcia-Herranz M, Moro E, Cebrian M, Christakis NA, Fowler JH. 2014 Using friends as sensors to detect global-scale contagious outbreaks. *PLoS ONE* **9**, e92413. (doi:10.1371/journal.pone.0092413)
133. Gallotti G, Valle F, Castaldo N, Sacco P, De Domenico M. 2020 Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nat. Hum. Behav.* **4**, 1285–1293. (doi:10.1038/s41562-020-00994-6)
134. Finkel EJ *et al.* 2020 Political sectarianism in America. *Science* **370**, 533–536. (doi:10.1126/science.abe1715)
135. Waller I, Anderson A. 2021 Quantifying social organization and political polarization in online platforms. *Nature* **600**, 264–268. (doi:10.1038/s41586-021-04167-x)
136. Lazer D, Kennedy R, King G, Vespignani A. 2014 The parable of Google flu: traps in big data analysis. *Science* **343**, 1203–1205. (doi:10.1126/science.1248506)
137. Wansink B, Sobal J. 2007 Mindless eating: the 200 daily food decisions we overlook. *Environ. Behav.* **39**, 106–123. (doi:10.1177/0013916506295573)
138. Gleeson JP, Cellai D, Onnela J-P, Porter MA, Reed-Tsochas F. 2014 A simple generative model of collective online behaviour. *Proc. Natl Acad. Sci. USA* **111**, 10 411–10 415. (doi:10.1073/pnas.1313895111)
139. Youngblood M, Stubbersfield JM, Morin O, Glassman R, Acerbi A. 2021 Cultural transmission bias in the spread of voter fraud conspiracy theories on Twitter during the 2020 US election. PsyArXiv (doi:10.31234/osf.io/2jksg).
140. Mauch M, MacCallum RM, Levy M, Leroi AM. 2015 The evolution of popular music: USA 1960–2010. *R. Soc. Open Sci.* **2**, 150081. (doi:10.1098/rsos.150081)
141. Andris C, Lee D, Hamilton MJ, Martino M, Gunning CE, Selden JA. 2015 The rise of partisanship and super-cooperators in the U.S. House of Representatives. *PLoS ONE* **10**, e0123507. (doi:10.1371/journal.pone.0123507)
142. Vidiella B, Carrignon S, Bentley RA, O'Brien MJ, Valverde S. 2022 A cultural evolutionary theory that explains both gradual and punctuated change. Figshare. (doi:10.6084/m9.figshare.c.6277167)