Prestige as a Driving Force in Cultural Transmission

Saar Egozi[1] and Yoav Ram[2,3,*]

[1]School of Computer Science, Reichman University, Herzliya 4610101, Israel
[2]School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
[3]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel
[*]Corresponding author: yoav@yoavram.com, ORCID 0000-0002-9653-4458

April 25, 2022

Abstract

Copying role-models can be an efficient method for acquiring knowledge. A common bias when choosing a role-model to copy is success bias: copying whoever appears more successful. This bias depends on the performance of the role-model alone, with no other factors. We propose an additional bias that may be prevalent in cultural transmission: influence bias, in which role-model choice is affected by the number of individuals that have already copied each potential role-model. We combine success and influence bias together to a "prestige bias" and analyze its effects on cultural-evolutionary dynamics using mathematical analysis and stochastic simulations. We find analytic approximations to our stochastic model, facilitating further mathematical analysis and reducing the computational complexity of simulations. We validate these approximations using simulations, and demonstrate their robustness to model assumptions. We also find approximations to the fixation probability and the fixation time of an invading advantageous cultural trait, in both a constant and changing environment, which resemble Kimura's classical formulas. These approximations show that success bias effectively plays the part of natural selection, whereas influence bias effectively reduces the population size. It also accelerates the evolutionary dynamics, as can be expected in a *rich-getting-richer* process. Our model may provide a better description of human cultural transmission, especially in the last years where social networks are very popular. Further work is needed to test if this model could predict various phenomena in human cultural evolution when extended with the effects of selection and innovation.

Introduction

Cultural transmission. In cultural transmission, individuals transmit cultural traits (i.e., behavior, beliefs, norms) to one another, typically via learning and demonstrating (Cavalli-Sforza and Feldman, 1981). Examples for cultural traits in humans are behavioral patterns, such as personalities and habits, transmitted via both verbally and by observations (O'Brien et al., 2010). Although cultural transmission is most common in humans, it is also observed in other animals such as chimpanzees (Horner et al., 2010; Kendal et al., 2015), dolphins and whales (Whitehead, 2017). In elephants, McComb et al. (2001) showed that once a matriarch is removed from the group, the group's survival instincts are inferior and that "the possession of enhanced discriminatory abilities by the oldest individual [matriarch] in a group can influence the social knowledge of the group as a whole." By playing audio recordings of African elephants, they showed that groups with a matriarch recognize and react better to hostile or friendly calls than groups without a matriarch. Battesti et al. (2012) showed that choice of oviposition site in fruit flies is culturally transmitted: inexperienced flies that spent some time with experienced flies chose the same type of oviposition site even without directly observing this behavior. How the information is transmitted is still an open question, but it has been suggested that flies may use olfactory cues like other animals, such as rodents and bees.

Direction of transmission. Similar to genetic transmission, culturally transmitted traits can be transmitted from parents to offspring, and their effects of can be physiological rather than behavioral. For example, parents can "teach" their children to be strong or tall, within some biological limits, by instructing them to maintain a specific diet and engage in physical activity. Contrary to genetic transmission, cultural transmission can be non-vertical, that is, traits may be transmitted via social learning from non-parental individuals, and even unrelated individuals such as teachers, leaders, media, or any stranger that interacts with the learning individual. Thus, cultural transmission may combine vertical transmission, where parents transmit to their offspring; oblique transmission, where adults transmit traits to unrelated offspring; and horizontal transmission, where peers from the same age cohort transmit to one another. Vertical transmission is also possible in the opposite direction: parents may copy traits from their offspring (Cavalli-Sforza and Feldman, 1981; Creanza et al., 2017).

Transmission biases. In social learning, transmission biases cause a trait to have a disproportionate probability to be transmitted compared to its frequency in the population. Although more common in cultural transmission, transmission biases do occur in genetic transmission. For example, *wtf genes* in yeast bias their transmission to the gamete by secreting a long life-expectancy poison together with a short life-expectancy antidote, so that a gamete without the gene will perish because the poison will outlive the antidote (Eickbush et al., 2019). Importantly, even when a trait is disfavored by natural selection, it may still spread in a population due to transmission biases that are strong enough to overcome selection (Boyd and Richerson, 1988, Ch. 8 pg. 279). Cohen et al. (2021) show that cooperative behavior can evolve in an individual due to horizontal transmission bias even when there is no benefit to it, or when it benefits its competitors.

Success bias. Rendell et al. (2010) have conducted a tournament between learning strategies. Each strategy defines when individuals observe and copy from others, and when they engage in individual learning, in which an individual learns a cultural trait on his own. The best strategies had a high frequency of social learning relative to individual learning, even when the transmission error was almost 50%. From these results we understand that all the winning strategies were mostly based on success biased social learning, meaning it contributed more to the general success of the population than individual learning. However, all winning strategies included individual learning to some degree, implying that success-biased learning alone isn't the best way to advance, even when success is measured accurately.

Evaluating success. Boyd and Richerson (1988, Ch. 5) suggest that the evaluation of success can be divided into three groups: *direct bias*, *indirect bias* and *frequency-dependent bias*. A direct bias occurs when a variation of a trait is more attractive than others, and is evaluated by *directly* testing the variation of the trait. For example, an individual observing a Ping-Pong match can attempt both of the observed paddle grips to determine which grip is better. An indirect bias occurs when an individual uses the value of one trait to determine the attractiveness of another, so it *indirectly* evaluates the attractiveness of the role-model. For example, an observer may copy the paddle grip of the Ping-Pong player who scored more points in the match, thus indirectly evaluating the grip by the points scored. A frequency-dependent bias occurs when an individual has a probability to copy a variant of the trait that higher or lower than trait's frequency among demonstrators. For example, when an individual is 80% likely to copy the common paddle grip even when only 60% of the population is using it, it is said to be frequency-biased, or in this case, conformist. Frequency bias could be negative, i.e., non-conformist bias. Conformity and non-conformity are well-known biases in cultural transmission (Molleman et al., 2013), and its effect on cultural evolution have been studied with both models (Denton et al., 2020a,b) and experiments (Aljadeff et al., 2020).

Prestige. Prestige means having a good reputation or high esteem, therefore does not directly signify success (although it may imply it), making it an indirect bias. Both Boyd and Richerson (1988, Ch. 8) and Fogarty et al. (2017) suggest that prestige biases are probably more common in humans than success biases. Boyd and Richerson (1988, Ch. 8) add that maladaptive traits may spread widely in a population if indirect biases are strong enough. They suggest that such biases could lead to a runaway process caused by a cultural equivalent of sexual selection (Andersson, 1994). On the other hand, Henrich and Broesch (2011) suggest that prestige biases, over generations, can lead to cultural adaptations, and that although prestige can lead to maladaptive traits spreading in the population, it can also accelerate the spread of adaptive traits. Prestige is often mentioned in the cultural-evolution literature, but seldom modeled.

Influence bias. Today, social media provides an easy way to estimate the influence individuals have over others, and therefore may have an effect on decision making. Online social networks such as *Facebook* or *Instagram* are known to affect the social influence of individuals (Anagnostopoulos et al., 2008; Peng et al., 2018; Diga and Kelleher, 2009). For example, specific marketing practices were invented following this realization (Lee et al., 2012), Here we propose an indirect bias that we

call *influence bias*, in which the choice of a role-model depends on the the choices made by other individuals that have already chosen a role-model. This bias depends on the state of a role-model rather than a trait, in contrast to frequency biases such as conformity, which depend on the frequency of a trait in the population or in a sample of role-models. We define a model for prestige bias that combines both success and influence biases, provide analytic approximations for this model, and analyze its dynamics.

Models and Methods

We begin with a continuous trait model with indirect bias suggested by Boyd and Richerson (1988), develop an extension with influence bias, and then develop a model with a dichotomous trait. We implemented our stochastic models and approximations, performed statistical analyses, and produced figures using Python (Van Rossum et al., 2007) with NumPy (Van Der Walt et al., 2011) and Matplotlib (Hunter, 2007).

Source code is available at https://github.com/yoavram-lab/PrestigeBias.

Continuous trait

We follow the model of Boyd and Richerson (1988), assuming only oblique transmission of the trait and omitting the indirect trait in order to reduce model complexity. We consider a population of N individuals, described by a single trait with a continuous value. Each generation, N naive individuals, or copiers, choose an individual from the previous generation, or role-models, from which they will copy their trait. Similar to a Wright–Fisher model, we assume non-overlapping generations such that the entire population is replaced in each generation. The population at time t can be described by $\mathbf{A}(t) = \left(A_1(t), \ldots, A_N(t)\right)$ where $A_i(t)$ is trait value of individual i at time t. We assume the initial population is drawn from a standard normal distribution, $\mathbf{A}(0) \sim N(0, 1)$. Cultural transmission is modeled by a function F such that

$$A_i(t+1) = F_i(\mathbf{A}(t)) \ . \tag{1}$$

Success bias. Boyd and Richerson (1988, Ch.8, p.247-249) describe a blended transmission algorithm by defining F as a weighted average of the traits of all role-models,

$$F_i(\mathbf{A}) = \sum_{j=1}^{N} G_{i,j} \cdot A_{i,j} \ , \tag{2}$$

where $G_{i,j}$ is the success bias of role-model j in the eyes of copier i,

$$G_{i,j} = \frac{\beta(A_{i,j})}{\sum_{k=1}^{N} \beta(A_{i,k})} \ , \tag{3}$$

$A_{i,j}$ is the absolute trait value copier i estimates for role-model j with some error $e_i \sim N(0, \eta^2)$,

$$A_{i,j} = A_j + e_i, \tag{4}$$

and $\beta(\cdot)$ is the bias function that quantifies the success bias of a role-model,

$$\beta(A_{i,j}) = b \cdot \exp\Big(- \frac{(A_{i,j} - \hat{A})^2}{2J} \Big), \tag{5}$$

with $\hat{A}$ as the arbitrary optimal trait value, and J and b as parameters that control the bias strength. Therefore, $G_{i,j}$ is a relative success score that copier i assigns to role-model j, equivalent to relative fitness in evolutionary-genetic transmission models.

Boyd and Richerson (1988) note that the deterministic blended transmission algorithm they use has alternatives. We thus develop a similar stochastic model with transmission from a single random role-model. Instead of eq. (2), we define the transmission function F as a random variable with its distribution given by

$$\Pr\big(F_i(\mathbf{A}) = A_j\big) = G_{i,j} \ , \tag{6}$$

such that $G_{i,j}$ gives the probability of copier i to choose to copy the trait of role-model j.

Influence bias. Here we introduce a new element to the model. We assume that in each generation, copiers choose their role-models one by one. We denote $K_{i,j}$ as the number of copiers that chose role-model j after copier i chose a role-model. Thus, i out of N copiers already chose a role-model, $\sum_{j=1}^{N} K_{i,j} = i$, and there are $N - i$ copiers that have yet to choose a role-model. The stochastic process of role-model choice,

$$\big\{\mathbf{K}_i = (K_{i,1}, \ldots, K_{i,N})\big\}_{i=1}^{N} \ , \tag{7}$$

is described by the recurrence equation

$$K_{i,j} = K_{i-1,j} + S_{i,j}, \quad i, j = 1, 2, \ldots, N \ , \tag{8}$$

where $S_{i,j} = 1$ if the i-th copier chose role-model j and 0 otherwise, and the initial state is $K_{0,j} = 0$.

Following eq. (6), the probability that the i-th copier chose role-model j is given by the prestige of role-model j in the eyes of copier i,

$$Pr(S_{i,j} = 1 \mid S_{1,j}, S_{2,j}, ..., S_{i-1,j}) = G_{i,j} \ . \tag{9}$$

The prestige $G_{i,j}$ of role-model j in the eyes of copier i is determined by the estimated biased trait value $\beta(A_{i,j})$ and the number of copiers that chose role-model j before copier i, $K_{i-1,j}$, replacing eq. (3) of Boyd and Richerson (1988) with

$$G_{i,j} = \frac{\alpha_j \cdot \beta(A_{i,j}) + (1 - \alpha_j) \cdot K_{i-1,j}}{W_i} \ . \tag{10}$$

Here, the bias weight α_j is a characteristic of role-model j that determines the relative significance of success and influence in his prestige, the trait value of role-model j estimates by copier i, $A_{i,j}$, remains

as in eq. (4), and W_i is a normalizing factor that sums the numerator over all role-models ($1 \le j \le N$) to ensure $\sum_{j=1}^{N} G_{i,j} = 1$.

Dichotomous trait

To allow for mathematical analysis of the model, we introduce a simplified version where the trait only has two phenotypes: the optimal phenotype $\hat{A}$ and the sub-optimal phenotype A. All role-models with the same phenotype will contribute to the probability of that phenotype to be transmitted, and thus influence is determined by the number of copiers that have already chosen a role-model with either phenotype. Therefore, the probability of the i-th copier to copy phenotype A is

$$G_{i,A} = \frac{(N-X)\alpha'\beta(A) + K_{i,A}}{i - 1 + (N-X)\alpha'\beta(A) + X\alpha'\beta(\hat{A})} = \frac{(N-X)\alpha'\beta(A) + K_{i,A}}{i - 1 + (N-X)\alpha'\beta(A) + \alpha' X} \tag{11}$$

where X is the number of role-models with trait $\hat{A}$ and $K_{i,A}$ is the number of copiers that already chose A when copier i chooses a role-model and α' is the odd ratio of the bias weight $\alpha' = \frac{\alpha}{1-\alpha}$.

We prove this equation later, based on the proof of the Dirichlet-Multinomial approximation.

Initially, we assume the population has a single individual with phenotype $\hat{A}$ and $N-1$ individuals with phenotype A. The rest of the details follow the continuous trait model.

Results

Approximations

Our model is defined by two nested stochastic processes. Change over multiple generations is described by the phenotype distribution at each generation, $\{\mathbf{A}(t)\}_t$. The transition from one generation to the next is described by the number of copiers each role-model has after i copiers have chosen a role-model, $\{\mathbf{K}_i\}_{i=1}^{N}$. We emphasize that the models are nested: $\mathbf{A}(t+1)$ can be computed from $\mathbf{A}(t)$ by evaluating $\mathbf{K}_N$. However, the former requires iterating over eqs. (8) and (9). Thus, we sought to find an equivalent stochastic process that has the same joint distribution as $\mathbf{K}_N$. We found two such approximations, summarized here and explained in detail below:

1. The number of copiers of a specific role-model at each step, $K_{i,j}$, follows the *generalized binomial distribution* (Drezner and Farnum, 1993) and therefore,

   (a) the expected number of copiers of role-model j equals its prestige in the eyes of the first copier, multiplied by the total number of copiers, that is, $\mathrm{E}[K_{N,j}] = N \cdot G_{1,j}$ if trait estimation error is uniform for all copiers ($e = e_i$ for $i = 1, \ldots, N$).

   (b) the expected number of copiers of each role-model equals its relative biased trait value, similar to the role of relative fitness in population-genetic models, that is, $\mathrm{E}[K_{N,j}] = \beta(A_j + e)/\overline{\beta}$ if the bias weight is uniform for all role-models ($\alpha = \alpha_j$ for $j = 1, \ldots, N$), where $\overline{\beta} = 1/N \sum_{j=1}^{N} \beta(A_j + e)$ is the population mean estimated trait value.

2. The role-model choice process, $\{\mathbf{K}_i\}_{i=1}^{N}$, is equivalent to a *Pólya urn* model if trait estimation error is uniform for all copiers ($e = e_i$ for all $i = 1, \ldots, N$), meaning there's no meaning for the order of copiers. Therefore, the number of copiers of all role-models, $\mathbf{K}_i = (K_{i,1}, \ldots, K_{i,N})$, follows a Dirichlet-Multinomial distribution, $\mathbf{K}_i \sim DM\big(N, \mathbf{G}_1\big)$, where $\mathbf{G}_1 = (G_{1,1}, \ldots, G_{1,N})$ is the prestige vector of all role-models in the "eyes" of the first copier. Note that here $G_{i,j}$ is only a function of the trait values A_j and the bias weights α_j, as the estimation error is uniform, meaning all copiers are identical.

Generalized binomial distribution

The generalized binomial distribution (GBD) emerges from a series of dependent Bernoulli trials (in contrast to the standard binomial distribution in which trials are independent) and is parameterized by $GBD(n, p, \theta)$ where n is the number of trials, p is the probability of success of the first trial, and θ is the correlation between trials ($\theta = 0$ gives the standard binomial distribution).

Result 1 (GBD approximation)**.** *The number of copiers of role-model j after i copiers have chosen a role-model follows the GBD, $K_{i,j} \sim GBD(i, \alpha_i \cdot \beta(A_j + e))$ if $e_i = e$ for all role-models $i = 1, \ldots, N$.*

Proof. We denote $Q_j(k, i) = P(K_{i,j} = k \mid K_{i-1,j})$ as the probability that exactly k out of i copiers choose role-model j given $K_{i-1,j}$ our of $i-1$ copiers chose role-model j. Using conditional probability and eq. (8),

$$Q_j(k, i) = P_j(S_{i,j} = 1|k-1, i-1) \cdot Q_j(k-1, i-1) + P_j(S_{i,j} = 0|k, i-1) \cdot Q_j(k, i-1), \quad (12)$$

where $S_{i,j} = 1$ when the i-th copier chooses role-model j. Equation (12) is equivalent to eq. (2.1) by Drezner and Farnum (1993), which completes the proof.

Corollary 1. $\mathrm{E}[K_{N,j}] = N \cdot G_{1,j}$.

Proof. Drezner and Farnum (1993, eq. (2.3)) show that $\mathrm{E}[k] = N \cdot Q_j(1, 1)$ (modified from their notation). $Q_j(1, 1)$ is the initial probability to choose role-model j, before any role-model choices are made. $Q_j(1, 1) = G_{1,j}$ by definition, which completes the proof.

Corollary 2. $\mathrm{E}[K_{Nj}] = \alpha_j \cdot \beta(A_j + e) / \overline{\alpha \cdot \beta(A + e)}$*, where the averaging in the denominator is over the role-models index, j.*

Proof. The initial prestige of role-model j based on eq. (10) is

$$G_{1,j} = \frac{\alpha_j \cdot \beta(A_j + e)}{\sum_{m=1}^{N} \alpha_m \cdot \beta(A_m + e)} \,. \quad (13)$$

The denominator of eq. (13) can also be formulated as:

$$\sum_{m=1}^{N} \alpha_m \beta(A_m + e) = N \cdot \overline{\alpha \cdot \beta(A + e)} \,, \tag{14}$$

where $\overline{\alpha\beta(A+e)}$ is the mean value of $\alpha_m \cdot \beta(A_m + e)$ for all m. Using eq. (14) and **Corollary 1** we get,

$$\mathrm{E}[K_{Nj}] = \alpha_j \cdot \beta(A_j + e) \Big/ \overline{\alpha \cdot \beta(A + e)} \,, \tag{15}$$

completing the proof.

The special case where the bias weight is uniform for all role-models ($\alpha = \alpha_j$ for $j = 1, \ldots, N$) is interesting, because we can evaluate the expected number of copiers using a linear equation

$$\mathrm{E}[K_{Nj}] = N \cdot \frac{\alpha \cdot \beta(A_j + e)}{\sum_{m=1}^{N} \alpha \cdot \beta(A_m + e)} = \beta(A_j + e) \Big/ \overline{\beta(A + e)} \,, \tag{16}$$

where the only variable is $A_j + e$, because $\overline{\beta(A+e)}$ is the mean of the distribution we draw the trait values from, modified by some constant parameters of β. We can then denote $L = 1/\overline{\beta(A+e)}$ and write

$$\mathrm{E}[K_{Nj}] = L \cdot \beta(A_j + e) \,. \tag{17}$$

Numerical validation. To validate our results we ran $1,000$ simulations of the full model, and compared the results with corollary 2. We compare the distribution of number of copiers by plotting the histograms of both our simulations results and the expected values based on corollary 2. We used the average of the $1,000$ simulations to eliminate drift errors, as can be seen in fig. 1.

Although basic, fig. 1 shows good fit of the GBD approximation. This validation is initial, and the more extensive validations we do on the Dirichlet-Multinomial approximation, because it is what we will use in our analysts.

Dirichlet-Multinomial distribution

Pólya urn model. This stochastic process consists of N draws from an urn with an initial amount of colored balls of M colors. When a ball is drawn, it is then placed back in the urn together with an additional new ball of the same color. Let $\mathbf{U_i} = \{u_{i,1}, u_{i,2}, ..., u_{i,M}\}$ where $u_{i,j}$ is the number of balls of the j-th color in the urn after i draws. Let $S_{i,j} = 1$ when drawing a j-colored ball on the i-th draw, and 0 otherwise. The probability that $S_{i,j} = 1$ given $\mathbf{U_{i-1}}$ is

$$P(S_{i,j} = 1 \mid \mathbf{U_{i-1}}) = \frac{u_{i-1,j}}{\sum_{m=1}^{M} u_{i-1,m}} = \frac{o_j + w_{i-1,j}}{\sum_{m=1}^{M} o_m + w_{i-1,m}} = \frac{o_j + w_{i-1,j}}{i - 1 + \sum_{m=1}^{M} o_m} \,, \tag{18}$$

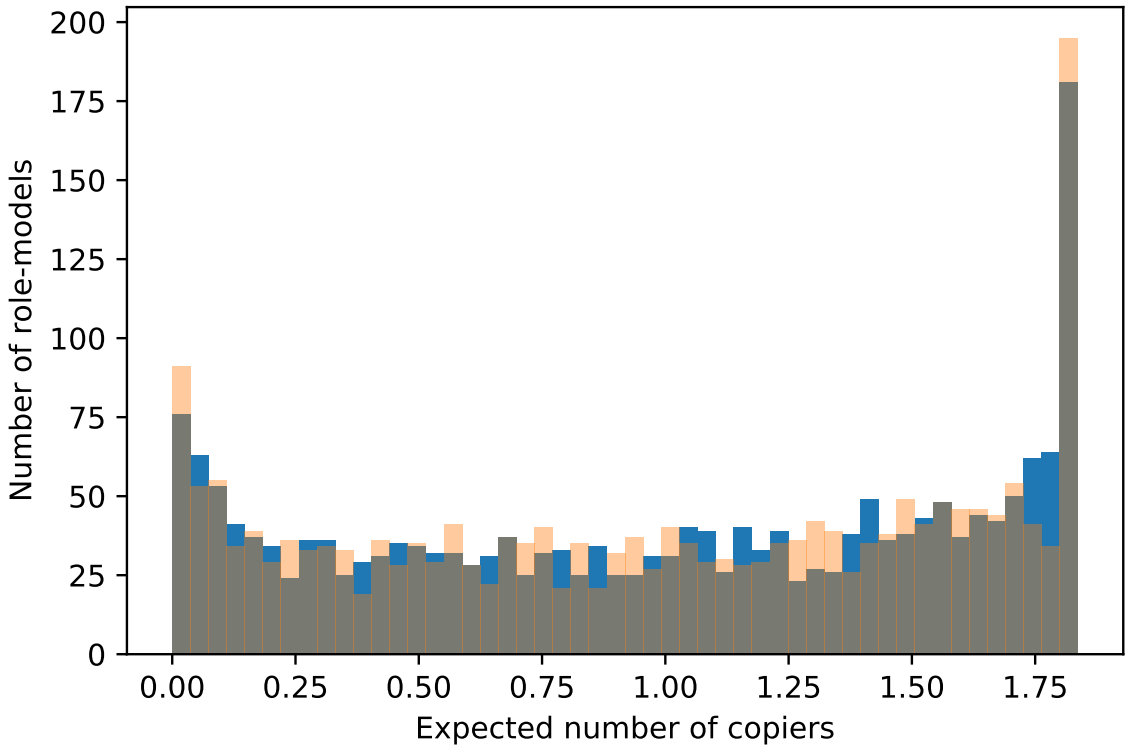

Figure 1: Numerical validation for the GBD approximation. The approximation (orange) fits simulation results (blue) well when using 1,000 simulations for both models. Here, population size, $N = 2,000$; bias weight, $\alpha = 0.1$; idea phenotype value, $\hat{A} = 1$; role-model traits $\mathbf{A} \sim N(0, 1)$; success bias value, $\beta(A) = 0.956$.

where o_j is the initial number of balls of the color j in the urn, and $w_{i,j}$ is the cumulative number of new balls that were added to the urn after i draws of the color j.

Result 2 (Pólya urn model)**.** *The role-model choice process,* $\{\mathbf{K}_i\}_{i=1}^{N}$*, is equivalent to a Pólya urn model if both trait estimation error and bias weight are uniform in the population,* $e = e_j$ *and* $\alpha = \alpha_j$ *for all* $j = 1, \ldots, N$.

Proof. Denote $\alpha' = \frac{\alpha}{1-\alpha}$ as the bias weight ratio, and $A'_j = A_j + e$. From eq. (10) and because $\sum_{j=1}^{N} K_{i,j} = i$, we have

$$G_{i,j} = \frac{\alpha'\beta(A'_j) + K_{i-1,j}}{\sum_{m=1}^{N} \alpha'\beta(A'_m) + K_{i-1,m}} = \frac{\alpha'\beta(A'_j) + K_{i-1,j}}{i - 1 + \sum_{m=1}^{N} \alpha'\beta(A'_m)} . \tag{19}$$

Substituting $M = N$, $o_j = \alpha'\beta(A'_j)$, and $w_{i,j} = K_{i,j}$ in eq. (18) gives eq. (19), thus completing the proof.

Frigyik et al. (2010, section 2) prove that the proportion of different colored balls in a *Pólya urn model* converges to the Dirichlet distribution as the number of draws approaches infinity, based on

the *Martingale Convergence Theorem* (Durrett, 1999). We therefore have an approximation for the relative prestige each role-model has when evaluated by copiers. Thus, choosing the role-models for all copiers is equivalent to drawing from a Multinomial distribution where the parameters are the modified weights from a Dirichlet distribution and we have the following corollary.

Corollary 3. *The number of copiers of each role-model follows a Dirichlet-Multinomial distribution,* $\mathbf{K_i} \sim DM(N, \mathbf{G_1})$*, under the conditions of Result 2.*

Numerical validation. To validate our analytical result (corollary 3) and test its sensitivity to the assumptions ($e_i = e$ and $\alpha_i = \alpha$ for $i = 1, \ldots, N$) we compare it to results of stochastic simulations of the full model. First, we computed an observed distribution of the number of copiers from the average empirical distribution of multiple simulations. We then compared this observed distribution with the expected theoretical DM distribution using Pearson's chi-squared test of goodness-of-fit. Thus, we can reject or accept the null hypothesis that simulation results are effectively drawn from a DM distribution.

Examining variable values of the bias weight parameter α, the test p-value was always 1, meaning that the null hypothesis cannot be rejected. For high α values (above 0.5), very few simulations are needed to reach p-value of 1 (less than five simulations for $\alpha = 0.9$, and less than 20 for $\alpha = 0.7$) For very low α values, which means very high influence weight, the number of simulations needed was 100. When testing our observed distributions against different distributions, for example the uniform distribution (all role-models have exactly one copier), the p-value was 0, so that the null hypothesis was rejected, as required.

Next, we examined the fixation probability and fixation time of an advantageous phenotype $\hat{A}$ when invading a population of phenotype A and compared results from the full model and the DM approximation. We find that the number of simulations needed to sufficiently approximate our model with the DM approximation is roughly $1,000$ (Figure 2). Next, we examined the robustness of the DM approximation to relaxing the approximation assumptions. First, we relaxed our assumption of no estimation error e. Estimation error in the original model was drawn from a normal distribution, and added to the trait value before evaluation of the bias ($A_{ij} = A_j + e_i$). When estimation error is applied, we sample J_i for each copier i from a normal distribution with varying scale (variance). Even when the standard deviation is 0.1, the fixation probability and time is similar (fig. 3). We also relaxed our assumption of a uniform bias weight α (i.e., $\alpha_i = \alpha$). We allowed α to vary in the population, drawing α_j for each role-model j from a normal distribution such that $\alpha_j \sim N(0.5, x)$ where $x \in [10^{-7}, 10^{-1}]$. We found again that results of the DM approximation are similar to those from stochastic simulations of the full model (fig. 4).

Fixation probability and time

After finding that the DM distribution is a good approximation of the (within-generations) role-model choice process, we turn our attention to the (between-generations) evolutionary dynamics. We focus

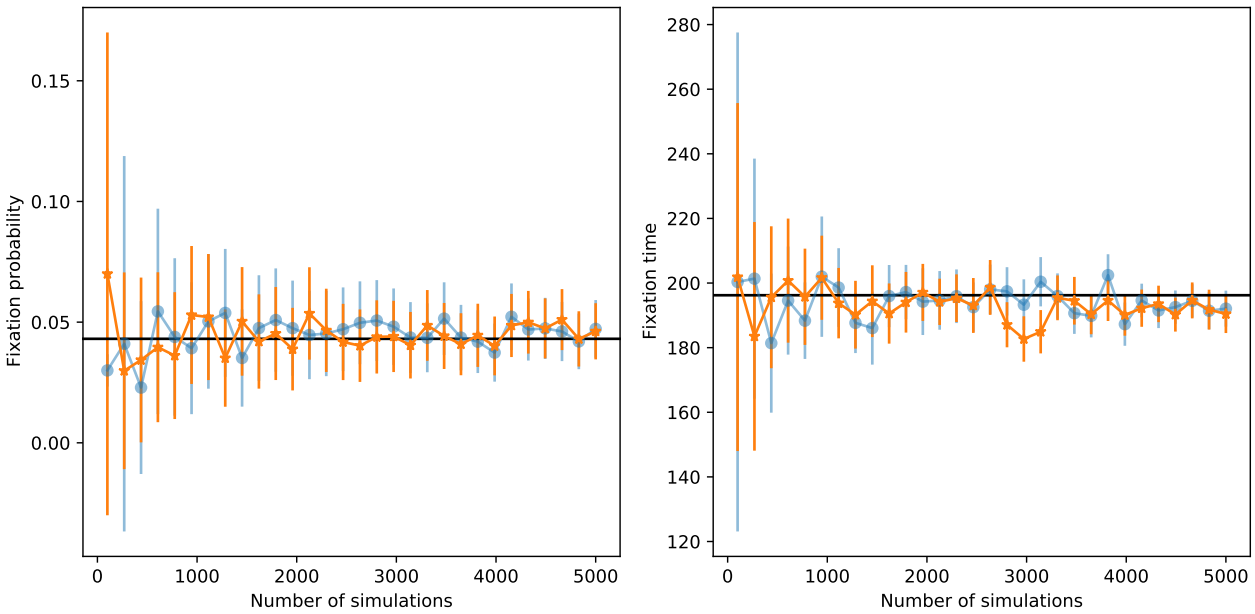

Figure 2: Minimum number of simulations required to get a good approximation. The approximation (orange) fits simulation results (blue) well when using 1,000 simulations for both models. The approximated value (black) is given by equation eq. (33). Markers for average value across simulations, error bars for 95% confidence interval. Here, population size, $N = 1000$; bias weight, $\alpha = 0.5$; phenotype values, $\hat{A} = 1$, $A = 0.7$; success bias value, $\beta(A) = 0.956$.

on the fixation probability and fixation time of an advantageous phenotype, similar to analyses in population-genetic models (Kimura, 1962; Kimura and Ohta, 1969; Otto and Whitlock, 2006).

We are mainly interested in the effect of the bias weight (α), which determines the relative effects of success and influence on prestige bias. For simplicity, we do not include estimation error in this analysis. As shown above, transmission in our model is approximately DM distributed with the parameters $\sim DM(\alpha' X, (N - X)\alpha'\beta(A))$ based on corollary 3 and eq. (19).

Result 3 (Effective selection coefficient)**.** *$1 - \beta(A)$ is equivalent to the selection coefficient s in the diffusion-equation approximation of the a classic Wright-Fisher model that approximate the fixation probability and fixation time of an advantageous allele.*

Proof. Let x be the frequency of type $\hat{A}$ in the population with N individuals. Let X be the number of individuals of type $\hat{A}$ so $x = X/N$. X' is the number of individuals with $\hat{A}$ in the next generation and x' their frequency. By definition $\beta(\hat{A}) = 1$, and for simplicity we'll denote $\beta(A) = \beta$ ($\beta < 1$).

The expected number of individuals of a DM distribution is:

$$E[X'] = N\frac{\alpha_1}{\alpha_1 + \alpha_2}, \tag{20}$$

where $\alpha_1 = \alpha' X$ and $\alpha_2 = \alpha'(N - X)\beta$, from eq. (11). We want to use frequencies instead of quantities to follow Durret's process so:

$$E[x'] = E[\frac{X'}{N}] = \frac{1}{N}E[X'] \tag{21}$$

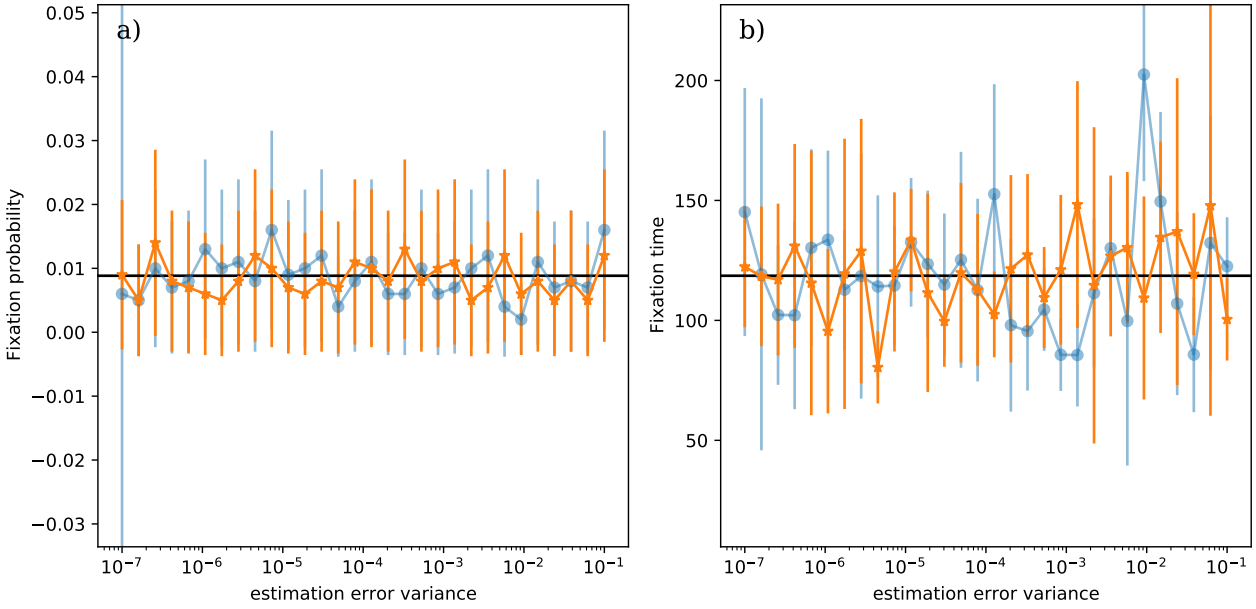

Figure 3: Robustness of DM approximations to inclusion of estimation error. Both the DM approximation (orange) and Kimura's equation (black line) fit the stochastic simulations (blue) well even with a high estimation error rate. Markers for average across simulations, error bars for 95% confidence intervals. 5,000 simulations per data point; population size, $N = 1000$; bias weight, $\alpha = 0.1$; phenotype values, $\hat{A} = 1$,$A = 0.7$; bias strength parameter $J \sim N(1, x^2)$ where $x \in [10^{-7}, 10^{-1}]$.

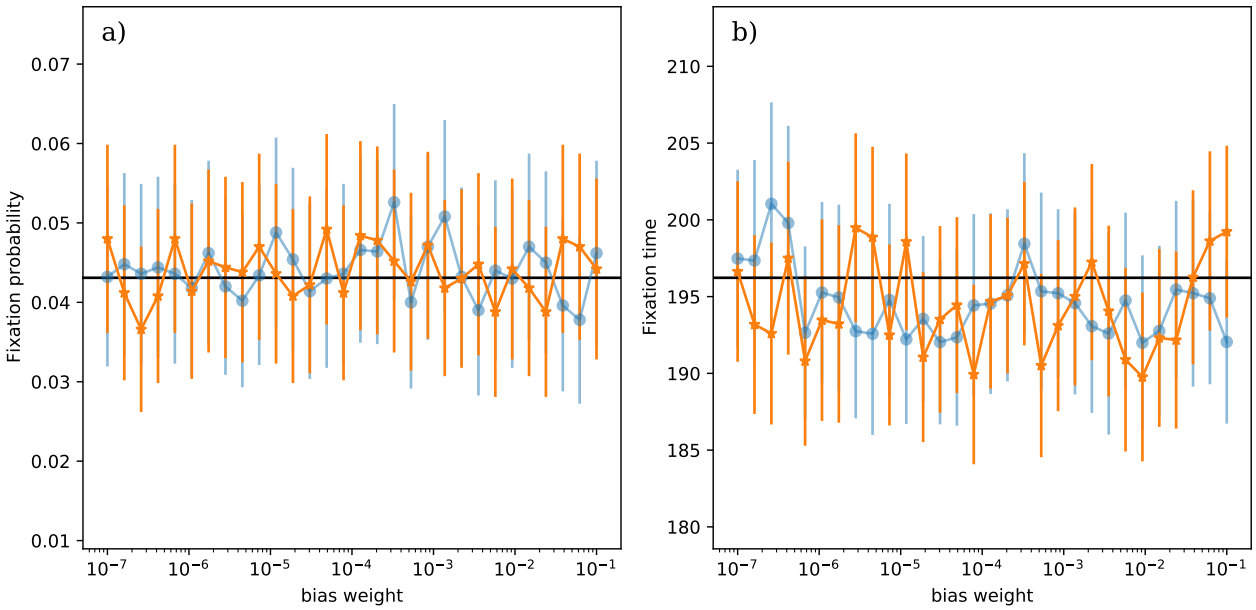

Figure 4: Robustness of DM approximations to variation in the bias weight α. High bias weight variance distances the approximation and the full model of generations to fixation from the Kimura's approximation, but not by much (confidence intervals still cover it). Error bars are 95% confidence intervals, and are less condensed (+- 0.03 probability and +-40 generations) 5000 simulations per data point; population size, $N = 1000$; trait bias normally distributed, $\alpha \sim N(0.5, x^2)$; phenotype values ,$\hat{A} = 1$,$A = 0.7$; success bias value, $\beta(A) = 0.956$.

Putting it together we get:

$$E[x'] = \frac{1}{N} N \frac{\alpha' x N}{\alpha' x N + \alpha' N (1-x)\beta} = \frac{x}{x + (1-x)\beta} \tag{22}$$

which is identical to the equation in the top of page 253, chap 7.2 in Durrett (2008). We therefore use the same approximation and say that:

$$\begin{aligned} E[x'] =& \frac{x}{x + (1-x)\beta} = \frac{x}{x + (1-x)(1-s)} = \\ &= x + x(1-x)s + o(s) \\ &= x + x(1-x)(1-\beta) + o(1-\beta) \end{aligned} \tag{23}$$

By definition, x is constant, so $E[x] = x$. We continue to calculate $E[x' - x]$:

$$E[x' - x] = E[x'] - E[x] = x(1-x)(1-\beta) + o(1-\beta) \tag{24}$$

where when substituting $1 - \beta$ with s, we get the same result as Durrett (2008) which is the desired result.

Result 4 (Effective population size)**.** *$N_e = \alpha N + (1 - \alpha)$, where Ne is the effective population size of our binary model.*

Proof. The variance of a DM distribution is:

$$V(X') = N \frac{\alpha_1}{\alpha_1 + \alpha_2} (1 - \frac{\alpha_1}{\alpha_1 + \alpha_2}) (\frac{N + \alpha_1 + \alpha_2}{1 + \alpha_1 + \alpha_2}) \tag{25}$$

And again, we want to use frequencies so:

$$V(\frac{X'}{N}) = \frac{1}{N^2} V(x') \tag{26}$$

Putting it together with our model's notations:

$$V(x') = \frac{1}{N^2} N \frac{x}{x + (1-x)\beta} (1 - \frac{x}{x + (1-x)\beta}) (\frac{N + \alpha' x N + \alpha' N (1-x)\beta}{1 + \alpha' x N + \alpha' N (1-x)\beta}) \tag{27}$$

Like Durrett, we'll use the zero order of the approximation when $\beta \approx 1$,so:

$$\frac{x}{x + (1-x)\beta} \approx x \tag{28}$$

and we also use $\beta \approx 1$ for the entire variance expression and get:

$$\begin{aligned} V(x') &\approx \frac{1}{N}x(1-x)(\frac{N+\alpha' xN+\alpha' N-\alpha' xN}{1+\alpha' xN+\alpha' N-\alpha' xN}) \\ &= x(1-x)(\frac{1+\alpha'}{1+\alpha' N}) \end{aligned} \tag{29}$$

Again following Durrett we want to calculate:

$$V(x'-x) = V(x') - V(x) \approx x(1-x)(\frac{1+\alpha'}{1+\alpha' N}) \tag{30}$$

because x is a constant so $V(x) = 0$

In our model, α' is the odds ratio of a parameter we called "bias weight", denoted in our model as α, so:

$$\alpha' = \frac{\alpha}{1-\alpha} \tag{31}$$

Combining eq. (30) and eq. (31) we get:

$$\begin{aligned} V(x'-x) &\approx x(1-x)(\frac{1+\frac{\alpha}{1-\alpha}}{1+\frac{\alpha}{1-\alpha}N}) \\ &= x(1-x)(\frac{\frac{1-\alpha+\alpha}{1-\alpha}}{\frac{1-\alpha+\alpha N}{1-\alpha}}) \\ &= x(1-x)(\frac{1}{1-\alpha(1-N)}) \\ &= x(1-x)(\frac{1}{\alpha N+(1-\alpha)}) \\ &= x(1-x)\frac{1}{N_e} \end{aligned} \tag{32}$$

Result 5 (Fixation probability approximation)**.** *The fixation probability is estimated by:*

$$P_{fix} = \frac{1-e^{-2(1-\beta)N_e x}}{1-e^{-2(1-\beta)N_e}} \tag{33}$$

where x is the initial frequency of the advantageous phenotype $\hat{A}$*.*

Proof. Using our substitute for a selection coefficient, $1-\beta$, and the effective population size N_e, we can estimate the fixation probability using the diffusion equation, $E[X'-X]$ (result 3), and $V(X'-X)$ (result 4).

Result 6 (Fixation time approximation)**.**

$$T_{fix} = \frac{1-P_{fix}}{1-\beta}\int_0^x \frac{e^{2(1-\beta)\xi}-1}{\xi(1-\xi)}d\xi + \frac{P_{fix}}{1-\beta}\int_x^1 \frac{1-e^{-2(1-\beta)(1-\xi)}}{\xi(1-\xi)}d\xi \tag{34}$$

Proof. Again using the diffusion equation and results 3 and 4, we get our result. the integrals cannot be solved in closed form, so we can only estimate them numerically.

Numeric validation. To validate our math we ran multiple simulations comparing our binary model with the classic Wright-Fisher model, using different α and β each time, and using the corresponding values of s and N_e for the WF simulations. In fig. 5 we changed α (and N_e accordingly) and used a constant β. In fig. 6 we changed β and used a constant α. In both cases we can see that the two models behave similarly, and both are approximated well by the Kimura's equations: eq. (33) and eq. (34).

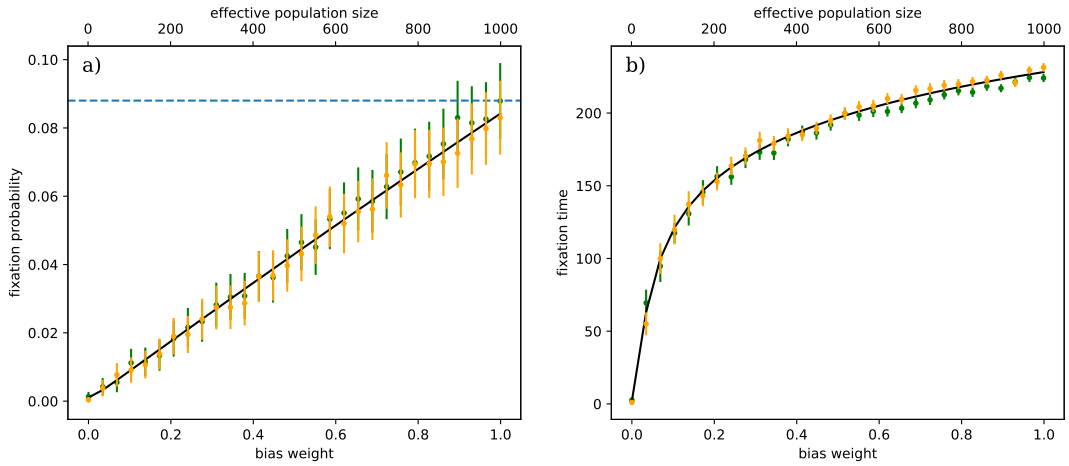

Figure 5: Goodness of fit of Kimura's approximation for varying bias weights. The approximation (black) fits DM simulations results (green) and Wright Fisher simulation results (orange) when using eq. (33). Fixation probability (a) is limited at approximately $2s$ (blue), like the classic WF model. Markers for average values across 5,000 simulations, error bars for 95% confidence intervals. Here, Population size, $N = 1,000$; phenotype values, $\hat{A} = 1$, $A = 0.7$; success coefficient, $1 - \beta = s = 0.044$.

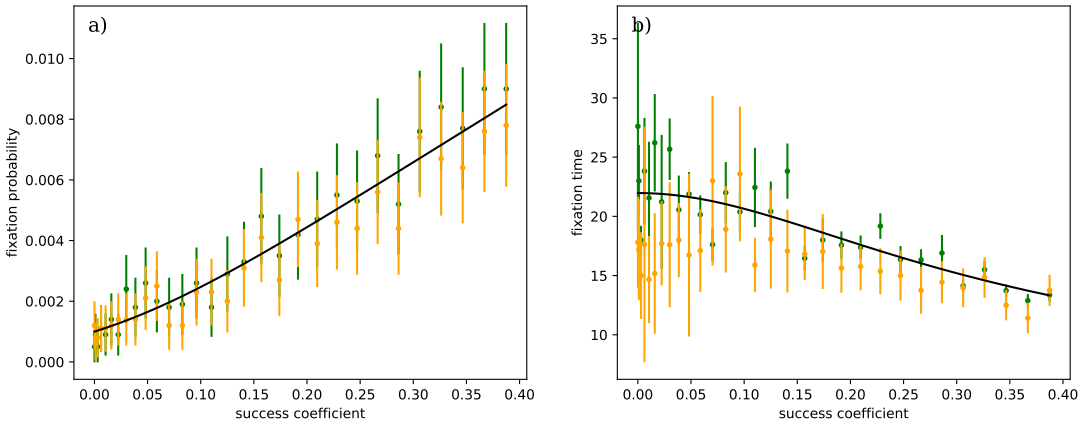

Figure 6: Goodness of fit of Kimura's approximation for varying success coefficients. The approximation (black) fits DM simulations results (green) and Wright Fisher simulation results (orange) when using eq. (33). Markers for average values across 5,000 simulations, error bars for 75% confidence intervals. Here, Population size, $N = 1,000$; phenotype values, $\hat{A} = 1$, $A = a \cdot \hat{A}, a \in [0.01, 0.99]$; bias weight, $\alpha = 0.01$.

After finding good estimations for our model in a constant environment, when the favorable trait is always $\hat{A}$, we want to find an estimation for our model in a changing environment.

For that we will find an expression for the expected and variance of the change in frequency between t generations. Let $s_t = N(1-\beta_t)$, and $S_n = \sum_{i=1}^{n} s_i$, where β_t is $\beta(A)$ at time/generation t.

Result 7 (Expected difference in frequency in a changing environment)**.** $E[\frac{X_t}{N} - x] \simeq \frac{1}{N} S_t x(1-x)$, *where x is the initial frequency of the favorable/invading trait and X_t is the number of individuals with the trait at time t.*

Result 8 (Variance of difference in frequency in a changing environment)**.** $V(\frac{X_t}{N}) \simeq \frac{1}{N_e} tx(1-x)$, *where x is the initial frequency of the favorable/invading trait and X_t is the number of individuals with the trait at time t.*

Proof. Following the proof of Ram et al. (2018), we prove by induction. We will prove results 7 and 8 together, because the induction assumptions are needed for both.

From eq. (24) we know that

$$\begin{aligned} E\left[\frac{X_{t+1}}{N} - \frac{X_t}{N}\middle| X_t\right] &= \frac{X_t}{N}\left(1 - \frac{X_t}{N}\right)(1-\beta_{t+1}) \\ &= \frac{1}{N}\frac{X_t}{N}\left(1 - \frac{X_t}{N}\right) s_{t+1} \end{aligned} \tag{35}$$

Also note that using the definition of $V(y) = E[y^2] - (E[y])^2$

$$\begin{aligned} E\left[\frac{X_t}{N}\left(1 - \frac{X_t}{N}\right)\right] &= E\left[\frac{X_t}{N} - \left(\frac{X_t}{N}\right)^2\right] \\ &= E\left[\frac{X_t}{N}\right] - E\left[\left(\frac{X_t}{N}\right)^2\right] \\ &= E\left[\frac{X_t}{N}\right] - V\left(\frac{X_t}{N}\right) - \left(E\left[\frac{X_t}{N}\right]\right)^2 \end{aligned} \tag{36}$$

we can now use the induction assumption of $V(\frac{X_t}{N})$ and get

$$E\left[\frac{X_t}{N}\left(1 - \frac{X_t}{N}\right)\right] \simeq E\left[\frac{X_t}{N}\right]\left(1 - E\left[\frac{X_t}{N}\right]\right) - \frac{1}{N_e} tx(1-x) \tag{37}$$

From eq. (35) we know that

$$\begin{aligned} E\left[\frac{X_{t+1}}{N} - \frac{X_t}{N}\right] &= \frac{1}{N} s_{t+1} E\left[\frac{X_t}{N}\left(1 - \frac{X_t}{N}\right)\right] \\ &\simeq \frac{1}{N} s_{t+1}\left(E\left[\frac{X_t}{N}\right]\left(1 - E\left[\frac{X_t}{N}\right]\right) - \frac{1}{N_e} tx(1-x)\right) \\ &\simeq \frac{1}{N} s_{t+1} \cdot E\left[\frac{X_t}{N}\right]\left(1 - E\left[\frac{X_t}{N}\right]\right) - \frac{1}{N_e N} s_{t+1} tx(1-x) \end{aligned} \tag{38}$$

Now we'll omit $O(\frac{1}{Ne \cdot N})$ and get

$$E\left[\frac{X_{t+1}}{N} - \frac{X_t}{N}\right] \simeq \frac{1}{N} s_{t+1} \cdot E\left[\frac{X_t}{N}\right]\left(1 - E\left[\frac{X_t}{N}\right]\right) \tag{39}$$

We'll now look at the induction assumption to see that

$$E\left[\frac{X_t}{N} - x\right] = E\left[\frac{X_t}{N}\right] - E[x] = E\left[\frac{X_t}{N}\right] - x, \tag{40}$$

so using the assumption we get

$$\begin{aligned} E\left[\frac{X_t}{N}\right] &\simeq \frac{1}{N} S_t x(1-x) + x \\ 1 - E\left[\frac{X_t}{N}\right] &\simeq 1 - \frac{1}{N} S_t x(1-x) + x \end{aligned} \tag{41}$$

we'll use both expressions in eq. (39) and get

$$\begin{aligned} E\left[\frac{X_{t+1}}{N} - \frac{X_t}{N}\right] &\simeq \frac{1}{N} s_{t+1} \left(\frac{1}{N} S_t x(1-x) + x\right)\left(1 - \frac{1}{N} S_t x(1-x) + x\right) \\ &\simeq \frac{1}{N} s_{t+1} \cdot x(1-x) \end{aligned} \tag{42}$$

after again omitting $O(\frac{1}{N^2})$ parts of the equation. To conclude our proof, we see that

$$E\left[\frac{X_{t+1}}{N} - x\right] = E\left[\frac{X_{t+1}}{N} - \frac{X_t}{N}\right] + E\left[\frac{X_t}{N} - x\right] \tag{43}$$

so again using the induction assumption, together with eq. (42) we get

$$\begin{aligned} E\left[\frac{X_{t+1}}{N} - x\right] &\simeq \frac{1}{N} s_{t+1} \cdot x(1-x) + \frac{1}{N} S_t \cdot x(1-x) \\ &\simeq \frac{1}{N} x(1-x)(S_t + s_{t+1}) \\ &\simeq \frac{1}{N} S_{t+1} x(1-x) \end{aligned} \tag{44}$$

which proves result 7.

For result 8, we'll use a property of variance:

$$V\left(\frac{X_{t+1}}{N}\right) = E\left[V\left(\frac{X_{t+1}}{N}\middle| X_t\right)\right] + V\left(E\left[\frac{X_{t+1}}{N}\middle| X_t\right]\right) \tag{45}$$

using eq. (35) we see that:

$$E\left[\frac{X_{t+1}}{N}\middle|X_t\right] - E\left[\frac{X_t}{N}\middle|X_t\right] = \frac{1}{N}s_{t+1}\frac{X_t}{N}\left(1-\frac{X_t}{N}\right)$$
$$E\left[\frac{X_{t+1}}{N}\middle|X_t\right] = \frac{X_t}{N} + \frac{1}{N}s_{t+1}\frac{X_t}{N}\left(1-\frac{X_t}{N}\right) \tag{46}$$

Using eq. (32) we get:

$$V\left(\frac{X_{t+1}}{N}\middle|X_t\right) = \frac{1}{N_e}\frac{X_t}{N}\left(1-\frac{X_t}{N}\right) \tag{47}$$

and using the equation $y'(1-y') \simeq y(1-y)$ on the first part of eq. (45) we get:

$$E\left[V\left(\frac{X_{t+1}}{N}\middle|X_t\right)\right] = \frac{1}{N_e}E\left[\frac{X_t}{N}\left(1-\frac{X_t}{N}\right)\right] \simeq \frac{1}{N_e}x(1-x) \tag{48}$$

and moving on to simplify the second part of eq. (45) using eq. (46):

$$V\left(E\left[\frac{X_{t+1}}{N}\middle|X_t\right]\right) = V\left(\frac{X_t}{N} + \frac{1}{N}s_{t+1}\frac{X_t}{N}\left(1-\frac{X_t}{N}\right)\right) \tag{49}$$

and now, because $\frac{X_t}{N}$ is a frequency, i.e $0 \leq \frac{X_t}{N} \leq 1$, we know that $V\left(\frac{X_t}{N}\left(1-\frac{X_t}{N}\right)\right) \leq \frac{1}{4}$. We therefore see that:

$$V\left(\frac{1}{N}s_{t+1}\frac{X_t}{N}\left(1-\frac{X_t}{N}\right)\right) \leq \frac{1}{4N^2}s_{t+1}^2 \tag{50}$$

and so it can be ignored. Combining our equations we get:

$$V\left(E\left[\frac{X_{t+1}}{N}\middle|X_t\right]\right) = V\left(\frac{X_t}{N}\right) + O\left(\frac{1}{N^2}\right) \simeq V\left(\frac{X_t}{N}\right) \tag{51}$$

Using the induction assumption and eq. (48):

$$V\left(\frac{X_{t+1}}{N}\right) \simeq \frac{1}{N_e}x(1-x) + \frac{1}{N_e}tx(1-x) \simeq \frac{1}{N_e}x(1-x)(t+1) \tag{52}$$

proving result 8.

Following our proof, we can say that after many cycles, we can use a modified version of our fixation probability:

$$P_{fix} = \frac{1-e^{-2\frac{S_n}{n}N_e x}}{1-e^{-2\frac{S_n}{n}N_e}} \tag{53}$$

where $\frac{S_n}{n} = \frac{k-l}{k+l}(1-beta)$, $n = k+l$. Put into words, we use the average selection coefficient of a cycle $(k+l)$ as the selection coefficient in our original equation. In our proof we showed that the expected change in frequency and variance is only manifested in the selection coefficient S_n, and that we can use those modified equation as a base for Kimura's equation.

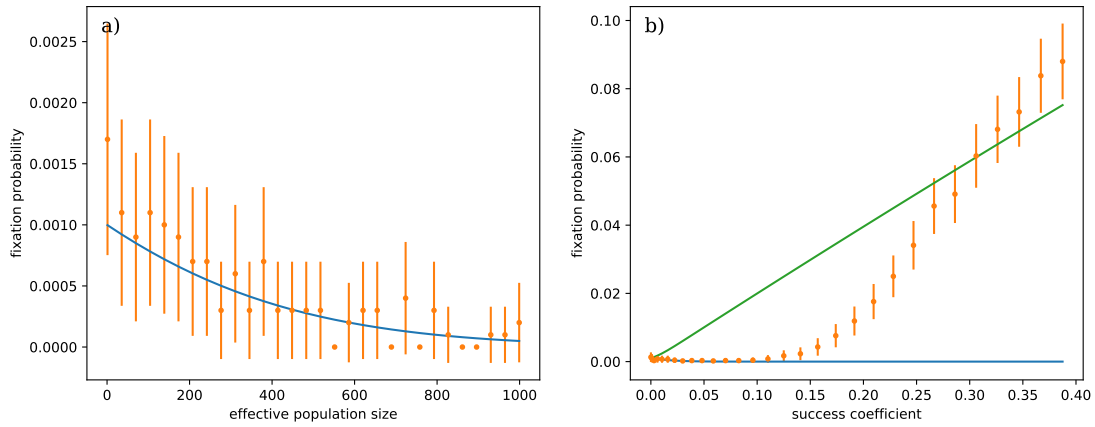

Figure 7: Robustness of approximation in a changing environment to varying success coefficient and success bias weight. Markers for average values across 10,000 simulations, error bars for (a) are 75% confidence intervals, and 95% for (b). The approximation (blue) given by eq. (53) fits simulation results (orange) well when the effective population size varies (a). On (b) we see that for high values of success bias ($s > 0.1$) will distance the simulations from the changing environment expected values. Very high values ($s > 0.35$) will even deviate from the constant environment expected values (green) given by eq. (33). This is expected because Kimura's approximation are only viable for low selection coefficient values. Here, population size, $N = 1,000$; phenotype values, $\hat{A} = 1$, $A = 0.9$; **(a)** success bias/selection coefficient is: $1 - \beta = s = 0.005$, **(b)** success bias weight is: $\alpha = 0.1$.

We validate our results using simulations. We find that α does not have a significant effect on the approximations (fig. 7). However, β does.

We suspect that when β is too large, there won't be many cycles in the simulations. This might happen if either the population reaches a high frequency of the fitter phenotype after only a few cycles, or the fitter phenotype get extinct very quickly. For the very high β values (0.35 and above), the fixation probability exceeds even the constant environment approximation (which is the upper limit), but it's to be expected, because Kimura's equations are only appropriate for weak fitness (i.e., low selection coefficients, s).

We also examined the effect of different choices of k and l while keeping a constant total cycle length, $n = k + l = 40$. We found that larger k-to-l ratio increases the agreement between the changing environment equation and the original estimation of the constant environment. When using higher values of n, the agreement between the approximation and simulation results is weaker. This is because our approximation is more precise when more cycles occur. When n is high, the cycles are longer. When the cycles are long, the settings are similar to a constant environment equation, and fixation/extinction is more likely to happen. This means there will be less cycles, which lowers our approximation accuracy.

Discussion

Summary. During cultural transmission, cultural traits such as attitudes, values, beliefs, and behavioral patterns are transmitted between individuals, for example via copying and social learning. Some

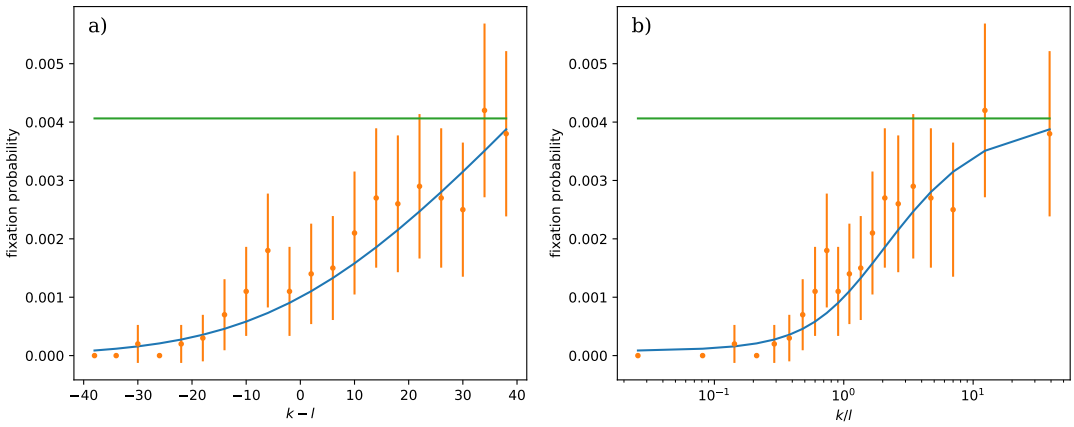

Figure 8: Robustness of approximation in a changing environment to varying sizes of the cycles. Markers for average values across 10,000 simulations, error bars are 95% confidence intervals. The approximation (blue) fits the simulations (orange) well for different sizes of the changing environment cycle, where k is the number of generations where the bias is advantageous, and l are the generations it is disadvantageous. When $k < l$ the approximation is good. When $k > l$, the approximation and the simulations are both very close to the constant environment approximation (green). Here, population size, $N = 1,000$; phenotype values, $\hat{A} = 1$, $A = 0.8$; success coefficient, $1 - \beta = s = 0.02$; success bias weight, $\alpha = 0.1$.

cultural traits or cultural role-models may be more likely to be copied due to transmission biases. A common one is success bias, in which copiers are more likely to copy a successful role-model. Many models assume that success can be correctly identified. However, it has been suggested (Fogarty et al., 2017) that because estimating success is hard, a more common bias is *prestige bias*—a bias towards role-models perceived to be successful by some other trait.

Inspired by a model by Boyd and Richerson (1988), we developed a cultural-evolution model with prestige bias, which included both success and influence biases, where the latter is a bias towards role-models with many copiers. We found approximations for the role-model choice process: the generalized binomial distribution that approximates the number of copiers of a single role-model, and the Dirichlet-Multinomial distribution that approximates the distribution of copiers in the entire population in a *rich-getting-richer* type of process. We then studied the effects of prestige bias on the fixation probability and time of an invading advantageous cultural phenotype in both a constant and a changing environment.

Prestige in the literature. According to King and Cowlishaw (2009), there are two main approaches to group decision-making in nature: leadership and consensus. Leaders may lead the group when traveling, decide on a nesting site, or choose foraging patches. They found that leadership is mostly observed when there are a lot of connections in the social network of the group, and when there are individuals that present leadership behavior. Leaders would usually be high-ranking members of the group: elders, individuals with many kin relations, or individuals possessing other dominant traits. When no individual posseses such traits, or when there are not enough connections and relations in the social network, decisions are more likely to be made by consensus. Leaders may carry selection costs, e.g. due to higher predation risk, poisoning when exploring new foraging patches. In some cases,

the leader may be the hungriest or the weakest individual, while the others would prefer to follow the leader and minimizing their risks and costs. However, leadership may also carry selection benefits. For example, given the route to the foraging site was successful, the leader and his closest followers would gain most or the best of the food, whereas in consensus decision making, the food would be shared more equally. Or, sexual selection may also provide a selection benefit advantage to leaders, who might gain more sexual partners due to their leadership of the group. King and Cowlishaw (2009) describe benefits for the closest associates of a dominant baboon, such as protection from predators. In some species, like the females of *Eulemur fulvus rufus* (Red lemur), leaders may arise due to nutritional needs, and not due to possessing superior traits (Erhart and Overdorff, 1999).

In humans, leadership also has its costs and benefits. Leaders can make decisions that would most benefit them and their closest followers, while still maintaining group cohesion. However, wrong decision making that harms the group could result in harm.

In modern society humans strive for the prestigious positions, as they may reap rewards greater than the risk and costs to achieve them, or due to individual personality and pressure/education from the family.

Van Vugt and Smith (2019) suggest a different view of leadership. They note that most discussions assume there is one type of leadership, as seen above, and so they differ in their definitions. Van Vugt and Smith (2019) suggest a way to solve said contradiction by defining two types of leaderships: prestige-based and dominance-based. They present classical views of leaderships by Confucius and Machiavelli. Confucius views leaders as role-models who exercise influence through possessing superior knowledge, skills, and (outstanding) personal qualities. This description is very similar to success bias in our model. By contrast, Machiavelli views leaders as rulers who exercise influence by imposing costs through (the threat of) punishment and violence. They say that these two opposing views are both partially supported by the available evidence but each one on its own offers an incomplete view into the complex and dynamic processes of leadership. Our model does not reflect these leadership styles, but several adjustments could be made in order to match it, such as assuming there is a correlation between phenotype to leadership style. The emerging cultural-evolutionary dynamics and their dependence on the costs and benefits are intriguing.

Henrich and Gil-White (2001) support said claim that there are two types of leadership, and also define the two as prestige based and dominance based leadership types. By their definition, the latter is defined by acquiring social status by using aggression, intimidation and violence. It is also more common than prestige in non humans. Their definition of prestige is somewhat synonymous with ours. According to their manuscript, prestige is composed both of estimation in the eyes of people (our success trait) and commanding position in people's minds, i.e number of copiers people think they have, which they define as *influence* (similar to our definition for influence). In their paper, they show that prestige evolved from natural selection, as an efficient process to extract reproductive benefit from the flow of socially transmitted information. Simply put, prestige is a natural step where social learning exists, due to saving costs of individual learning. It could be interesting in the future to expand our model using this idea: observing the copier trait of *evaluation*, rather than only observing the

evolution of the trait copied. Henrich and Gil-White (2001) suggest that the most skilled role-models will, on average, end up with most copiers. Their research, definitions and results, is consistent with ours.

So far we presented the theory behind prestige, and it's appearance in nature. The following will show the appearances of prestige biases in humans, and in recent times.

Chudek et al. (2012), for example, tested the existence of prestige in young children. Chudek et al. (2012) report the first direct tests in children that suggest the existence of *prestige bias*, a tendency to learn from individuals to whom others have preferentially attended, learned or deferred. Their definition of prestige is similar to our *influence bias*, and brings concrete proof of its existence and effects. Their study showed that the odds of 3-4 years-old children learning from an adult model to whom bystanders had previously preferentially attended for 10 seconds were over twice those of their learning from a model whom bystanders ignored. In addition to this first study, they also discovered prestige effects are domain-sensitive. They saw that prestigious models were listened to by most when demonstrating artifact-use, but not as much as when presenting food preferences. It lead Chudek et al. (2012) to believe that when the trait is costly to learn individually, prestige will have a higher effect. It would be interesting to include costs in our model to try and observe these effects and their dynamics in the simulations of a larger population than this study.

Henrich and Broesch (2011) researched Fijian villages, looking for evidence of social learning biases and their origins. They mention that:

> evolutionary theorists propose that natural selection has favored the emergence of psychological biases for learning from those individuals most likely to possess adaptive information.

Their goal is to bridge from the laboratory to the field by examining if and how these biases emerge in a small-scale society. During their research they found that:

> Fijian villagers (ages 10 and up) are biased to learn from others perceived as more successful/knowledgeable, both within and across domains (prestige effects).

Their research shows promising evidence for our prestige model, suggesting that copying from others who are *perceived* as successful, rather than actually are successful. In their paper, they show that the social networks representing copier-role-model relationships are centralized, suggesting:

> This degree of centralization is consistent with the prediction that people substantially share notions about who is a good cultural model (network centrality), but that individuals' model selections are influenced by multiple factors.

We see here support for both our trait and our influence bias in their data.

Aside from children's learning biases and small villages in a relatively primitive population, we can see prestige in more advanced domains as well, like western medicine.Norredam and Album (2007) present a specific and important effect of prestige - its significance for medical specialties and diseases.

They examined literature from 1950 to 2005 regarding the effects of prestige on medicinal practices. They discovered that active, specialized, biomedical, and high-technological types of medicine on organs in the upper part of the bodies of young and middle-aged people were accorded high levels of prestige, while medicine with opposite characteristics had low levels of prestige. They have concluded that such differences in prestige bear consequences for actual priority setting in healthcare systems. They discovered that surgery counts as the most prestigious specialty, while psychiatry is the less prestigious. In addition, doctors tend to rank practices that require more time to master as more prestigious, while other procedures that are considered *easier* are less prestigious. Simply put, they found that the advance in technology and research was in accordance with the prestige rankings. This means that there may be very important practices that are neglected due to the prestige bias.

As we seen so far, prestige can explain many behaviors and evolution of cultural traits. It is a tool to cheaply estimate and acquire knowledge, which helps an individual to survive and breed. However, it is not always the case, and there could be negative repercussions to this bias, such as invasion of maladaptive traits.

Takahashi and Ihara (2019) mention that social learning not only takes the form of random copying of other individuals, but also involves learners' choice of what to learn and from whom to learn. They suggest a best-of-K model where an individual samples k role-models and choose the one he deems most "successful". They mentioned that a previous mathematical analysis has shown that it may sometimes result in maladaptive cultural evolution when the payoffs associated with cultural variants vary stochastically. In such a case, learners may be selectively disfavored and in the long run replaced by unbiased learners, who simply copy someone chosen at random. They develop new mathematical models that are simpler and mathematically tractable. They found that best-of-k learning, unlike unbiased learning, can facilitate the invasion of an on average inferior variant that sometimes gives a very high payoff. Our model, which includes influence bias, is consistent with that claim. When a maladaptive trait is "piggybacking" a role-model with high influence, said trait could spread in the population, as mentioned. In addition, they show that best-of-k learning can be stable against invasion by unbiased learning if social learning is sometimes combined with individual learning. Our model is based on copying based learning only, but it could be interesting to combine it with individual learning and see how it affects the dynamics of the population.

We discussed prestige in depth, and provided several proofs for its existence in nature, humans, and even medicine. We saw it could aid invasion of maladaptive traits, or neglect of important medicinal specialities. But, it can also accelerate reversal of harmful traditions. Harmful traditions can be child marriage, open defecation, and domestic violence, to name a few. Efferson et al. (2020) suggest a mechanism called *spillover*. By their definition, a spillover is when an intervention affect a large enough group in a target population, so that others not included in the intervention starts changing their behavior as well. In their research, they found that there are individuals who act as *agents*, who are often looked upon, and therefore they are ideal targets for interventions. This is the same concept as our role-models, where a more prestigious individual will be copied more, therefore spreading his trait wider in the population. Their research support therefore in our assumption that there are social

biases, conformist influence specifically. They also suggest a way to use this phenomena to change existing traditions in a population. It is very clear however, that just as it can be used to end harmful traditions, the same agents could be used for any negative way that comes to mind.

Up until now, we showed that cultural transmission is a process that manifests in many species, with emphasis on humans. We also displayed similarities between this process and genetic transmission, while presenting differences between them, specifically selective biases such as influence and prestige. We also presented examples of good and bad usages of such biases. All of these are mainly presented as a parallel process to the natural selection process in regards to physical anatomy, or at least have an indirect effect on it. Muthukrishna and Henrich (2016) offer a take on prestige as a factor of human physical evolution directly. They present a concept called *cultural brains* | brains that evolved primarily for the acquisition of adaptive knowledge. They build on the hypothesis of Dunbar (2009) that shows that larger, more complex brains can store and manage more information and in turn, this information can support the costs of a larger brain. Muthukrishna and Henrich (2016) built a model that predicts a strong relationship between brain size and group size, because group size also provides access to more adaptive knowledge. They later present their *cumulative cultural brain* hypothesis, an approach which proposes that human brains have evolved with an ability and proclivity for selective, high fidelity social learning. As part of this process, there are a variety of strategies and biases that have evolved to hone in on the most adaptive knowledge. These strategies and biases include direct and indirect cues of the popularity of cultural traits (e.g. success and prestige biases). In short, they suggest that some of the reasons for the extreme increase in brain size in humans, are the ability to "cheaply" acquire adaptive knowledge, i.e transmission biases, such as prestige.

Further work. Our model can be expanded in various ways: observing the effects of different bias functions, including errors when estimating influence or costs when copying from an influential role-model, and inclusion of multiple, rather than a single, optimal phenotype.

Acknowledgements

We thank Marcus Feldman, Martin Pontz, and Tal Simon for discussions and comments. This work was supported in part by the Israel Science Foundation (YR 552/19), Minerva Stiftung Center for Lab Evolution (YR), and John Templeton Foundation (YR 61809).

References

Aljadeff, N., Giraldeau, L.-A., and Lotem, A. (2020). Competitive advantage of rare behaviours induces adaptive diversity rather than social conformity in skill learning. *Proceedings of the Royal Society B*, 287(1933):20201259.

Anagnostopoulos, A., Kumar, R., and Mahdian, M. (2008). Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 7–15.

Andersson, M. B. (1994). *Sexual selection*. Princeton University Press.

Battesti, M., Moreno, C., Joly, D., and Mery, F. (2012). Spread of social information and dynamics of social transmission within drosophila groups. *Current Biology*, 22(4):309 – 313.

Boyd, R. and Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.

Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press.

Chudek, M., Heller, S., Birch, S., and Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, 33(1):46–56.

Cohen, D., Lewin-Epstein, O., Feldman, M. W., and Ram, Y. (2021). Non-vertical cultural transmission, assortment and the evolution of cooperation. *Proceedings of the Royal Society B*, 288(1951):20203162.

Creanza, N., Kolodny, O., and Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *Proceedings of the National Academy of Sciences*, 114(30):7782–7789.

Denton, K. K., Ram, Y., Liberman, U., and Feldman, M. W. (2020a). Cultural evolution of conformity and anticonformity. *Proceedings of the National Academy of Sciences*.

Denton, K. K., Ram, Y., Liberman, U., and Feldman, M. W. (2020b). Cultural evolution of conformity and anticonformity. *Proceedings of the National Academy of Sciences*, 117(24):13603–13614.

Diga, M. and Kelleher, T. (2009). Social media use, perceptions of decision-making power, and public relations roles. *Public Relations Review*, 35(4):440–442.

Drezner, Z. and Farnum, N. (1993). A generalized binomial distribution. *Communications in Statistics - Theory and Methods*, 22(11):3051–3063.

Dunbar, R. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36(5):562–572. PMID: 19575315.

Durrett, R. (1999). *Essentials of stochastic processes*, volume 1. Springer.

Durrett, R. (2008). *Probability models for DNA sequence evolution*, volume 2. Springer.

Efferson, C., Vogt, S., and Fehr, E. (2020). The promise and the peril of using social influence to reverse harmful traditions. *Nature human behaviour*, 4(1):55–68.

Eickbush, M. T., Young, J. M., and Zanders, S. E. (2019). Killer meiotic drive and dynamic evolution of the wtf gene family. *Molecular biology and evolution*, 36(6):1201–1214.

Erhart, E. M. and Overdorff, D. J. (1999). Female coordination of group travel in wild propithecus and eulemur. *International Journal of Primatology*, 20(6):927–940.

Fogarty, L., Wakano, J. Y., Feldman, M. W., and Aoki, K. (2017). The driving forces of cultural complexity. *Human Nature*, 28(1):39–52.

Frigyik, B. A., Kapila, A., and Gupta, M. R. (2010). Introduction to the dirichlet distribution and related processes. *Department of Electrical Engineering, University of Washignton, UWEETR-2010-0006*, (0006):1–27.

Henrich, J. and Broesch, J. (2011). On the nature of cultural transmission networks: evidence from fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1139–1148.

Henrich, J. and Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and human behavior*, 22(3):165–196.

Horner, V., Proctor, D., Bonnie, K. E., Whiten, A., and de Waal, F. B. (2010). Prestige affects cultural learning in chimpanzees. *PloS one*, 5(5):e10625.

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in science & engineering*, 9(03):90–95.

Kendal, R., Hopper, L. M., Whiten, A., Brosnan, S. F., Lambeth, S. P., Schapiro, S. J., and Hoppitt, W. (2015). Chimpanzees copy dominant and knowledgeable individuals: implications for cultural diversity. *Evolution and Human Behavior*, 36(1):65–72.

Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713.

Kimura, M. and Ohta, T. (1969). The average number of generations until fixation of a mutant gene in a finite population. *Genetics*, 61(3):763.

King, A. J. and Cowlishaw, G. (2009). Leaders, followers, and group decision-making. *Communicative & Integrative Biology*, 2(2):147–150.

Lee, W., Xiong, L., and Hu, C. (2012). The effect of facebook users' arousal and valence on intention to go to the festival: Applying an extension of the technology acceptance model. *International Journal of Hospitality Management*, 31(3):819–827.

McComb, K., Moss, C., Durant, S. M., Baker, L., and Sayialel, S. (2001). Matriarchs as repositories of social knowledge in african elephants. *Science*, 292(5516):491–494.

Molleman, L., Pen, I., and Weissing, F. J. (2013). Effects of conformism on the cultural evolution of social behaviour. *PloS one*, 8(7):e68153.

Muthukrishna, M. and Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1690):20150192.

Norredam, M. and Album, D. (2007). Prestige and its significance for medical specialties and diseases. *Scandinavian journal of public health*, 35(6):655–661.

O'Brien, M. J., Lyman, R. L., Mesoudi, A., and VanPool, T. L. (2010). Cultural traits as units of analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1559):3797–3806.

Otto, S. P. and Whitlock, M. C. (2006). Fixation probabilities and times. *Encyclopedia of Life Sciences*.

Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J., and Jia, W. (2018). Influence analysis in social networks: A survey. *Journal of Network and Computer Applications*, 106:17–32.

Ram, Y., Liberman, U., and Feldman, M. W. (2018). Evolution of vertical and oblique transmission under fluctuating selection. *Proceedings of the National Academy of Sciences*, 115(6):E1174–E1183.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T., and Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science*, 328(5975):208–213.

Takahashi, T. and Ihara, Y. (2019). Cultural and evolutionary dynamics with best-of-k learning when payoffs are uncertain. *Theoretical Population Biology*, 128:27–38.

Van Der Walt, S., Colbert, S. C., and Varoquaux, G. (2011). The numpy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2):22–30.

Van Rossum, G. et al. (2007). Python programming language. In *USENIX annual technical conference*, volume 41, pages 1–36.

Van Vugt, M. and Smith, J. E. (2019). A dual model of leadership and hierarchy: Evolutionary synthesis. *Trends in Cognitive Sciences*, 23(11):952–967.

Whitehead, H. (2017). Gene–culture coevolution in whales and dolphins. *Proceedings of the National Academy of Sciences*, 114(30):7814–7821.