

# A GENERAL SOLUTION TO THE COMBINED EFFECTS OF SELECTION AND BIASED MUTATION IN CULTURAL EVOLUTION

Vanessa Ferdinand

Melbourne School of Psychological Sciences, University of Melbourne, Australia  
Santa Fe Institute, New Mexico, USA

The evolution of culture is shaped by two primary forces of directional change: *biased transmission*, which is similar to natural selection, and *biased reconstruction*, which is similar to non-random recurrent mutation. The relative strength of these two forces has been the subject of much debate and formal models to date provide conflicting results with no definitive resolution. In this paper, I derive an analytical solution to the combined effects of selection and biased mutation when each are formalized as stochastic processes (e.g. Griffiths & Kalish, 2007). An important observation that immediately appears in the solution is that the temporal order in which these forces operate causally affects the evolutionary dynamics of the system. This order effect explains conflicting results in the literature to date: models in which selection appears to be stronger allow selection to operate last in the generational structure (e.g. Henrich & Boyd, 2002) and models in which biased mutation appears to be stronger allow mutation to operate last (e.g. Griffiths, Kalish, & Lewandowsky, 2008). In cultural evolution, various combinations of biasing forces are possible and these combinations are likely to vary across empirical domains. The solutions in this paper offer a shortcut to calculating the combined effects of any arbitrary number and ordering of selection and biased mutation processes, which should facilitate a more complete characterization of the space of dynamics that cultural evolution is capable of and improve our understanding of the relative effects of selection and biased mutation processes in culturally evolving systems.

## 1. Introduction

The evolution of culture is shaped by two primary forces of directional change. The first type is selective in nature; it determines which kinds of cultural entities will be copied by people (winning these variants fecundity and “offspring”) or ignored (effectively pronouncing these variants “dead”). A variety of selective forces operate on culture, such as natural selection, social learning strategies, and biased transmission processes (Boyd & Richerson, 1985). The second type is mutational in nature; it transforms cultural entities from one type to another while these entities are being copied and is capable of injecting novel variation into culture, often in biased ways. A variety of mutation-like processes operate on culture and fall under the umbrella of biased reconstruction processes (Sperber, 1996). Both of these forces are capable of creating directional changes that increase the frequency of certain cultural variants relative to others.

There has been a great deal of debate over the relative importance of each of these forces in shaping culturally evolving systems (e.g. Acerbi & Mesoudi, 2015). Although much of the discussion focusses on the fidelity of the cultural copying process, such as high vs low rates of random mutation, researchers are also interested in pitting biased mutation against selective forces and observing which force is more causally potent. This latter endeavour has produced several idiosyncratic models with conflicting results, leading Acerbi and Mesoudi (2015) to suggest the debate be moved to the empirical realm.

However, there is still conceptual ground to be gained in the modelling realm. What is lacking is a sufficiently general framework that allows a wide range of fitness functions to be combined with a wide range of biased mutation functions and straightforward analytical solutions to each of their combinations. In this paper, I adopt a general stochastic process modeling framework for cultural evolution and derive an analytical solution for the combined effects of any arbitrary fitness values and any arbitrary biased recurrent mutation process.

## 2. Modelling framework

This section proposes a stochastic process model (e.g. Lawler, 2018) for biased cultural mutation. Stochastic process modelling is commonly used in complex systems science and was first applied to cultural evolution in the domain of language evolution (Niyogi & Berwick, 1997; Nowak, Komarova, & Niyogi, 2001, 2002) and expanded with the development of the Bayesian iterated learning model (Griffiths & Kalish, 2007). The main representational device in this framework is a stochastic matrix called a *transition matrix* which contains the probability,  $p_{ij}$ , that a variable in state  $v_i$  transitions to state  $v_j$ :

$$\mathbf{M} = \begin{matrix} & \begin{matrix} v_1 & v_2 & \dots & v_n \end{matrix} \\ \begin{matrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{matrix} & \begin{pmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,j} \\ p_{2,1} & p_{2,2} & \dots & p_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ p_{i,1} & p_{i,2} & \dots & p_{i,j} \end{pmatrix} \end{matrix} \quad (1)$$

For the case of biased mutation,  $p_{ij}$  is the probability that a cultural variant of type  $v_i$  mutates into type  $v_j$  and we will refer to the transition matrix as a *mutation matrix*,  $\mathbf{M}$ . Transition matrices are also used in the field of mathematical population genetics to represent non-random recurrent mutation processes (e.g. Ewens, 2004, p. 174), but these are generally only 2-allele models with  $2 \times 2$  matrices. Cultural entities, however, typically exhibit several variants. For example, if you go to the hardware store to decide which cultural variant of “door handle” you’d like your kitchen cabinets to have, there won’t be only two options, but

dozens of variants to choose from. Equation 1 accommodates this by generalizing to an arbitrarily large number of variants,  $V = \{v_1, v_2, \dots, v_n\}$ .

Matrices are useful organizational devices for visualizing different patterns of mutation flow. The off-diagonal values on the  $i$ th row show  $v_i$ 's losses under the mutation process and the off-diagonal values on the  $j$ th column show  $v_j$ 's gains. The probabilities on each row sum to one, indicating that all of the mutational events are accounted for, but the columns may or may not sum to one. When the column sums do not all sum to one, the mutation process is *biased*, with certain variants receiving higher gains under the mutation process than others.

Representing biased mutation by a stochastic matrix is also convenient because it puts the tools of linear algebra at our disposal for calculating things such as the evolutionary trajectories of populations and their equilibrium frequencies. To make use of these tools, we represent the population as a stochastic vector,  $\vec{q}_t$ , where  $\vec{q}_t(v_i)$  is the relative frequency of  $v_i$  in the population at time  $t$ . For example,  $\vec{q}_t = (.3, .5, .2)$  means that  $v_1$  constitutes 30% of the population at time  $t$  whereas  $v_2$  constitutes 50% and  $v_3$  constitutes 20%. We will refer to  $\vec{q}_t$  as the *population vector*. The expected frequency of each variant in two successive generations is  $\vec{q}_{t+1} = \vec{q}_t \mathbf{M}$  and for any arbitrary number of generations into the future,  $g$ , it is  $\vec{q}_{t+g} = \vec{q}_t \mathbf{M}^g$ . The equilibrium frequencies are found by solving for the stationary distribution(s) of  $\mathbf{M}$ , as in  $\pi = \pi \mathbf{M}$ , or by solving  $\pi = \vec{q}_t \mathbf{M}^\infty$ , using a sufficiently large numeric stand-in for  $\infty$ . Example solutions are shown in Figure 1 (leftmost panel) for the example mutation matrix

$$\mathbf{M} = \begin{pmatrix} .8 & .1 & .1 \\ .2 & .6 & .2 \\ .3 & .3 & .4 \end{pmatrix} \quad (2)$$

where each black line represents the expected frequency of  $v_i$  over the course of 50 generations and they converge to the unique stationary distribution of  $\mathbf{M}$ , which is  $\pi = (.55, .27, .18)$ . Here we see that  $\mathbf{M}$  is a biased process that causes  $v_1$  to be the most plentiful variant, constituting 55% of the population at equilibrium.

The stochastic process model of cultural mutation described here is more general than the Bayesian iterated learning model because it does not require cultural mutation to be implemented by Bayesian rational learners, and it is more restricted than the Evolutionary Causal Matrix developed by Claidière, Scott-Phillips, and Sperber (2014) because 1) the entries in the matrix are probabilities and 2) it is meant to model a non-random recurrent mutation *process*, rather than model cultural attraction theory more broadly.

### 3. Combined effects solution

One approach to deriving an analytical solution to the combined effects of selection and biased mutation entails finding a way to incorporate fitness values into the mutation matrix, creating a new combined effect matrix,  $\mathbf{C}$ , and then applying

our linear algebra solutions to  $\mathbf{C}$ . In this section, I show how this can be done for a general model of natural selection, the replicator dynamics (Nowak, 2006).<sup>1</sup> The discrete-time replicator equation can be written as  $q_{t+1}(v_i) \propto q_t(v_i)f(v_i)$ , where  $f(v_i)$  is the fitness of the  $i$ th variant. Fitness values can be represented in several ways: as each type's expected number of children, e.g.  $f(v_1, v_2) = (10, 5)$ , as each type's fitness relative to one a reference type, e.g.  $f(v_1, v_2) = (1, \frac{1}{2})$ , or as each type's fitness relative to the population's total fitness, e.g.  $f(v_1, v_2) = (\frac{2}{3}, \frac{1}{3})$ , but the choice of representation does not affect the solution for the relative frequency  $q_{t+1}(v_i)$ . We will work with the latter representation where fitness is a stochastic vector and refer to it as the *fitness vector*. Figure 1 (rightmost panel) shows the dynamics for example fitness values  $f(v_i) = (.5, .4, .1)$ . Here we see that the variant with the highest relative fitness,  $v_1$ , goes to fixation, meaning it takes over the population and drives all other variants to extinction.

The combined effects matrix,  $\mathbf{C}$ , can be obtained by converting the fitness vector into a diagonal matrix,  $\mathbf{F}$ , where

$$\mathbf{F} = \text{diag}(f_1, \dots, f_n) := \begin{bmatrix} f_1 & & \\ & \ddots & \\ & & f_n \end{bmatrix} \quad (3)$$

and then multiplying  $\mathbf{F}$  by the mutation matrix,  $\mathbf{M}$ . It is important to note that matrix multiplication is not commutative, meaning  $\mathbf{F} \times \mathbf{M} \neq \mathbf{M} \times \mathbf{F}$ . Therefore, there are clearly two solutions to  $\mathbf{C}$ :  $\mathbf{C}_{fm} = \mathbf{F} \times \mathbf{M}$ , where selection acts before mutation in the generational structure of the cultural system, and  $\mathbf{C}_{mf} = \mathbf{M} \times \mathbf{F}$ , where mutation acts before selection.

When selection occurs first,  $\mathbf{F} \times \mathbf{M}$  implements element-wise multiplication of the fitness vector by each *column* in  $\mathbf{M}$  and can be understood as incorporating the fitness vector into the *gains* or *input* values of the mutation matrix, allowing selection to act before the mutation process occurs:

$$\mathbf{C}_{fm} = \begin{bmatrix} f_1 p_{1,1} & f_1 p_{1,2} & \cdots & f_1 p_{1,j} \\ f_2 p_{2,1} & f_2 p_{2,2} & \cdots & f_2 p_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ f_n p_{n,1} & f_n p_{n,2} & \cdots & f_n p_{n,j} \end{bmatrix} \quad (4)$$

When mutation occurs first,  $\mathbf{M} \times \mathbf{F}$  implements element-wise multiplication of the fitness vector by each *row* in  $\mathbf{M}$  and can be understood as incorporating

---

<sup>1</sup>Even though it contains the word “replicator”, the replicator dynamics equation itself neither implies nor requires that the variants in question be high-fidelity replicators. The equation simply assigns fitness values to variants based on their type at the moment the selection process acts and can be combined with any model of mutation.

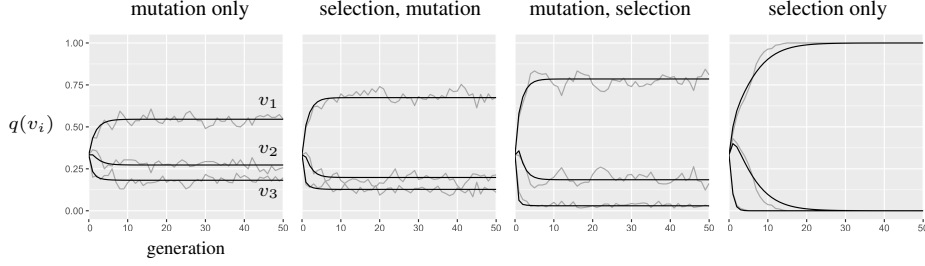


Figure 1. Evolutionary dynamics of four example combinations of selection and biased mutation, using fitness values  $f(v_i) = (.5, .4, .1)$ , mutation process  $\mathbf{M}$  from Equation 2, and initial population  $q(v_i) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ . Each panel (from left to right) shows the result when: mutation acts alone, selection occurs before mutation in the generational structure, mutation occurs before selection, and selection acts alone. Black lines show the expected frequency of each variant in the population at the end of each generation:  $v_1$  (top),  $v_2$  (middle),  $v_3$  (bottom). Grey lines show the result of a simulated trajectory for a finite population size of 300 variants. The expected trajectories converge to the stationary frequency of each variant,  $\pi(v_i)$ .

the fitness vector into the *losses* or *output* values of the mutation matrix, allowing selection to act after the mutation process occurs:

$$\mathbf{C}_{mf} = \begin{bmatrix} f_1 p_{1,1} & f_2 p_{1,2} & \cdots & f_n p_{1,j} \\ f_1 p_{2,1} & f_2 p_{2,2} & \cdots & f_n p_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ f_1 p_{i,1} & f_2 p_{i,2} & \cdots & f_n p_{i,j} \end{bmatrix} \quad (5)$$

Figure 1 shows the results for  $\mathbf{C}_{fm}$  (second panel) and  $\mathbf{C}_{mf}$  (third panel) for identical  $\mathbf{F}$  and  $\mathbf{M}$ . Here we see that the dynamics of these two processes differ. Process  $\mathbf{C}_{fm}$  more closely resembles the dynamics of mutation acting on its own, whereas  $\mathbf{C}_{mf}$  seems to be more influenced by selection. The analytical solutions  $\pi \mathbf{C}_{fm} = (.67, .20, .13)$  and  $\pi \mathbf{C}_{mf} = (.79, .18, .03)$  confirm that these are, indeed, identifiably different stochastic processes.

#### 4. Discussion of order effects

In the solution above, we saw that order matters because matrix multiplication is not commutative. To understand why order matters, let's consider the following thought experiment. Imagine that you and I are standing in front of a ball of clay. What we're going to do is take turns punching the ball of clay: first I throw a punch, then you throw a punch, and we repeat this a few more times. Now let's both inspect the clay and ask, whose handprint can we see most clearly? The one of us whose punch is the strongest? Perhaps. Or what about the one of us who punched the clay most recently? If we introduce a consistent bias into our measurement system by only inspecting the clay after I punch it, then we are

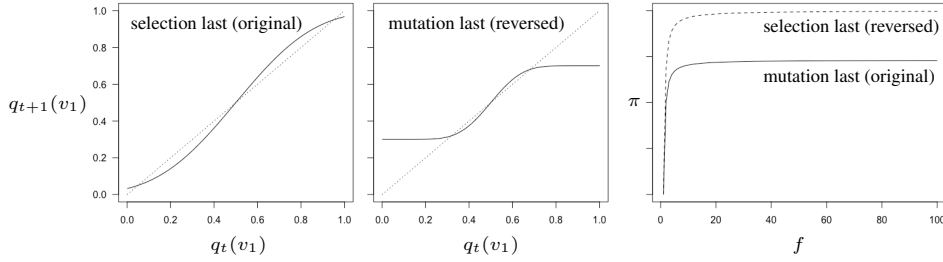


Figure 2. Replications of Henrich and Boyd (2002) (left, middle) and Griffiths et al. (2008) (right) alongside results that reverse the order of selection and mutation in their original models.

probably going to walk away thinking I packed the stronger punch. But if we only inspect it after you punch it, then we'll make the opposite inference.

Many formal models of evolution assume discrete generations and those that deal with multi-process evolutionary systems normally organize the processes into a set linear order within the generational structure. Furthermore, population frequencies are typically measured at the end of the generation, rather than after each force acts. This particular set of modeling choices is common, but it is also a recipe for over-representing the effects of each generation's final force.

Two examples of this ordering effect bias from the cultural evolution literature are found in models by Henrich and Boyd (2002) and Griffiths et al. (2008). Both of these papers address the relative potency of selection vs mutation in cultural evolution, but their models yield opposite results. Henrich and Boyd (2002) combine a model of conformity copying (a type of selection) with high mutation rates and find that the resulting evolutionary dynamics are barely affected by mutation. Figure 4 (left) replicates their result for a mutation rate of 30% and a population size of 20 (see their Figure 5). The overall shape of this curve resembles the typical s-shaped curve of conformity copying, which would have stable equilibria at coordinates  $q(v_1) = 1$  and 0. When combined with high mutation rates, the stable equilibria remain near  $q(v_1) = 1$  and 0 and do not display the 30% mutational load that would be expected from a 30% mutation rate. On the basis of this model, the authors conclude that error-prone cultural evolution is "corrected" by conformity biases. However, in this model, mutation operates first in the generational structure, conformity copying operates last, and frequencies are measured at the end of each generation. Figure 4 (middle) shows the results when the order of the selection and mutation processes are reversed. Here, the effects of mutation are much more pronounced and we observe the expected mutational load of 30%, pushing the equilibria to  $q(v_1) = 0.7$  and 0.3. This new perspective suggests a win for mutation over conformity copying.

Griffiths et al. (2008) combine a model of cultural mutation, resulting from a Bayesian learner with a prior bias favoring one cultural variant ( $v_1$ ) over another

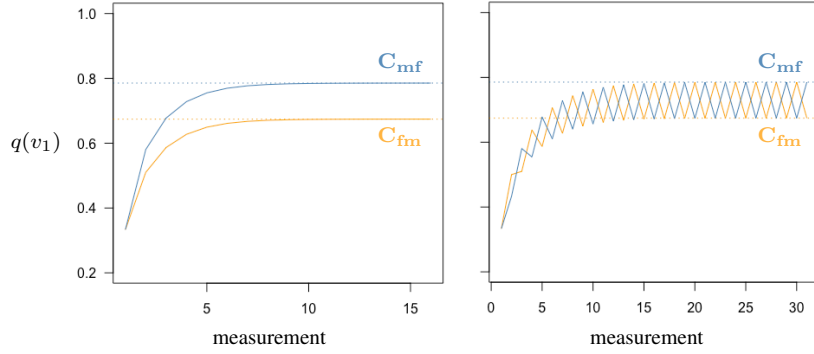


Figure 3. Replot of the trajectories defined by process  $C_{mf}$  (blue) and  $C_{fm}$  (orange) showing the frequency of  $v_1$  over 15 generations when measured at the end of each generation (left) and when measured twice per generation after each force acts (right). Both processes begin with the same uniform initial population composition,  $\vec{q}_{t=0} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .

( $v_2$ ), with a selection pressure causing differential fitness among the variants. This model also assumes a fixed generational structure, but one where selection operates first and mutation operates last. Figure 4 (right, solid line) replicates their result for a 20% prior bias on  $v_1$  and a 10% noise rate (corresponding to parameters  $\pi = 0.2$  and  $\epsilon = 0.1$  in their notation, see their Figure 5). Here, the equilibrium frequency of  $v_1$  ( $x$ -axis) goes up as the fitness of  $v_1$  increases ( $y$ -axis), but plateaus around 78% due to a mutational load of about 22%. This result was interpreted as a win for mutation over selection, because even infinitely strong selection is incapable of pushing  $v_1$  to fixation. The dashed line, however, reverses the order of these two processes. In this case,  $v_1$  goes to fixation as fitness goes to infinity and this scenario could be interpreted as a win for selection over mutation.

How should we deal with these ordering effects when evaluating the relative effects of selection and biased mutation in cultural evolution? One solution is to simply measure the population after every process acts. Referring back to Figure 1, the data in these plots show the population frequencies at the end of each generation only. Figure 4 replots the evolutionary trajectories of  $v_1$  for process  $C_{fm}$  and  $C_{mf}$ , where the system is measured once at the end of each generation (left) and once after each force acts (right). The right panel clearly shows that each force affects the frequency of  $v_1$  when it acts and that the system reaches an oscillating attractor where the two forces kick the same population frequencies back and forth between the equilibrium frequency solutions for  $C_{fm}$  and  $C_{mf}$ .

A second solution is to give a concrete or theoretically-justified interpretation to the ordering of selection and mutation processes in our models. This is more challenging as it requires some understanding about the “life-cycle” of a cultural variant, including when specific selection and mutation forces typically act upon

it. For example, biases in attention, memory, and the pragmatic usage of language may occur in a set order that is causally relevant to the evolution of language. Future research into the relative importance of selection and biased mutation in cultural evolution should either control for order effects or explicitly address the causal role they play in determining population frequencies over time.

### Acknowledgements

This work was completed with support from the University of Melbourne's Research Fellowship in Computational Cognitive Science, under the supervision of Andrew Perfors and Yoshihisa Kashima, and the Omidyar Fellowship at the Santa Fe Institute. Thanks to three anonymous reviewers for their helpful comments.

### References

- Acerbi, A., & Mesoudi, A. (2015). If we are all cultural darwinians what's the fuss about? clarifying recent disagreements in the field of cultural evolution. *Biology & philosophy*, 30(4), 481–503.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- Claidière, N., Scott-Phillips, T. C., & Sperber, D. (2014). How darwinian is cultural evolution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(20130368).
- Ewens, W. J. (2004). *Mathematical population genetics: theoretical introduction* (Vol. 1). Springer.
- Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with bayesian agents. *Cognitive science*, 31(3), 441–480.
- Griffiths, T. L., Kalish, M. L., & Lewandowsky, S. (2008). Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3503–3514.
- Henrich, J., & Boyd, R. (2002). On modeling cognition and culture. *Journal of Cognition and Culture*, 2(2), 87–112.
- Lawler, G. F. (2018). *Introduction to stochastic processes*. Chapman and Hall.
- Niyogi, P., & Berwick, R. C. (1997). Evolutionary consequences of language learning. *Linguistics and Philosophy*, 20(6), 697–719.
- Nowak, M. A. (2006). *Evolutionary dynamics: exploring the equations of life*. Harvard university press.
- Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science*, 291(5501), 114–118.
- Nowak, M. A., Komarova, N. L., & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, 417(6889), 611–617.
- Sperber, D. (1996). Explaining culture: A naturalistic approach. *Cambridge, MA: Cambridge*, 1101.