

WHEN DOES SELECTION FAVOR LEARNING FROM THE OLD? SOCIAL LEARNING IN AGE-STRUCTURED POPULATIONS

DOMINIK DEFFNER AND RICHARD MCELREATH

ABSTRACT. Culture and demography jointly facilitate flexible human adaptation, yet it still remains unclear how social learning operates in populations with age structure. Specifically, how do demographic processes affect the adaptive value of culture, cultural adaptation and population growth and when does selection favor copying the behavior of older vs. younger individuals? Here, we develop and analyze a mathematical model of the evolution of social learning in a population with different age classes. We find that adding age structure alone does not resolve Rogers’ paradox, i.e. the finding that social learning can evolve without increasing population fitness. Cultural transmission in combination with demographic filtering, however, can lead to much higher adaptation levels. This is because by increasing proportions of adaptive behavior in older age classes, demographic filtering constitutes an additional adaptive force that social learners can benefit from. Moreover, older age classes tend to have higher proportions of adaptive behavior when the environment is relatively stable and adaptive behavior is hard to acquire but confers large survival advantages. Through individual-based simulations comparing temporal and spatial variability in the environment, we find a “copy older over younger models”-strategy only evolves readily when social learning is erroneous. The opposite “copy the younger”-strategy is adaptive when the environment fluctuates frequently but still maintains large proportions of social learners. Our results demonstrate that age structure can substantially alter cultural dynamics and should be addressed in further theoretical and empirical work.

November 20, 2020

1. INTRODUCTION

When should individuals copy the behavior and norms of older generations? In her 1970 book “Culture and Commitment”, anthropologist Margaret Mead considered relations between generations based on the prevalent rate of social change [1]. When this rate is slow, juveniles grow up in an environment that resembles that of their parents and older generations serve as valued models and authorities. Under such circumstances being old may signal adaptive behavior as one has managed to survive to old age. In so called gerontocratic societies, which include the Ancient Greek city state of Sparta and contemporary East-African pastoralists such as the Kenyan *Samburu* [2; 3], social stratification is predominantly based on age-classes and opinions of older individuals tend to be highly valued. When societal change is more rapid, on the other hand, being old may predict being out of date as relevant environmental conditions may have changed since older generations have acquired their norms and behaviors. Starting in the 1960s, attitudes associated with older generations are often seen as backward and outdated in Western societies (“OK Boomer” is a recent example) and young individuals regularly attend to their peers instead. Which demographic and cultural factors could explain these opposing views on the value of information provided by older individuals?

Children across societies need to learn many essential skills to become competent adult members of their communities. The transmission and gradual modification of cultural information over generations is often considered key to human evolutionary success as it allows us to flexibly adjust to vastly different environments through locally adapted tools, beliefs and institutions [4; 5; 6]. Past theoretical work has investigated the conditions that favor the evolution of social learning and the way individuals should combine individual and social information strategically [7; 8; 9].

One starting point was Alan Rogers’ equally simple and elegant model that demonstrates how pure social learning does not increase mean population fitness because its adaptive value is strongly frequency-dependent [10]. With few other social learners, chances are high to copy an individual learner, who has learned from direct interactions with the environment, providing adaptive information without costly trial-and-error learning. As the proportion of social learners increases, fewer and fewer individuals track the state of the environment, which can result in maladaptive information cascades [11]. At equilibrium, there tends to be a mix of individual and social learners that have the same mean fitness as a population entirely comprised of individual learners, an observation termed “Rogers’ paradox” in the literature.

This and most following models omit age structure and study the evolution of learning in populations with discrete, non-overlapping generations. Individuals are assumed to be born, learn from the previous generation, reproduce and then all die at the same time. Call this the “annual grass model” of culture. Social learning and life history, however, interact in human adaptation and we need further modeling work that incorporates age structure and different population dynamics [12]. Age structure is an undeniable feature of human (and other animal) populations, can have profound consequences for evolutionary dynamics [13; 14] and is also expected to shape the informational environment learning is responding to. Thus, to address when juveniles should learn from older generations but also how culture facilitates human adaptation in general, we need a formal theory of the evolution of social learning in age-structured populations.

The goal in this paper is to develop the simplest possible model of the classic adaptive value of culture, but in an age-structured population. So call this, “Rogers’ model with age classes”. As such, it isn’t meant to represent any specific organism. Rather it represents the structure of an argument. How do demographic and cultural forces affect adaptation and population growth? And what are the minimal conditions for such forces to make it worthwhile to copy older individuals? We first formulate a mathematical model of the population and learning dynamics and demonstrate that adding age does not resolve Rogers’ paradox but can lead to more social learning and higher adaptation levels. We then solve for the conditions that favor learning from older vs. younger individuals. Finally, we confirm and extend the analytical results through individual-based simulations and investigate how age-biased social learning evolves in temporally and spatially varying environments.

2. THE MODEL

Consider a large population of perennial organisms that live in a temporally varying environment. The environment can take on one of a very large number of states. Each state has a corresponding adaptive response. An adaptive response may increase both survival and fertility of adults. The environment may change with probability u each time step, rendering all previously adaptive behavior non-adaptive.

Individuals are born as *juveniles* who cannot yet reproduce, but may learn about the environment. After one time step, a juvenile who survives transitions to adulthood. Adults do not learn, but do reproduce. Let s_j be the probability of survival of a juvenile or an adult who practices response j , where $j = 1$ indicates adaptive behavior and $j = 0$ indicates non-adaptive behavior. Let b_j be the fertility of an adult with the adaptive response. Juveniles have zero fertility always. They must survive once to have any chance of producing descendants.

The organism's challenge is to find the currently adaptive response. Juveniles may learn either socially or individually. Social learners sample two models and then may choose among them based upon relative age. Individual learners instead reduce their probability of survival to adulthood by a factor $c < 1$ for a chance z of innovating an adaptive response to current conditions.

2.1. Recursions. The state of the population is given by the state variables n_{ij} , where i is an age class and $j = 1$ indicates an individual with adaptive response. The number of juveniles n_{10} (all are born non-adapted) is regenerated each time period by fertility from all adult age classes:

$$(1) \quad n'_{11} = p \sum_{i=2}^{\infty} \sum_{j=0}^1 n_{ij} b_j$$

$$(2) \quad n'_{10} = (1 - p) \sum_{i=2}^{\infty} \sum_{j=0}^1 n_{ij} b_j$$

where p is the probability a juvenile acquires currently adaptive response through learning. We define p as a function of heritable strategy in the next section. The number in each adult age class $i > 1$ is given instead by:

$$(3) \quad n'_{i0} = s_0 n_{i-1,0} + u s_1 n_{i-1,1}$$

$$(4) \quad n'_{i1} = (1 - u) s_1 n_{i-1,1}$$

Since adults do not learn in this model, these recursions are very simple. However, when the environment changes with probability u , all adapted adults are rendered forever non-adapted. There is a special case for $i = 2$ that includes the cost of innovation by juveniles c :

$$(5) \quad n'_{20} = (s_0 n_{1,0} + u s_1 n_{1,1})(1 - c\pi)$$

$$(6) \quad n'_{21} = ((1 - u) s_1 n_{1,1})(1 - c\pi)$$

The symbol π gives the probability of individual learning and is further defined in the next section.

2.2. Learning. Let p be the probability a juvenile acquires currently adaptive behavior. We define this as a function of a strategy vector $\{\pi, \phi\}$, where π is the probability of individual learning and ϕ determines the direction and strength of any bias for age of model to socially learn from. Specifically:

$$(7) \quad p(\pi, \phi) = \pi z + (1 - \pi)Q(\phi)$$

where z is the probability an individual learner acquires adaptive behavior. Q is a function for the probability of acquiring adaptive response by social learning. We define ϕ as the proportional odds of copying the older of two individuals. Let R be the probability of copying the older of two individuals. This implies that the probability of copying the older individual is given by solving this expression for R :

$$(8) \quad \phi = \frac{R}{1 - R}$$

yielding $R = \phi/(\phi + 1)$. To compute Q , we need to average over the possible pairs of social models. Let q_i be the proportion of individuals in age class i with adaptive behavior. Let a_i be the proportion of adults ($i > 1$) in age class i . Then the probability a juvenile acquires adaptive behavior through social learning is:

$$(9) \quad Q(\phi) = \sum_{i=2}^{\infty} a_i^2 q_i + \sum_{i=3}^{\infty} \sum_{j=2}^{i-1} 2a_i a_j \left(q_i \frac{\phi}{\phi+1} + q_j \frac{1}{\phi+1} \right)$$

90 In the above expression, the first summation is all pairs with tied ages. In these pairs, there is no age asymmetry. So the learner imitates at random and acquires adaptive response when a random individual of age class i has an adaptive response. In the second summation, the double one, we sum over asymmetric pairs where an individual in age class i is older than one in age class j . There is a probability $2a_i a_j$ of such a pair, allowing for both orderings, and then the learner acquires adaptive response according to which model is imitated and whether an individual of that age class has an
 95 adaptive response. When $\phi = 1$, the above simplifies to random copying of any adult:

$$(10) \quad Q(1) = \sum_{i=2}^{\infty} a_i q_i$$

3. ANALYSIS

3.1. Does adding age structure resolve Rogers paradox? As a first step, we investigate whether age structure resolves ‘‘Rogers’ paradox’’, the finding that social learning can invade a
 100 population of individual learners without increasing mean population fitness. We compare a baseline model with discrete, non-overlapping generations (i.e. the original Rogers model) to the age-structured learning model described above. For this analysis, we assume individual learning always produces adaptive behavior (i.e. $z = 1$) and social learners copy adults randomly (i.e. $\phi = 1$). Relaxing these assumptions does not change the results. Fig. 1 shows proportions of social learners (Fig. 1A,B) and fitness (i.e. lineage growth rate at equilibrium, λ , Fig. 1C,D) for the original Rogers
 105 model and the age-structured version. In both models, social learning evolves more readily as individual learning is more costly and the environment changes less frequently. Lines in Fig. 1C,D represent mean fitness in populations with social learners, points give corresponding values in populations of only individual learners. Fitness in both models is determined by the fitness of individual
 110 learners and does not change with the introduction of social learners. Age structure, thus, does not resolve Rogers paradox. As a direct consequence of age structure and the assumption that only juveniles learn, fitness in the age-structured model decreases as the environment becomes more variable: With discrete generations, the whole population gets replaced each time step, so that fitness of individual learners does not depend on environmental stability. With age structure, individual
 115 learners can lose adaptive behavior during their lifetime through environmental change, resulting in the decline in population fitness.

3.2. Demographic filtering and its interplay with social learning. Next, we investigate the interplay between social learning and demography and their joint effect on adaptation. With viability selection, i.e. differential chances of survival, the demographic filtering of adaptive behavior in older
 120 age classes might constitute a second adaptive force in addition to individual learning that social learners could benefit from.

Individual Learning only. We start with a population of only individual learners and compare the behavior of the model with a stochastically changing environment under pure fecundity and pure viability selection. Fig. 2A illustrates the effect of demographic filtering on the distribution of
 125 adaptive behavior across age classes. If there is no survival advantage to adaptive behavior (solid line), older age classes will always have lower proportions of adaptive behavior as environmental change periodically renders their behavior out of date. If being adapted confers survival benefits (dashed line), only those possessing adaptive behavior will be likely to survive to certain ages making

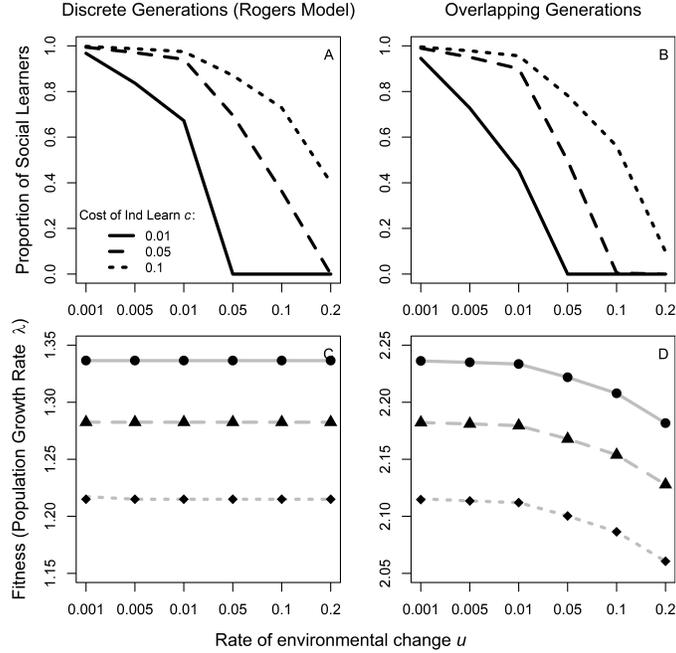


Figure 1. Proportions of social learners (A,B) and mean population fitness (C,D, lineage growth rate at equilibrium, λ) conditional on the rate of environmental change u and the cost of individual learning c . Results for the original Rogers model with discrete, non-overlapping generations are shown on the left, results for the age-structured model with overlapping generations on the right. Lines in plots C and D represent populations with social learners, points give corresponding values for populations of only individual learners. Other parameter values were: $z = 1$, $\phi = 1$, $b_0 = 1$, $b_1 = 1.5$, $s_0 = 0.9$, $s_1 = 0.9$.

age a predictor of adaptive behavior. The probability that an adult in age class i possesses adaptive
 130 behavior can be calculated as follows:

$$(11) \quad P_{Adapt,i} = (1-u)^{i-1} \frac{s_1^{i-1} z}{s_1^{i-1} z + s_0^{i-1} (1-z)}$$

A first prerequisite is that the environment has not changed since an individual has learned which
 is given by the first part of the equation. This term is multiplied by the relative probability adults
 have acquired adaptive behavior as juveniles and survived up to age i . Fig. S1 in the ESM plots
 135 resulting curves for different values of u , $\sigma = s_0/s_1$, and z . Expectedly, more rapid changes in
 the environment result in lower values and earlier peaks of adaptive behavior. Stronger viability
 selection leads to earlier peaks in adaptive behavior, as even surviving to younger ages indicates
 being adapted when selection acts strongly on chances of survival. Resulting from such filtering,
 populations under viability selection reach higher adaptation levels compared to fecundity selection,
 but still plateau at a value determined by the success rate of individual learning as adaptive filtering
 140 cannot affect juvenile learning (Fig. 2C).

Individual and Social Learning. Now we also let social learners evolve. Demographic filtering
 increases the adaptive value of culture and leads to substantially higher frequencies of social learners
 in the population (means of 98% and 38% for viability and fecundity selection, respectively). By
 increasing proportions of adaptive behavior in older age classes, demographic filtering improves the
 145 quality of social information and thereby results in more copying among juveniles. As social learn-
 ing effectively drains accumulated adaptive information from older age classes, it lowers the peak

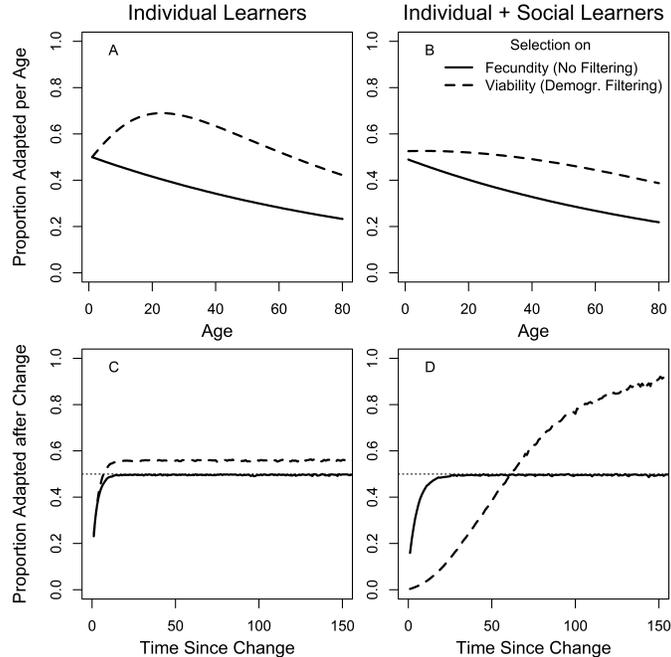


Figure 2. Proportions of social learners per age (A,B) and proportion of adapted individuals per time after an environmental change (C,D). Results for populations with only individual learners are shown on the left, results for populations comprising both individual and social learners on the right. Solid lines represent dynamics under fecundity selection (i.e. without demographic filtering), dashed lines represent dynamics under viability selection (i.e. with demographic filtering). Results are shown for $z = 0.5$ which is also indicated by the dotted horizontal line in the bottom. Other parameter values were chosen to keep mean fitness constant at $\lambda \approx 1.20$ across different scenarios: $u = 0.01, c = 0.01, \phi = 1, b_0 = 0.35, b_1 = 0.5, s_0 = 0.85, s_1 = 0.93$.

in adaptive behavior (Fig. 2B). After an environmental shock (Fig. 2D), populations under viability selection take relatively long to recover because few individual learners track the state of the environment. Without filtering, populations recover faster but are limited to the adaptation levels individual learners can obtain, as innovation is the only adaptive force in the system. With viability selection, adaptation levels exceed those of individual learners and increase until the whole population is adapted. Combining demographic filtering and social learning, the accumulated benefits of adaptive filtering can feed back on juveniles resulting in an adaptive interplay between cultural and demographic forces that enables much higher adaptation levels.

3.3. When are older individuals better adapted? After considering the general impacts of age structure, we now turn to our primary question: When will selection favor a positive ($\phi > 1$) or negative ($\phi < 1$) age bias? To address this question, we require expressions for the q_i state variables, the proportion of each age class i with adaptive behavior. Specifically, we would like to know, under a separation of genetic and demographic time scales, the demographic environment that selection of ϕ responds to. This means we seek expressions for the steady state \hat{q}_i values corresponding to a stable age distribution.

Since this model contains no population regulation, the population will either increase to infinity or decline to zero (see [12] for effect of population regulation on social learning). We are interested in the former case. In that case, the frequency of individuals in each age-and-behavior class ij will eventually stabilize, even though the number of individuals in the population will continue to

grow. Therefore we analyze the frequencies $\hat{f}_{ij} = n_{ij}/N$ that define the stable age-and-behavior distribution. These frequencies are defined by the recurrence equations in the previous section, normalized by N . This normalization introduces a population growth adjustment. To see this, consider the simplest case:

$$(12) \quad \hat{f}_{i,1} = \frac{n'_{i,1}}{N'} = \frac{1}{N'}(1-u)s_1 n_{i-1,1}$$

170 Since we consider now the stable age-class distribution, we divide both sides by N so we relate stable frequencies:

$$(13) \quad \frac{\hat{f}_{i,1}}{N} = \frac{1}{N'}(1-u)s_1 \frac{n_{i-1,1}}{N}$$

$$(14) \quad = \frac{1}{N'}(1-u)s_1 \hat{f}_{i-1,1}$$

Finally, move the N^{-1} on the left to the right and note that $N'/N = \lambda$, the stable population growth rate:

$$(15) \quad \hat{f}_{i,1} = \frac{N}{N'}(1-u)s_1 \hat{f}_{i-1,1} = \frac{1}{\lambda}(1-u)s_1 \hat{f}_{i-1,1}$$

The same approach yields the recurrence for non-adapted individuals at age i :

$$(16) \quad \hat{f}_{i,0} = \frac{1}{\lambda} \left(s_0 \hat{f}_{i-1,0} + u s_1 \hat{f}_{i-1,1} \right)$$

175 Note that when the population grows slowly, due to density-dependent regulation of fertility, for example, then $\lambda \approx 1$. This is a very common assumption in much life history analysis. But we will attempt to keep λ general, in case some further insights about the value of learning arise from the generality. Since much of human cultural evolution was spent in expanding populations, this generality is not irrelevant. These equations can be solved explicitly, yielding formulas in terms of
180 only initial conditions and parameters. For $i > 1$:

$$(17) \quad \hat{f}_{i,0} = \hat{f}_{1,0} \left(\frac{s_0}{\lambda} \right)^{i-1} + \hat{f}_{1,1} u s_1 \frac{\left(\frac{s_0}{\lambda} \right)^{i-1} - s_1^{i-1} (1-u)^{i-1}}{s_0 - \lambda s_1 (1-u)}$$

$$(18) \quad \hat{f}_{i,1} = (\lambda^{-1}(1-u)s_1)^{i-1} \hat{f}_{1,1}$$

These were solved using the honorable *Ansatz* method, but Mathematica confirms the solutions via `RSolve`. The stable proportion of individuals at age i with adaptive behavior is defined as:

$$(19) \quad \hat{q}_i = \frac{\hat{f}_{i,1}}{\hat{f}_{i,1} + \hat{f}_{i,0}}$$

This expression depends implicitly upon \hat{p} , the stable probability that a juvenile acquires adaptive behavior. But \hat{p} depends in turn upon the \hat{q}_i values. This does not stop our inferential progress
185 (yet), because we do not need to determine \hat{p} at steady state. We just need to use it implicitly in our solution. Specifically, note that the expressions for $n'_{1,1}$ and $n'_{1,0}$ in the previous section imply:

$$(20) \quad \hat{p} = \frac{\hat{f}_{1,1}}{\hat{f}_{1,1} + \hat{f}_{1,0}}$$

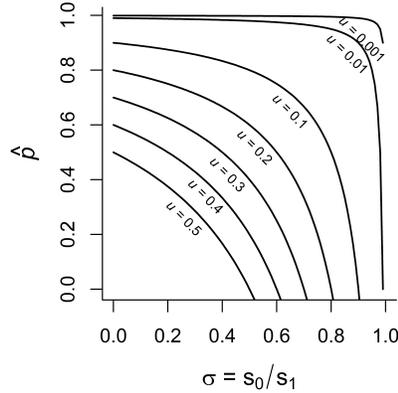


Figure 3. Threshold value of \hat{p} , as defined by expression 24. Each curve shows the threshold \hat{p} for a different value of u . Areas underneath the curves indicate parameter combinations with higher proportion of adaptive behavior in older age classes.

Equivalently:

$$(21) \quad \frac{\hat{p}}{1 - \hat{p}} = \frac{\hat{f}_{1,1}}{\hat{f}_{1,0}}$$

This allows us to substitute $\hat{f}_{1,1} = \hat{f}_{1,0} \hat{p} / (1 - \hat{p})$, yielding (for $\lambda \approx 1$):

$$(22) \quad \hat{q}_i = \frac{ps_0s_1^i(s_0 - s_1(1 - u))(1 - u)^i}{s_0^i s_1(1 - u)(us_1 - (1 - p)(s_1 - s_0)) - ps_0(s_1 - s_0)s_1^i(1 - u)^i}$$

We also have the above expression for general λ , but it is very messy.

Now we can ask when $\hat{q}_i > \hat{q}_{i-1}$, yielding the condition (again for $\lambda \approx 1$):

$$(23) \quad \hat{p} < 1 - u \frac{s_1}{s_1 - s_0}$$

Or equivalently, letting $\sigma = s_0/s_1$:

$$(24) \quad \hat{p} < 1 - \frac{u}{1 - \sigma}$$

190 This is the condition for older age classes to have larger proportions of adaptive behavior and therefore the precondition for selection to favor $\phi > 1$. Neither copying the old nor the young is always beneficial. Rather, whether or not older individuals are better adapted depends upon a balance of forces. The force of environmental change u is to reduce the proportion of adaptive behavior in older age classes—as u increases, the condition above becomes increasingly difficult to
 195 satisfy. The force of a survival advantage to adaptive behavior, in contrast, is to increase adaptive behavior in older age classes. As σ decreases, corresponding to a larger advantage to adaptive behavior, the condition is easier to satisfy.

The condition above confirms the intuition behind the model. This is useful, because intuition is often—if not usually—wrong. But it also provides quantitative guidance on the relative strength
 200 of these forces. In Figure 3, we plot the threshold value of \hat{p} across all possible values of σ , for different values of u . The region below each curve corresponds to combinations of \hat{p} and σ that lead to older age classes having higher proportions of adaptive behavior. When environments change

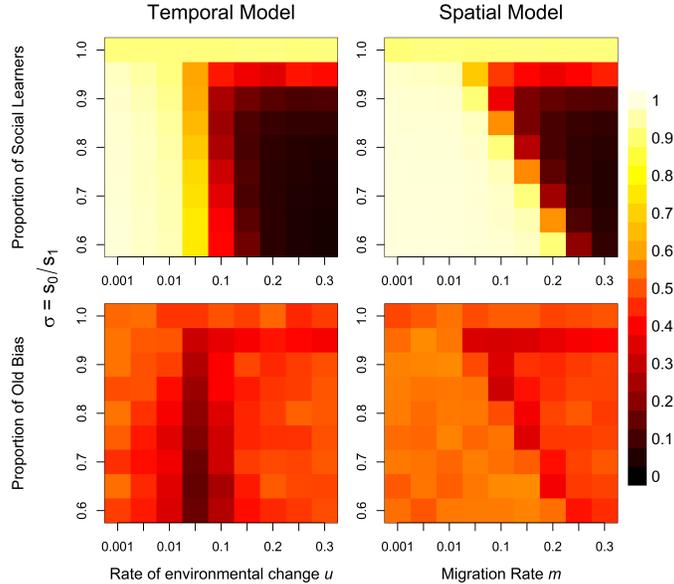


Figure 4. Proportions of social learners (top) and old bias (bottom) for the temporal (left) and spatial (right) individual-based model. Lighter colors indicate higher proportions, darker colors lower proportions. Results are shown for $c = 0.05$, $z = 0.5$, $s_1 = 0.9$, $\mu = 0.005$.

very rapidly, the adaptive behavior must be hard to acquire for a juvenile and must provide large survival benefits for older individuals to be better adapted. On the other hand, when environments
 205 are very stable, the above condition is almost always fulfilled, so that juveniles should copy older individuals irrespective of other factors.

3.4. The evolution of age-biased social learning in temporally and spatially varying environments. Juveniles should be more likely to copy the older of two individual if the environment is relatively stable, adaptive behavior is hard to acquire and there is a large survival advantage
 210 to adaptive behavior. Which of these forces is strongest depends not only on their parameters, $\sigma = s_0/s_1$ and u , but also on the frequency of adaptive behavior in the population. To confirm and extend the analytical results, we construct and analyze a stochastic individual-based version of the model and explore how readily age-biased social learning evolves in both temporally and spatially
 215 varying environments (see ESM for simulation details).

Fig. 4 shows proportions of social learners and those biased towards older individuals (“Old Bias”) for both the temporal (left) and spatial model (right). Lighter colors indicate higher proportions. In the temporal model, social learning evolves when the environment is relatively stable or there is very little advantage to possessing adaptive behavior. Intriguingly, there are no regions in parameter space that clearly favor copying the old. Instead, when the environment changes frequently (but
 220 not rapidly enough to favor large numbers of individual learners), evolution favors a pronounced “copy-the-young” bias. In the spatial model, high proportions of social learning can also occur with much more unstable environments (i.e. high migration rates) in simulations with strong viability selection. This is because selection weeps out non-adaptive variation brought in by migrants, such that most group members possess adaptive behavior. Here, selection favors a slight “copy-the-old”
 225 bias in regions with stable environments and/or large advantages to adaptive behavior, as predicted by the analytical model.

While u and σ are parameters that we can directly manipulate, \hat{p} , the third variable in expression 24, arises endogenously from different interacting forces within the model. The reason no stronger “copy-the-old” bias can evolve in either model, is that in stable environments with large benefits to adaptive behavior (conditions that would favor old bias), essentially everyone will be adapted, so either social learning strategy will result in adaptive behavior and there is a strong ceiling effect on demographic filtering. So far, we have assumed there is no copying error, i.e. learners can perfectly reproduce the behavior they observe in a model. This is a rather strong assumption and in reality there might be many circumstances where juveniles fail to acquire a trait through social learning. Fig. S2 in the ESM shows the results for simulations where social learning results in the successful adoption of adaptive behavior in only 70% of learning events. By lowering \hat{p} , this slight modification leads to the evolution of large proportions of “copy-the-old” bias in both the temporal and spatial model. When social learning is imperfect, there is room for demographic filtering to increase proportions of adaptive behavior in older age classes making them preferable models.

Finally, we calculate \hat{p} from simulation results and plot it with the proportion of old bias for different values of u and σ . Fig. S3 in the ESM shows that individuals are less likely to copy older individuals when it is relatively easy to acquire adaptive behavior as a juvenile as implied by expression 24. This holds for different combinations of u and σ .

4. DISCUSSION

Culture and demography jointly facilitate flexible human adaptation, yet most previous models decided to leave out demographic complexity and studied the evolution of learning and cultural dynamics in isolation. Making simplifying assumptions is of course important but even relatively basic questions, such as “When should juveniles learn from older individuals?”, cannot be answered or even formally asked unless age structure is taken into account. Here we develop and analyze the age-structured version of Rogers’ model of the adaptive value of culture [10] and ask under which conditions juveniles should preferentially attend to older vs. younger individuals. As such, our model also omits a good deal of real-world complexity. The goal of theoretical modeling is not to approximate the natural world as closely as possible but to isolate the fundamental structure of an overly complex reality and apply formal logic to deduce the basic causal forces [15; 16; 17].

First, we found that adding age structure does not resolve Rogers’ paradox. In populations with age classes pure social learning still does not increase mean population fitness compared to individual learners alone. Theory has shown that for culture to increase population fitness, it must make individual learning either more accurate or less costly [18]. [19], for instance, showed that a strategy of “critical social learning”, where individuals switch to individual learning if social learning proves unsatisfactory, outcompetes pure individual and social learning strategies (similarly for the reverse “conditional social leaning”). In the present model, age structure can increase the adaptive value of culture, such that a population can maintain higher proportions of copying, but social learning cannot increase the fitness of individual learners. As both types of learners must per definition have the same fitness at equilibrium, social learning cannot increase mean population fitness. In a stochastic model with spatial population structure, [20] found that under some circumstances pure social learning can increase average fitness. They also show that spatial structure can introduce a new paradox with social learning spreading even when it decreases average fitness below that of individual learners. Individuals in this model occupied different cells within a square toroidal environment with both dispersal and learning occurring locally. This creates an “edge effect” where social learners can spread because they have greater fitness in contact zones between genotypes (due to higher chances to copy individual learners), even though their average fitness is lower. Future work should study the effects of different kinds of population structure and their interaction with each other and cultural dynamics.

275 Even though social learning did not increase mean population fitness, in combination with de-
mographic filtering it enabled the population to reach much higher adaptation levels. Demographic
filtering increases proportions of adaptive behavior in older age classes and thus constitutes a sec-
ond adaptive force in addition to individual learning. While this source of adaptive information is
unavailable to populations of pure individual learners, social learners can utilize it by copying older
generations. Thereby, they can spread to higher frequencies and simultaneously increase adaptation
280 levels compared to situations without such filtering (i.e. pure fecundity selection). These higher
adaptation levels are counteracted by longer recovery periods after environmental changes due to
less individual learners tracking the environment. Previous work has investigated how social learning
and culture help organisms adapt to changing environments, but cultural adaptation is intimately
related to the way organisms grow, reproduce and die [21; 22]. Humans show highly developed
285 abilities to learn from others, but we also exhibit a prolonged childhood and juvenile period, shorter
intervals between births, and a significant post-reproductive lifespan. To explain how these conspic-
uous features might have coevolved with our unique reliance on culture, we need more modeling
work on the adaptive interplay between social learning and different life history dynamics [12].

Our primary question was when natural selection would favor learning from older vs. younger
290 individuals. This question has attracted considerable empirical attention [23], yet we still lack a
principled theoretical framework to unify existing findings and generate new predictions. In devel-
opmental psychology, 15-month-olds were reported to be more likely to copy behaviors performed
by an adult versus a two-year-old child [24] and three- and four-year-olds preferentially attended to
information provided by an adult over a child in a object labelling task [25]. Studies also showed
295 that children faithfully copy both relevant and irrelevant actions demonstrated by adult models but
only relevant actions demonstrated by children [26; 27]. Research in anthropology suggests that
hunter-gatherer social learning is primarily vertical under age 5 and oblique and horizontal between
the ages of 6 and 12 [28]. Taken together, these studies seem to imply that under most circum-
stances children preferentially attend to older generations. Our modeling results show that this can
300 be a good strategy, but only when the environment is relatively stable, adaptive behavior is hard to
acquire and confers large survival advantages. Moreover, we found through simulation that due to
ceiling effects in adaptive behavior a “copy older over younger models” strategy can only reliably
evolve when social learners occasionally fail to copy adaptive behavior. The opposite “copy younger
over older models” strategy can be advantageous when the environment fluctuates frequently but
305 still maintains large proportions of social learners. Similar to our findings, the best strategies in the
second round of the social learning strategies tournament copied both successful and young demon-
strators [29; 30]. Individuals in their simulations could update behavior throughout their lifetime, so
one might expect older individuals would be better adapted due to more learning opportunities. At
each point in time, individuals could either exploit known behaviors to obtain their rewards or learn
310 a new behavior. As a consequence, individuals mostly learned right after birth or after a change in
the environment when they experienced a drop in payoffs. Because births were random with respect
to changes in the environment, juveniles were more likely to acquire adaptive behavior after birth
compared to adults learning when the environment has just changed. This shows that selection can
favor learning from the young as long as there is some exploration-exploitation trade-off between
315 investing in learning and reproduction and not only when learning is restricted to juveniles which
represents the extreme form of this tradeoff.

In this paper, we developed and analyzed an age-structured model of the evolution of social
learning. Even in such simple models, we find intricate interactions between culture and demography
that change our understanding of how cultural organisms learn and adapt. We are just beginning
320 to understand how such joint culture-demography-systems might behave in general and further
bodies of theory including strategic learning and cumulative culture are necessary to untangle the
coevolutionary relationships between demography and culture.

Ethics

This article does not present research with ethical considerations.

325 Data Accessibility

This article does not contain any new data. Simulation and plotting code can be found on GitHub:
<https://github.com/DominikDeffner/Age-structured-Social-Learning>.

Authors' Contributions

DD and RM jointly designed the model. RM derived the analytical solution. DD wrote and con-
 330 ducted numerical and individual-based simulations, as well as wrote the first draft of the manuscript.

Competing interests

We declare we have no competing interests.

Funding

This research has been supported by the Max Planck Society.

335

REFERENCES

- [1] Margaret Mead. *Culture and commitment: A study of the generation gap*. New York: John Wiley, 1970.
- [2] Paul Spencer. *The Samburu: A study in geocentracy*. Routledge, 2004.
- 340 [3] Carolyn K Lesorogol. Bringing norms in: the role of context in experimental dictator games. *Current anthropology*, 48(6):920–926, 2007.
- [4] Joseph Henrich and Richard McElreath. The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 12(3):123–135, 2003.
- [5] Robert Boyd, Peter J Richerson, and Joseph Henrich. The cultural niche: Why social learning
 345 is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10918–10925, 2011.
- [6] Kevin N Laland. *Darwin's unfinished symphony: how culture made the human mind*. Princeton University Press, 2018.
- [7] Luigi Luca Cavalli-Sforza and Marcus W Feldman. *Cultural transmission and evolution: A
 350 quantitative approach*. Number 16. Princeton University Press, 1981.
- [8] Robert Boyd and Peter J Richerson. *Culture and the evolutionary process*. University of Chicago press, 1985.
- [9] Kenichi Aoki and Marcus W Feldman. Evolution of learning strategies in temporally and spatially variable environments: a review of theory. *Theoretical population biology*, 91:3–19,
 355 2014.
- [10] Alan R Rogers. Does biology constrain culture? *American Anthropologist*, 90(4):819–831, 1988.
- [11] Luc-Alain Giraldeau, Thomas J Valone, and Jennifer J Templeton. Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1427):1559–1566, 2002.
- 360 [12] Dominik Deffner and Richard McElreath. The importance of life history and population regulation for the evolution of social learning (in print). *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 2020.
- [13] Hal Caswell. Matrix population models. *Encyclopedia of Environmetrics*, 3, 2006.
- [14] Hal Caswell. *Sensitivity Analysis: Matrix Methods in Demography and Ecology*. Springer, 2019.
- 365 [15] Richard McElreath and Robert Boyd. *Mathematical models of social evolution: A guide for the perplexed*. University of Chicago Press, 2008.
- [16] Sarah P Otto and Troy Day. *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press, 2011.
- [17] Paul E Smaldino. Models are stupid, and we need more of them. *Computational social psy-
 370 chology*, pages 311–331, 2017.

- [18] Robert Boyd and Peter J Richerson. Why does culture increase human adaptability? *Ethology and sociobiology*, 16(2):125–143, 1995.
- [19] Magnus Enquist, Kimmo Eriksson, and Stefano Ghirlanda. Critical social learning: a solution to rogers’s paradox of nonadaptive culture. *American Anthropologist*, 109(4):727–734, 2007.
- 375 [20] Luke Rendell, Laurel Fogarty, and Kevin N Laland. Rogers’paradox recast and resolved: Population structure and the evolution of social learning strategies. *Evolution: International Journal of Organic Evolution*, 64(2):534–548, 2010.
- [21] Hillard S Kaplan, Kim Hill, Jane Lancaster, and A Magdalena Hurtado. A theory of human life history evolution: diet, intelligence, and longevity. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 9(4):156–185, 2000.
- 380 [22] Mauricio González-Forero and Andy Gardner. Inference of ecological and social drivers of human brain-size evolution. *Nature*, 557(7706):554, 2018.
- [23] Lara A Wood, Rachel L Kendal, and Emma G Flynn. Whom do children copy? model-based biases in social learning. *Developmental Review*, 33(4):341–356, 2013.
- 385 [24] Sabine Seehagen and Jane S Herbert. Infant imitation from televised peer and adult models. *Infancy*, 16(2):113–136, 2011.
- [25] Vikram K Jaswal and Leslie A Neely. Adults don’t always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, 2006.
- [26] Emma Flynn. Investigating children as cultural magnets: do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509):3541–3551, 2008.
- 390 [27] Lara A Wood, Rachel L Kendal, and Emma G Flynn. Context-dependent model-based biases in cultural transmission: children’s imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*, 33(4):387–394, 2012.
- 395 [28] Barry S Hewlett, Hillary N Fouts, Adam H Boyette, and Bonnie L Hewlett. Social learning among congo basin hunter–gatherers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1168–1178, 2011.
- [29] James P Ounsley. *The diffusion of culture: computational and statistical models of social learning and cultural transmission*. PhD thesis, University of St Andrews, 2017.
- 400 [30] Luke Rendell, Robert Boyd, Daniel Cownden, Marquist Enquist, Kimmo Eriksson, Marc W Feldman, Laurel Fogarty, Stefano Ghirlanda, Timothy Lillicrap, and Kevin N Laland. Why copy others? insights from the social learning strategies tournament. *Science*, 328(5975):208–213, 2010.

**SUPPLEMENTARY MATERIAL FOR
“WHEN DOES SELECTION FAVOR LEARNING FROM THE OLD?
SOCIAL LEARNING IN AGE-STRUCTURED POPULATIONS”**

DOMINIK DEFFNER¹ & RICHARD MCELREATH¹

¹ *Max Planck Institute for Evolutionary Anthropology,
Department of Human Behavior, Ecology and Culture.*

Appendix (A) describes details of individual-based simulations and (B) shows further results and illustrations.

A. Individual-based simulations

Temporal Variation. In the temporal model we consider a population of fixed size with $N = 1000$ individuals. After birth, each juvenile engages in either individual or social learning according to a dichotomous genotype. Individual learners acquire adaptive behavior with probability z and pay a recruitment cost c . Social learners randomly sample two interaction partners from the adult population and copy either the older or younger one depending on an inherited learning strategy. Individuals acquire adaptive behavior from an adapted interaction partner with probability $1 - e_{SL}$, where e_{SL} represents the degree of copying error. All individuals survive to the next age class with probability s_0 (non-adapted) or s_1 (adapted). We assume asexual, haploid reproduction. All adults have an equal opportunity to give birth to an offspring to fill one of the empty spots in the population (including fecundity selection does not affect the results). Juveniles inherit the genotype governing learning strategies from their parents with a small probability ($\mu = 0.005$) of mutation for both loci. Each time step, there is a probability u that the environment changes. When environmental change occurs, all variants in the population become non-adaptive.

Spatial Variation. For the spatial model, we consider a population of fixed size with $N = 2000$ individuals that are divided into 4 equal sub-populations linked by migration. Life cycle and learning process are identical to the temporal model with the only exception that instead of temporal variation, there is constant migration between different habitats that differ in important respects such that each is characterized by a different adaptive response. Specifically, per time step each adult has a probability m to migrate into another habitat which results in a loss of adaptive behavior. Migration is implemented in a way to keep group sizes constant.

Results for both models are averaged over the last 5000 time steps of 10 independent 7000 time-step simulations per parameter combination.

B. Extended Results

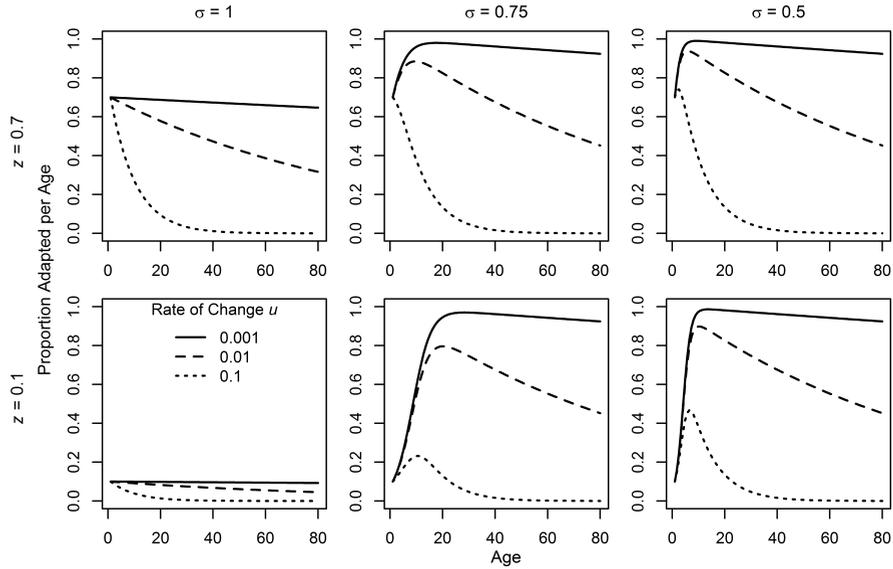


FIGURE S1. Proportion of adaptive behavior conditional on age in population of individual learners, as defined by expression 11, for different values of z (top row: $z = 0.7$, bottom row: $z = 0.1$), σ (left: $\sigma = 1$, center: $\sigma = 0.75$, right: $\sigma = 0.5$) and u (solid: $u = 0.001$, dashed: $u = 0.01$, dotted: $u = 0.1$).

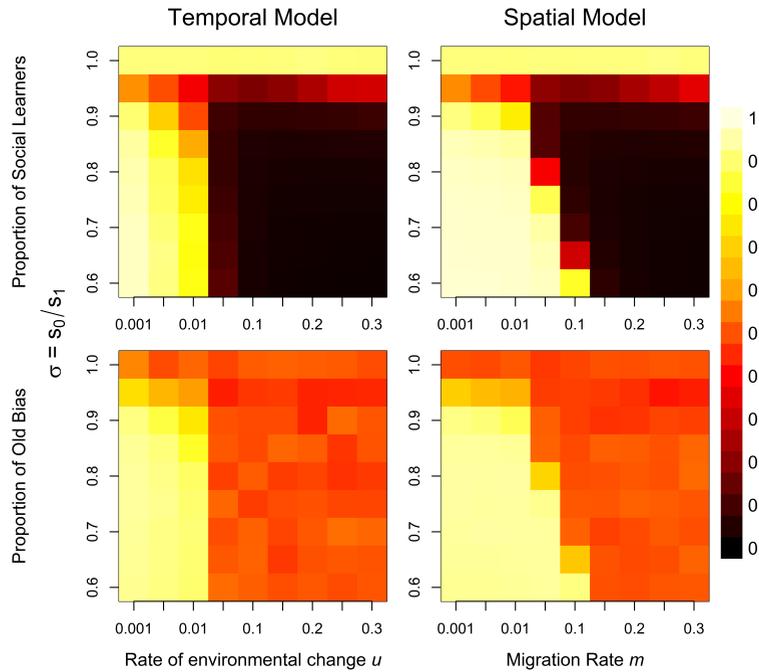


FIGURE S2. Proportions of social learners (top) and old bias (bottom) for the temporal (left) and spatial (right) individual-based model with social learning error ($e_{SL} = 0.3$). Lighter colors indicate higher proportions, darker colors lower proportions. Results are shown for $c = 0.05$, $z = 0.5$, $s_1 = 0.9$, $\mu = 0.005$.

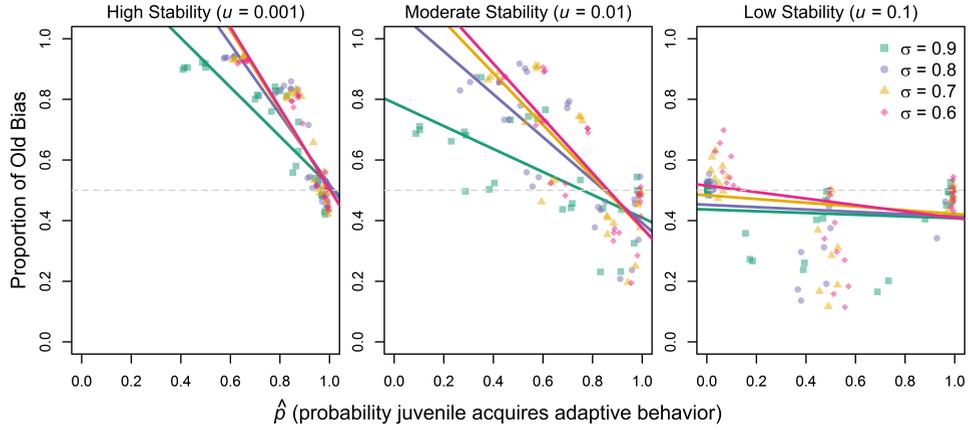


FIGURE S3. Results for temporal individual-based model. Proportions of old bias conditional on \hat{p} , the probability juveniles acquire adaptive behavior, for high (left), moderate (middle) and low (right) environmental stability. Each point represents mean across last 5000 timesteps of 10 independent 7000 timestep simulations per parameter combination. Colors/shapes indicate different values of σ , the survival advantage to adaptive behavior. In addition to u and σ , we also varied the cost of individual learning ($c = 0.05, c = 0.1, c = 0.2$), the success rate of individual learning ($z = 0.01, z = 0.5, z = 1$) and the error rate of social learning ($e_{SL} = 0, e_{SL} = 0.1, e_{SL} = 0.3$). Lines represent best linear fit.