

# Social transmission is conformist, variation is guided, variance is proportional

Pavína Hillerová<sup>1</sup>, Petr Tureček<sup>1,2</sup>

<sup>1</sup> Department of Philosophy and History of Science, Faculty of Science, Charles University, Viničná 7, Prague 2, 128 00, Czech Republic

<sup>2</sup> Center for Theoretical Study, Charles University and Czech Academy of Sciences, Jilská 1, Prague 1, 110 00, Czech Republic

**Keywords:** cultural evolution, cultural transmission, Galton-Pearson model, probability density function, Bayesian analysis,

## Abstract:

While cultural evolution has matured into a fully-fledged field, the science of cultural transmission remains underdeveloped. Most existing models still rely on naive assumptions of discrete trait imitation, while more complex transmission models for continuous traits are virtually non-existent. This paper fills the void by introducing a continuous transmission model that integrates biased transmission, conformity, and three distinct variance components into a single probability density function; one that still draws inspiration from the original biometric models that shaped early thinking about biological inheritance at the end of the 19th century.

Participants played a trivia-style guessing game in an app, being exposed to the answers of previous participants (“cultural parents”) and providing numeric estimates (“cultural offspring”) that serve as a continuous heritable trait for analysis.

This approach yields clear evidence for three key features long hypothesized to distinguish cultural from biological inheritance: people conform to majority input, aligning with mutually similar social cues; they systematically adjust their estimates toward the correct answer (guided variation); and their estimate variance is not constant, but scales with the variance of presented cues. All three effects can be reliably estimated within a single generative model.

## Significance statement:

Understanding how people learn from others is key to explaining cultural change. While most models assume discrete trait copying, real-world cultural transmission often involves continuous traits underlying preferences, properties, and propensities. We introduce a trivia guessing game and analyze participant responses using a generative function that captures how people integrate social information, conform to similar inputs, and adjust their answers toward better ones, while scaling uncertainty with the diversity of observed cues. These features, absent in biological inheritance, allow cultural systems to flexibly balance stability and exploration. Our results show that a single probability distribution can formalize these effects, offering a powerful tool for studying cultural transmission.

## Introduction:

The ability to learn from others is a key driver of adaptive behaviour across species (Heyes, 2012; Hoppitt & Laland, 2013; Whiten et al., 2017). It has been documented across a wide range of taxa; from birds using social cues to locate food (Galef & Giraldeau, 2001) to primates and cetaceans exhibiting cultural transmission of foraging techniques (Krützen et al., 2005; Whiten et al., 2005). Even fruit flies exhibit capacity for imitating preferences of conspecifics (Danchin et al., 2018). One species, however, stands out: human cultural transmission appears to be especially faithful and forms the foundation of our evolutionary success (Boyd & Richerson, 1996; Laland, 2017; Tennie et al., 2009; Tomasello, 1999). Socially acquired information enables us to navigate new, complex and unstable environments efficiently (Tomasello, 1999).

Cultural inheritance shares several features with genetics: we can describe distinct variants and track changes in their frequency due to selection or biased transmission. Both systems have been proposed as robust substrates for cumulative and open-ended evolution (Morgan & Feldman, 2024). Early researchers of cultural evolution leveraged this similarity by adapting models from population genetics. However, despite obvious parallels, cultural evolution exhibits several unique features that depart from biological evolution, features that traditional genetic models are not designed to capture.

At first glance, cultural evolutionists may appear to be just rogue biologists that try to explain cultural change through gene-like entities and selection pressures (Blackmore, 2000; Dawkins, 1989). However, more nuanced approaches exist. Dual inheritance theory views culture as an integral part of human biology; not as a detachable add-on to our genes, but as a co-evolved feature that has influenced survival, reproduction, and fitness across generations (Boyd & Richerson, 1985; Henrich, 2016; Laland et al., 2001). The concept of culture in question extends beyond the narrow sense of high art or symbolic expression. It includes domestication, tool-making, shelter-building, and clothing – practices with direct consequences for survival and reproduction (Richerson & Boyd, 2005; Sterelny, 2012). This broader perspective does not mean dismissing *Hamlet*, opera, or viral dances. On the contrary, understanding why such phenomena spread and persist can offer valuable insights into cultural transmission (Acerbi & Mesoudi, 2015; Mesoudi, 2011). Yet, all these examples suggest a discrete framing of cultural manifestations, reflecting a common bias in how culture is described and modeled in contemporary literature (Yeh et al., 2019). Continuous cultural traits – such as artifact dimensions (lengths, proportions, or continuous embeddings of co-occurring discrete traits) – and their evolution deserve no less attention.

Early works on cultural evolution proposed both discrete and continuous models of transmission. Cavalli-Sforza and Feldman (1981) or Boyd and Richerson (1988) gave considerable space to both approaches. Over time, however, the field has come to rely heavily on discrete models (Henrich & Boyd, 2002; Mesoudi et al., 2006). This makes sense: discrete models are easier to formalize, simulate, and analyze – especially given the computational tools that were at hand in the 1980s. They also align with the forms of inheritance and transmission we understand best, such as genes, alleles, and the spread of viruses. Yet this convenience has come at a cost: we may be overfitting our theories to the kinds of traits that are easiest to model, rather than to those that actually matter (Acerbi & Mesoudi, 2015; Claidière et al., 2014).

Occasionally, efforts are made to “continuize” models of cultural evolution. Morin’s (2013) analysis of the evolution of Renaissance portraiture being a notable example. Although the central hypothesis in his study concerns a binary outcome (whether the individual depicted in a painting averts their gaze from the viewer), the underlying variable is clearly continuous.

The model ultimately produces a probability of gaze aversion in portraits across different artists and stages of their careers. In short, while data are categorical, the explanatory mechanism is inherently continuous.

Similar tendencies are clearly present in the work of Dan Sperber, a major theorist of cultural evolution and Morin’s teacher. Sperber often invokes the idea of a “space of possibilities”, essentially a design space (much like, Liu et al., 2021; O’Brien et al., 2016, in which cultural variants occupy positions within a multidimensional coordinate system. Many works inspired by Sperber’s approach go on to develop models that incorporate continuous dimensions (e.g., [insert relevant references here]). Such concepts encourage us to think about cultural change in terms of gradual variation, not just discrete categories.

A notable exception that tends to employ continuous models is research quantifying the cumulateness of cultural evolution through transmission-chain (Caldwell & Millen, 2008, 2009; Kirby et al., 2008). These studies often track how the success of cultural products – typically artifacts such as paper planes or spaghetti towers – changes over time. However, they frequently focus on utility, as if “performance” were the only relevant trait, rather than examining the traits of the artifacts themselves. Moreover, they typically do not account for who learned from whom. They do not study the cultural equivalent of inheritance, just evolution, the change in performance across generations. Without parent–offspring links, it is difficult to study inheritance properly; even Mendel needed to know which plants were crossbred to produce offspring to uncover the laws of genetics. Models that could take advantage of such links already exist and deserve further development.

What we find in early works of Cavalli-Sforza and Feldman is some version of what they refer to as “Galton-Pearson” model (see, for instance the chapter *Cultural Transmission for a Continuous Trait* in Cavalli-Sforza and Feldman, 1981), something like

$$t_o \sim N(\text{mean}(t_p), \eta), \quad (1)$$

where  $t_o$  is the offspring’s trait value, drawn from normal distribution ( $N$ ) centred on the mean of the parental trait vector  $t_p$  with a constant standard deviation  $\eta$  (implying a trait variance of  $\eta^2$ ).

This is nothing else than an approximation of polygenic continuous trait transmission. Ronald Fisher famously bridged the gap between this biometric model and Mendelian genetics by decomposing offspring trait variance into three components: (i) mutation at the level of each allele, (ii) recombination and segregation of alleles at the level of parental gametes, and (iii) random modification at the level of the phenotype. The use of a constant variance is unsurprising, as stochasticity at each of these levels is isolated from – and insensitive to – other sources of variability. The total variance is thus simply the sum of contributions from (potentially) polymorphic loci and phenotypic plasticity.

Cavalli-Sforza and Feldman were well aware that Fisher’s variance decomposition should not be applied directly to culture. After all, culture is not stored on paired chromosomes, and there is no cultural transmission without cultural expression (unlike in genetics, where an eroded gene for tooth enamel can be passed on for millions of bird generations without ever affecting the fitness of its carriers (Sire et al., 2008)).

Stepping back to a model once used by prominent biometricians to capture the inheritance of traits like human height or sweet-pea seed size (see Eq. 1) may be a reasonable starting point but it is certainly not the end. Unfortunately, the past forty years have seen little follow-up work in this direction, as testified by a recent review (Fogarty & Kandler, 2025).

There are at least three key aspects in which a model aiming to capture cultural transmission must depart from Galton–Pearson inheritance:

**1. Conformity.** In sexual reproduction, two parents each contribute half of their genetic material to the offspring. In cultural transmission the number of “parents” (i.e., sources of influence) can exceed two, which invites weighting each parent’s contribution by their similarity to other parents. We speak about conformist transmission when individuals disproportionately follow mutually similar inspiration sources.

**2. Guided variation.** While biological mutations are typically modeled as random with respect to trait values, it has been proposed that humans can non-randomly improve (or at least systematically bias) imitated cultural manifestations. This challenges the assumption of equality between the mean trait value of parents and that of the offspring

**3. Variance proportionality.** As noted earlier, the stochastic term in approximation of polygenic trait inheritance does not integrate information from multiple sources at any point. In contrast, cultural transmission inherently involves pooling information across multiple observations or individuals. Therefore, models of cultural inheritance must accommodate variance-proportional variance. Even if precise imitation of the mean is the goal, as in descriptive statistics, standard error ( $SE$ ) of the estimate grows with the sample standard deviation  $\sigma_x$ , following the formula  $SE = \frac{\sigma_x}{\sqrt{n}}$ , where  $n$  is the sample size. An analogous process is at play in cultural transmission: even when imitating a single model, that model’s behaviour is typically observed multiple times under noisy conditions effectively turning the process into statistical sampling.

Moreover, error (analogous to mutation) is not the only potential source of variance proportionality in cultural transmission. Humans may adjust their willingness for exploration to the observed range of successful strategies. If the herd sizes of neighbors are tightly clustered (e.g., 62, 63, 63), people will likely stick to such numbers. But if they are widely spread (e.g., 10, 62, 116), people may feel freer and safer to deviate from the average in either direction. Models with fixed variance cannot account for this kind of flexibility.

All three listed deviations from biometric inheritance have profound implications for cultural evolution. Conformity, in particular, has been proposed as a key mechanism for maintaining sufficient fidelity in cultural transmission, thereby enabling cumulative culture.

Guided variation may accelerate the tempo of cultural adaptation by enabling faster intergenerational change than would be expected based solely on population variance and the strength of selection pressures. The extent to which cultural variation is genuinely “guided” rather than merely biased remains an open question. It has been shown that technologies can evolve through the accumulation of beneficial modifications to existing solutions, even when the individuals involved in cultural exchange are entirely unaware of why one solution is superior to another .

Proportionality between input and output variance may also accelerate cultural adaptation by increasing population variance under favorable conditions, thereby allowing for larger generational leaps under directional selection. While biological adaptation underpinned by constant population variance tends to progress linearly, adaptation in a system where offspring standard deviation is proportional to parental standard deviation can proceed exponentially toward the optimum. However, the likelihood of adaptation in a system with proportional variance depends critically on population size. In biological evolution, where each parental pair produces offspring with constant variance, the speed of adaptation is expected to be similar across small and large populations (aside from a slight disadvantage in smaller populations due to stronger genetic drift). Indeed, substantial evolutionary change has

occurred on isolated islands with small populations. In contrast, technological adaptation under variance-proportional dynamics is expected to be far less likely in small populations.

Some constant variance models can achieve this through right-skewed Gumbel distribution, which makes the probability that one member of a new generation surpasses the best individual of the previous contingent on population size. Such models, however, produce a critical-threshold behaviour (e.g. all symbolic revolutions expected at identical population densities), which motivates criticism.

Some constant-variance models can achieve similar prediction with right-skewed Gumbel distribution. It makes the probability that at least one individual in the new generation surpasses the previous generation's best contingent on population size (Henrich, 2004; Powell et al., 2009). However, such models tend to produce critical-threshold behavior (for example, predicting that all symbolic revolutions occur at the same population density) which has drawn considerable criticism (Vaesen et al., 2016).

Models with proportional variance exhibit a broad buffer zone in which markedly different outcomes – rapid adaptation or stasis – are similarly likely, even under identical initial conditions. Any variance expended through directional selection to shift the population mean closer to the optimum in one generation is no longer available to boost variance – and thus the expected adaptive step – in the next. This inherent negative feedback introduces a level of uncertainty that is absent from Galton–Pearson or Gumbel-distribution-based models of cultural evolution.

In a multidimensional “culturespace,” proportional variance between inputs and outputs (e.g. the positions of interaction partners before and after opinion exchange) facilitates the formation of subcultures, i.e., distinct clusters of mutually similar individuals.

The three phenomena – conformity, guided variation, and variance proportionality – have never been combined in a single model fitted to empirical data. Moreover, conformity and guided variation have rarely been translated from narrative descriptions into formal parameters within models of continuous trait transmission. In this study, we observe how these mechanisms co-define the distribution of “offspring” positions from socially transmitted “parental” cues along a single continuous dimension.

To collect an optimal dataset for this purpose, we developed an app called *Tip for Science*, styled as a trivia-guessing game. Each player answers a series of quantitative questions (e.g., “How high is Mount Uluru from base to top?”; see Supplement 1 for the full list), aiming to be as precise as possible to earn points. In most rounds, 1 to 5 guesses from previous players are displayed as potential hints.

A similar approach was introduced in the BEAST (Berlin Estimate AdjuStment Task) paradigm (Molleman et al., 2019), in which participants estimate the number of animals in a picture – a classic wisdom-of-crowds task – and later have the opportunity to adjust their guesses based on estimates provided by others. While BEAST focuses on individual variation in the use of social information, our framework is primarily concerned with how people integrate information from multiple sources. The game features a broader range of questions and a backend composed of multiple transmission chains with varying levels of selection and population size.

We model the likelihood of each submitted estimate (cultural offspring) as a function of the presented previous tips (cultural parents), using parameters that capture conformity, guided variation, and proportionality between parental and offspring standard deviation (SD). We also introduce a parameter that reflects proportionality between the parental mean and

offspring SD, based on the idea that people tend to vary the mean by a percentage (e.g. 10%) rather than by a fixed unit (e.g. 10 centimetres).

We compare models that assume humans integrate social information into a single unimodal distribution with benchmark models that assume the selection of a single parent followed by random modification (“kernel inheritance models”). All models account for variation in parameter values across participants and questions. The models presented in the main text are fitted on the decadic logarithm scale (see Supplement 3 for alternative).

## Methods

### The app

*Tip for Science* is a single-page web application developed in TypeScript (Bierman et al., 2014) and backed by the non-relational database DynamoDB.

Each question screen includes the question text, an illustrative image, 0–5 answers from previous participants, an answer field, and a submit button that also serves as a countdown timer. Players have 20 seconds to respond, a limit intended to discourage looking up answers online. After submitting a response, participants are shown the correct answer, their score (up to 100 points for a precise estimate, with partial credit awarded for answers within half to double the correct value), and a fun fact related to the question. If a player’s guess is highly accurate (within  $\pm 10\%$  of the correct answer), they are asked whether they already knew the correct answer. Points are awarded in both cases, but if the answer was not a guess, it is excluded from the analysis.

During registration, participants are informed that anonymized data will be used for scientific research and are asked to provide informed consent before proceeding. A short tutorial game then introduces the game mechanics. The questions are randomized in two blocks, since players may play for as long as they wish and are not required to answer all questions. Questions 1–15 are randomized independently and displayed first, ensuring that most participants answer at least this subset. The remaining questions are randomized in a second block.

The game backend is organized into *runs*. Each run is initialized with a specific population size per generation and a defined level of selection pressure. Initial tips for each participant in the first generation are either around  $3\times$  the correct answer or  $1/3$  of the correct answer. The variance of these tips is set at one of three levels – small, medium, or large – corresponding to a relative SD of 0.1, 0.2, or 0.3, respectively. For example, if the variance is medium (SD=0.2) and the mean is  $1/3$  of the correct answer, two parents would be positioned at 0.233 and 0.433 times the correct value.

After each generation, selection pressure is applied to the answers. This pressure is randomized across runs, ranging from 0% to 100%. The upper extreme (100%) corresponds to a situation where no parental tips are provided, and each player must estimate the quantity independently. For each generation,  $n$  guesses are obtained, and  $p$  percent of those furthest from the correct answer are eliminated from the pool of possible parents for the next cultural generation. The analysis of individual runs or adaptation dynamics within them is not a focus of this paper.

### Participants

Data were collected between August 2022 and May 2025. The game was freely accessible online to anyone, but most of the data come from two main waves. The first wave (August–December 2022; 32.8% of the data) used a snowball sampling method to recruit participants.

The second wave (two periods in August 2024 and April–May 2025; 60.9% of the data) was conducted under controlled laboratory conditions as part of a larger data collection project. Participants in this second wave were financially incentivized to achieve a high score, earning 1 Czech crown for every 50 points. The maximum reward was 284 CZK (approximately \$14) for completing the full set of 142 questions.

As announced in the preregistration available at <https://osf.io/ag5sy>, we excluded participants who answered fewer than 5 questions (27 participants), those who answered more than 20 questions but failed to earn any points (1 participant), responses tagged as "I knew this" (557 cases), and missing values (1 002 cases) where participants did not respond within the time limit. Additionally, we excluded 165 tips with a value of 0 (which corresponds to  $-\infty$  on the log scale) and 4 extreme tips. Three of these exceeded  $10^{21}$ , one that exceeded  $10^{15}$  was also removed as it was submitted in the question of average human lifespan in hours (implying a lifespan longer than the age of the Earth). None of these exclusions influenced the conclusions of our analysis. The final dataset consists of 42 048 unique submitted estimates.

In this study, we focus on data where two or more previous participants' tips with non-zero variance are displayed. This allows us to evaluate variance proportionality (with 2 or more parents) and conformity (with 3 or more parents). In the main article, we analyze data from both major waves together, regardless of whether participants were financially motivated. However, we also include a supplementary analysis comparing laboratory conditions with financial incentives to playing the game "in the wild."

This results in 33 814 unique submitted estimates from 492 participants across 140 questions (mean = 68.73 questions per participant, SD = 41.55). One question was excluded because it was numeric but not continuous ("What is the international dialling code for Yemen?"; see Supplement 6), and another was removed because it contained only a single long transmission chain with one surviving parent per generation. Providing information about gender or age was not required.

## Data analysis

All analyses were conducted in R using the *rethinking* package (McElreath, 2020) for Bayesian statistics and the Stan Hamiltonian Monte Carlo (HMC) sampling infrastructure. Posterior samples were extracted, visualized, and used to compute mean parameter estimates and 90% Compatibility Intervals (CIs, calculated as percentile intervals with 5% of posterior samples falling below and above the interval). Each new tip  $t_o$  was modeled as drawn from a distribution extending the Galton–Pearson model described in Eq. 1.

## Conformity

In most experiments studying conformity – that is, the tendency to preferentially imitate the majority (Bond & Smith, 1996; Cialdini & Goldstein, 2004) – participants choose from two or a few discrete options (Toyokawa & Gaissmaier, 2022; van Leeuwen et al., 2018). Here, we introduce a convenient parametrization of conformity for a continuous trait, captured by the parameter  $\mathbb{C}$ .

We replace the simple arithmetic mean of parental values with a function that weights parental values based on their mutual similarity (Figure 1).

Specifically, emphasis  $z_i$  on parental value  $t_{pi}$  is computed as the negative product of the conformity parameter  $\mathbb{C}$  and the total distance  $d_i$  from all other parental values  $t_{pj}$  normalized by the sum of all parental values  $\sum_{j=1}^M t_{pj}$ . The emphasis values  $z_i$  are then transformed into weights  $w_i$  using the softmax function (weights sum up to 1). In equations, this can be expressed as



$$d_i = \sum_{j=1}^M |t_{pi} - t_{pj}|, z_i = -\frac{\mathbb{C} d_i}{\sum_{j=1}^M t_{pj}}, w_i = \frac{e^{z_i}}{\sum_{j=1}^M e^{z_j}}. \quad (2)$$

The weighted mean is then defined as  $\mu_w(t_{pi}) = \sum_{i=1}^M w_i t_{pi}$ .

This formulation ensures that when  $\mathbb{C} = 0$  the weighted mean reduces to the arithmetic mean; when  $\mathbb{C} > 0$ , the cultural offspring's expected value is biased toward clusters of mutually similar parents (conformity); and when  $\mathbb{C} < 0$  the expected value shifts toward more isolated, less typical parents (anti-conformity).

This approach provides much finer resolution than, for example, the probability of copying the more frequent cultural variant (Kendal et al., 2009). It also offers several advantages over alternative continuous formulations we considered:

- (i) The conformity parameter  $\mathbb{C}$  is independent of offspring variance and exclusively captures the balance between conformist bias and the preferential imitation of rare cultural variants (Grabisch et al., 2019).
- (ii) The weighted mean scales proportionally with the magnitude of the parental values – the outcome is invariant under multiplication (e.g., parents 20, 21, 30 in Figure 1A yield the same relative outcome as 200, 210, 300 in Figure 1B).
- (iii) The formula is robust to the inclusion of an additional parent (e.g., 205 in Figure 1C): it does not overreact to such additions, yet is sensitive enough to register their influence.

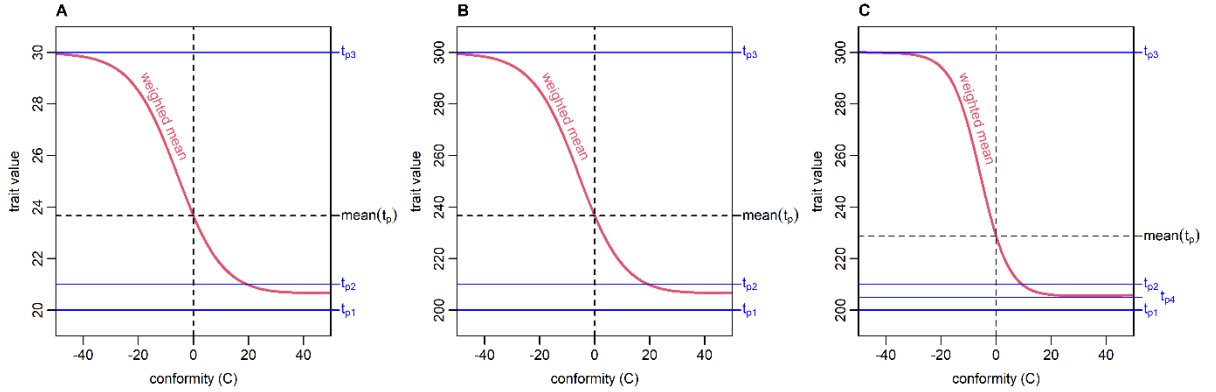


Figure 1. Visualization of the weighted mean of the offspring distribution for different values of the conformity parameter  $\mathbb{C}$ . A: Two mutually similar parents ( $t_{p1}, t_{p2}$ ) and one parent differing from both ( $t_{p3}$ ). B: Same configuration as in A, but each value is multiplied by 10. C: A fourth parent ( $t_{p4}$ ), similar to the first two, is added. The resulting weighted mean function becomes steeper.

## Guided Variation

Continuous models can relax the assumption of unbiased transmission by replacing the mean (or weighted mean) with a function  $f(t_p, \theta)$ , where  $\theta$  represents a cognitive attractor, or an optimal technological solution (Foster, 1986; Klepper & Simons, 2000; Lake & Venti, 2009). In the latter case, we refer to guided variation in the narrow sense.

The most straightforward implementation of guided variation computes the vector from the (weighted) parental mean to the optimum,  $\theta - \mu(t_p)$ , and adds a fraction  $a$  of it to the original mean.



$$\mu(t_o) = \mu(t_p) + a(\theta - \mu(t_p)). \quad (3)$$

If  $0 < a < 1$ , the cultural offspring tend to move toward the optimum. If  $a < 0$ , cultural offspring move away from the optimum and other forces such as cultural selection must maintain the course of adaptation. If  $a > 1$ , offspring are most likely to overshoot the optimum, landing on the opposite side from the starting position defined by the (weighted) mean of the cultural parents.

### Proportionality between input and output variance

If people adjust how much they deviate from cultural input based on the amount of variation they observe in others (as suggested, e.g., by Lorenz et al., 2011), the constant standard deviation  $\eta$  in Eq. 1 should be replaced or complemented by a term  $v\text{sd}(t_p)$ , a product of additional model parameter  $v$  and the standard deviation of the parental values. We treat  $\text{sd}(t_p)$  as the “population standard deviation” of the given set of parents (as opposed to “sample standard deviation”, the R’s default, that divides the sum of squares in the variance formula by  $n - 1$  instead of the number of parents  $n$ ). This definition ensures that for two parents,  $\text{sd}(t_p)$  is equal to half of the distance between them.

Adding this variance-dependent term introduces another source of variation independent of the constant. The resulting offspring trait  $t_o$  is modeled as the sum of two normally distributed random variables

$$t_o \sim N(f(t_p), \eta) + N(0, v\text{sd}(t_p)). \quad (4)$$

Thanks to the additivity of variance – the property that the variance of the sum of two independent normal random variables is the sum of their variances (Fisher, 1918) – this can be expressed as a single normal distribution

$$t_o \sim N\left(f(t_p), \sqrt{\eta^2 + v^2 \text{sd}^2(t_p)}\right) \quad (5)$$

### Proportionality between input mean and output variance

It is difficult to dismiss the intuition that humans are more likely to modify a trait by a percentage than by a fixed number of units. This tendency toward relative (rather than absolute) modification can be captured by introducing another stochastic term in which offspring variance is proportional to the parental mean.

Galton himself originally considered such a model alongside the model with constant standard deviation, so we label it Galton 2. Later, he also likely considered and rejected – from how data on heritability of human height are reported (Galton, 1877, 1886) – a model of biological inheritance with proportional variance between inputs and outputs.

In biological inheritance, the parental mean itself – unlike the alleles that constitute it – has no causal influence. Even if alleles affect traits in a multiplicative rather than additive manner, applying a logarithmic transformation to the trait removes any relationship between the parental mean and offspring variance. This assumption, however, does not necessarily hold for culturally transmitted traits.

Accordingly, the full likelihood function predicting an individual estimate (tip) in the *Tip for Science* game is defined as

$$N\left(f(t_p, \theta, a, \mathbb{C}), \sqrt{\eta^2 + \zeta^2 \text{mean}^2(t_p) + \nu^2 \text{sd}^2(t_p)}\right). \quad (6)$$

Here the function  $f(t_p, \theta, a, \mathbb{C})$  calculates the weighted mean of previous tips  $t_p$  based on the conformity parameter  $\mathbb{C}$ , and then shifts that mean by a fraction  $a$  toward the correct answer  $\theta$ .

We used weakly regularizing, unbiased priors based on maximum entropy criteria. The SD-defining parameters  $\eta$ ,  $\zeta$ , and  $\nu$ , which must be positive to ensure valid variances, were assigned exponential priors with rate = 1. The mean-modifying parameters  $a$  and  $\mathbb{C}$  were drawn from a normal distribution with mean 0 and standard deviation 1.

### Varying effects

All model parameters were assigned varying effects by question (since some questions may be inherently easier or harder to estimate) and by participant (to account for individual differences in strategy or reliance on previous tips). The standard deviation for each parameter at each varying-effect level was given an exponential prior with rate = 1. To improve model convergence, all varying effects were coded as non-centered and uncorrelated. Parameters constrained to be strictly positive were varied on the log scale.

For example, the varying effect structure for the parameter  $\eta$  is defined as:

$$\exp(\ln(\eta) + z_{q,\eta}[q]\sigma_{q,\eta} + z_{p,\eta}[p]\sigma_{p,\eta}) \quad (7)$$

where  $z_{l,\pi}$  is a standardized (mean = 0, SD = 1) varying-effect vector for level  $l$  (participant  $p$  or question  $q$ ) and parameter  $\pi$ , and  $\sigma_{l,\pi}$  is the corresponding standard deviation for that level and parameter. Mean modifiers that can take both positive and negative values were varied by simple addition.

### Model comparison

The full model (Eq. 6) is compared to simpler alternatives in which some of the inheritance parameters are fixed at 0. For example, setting  $a$ ,  $\mathbb{C}$ ,  $\zeta$ , and  $\nu$  to 0 reduces the likelihood function to the basic Galton-Pearson (GP) model (Eq. 1). Other model abbreviations used throughout the text are G2 (Galton 2, free  $\zeta$ ), VD (Variance-Dependent, free  $\nu$ ), A (Attraction or Guided variation, free  $a$ ), and C (Conformity, free  $\mathbb{C}$ ).

All combinations of free parameters are possible. This results in  $2^3 - 1 = 7$  configurations of SD-controlling parameters ( $\eta$ ,  $\zeta$ ,  $\nu$ , with at least one non-zero), and  $2^2 = 4$  configurations of mean-modifiers ( $a$  and  $\mathbb{C}$ ), giving the total of  $7 \times 4 = 28$  distinct models. These models are compared based on their expected out-of-sample performance using Widely Applicable Information Criterion (WAIC).

Because the responses span several orders of magnitude, and all questions assume positive numeric answers, we transform both tips and correct answers using the decadic logarithm before analysis. This transformation makes the offspring likelihood function effectively log-normal, which is the appropriate two-parameter maximum entropy distribution for strictly positive values.

We acknowledge that using the log scale does not allow to consider the naive Galton-Pearson model with constant standard deviation on the original scale. We therefore include comparisons with models fitted on the original (non-log) scale (most of them showing higher WAICs than even poorly performing log-scale models) in the Supplement.

## Kernel Inheritance models

It is an empirical question whether the mode of the resulting "offspring tip" distribution coincides with a (weighted) mean (shifted towards or away from the optimum) or whether a multimodal distribution provides a better fit. A process, where a participant selects a single parental tip to imitate (potentially influenced by conformity) and then draws an estimate around that value (optionally shifted toward the optimum) would naturally produce a proportionality between parental and offspring variance, without requiring a dedicated parameter like  $\nu$ . As parental values become more dispersed, the distribution of offspring in those models increases variance, then plateaus, and eventually produces a dip between multiple modes.

We refer to such multimodal mixture-distribution models as kernel inheritance models, due to their similarity to kernel density estimation (Chen, 2017). These models are straightforward to implement in a Bayesian framework and can be compared to the unimodal distribution models using the Widely Applicable Information Criterion (WAIC).

We identified three possible types of kernel inheritance models: (i) Galton-Pearson, i.e. constant SD kernels  $N(t_{p_i}, \eta)$ , (ii) Galton 2, where each parental kernel has standard deviation proportional to the trait value of that parent  $N(t_{p_i}, \zeta t_{p_i})$ , and (iii) Galton 2M, where each kernel's SD is proportional to the mean of all parental values  $N(t_{p_i}, \zeta \text{mean}(t_p))$ .

Strictly speaking, the third type – and any model incorporating conformity – already involves integration of information from multiple sources and therefore should not be considered uniparental (which is, unfortunately, the assumption of most theoretical models to date). Our main interest here, however, is in the contrast between unimodal and multimodal offspring distributions, and in the relevance of the proportional-variance parameter  $\nu$ , given its implications for adaptation probability, adaptation speed, and the potential role it plays in subculture formation.

All models were run in 4 chains for 1 000 iterations each (with each chain producing 500 posterior samples, since the first 500 iterations are allocated to warm-up). All data and code necessary to reproduce the results are available at: <https://osf.io/h6qz4/>.

## Results:

Models that assume integration of information from parental tips into a single unimodal distribution consistently outperformed multimodal kernel inheritance models (Table 1). The full model on the logarithmic scale showed the best out-of-sample performance overall (Figure 2). However, the ordering of the top models was not stable, as posterior distributions were highly multimodal and sampling effectiveness was low. Repeating the same Bayesian model comparison can occasionally result in a simpler model (containing only the constant and variance-proportional components, second best in Table 1) being selected as best (sampling works fine in this model, see Supplementary Figure S3).

The log transformation increases the interchangeability of  $\eta^2$  and  $\zeta^2 \text{mean}^2(t_p)$  in the variance term, because even a constant standard deviation on the log scale translates to greater dispersion around higher values on the original scale. This creates difficulties particularly in the by-question varying effects, where all responses to a given question tend to be similar in order of magnitude. As a result, in models that include both these variance components, it is difficult to determine whether questions (and participants, see Figure 2) vary more in their constant or mean-proportional variance.

The issue becomes even more pronounced in models on the original scale (Supplement 3), where higher means are typically associated with higher variances between parental tips. To some extent, this leads to confounding between  $\zeta$  and  $\nu$ . Likewise,  $\mathbb{C}$  and  $a$  can become slightly confounded with the variance terms, as both introduce greater offsets of the offspring tip from the parental mean (this problem is also more apparent on the original scale).

Details on how to address model convergence issues and multimodal posteriors can be found in Supplement 4.

**Table 1. Model comparison (top 10 models)**

Rank	Model	WAIC	SE	pWAIC	pSE	dWAIC	weight
1	m_GPG2VD_CA_log	32640.60	674.27	6628.74	140.23	0.00	1.00
2	m_GPVD_CA_log	33342.50	688.44	3235.04	116.97	701.90	0.00
3	m_G2VD_CA_log	33490.90	695.62	3390.36	124.10	850.30	0.00
4	m_GPG2VD_C_log	33890.95	696.48	6845.89	156.26	1250.35	0.00
5	m_GPVD_C_log	33947.40	689.47	2967.60	106.89	1306.80	0.00
6	m_G2VD_C_log	34116.26	697.26	3102.49	115.48	1475.66	0.00
7	m_GPVD_log	34120.77	691.10	2965.16	110.82	1480.17	0.00
8	m_G2VD_log	34291.22	695.31	3026.51	113.04	1650.62	0.00
9	m_kGP_CA_log	34334.15	687.12	2454.97	98.53	1693.55	0.00
10	m_GPG2VD_log	34391.99	706.97	6971.09	166.21	1751.39	0.00

...

*WAIC = Widely Applicable Information Criterion, SE = standard error, pWAIC = penalty term (average variance in lppd), pSE = standard error of the penalty term, dWAIC = difference from the best model, weight = relative model weight calculated as  $\exp(-0.5 \times dWAIC) / \sum \exp(-0.5 \times dWAIC)$ . GP = Galton-Pearson ( $\eta$ ), G2 = Galton 2 ( $\zeta$ ), VD = variance-dependent ( $\nu$ ), C = conformity ( $\mathbb{C}$ ), A = attraction ( $a$ ), k = kernel model. Comparison of all 40 models can be found in Supplementary Table S2*

Models that include variance-proportional variance, conformity, and guided variation consistently outperform those that omit them. Posterior estimates of these parameters and their varying effects are also unimodal and stable on the log-scale. Moreover, selection of a particular model does not influence these estimates substantially. According to the best-performing log scale model, a one-order-of-magnitude increase in parental standard deviation leads to an expected increase in offspring standard deviation of  $\nu = 0.85$  [90% CI: 0.78, 0.92]. The conformity parameter is reliably positive, with  $\mathbb{C} = 2.78$  [90% CI: 2.09, 3.53]. The attraction parameter indicates weak but detectable guided variation toward the optimum. On the log scale, offspring estimates are expected to shift between 2% and 6% of the distance from the parental mean toward the optimum, with  $a = 0.04$  [90% CI: 0.02, 0.06]. The implications of values of mean-modifying parameters are visualized in Figure 3.

Mean-modifying parameter values tend to vary more between questions than between participants. In some questions, participants conform more (variance in  $\mathbb{C}$ ), perhaps because they feel less confident about their own estimates, and answers in some questions can be more easily improved than in others (variance in  $a$ ).

These effects hold in both prominent waves when analyzed separately. When people are financially motivated to achieve higher scores, however, they become more risk-averse and rely more on social cues: they show higher conformity, and their variance-proportional parameter  $\nu$  is closer to 1, primarily at the expense of the mean-proportional variance term. However, lower independence from parental tips also limits the ability to constructively shift

the mean toward the optimum, which is reflected in a lower value of the guided variation parameter  $a$  (see Supplement 8).

Based on the comparison between the GPVD\_CA ( $\eta = 0.33$  [90% CI: 0.29, 0.37]) and GP\_CA ( $\eta = 0.42$  [90% CI: 0.38, 0.46]) models, where  $\eta$  captures all variance-independent components of offspring variance and serves as a measure of residuals, the variance-proportional term  $v$  and its varying effects uniquely account for approximately 38% of the total variance is participant guesses.

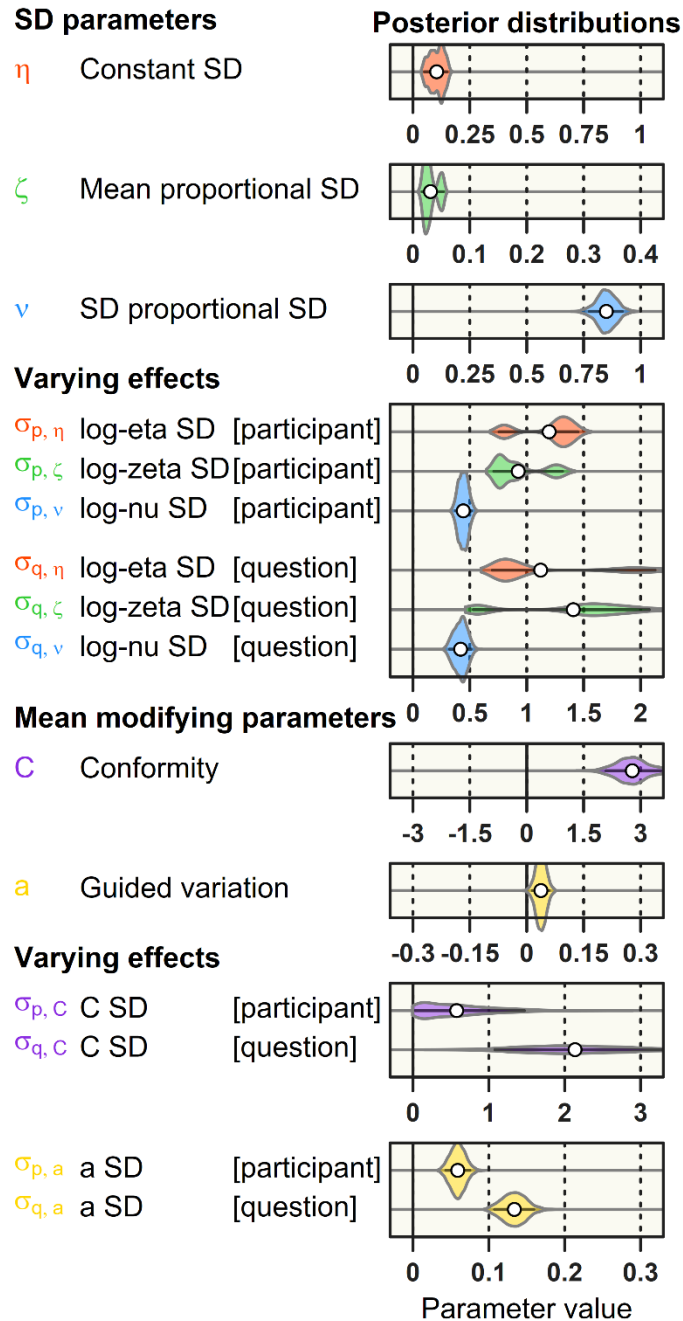
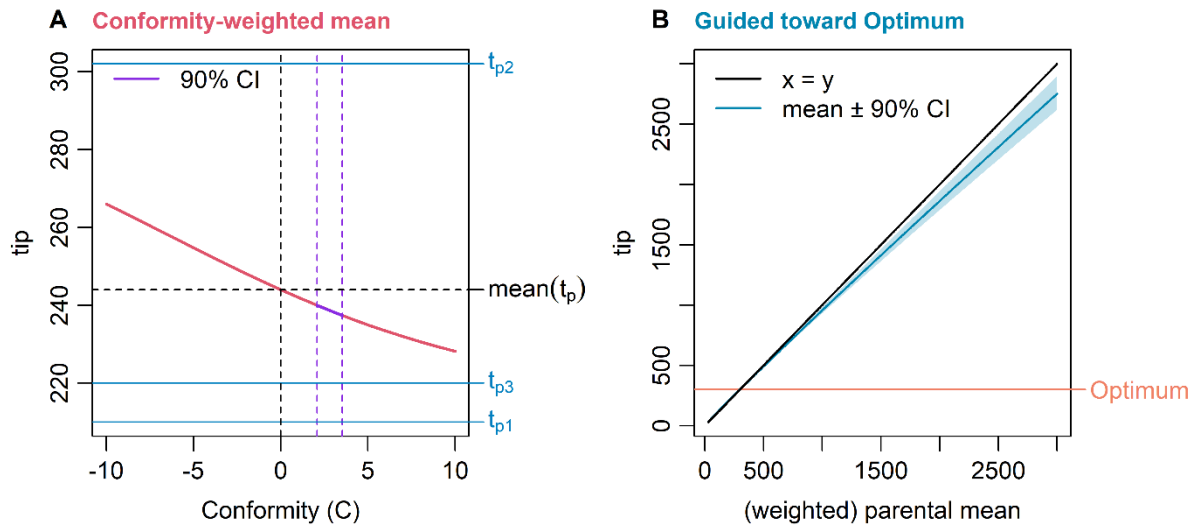


Figure 2 Posterior distribution of parameter values from the full model on the  $\log_{10}$  scale. The height of each coloured polygon corresponds to posterior probability density. White dots indicate the mean of the sampled parameter values; black lines represent the 90% Compatibility Intervals (CIs).



*Figure 3 Implications of mean-modifying parameters  $\mathbb{C}$  and  $\alpha$ . **A:** In a case with two mutually similar parental tips (210, 220) and one different (302), the expected offspring mean shifts from the unweighted parental mean (244) towards the two similar parents. With 90% probability, the offspring mean is expected to lie between 237 and 240 (violet segment of the  $\mathbb{C}$ -dependent weighted mean function; violet dashed lines indicate the 90% CI of the  $\mathbb{C}$  parameter itself). **B:** Further adjusted of offspring mean by guided variation. The black line shows the unmodified (weighted) mean, while the blue band indicates the 90% CI of the expected value after a 2–6% shift toward the optimum on the log scale. (For equivalent log-scale visualization, see Supplement 7.)*

## Discussion

Our results highlight the importance of unique features in cultural transmission: conformity, guided variation, mean proportionality, and variance proportionality. These factors illustrate how cultural inheritance departs from the Galton–Pearson model that successfully describes inheritance of continuous biological traits.

Conformity is robustly supported by our data: individuals tend to align with clusters of mutually similar answers rather than weighting isolated guesses equally. The formalization of conformity for continuous traits introduced in this manuscript goes beyond modelling guessing games. It offers a rigorous framework for studying conformity in latent spaces underlying discrete cultural preferences, values, and norms; domains in which people may also disproportionately follow mutually similar models.

Guided variation component shows that cultural adaptation may be feasible even without selection – participants systematically shift their estimates toward correct values by a small fraction, supporting the view that human learners do not copy blindly but refine information in the direction of improvement.

Guided variation shows that cultural adaptation may be feasible even in the absence of selection: participants systematically shift their estimates toward correct values by a small fraction, supporting the view that human learners do not copy blindly but refine information in the direction of improvement. However, the attraction force toward optimum is very small so selective forces are likely responsible for the fast pace of cultural adaptation.

Mean proportionality captures the intuition that people tend to adjust traits by percentages rather than fixed units: a 10% tweak leads to small deviations around small values and large deviations around large ones. On a logarithmic scale, this results in approximately equivalent to constant standard deviation, which explains why models with constant and mean-proportional variance can often be used interchangeably on transformed data. Notably, many studies of genetically inherited traits log-transform their data without acknowledging that this implies a multiplicative, rather than additive, model of inheritance (Couto Alves et al., 2019; Sinnott-Armstrong et al., 2021; Stopher et al., 2012; Van Rijn et al., 2007; Vattikuti et al., 2012). The mean–variance link helps explain why cultural change can scale smoothly across orders of magnitude. It also clarifies why models that fail to accommodate the extreme right skew of untransformed data often show convergence issues.

A distinctive feature of cultural inheritance – unparalleled in genetics – is variance proportionality. Unlike biological mutation, the “error” in imitation depends on how variable the inputs are: people explore more widely when others disagree and narrow their estimates when others show high agreement. This proportionality allows cultural systems to scale their exploratory range flexibly, enabling bursts of rapid change or periods of stability depending on population diversity (Tureček et al., 2019).

Distinguishing between kernel density models and unimodal models with proportional variance between inputs and outputs is difficult without a large dataset. When parental values are closely spaced, the resulting distributions may appear nearly identical. In our dataset, kernel models perform worse. Social transmission should not – and does not – opt out of integrating information from multiple sources. However, in real-world settings, the distinction may be less clear, as people likely filter out irrelevant information first. Future work in more messy, non-experimental contexts may require more complex approaches in which some potential cultural parents are explicitly ignored (as in kernel inheritance models) but the number of inspiration sources remains larger than one (as in unimodal models). Even in our experimental setting, the uniparental model – where each learner selects a single source and relies exclusively on that source – appears unrealistic (aligning with Enquist et al., 2010) despite the enduring popularity of such models in theoretical work and simulations (see e.g., Bentley et al., 2007; Kandler et al., 2023).

Comparison of participants from the two major waves suggests that financial incentives make people more risk-averse, a pattern consistently reflected across distribution function parameters (higher conformity, less constant variance, more variance-proportionate variance). Overall, they rely more on social information, as indicated by multiple model components. However, their reduced willingness to vary independently – regardless of social cues – also limits their ability to guide variation toward the correct answer (guided variation parameter  $\alpha$  is closer to 0). This effect persists even when parental mean distance from the optimum, parental variance, and their combination are held constant between the waves.

Together, our results suggest that social learning operates as a dynamic estimation process, which can, however, be adequately formalized using a parametric probability density function that involves multiple variance terms, conformity, and guided variation.



## References:

- Acerbi, A., & Mesoudi, A. (2015). If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution. *Biology and Philosophy*, 30(4), 481–503. <https://doi.org/10.1007/s10539-015-9490-2>
- Bentley, R. A., Lipo, C. P., Herzog, H. A., & Hahn, M. W. (2007). Regular rates of popular culture change reflect random copying. *Evolution and Human Behavior*, 28(3), 151–158. <https://doi.org/10.1016/j.evolhumbehav.2006.10.002>
- Bierman, G., Abadi, M., & Torgersen, M. (2014). Understanding TypeScript. In R. Jones (Ed.), *ECOOP 2014 – Object-Oriented Programming* (Vol. 8586, pp. 257–281). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-44202-9\\_11](https://doi.org/10.1007/978-3-662-44202-9_11)
- Blackmore, S. (2000). *The Meme Machine*. Oxford University Press.
- Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, 119(1), 111–137. <https://doi.org/10.1037/0033-2909.119.1.111>
- Boyd, R., & Richerson, P. (1996). Why Culture is Common, but Cultural Evolution is Rare. *Proceedings of the British Academy*, 88, 77–93.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.
- Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/C/bo5970597.html>
- Caldwell, C. A., & Millen, A. E. (2008). Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior*, 29(3), 165–171. <https://doi.org/10.1016/j.evolhumbehav.2007.12.001>

- Caldwell, C. A., & Millen, A. E. (2009). Social Learning Mechanisms and Cumulative Cultural Evolution: Is Imitation Necessary? *Psychological Science*, 20(12), 1478–1483.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.
- Chen, Y.-C. (2017). A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, 1(1), 161–187. <https://doi.org/10.1080/24709360.2017.1396742>
- Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(Volume 55, 2004), 591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015>
- Claidière, N., Scott-Phillips, T. C., & Sperber, D. (2014). How Darwinian is cultural evolution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1642). <https://doi.org/10.1098/rstb.2013.0368>
- Couto Alves, A., De Silva, N. M. G., Karhunen, V., Sovio, U., Das, S., Taal, H. R., Warrington, N. M., Lewin, A. M., Kaakinen, M., Cousminer, D. L., Thiering, E., Timpson, N. J., Bond, T. A., Lowry, E., Brown, C. D., Estivill, X., Lindi, V., Bradfield, J. P., Geller, F., ... Early Growth Genetics (EGG) Consortium. (2019). GWAS on longitudinal growth traits reveals different genetic factors influencing infant, child, and adult BMI. *Science Advances*, 5(9), eaaw3095. <https://doi.org/10.1126/sciadv.aaw3095>
- Danchin, E., Nöbel, S., Pocheville, A., Dagaëff, A.-C., Demay, L., Alphand, M., Ranty-Roby, S., van Renssen, L., Monier, M., Gazagne, E., Allain, M., & Isabel, G. (2018). Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-

- choice traditions. *Science*, 362(6418), 1025–1030.  
<https://doi.org/10.1126/science.aat1590>
- Dawkins, R. (1989). *The selfish gene*. New edition. Oxford ; New York : Oxford University Press, 1989. <https://search.library.wisc.edu/catalog/999612476002121>
- Dere, M., Bonnefon, J. F., Boyd, R., & Mesoudi, A. (2019). Causal understanding is not necessary for the improvement of culturally evolving technology. *Nature Human Behaviour*, 3(5), 446–452. <https://doi.org/10.1038/s41562-019-0567-9>
- Enquist, M., Strimling, P., Eriksson, K., Laland, K., & Sjostrand, J. (2010). One cultural parent makes no culture. *Animal Behaviour*, 79(6), 1353–1362.  
<https://doi.org/10.1016/j.anbehav.2010.03.009>
- Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52(2), 399–433.  
<https://doi.org/10.1017/S0080456800012163>
- Fogarty, L., & Kandler, A. (2025). Modelling Cultural Transmission. In J. J. Tehrani, J. Kendal, & R. L. Kendal (Eds), *Oxford Handbook of Cultural Evolution* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198869252.013.2>
- Foster, R. N. (1986). *Innovation: The Attacker's Advantage*. Summit Books.
- Galef, B. G., & Giraldeau, L.-A. (2001). Social influences on foraging in vertebrates: Causal mechanisms and adaptive functions. *Animal Behaviour*, 61(1), 3–15.  
<https://doi.org/10.1006/anbe.2000.1557>
- Galton, F. (1877). Typical laws of heredity. *Proceedings of the Royal Institution*, 282–301.  
[https://galton.org/bib/JournalItem.aspx\\_action=view\\_id=80](https://galton.org/bib/JournalItem.aspx_action=view_id=80)

- Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246.  
<https://doi.org/10.2307/2841583>
- Grabisch, M., Poindron, A., & Rusinowska, A. (2019). A model of anonymous influence with anti-conformist agents. *Journal of Economic Dynamics and Control*, 109, 103773. <https://doi.org/10.1016/j.jedc.2019.103773>
- Henrich, J. (2004). Demography and cultural evolution: How adaptive cultural processes can produce maladaptive losses—The Tasmanian case. *American Antiquity*, 69(2), 197–214.
- Henrich, J. (2016). *The Secret of Our Success*. Princeton University Press; JSTOR.  
<https://doi.org/10.2307/j.ctvc77f0d>
- Henrich, J., & Boyd, R. (2002). On Modeling Cognition and Culture representations. *Journal of Cognition and Culture*, 2(2), 87–112.  
<https://doi.org/10.1163/156853702320281836>
- Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology*, 126(2), 193–202. <https://doi.org/10.1037/a0025180>
- Hoppitt, W., & Laland, K. N. (2013). *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton University Press.  
<https://www.jstor.org/stable/j.ctt2jc8mh>
- Kandler, A., Fogarty, L., & Karsdorp, F. (2023). The interplay between age structure and cultural transmission. *PLOS Computational Biology*, 19(7), e1011297.  
<https://doi.org/10.1371/journal.pcbi.1011297>

- Kendal, J., Giraldeau, L. A., & Laland, K. (2009). The evolution of social learning rules: Payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*, 260(2), 210–219. <https://doi.org/10.1016/j.jtbi.2009.05.029>
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10681–10686. <https://doi.org/10.1073/pnas.0707835105>
- Klepper, S., & Simons, K. L. (2000). The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the U.S. Tire Industry. *Journal of Political Economy*, 108(4), 728–760. <https://doi.org/10.1086/316100>
- Krützen, M., Mann, J., Heithaus, M. R., Connor, R. C., Bejder, L., & Sherwin, W. B. (2005). Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences*, 102(25), 8939–8943. <https://doi.org/10.1073/pnas.0500232102>
- Lake, M. W., & Venti, J. (2009). Quantitative Analysis of Macroevolutionary Patterning in Technological Evolution: BICYCLE DESIGN FROM 1800 TO 2000. In *Pattern and Process in Cultural Evolution* (1st edn). University of California Press.
- Laland, K. N. (2017). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton University Press. <https://www.jstor.org/stable/j.ctv39x6r5>
- Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2001). Cultural niche construction and human evolution. *Journal of Evolutionary Biology*, 14(1), 22–33. <https://doi.org/10.1046/j.1420-9101.2001.00262.x>
- Liu, L., Dehmamy, N., Chown, J., Giles, C. L., & Wang, D. (2021). Understanding the onset of hot streaks across artistic, cultural, and scientific careers. *Nature Communications*, 12, 5392. <https://doi.org/10.1038/s41467-021-25477-8>

- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22), 9020–9025. <https://doi.org/10.1073/pnas.1008636108>
- McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd edn). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429029608>
- Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226520452.001.0001>
- Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29(4), 329–347. <https://doi.org/10.1017/S0140525X06009083>
- Molleman, L., Kurvers, R. H. J. M., & van den Bos, W. (2019). Unleashing the BEAST: a brief measure of human social information use. *Evolution and Human Behavior*, 40(5), 492–499. <https://doi.org/10.1016/j.evolhumbehav.2019.06.005>
- Morgan, T. J. H., & Feldman, M. W. (2024). Human culture is uniquely open-ended rather than uniquely cumulative. *Nature Human Behaviour*, 9(1), 28–42. <https://doi.org/10.1038/s41562-024-02035-y>
- Morin, O. (2013). How portraits turned their eyes upon us: Visual preferences and demographic change in cultural evolution. *Evolution and Human Behavior*, 34(3), 222–229. <https://doi.org/10.1016/j.evolhumbehav.2013.01.004>
- O'Brien, M. J., Boulanger, M. T., Buchanan, B., Bentley, R. A., Lyman, R. L., Lipo, C. P., Madsen, M. E., & Eren, M. I. (2016). Design Space and Cultural Transmission: Case Studies from Paleoindian Eastern North America. *Journal of Archaeological Method and Theory*, 23(2), 692–740. <https://doi.org/10.1007/s10816-015-9258-7>

- Osiurak, F., Lasserre, S., Arbanti, J., Brogniart, J., Bluet, A., Navarro, J., & Reynaud, E. (2021). Technical reasoning is important for cumulative technological culture. *Nature Human Behaviour*, 5(12), 1643–1651. <https://doi.org/10.1038/s41562-021-01159-9>
- Powell, A., Shennan, S., & Thomas, M. G. (2009). Late pleistocene demography and the appearance of modern human behavior. *Science*, 324(5932), 1298–1301. <https://doi.org/10.1126/science.1170165>
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.
- Sinnott-Armstrong, N., Naqvi, S., Rivas, M., & Pritchard, J. K. (2021). GWAS of three molecular traits highlights core genes and pathways alongside a highly polygenic background. *eLife*, 10, e58615. <https://doi.org/10.7554/eLife.58615>
- Sire, J.-Y., Delgado, S. C., & Girondot, M. (2008). Hen's teeth with enamel cap: From dream to impossibility. *BMC Evolutionary Biology*, 8(1), 246. <https://doi.org/10.1186/1471-2148-8-246>
- Sperber, D. (1996a). *Explaining culture*. Blackwell.
- Sperber, D. (1996b). *Explaining Culture: A Naturalistic Approach*. Blackwell Publishers.
- Sterelny, K. (2012). *The Evolved Apprentice*. The MIT Press; JSTOR. <http://www.jstor.org/stable/j.ctt9qf9t0>
- Stopher, K. V., Walling, C. A., Morris, A., Guinness, F. E., Clutton-Brock, T. H., Pemberton, J. M., & Nussey, D. H. (2012). Shared Spatial Effects on Quantitative Genetic Parameters: Accounting for Spatial Autocorrelation and Home Range Overlap Reduces Estimates of Heritability in Wild Red Deer. *Evolution*, 66(8), 2411–2426. <https://doi.org/10.1111/j.1558-5646.2012.01620.x>



- Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2405–2415. <https://doi.org/10.1098/rstb.2009.0052>
- Tomasello, M. (1999). *The cultural origins of human cognition* (pp. vi, 248). Harvard University Press.
- Toyokawa, W., & Gaissmaier, W. (2022). Conformist social learning leads to self-organised prevention against adverse bias in risky decision making. *eLife*, 11, e75308. <https://doi.org/10.7554/eLife.75308>
- Tureček, P., Slavík, J., Kozák, M., & Havlíček, J. (2019). Non-particulate inheritance revisited: Evolution in systems with Parental Variability-Dependent Inheritance. *Biological Journal of the Linnean Society*, 127(2), 518–533. <https://doi.org/10.1093/biolinnean/blz041>
- Vaesen, K., Collard, M., Cosgrove, R., & Roebroeks, W. (2016). Population size does not explain past changes in cultural complexity. *Proceedings of the National Academy of Sciences*, 113(16), E2241–E2247. <https://doi.org/10.1073/pnas.1520288113>
- van Leeuwen, E. J. C., Cohen, E., Collier-Baker, E., Rapold, C. J., Schäfer, M., Schütte, S., & Haun, D. B. M. (2018). The development of human social learning across seven societies. *Nature Communications*, 9(1), 2076. <https://doi.org/10.1038/s41467-018-04468-2>
- Van Rijn, M. J. E., Schut, A. F., Aulchenko, Y. S., Deinum, J., Sayed-Tabatabaei, F. A., Yazdanpanah, M., Isaacs, A., Axenovich, T. I., Zorkoltseva, I. V., Zillikens, M. C., Pols, H. A., Witteman, J. C., Oostra, B. A., & Van Duijn, C. M. (2007). Heritability of blood pressure traits and the genetic contribution to blood pressure variance

- explained by four blood-pressure-related genes. *Journal of Hypertension*, 25(3), 565–570. <https://doi.org/10.1097/HJH.0b013e32801449fb>
- Vattikuti, S., Guo, J., & Chow, C. C. (2012). Heritability and Genetic Correlations Explained by Common SNPs for Metabolic Syndrome Traits. *PLoS Genetics*, 8(3), e1002637. <https://doi.org/10.1371/journal.pgen.1002637>
- Whiten, A., Ayala, F. J., Feldman, M. W., & Laland, K. N. (2017). The extension of biology through culture. *Proceedings of the National Academy of Sciences*, 114(30), 7775–7781. <https://doi.org/10.1073/pnas.1707630114>
- Whiten, A., Horner, V., & de Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437(7059), 737–740. <https://doi.org/10.1038/nature04047>
- Yeh, D. J., Fogarty, L., & Kandler, A. (2019). Cultural linkage: The influence of package transmission on cultural dynamics. *Proceedings of the Royal Society B: Biological Sciences*, 286(1916), 20191951. <https://doi.org/10.1098/rspb.2019.1951>