Repeated learning makes cultural evolution unique

Pontus Strimling[a], Magnus Enquist[a], and Kimmo Eriksson[a,b,1]

[a]Centre for the Study of Cultural Evolution, Stockholm University, SE-106 09 Stockholm, Sweden; and [b]School of Education, Culture and Communication, Mälardalen University, SE-721 23 Västerås, Sweden

Although genetic information is acquired only once, cultural information can be both abandoned and reacquired during an individual's lifetime. Therefore, cultural evolution will be determined not only by cultural traits' ability to spread but also by how good they are at sticking with an individual; however, the evolutionary consequences of this aspect of culture have not previously been explored. Here we show that repeated learning and multiple characteristics of cultural traits make cultural evolution unique, allowing dynamical phenomena we can recognize as specifically cultural, such as traits that both spread quickly and disappear quickly. Importantly, the analysis of our model also yields a theoretical objection to the popular suggestion that biological and cultural evolution can be understood in similar terms. We find that the possibility to predict long-term cultural evolution by some success index, analogous to biological fitness, depends on whether individuals have few or many opportunities to learn. If learning opportunities are few, we find that the existence of a success index may be logically impossible, rendering notions of "cultural fitness" meaningless. On the other hand, if individuals can learn many times, we find a success index that works, regardless of whether the transmission pattern is vertical, oblique, or horizontal.

cultural fitness | diffusion | retention

I n this paper, we will address the popular suggestion that biological and cultural evolution can be understood in similar terms (1–8).

Arguably, the most popular recent evolutionary hypothesis about culture is memetics, which maintains that cultural evolution is the playing field of selfish memes; simply put, the idea is that the success of cultural traits is determined by their inherent power to spread between human minds (1, 4, 9). This is a powerful idea that potentially explains observations of irrational behavior and strange beliefs and is in line with human susceptibility to persuasion and irrelevant stimuli (such as that in advertisements). Rational-choice theories of human behavior lend more power to the individual, who is assumed to be capable of making rational choices among existing cultural alternatives; theorists in this tradition tend to seek rational explanations to seemingly irrational elements of culture (10).

Here we suggest that both memetic and rational aspects can be accounted for in a theory of cultural evolution that considers in detail the consequences of repeated cultural learning. As we shall see, this also implies that cultural evolutionists must abandon the idea that the success of cultural traits can be captured with a single variable analogous to biological fitness. In contrast to standard models of cultural evolution (2, 11–14), we need to consider that not only naive but also experienced individuals take part in cultural learning. If exposed to a new cultural variant, an experienced individual may abandon the variant she is currently using (15). Such competition between cultural variants seems to occur in many domains (e.g. consumer products, fashion, beliefs), and the question for cultural evolution is which variants will survive in the long term (16). On the micro-level, whether an individual will abandon a currently used variant will depend not only on how visible, persuasive, or attractive the alternative is but also on how happy the individual is with her current variant, or how addicted she is to it. For instance, compared with an individual who currently lacks a working solution to a problem, an individual who

has a working solution already may be less prone to switch to a newly observed alternative.

To capture this in a simple model, we describe the competitiveness of cultural variants in terms of 2 values ranging between zero and 1: the "diffusion potential" d and the "retention potential" r. These values summarize the properties of a cultural variant that make it likely to become acquired by new individuals (d) or, when held by an individual, resist being replaced (r). As just 1 illustrative example, consider competing consumer products. Throughout life, people will have multiple opportunities to learn about various products and, possibly, try them. As studied in consumer psychology (17), the products that people try will be determined by certain factors (e.g. commercials), whereas partly different factors (e.g. satisfaction) will determine the products that products people will keep on using. The values of d and r would summarize the former and latter factors, respectively. Observe that both values may incorporate aspects of irrationality: A new variant may be very tempting without actually being good, whereas an individual addicted to something will have difficulties abandoning it, even in the face of an objectively better alternative.

To explore the consequences of these aspects of repeated learning, we here develop the simplest possible mathematical model. Assume that when an individual using variant A observes another variant B, the probability that she changes to the new variant is

$$d_B(1 - r_A). \qquad [1]$$

Intuitively, d_B is the probability that a social learner will be able to learn B in this interaction, and r_A is the probability that she will then keep variant A anyway. For simplicity, these probabilities are assumed to be constant over time and between individuals. We further assume that individuals of the new generation are born with no cultural variant (modeled as the "naive" variant with $d_{naive} = r_{naive} = 0$). Each new individual then interacts a number of times with a random subset of role models from the parent generation.

Results

By mathematical analysis, we can investigate the course of cultural evolution in this model and determine how it is influenced by variation in the number of learning opportunities as well as in the pattern of transmission. Fig. 1 illustrates how the frequencies of 2 cultural variants in a given generation change during repeated learning. Note that the impact of further learning opportunities is initially very large but that the frequencies will eventually stabilize. In particular, the frequency of naive individuals tends to zero. Note also that individuals will tend to adopt traits with high diffusion potential at first, and later replace them with traits with higher retention potential.
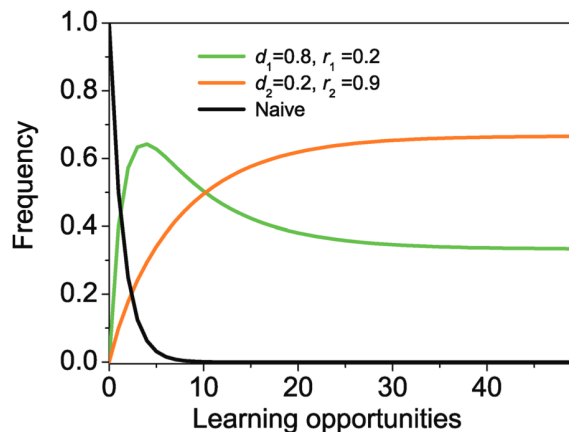
By letting the children of one generation take the role of parents for the next generation and repeating the just described "within-generation" process, we can study the outcome of cultural

Fig. 1. Frequencies of 2 cultural variants in a population as a function of the number of learning opportunities (oblique inheritance), based on the following parameter values: Variant 1 has higher diffusion potential than variant 2 ($d_1 = 0.8$, $d_2 = 0.2$) but lower retention potential ($r_1 = 0.2$, $r_2 = 0.9$), and the 2 variants are equally prevalent in the parent generation. Note that the frequencies eventually stabilize and that variant 1 initially spreads faster than variant 2 but ends at a lower frequency.

evolution over many generations. This intergenerational process tends to a unique equilibrium. Reflecting the importance of repeated learning for a single generation, the intergenerational equilibrium depends heavily on the number n of learning opportunities per individual. Fig. 2A–C illustrates this for the same 2 cultural variants as in Fig. 1. For a small n, the variant with the higher diffusion potential will take over (Fig. 2A); for moderate values of n, we obtain a mixed equilibrium of both cultural variants (Fig. 2B); when n gets sufficiently large, the other variant takes over instead (Fig. 2C). Fig. 2D and E illustrates how—if one trait is given fixed values of the diffusion and retention potentials—the equilibrium outcome varies with the potentials of the competing trait. For a small n, the outcome is mainly determined by the value of the diffusion potential d (Fig. 2D). When n is large, the 2 potentials are equally important (Fig. 2E).
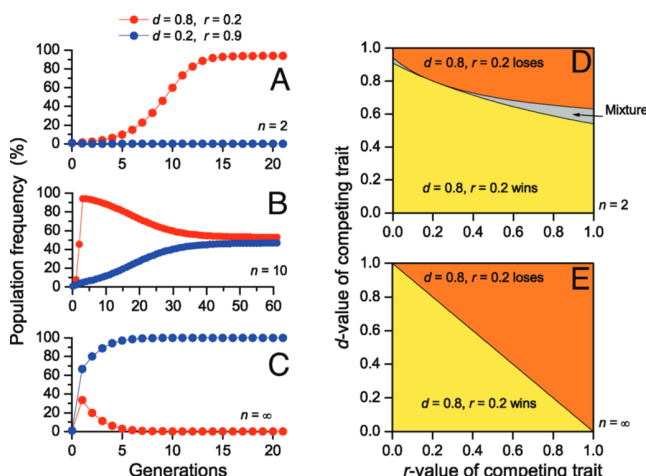
Fig. 2. Long-term cultural evolution (oblique inheritance). Frames A–C illustrate how frequencies change over generations, for 3 different values of the number n of learning opportunities. (A) $n = 2$; variant 1, which has the highest value of d, takes over. (B) $n = 10$; a mixed equilibrium of both variants. (C) $n = \infty$; variant 2, which has the largest value of $d/(1 - r)$, takes over. Frames D and E illustrate which combinations of d and r would outcompete variant 1 for 2 values of n. In D where $n = 2$, the outcome is mainly determined by d and mixed equilibria may occur. In E where $n = \infty$, d and r has equal impact and no mixed equilibria occur.

Table 1. Success index determining the winner among competing traits in cultural evolution

Transmission	Number of learning opportunities (n)		
	1	Few	Many
Vertical ($m = 2$)	d	No index	$\frac{d}{1-r}$
Oblique (m = population size)	d	No index	$\frac{d}{1-r}$
Horizontal (m = population size)	n/a	n/a	$\frac{d}{1-r}$

In biological evolution, the net effect of natural selection can be captured by a single success index called fitness, measuring reproductive success. Our model gives important insights into whether there may exist an analogous index that can predict long-term success in cultural evolution. In other words, the question is whether we can assign an index value w to each cultural variant, such that $w_A > w_B$ implies that variant A will outcompete trait B, whereas $w_A = w_B$ implies that the variants will coexist.

Depending on the number of learning opportunities, we obtain either positive or negative results concerning the existence of such an index. Our results are summarized in Table 1.

Beginning with the positive results, consider the case of a large n. By "large", we here mean that each generation has enough learning opportunities for the frequencies to become essentially stable (e.g., in Fig. 1 this occurs for $n > 20$). As illustrated in Fig. 2E, in the case of large n the diffusion and retention potentials have equal impact on which cultural variant will eventually win. Specifically, the quantity

$$w = d/(1 - r) \qquad [2]$$

will work as a predictive success index, such that the variant with the largest value of this index will take over. Further, this result holds regardless of whether transmission is "vertical" (each individual has only potential 2 role models) or "oblique" (all members of the entire parent population are potential role models). Cultural evolution scholars have also considered "horizontal" transmission, i.e., transmission within the same generation (2, 11), and it has been argued that this is what makes cultural evolution different (18). We think that the emphasis on mode of transmission stems from comparisons between horizontal transmission with many learning opportunities and vertical or oblique transmission with a limited number of learning opportunities per individual. If we introduce horizontal transmission into our model, so that individuals will act both as models and observers, it allows for faster (i.e. within-generation) evolution towards equilibrium, but the same success index will apply. Thus, any combination of within-generation (horizontal) and between-generation (vertical or oblique) transmission can be predicted by this success index, as long as the number of learning opportunities is sufficiently large.

The success index Eq. 2 does not apply to cases with a limited number of learning opportunities. The extreme case wherein there is just a single learning opportunity is easily analyzed: In this case, there is no role for retention and hence diffusion potential alone predicts success (so we can take $w = d$). An altogether different phenomenon arises in the intermediate case wherein $n > 1$ but not large: Mixed equilibria occur for entire sectors of the space of diffusion/retention-potential combinations (Fig. 2D). When this phenomenon occurs, it is logically impossible to find a success index that predicts the winner of cultural evolution. To see why this is so, we may consider 3 cultural variants (1, 2 and 3), where variant 1 is the fixed variant in Fig. 2D. Let 2 and 3 be any 2 variants in the gray area of mixed equilibria, such that they have equal retention potential but variant 3 has superior diffusion potential. That a success index exists means that we can rank alternatives and thereby predict the winner. Among the variants we are considering, 1 and 2 form a mixed equilibrium, so any potential success index w must satisfy $w_1 = w_2$. For the same reason, we must have

$w_1 = w_3$. Together, these relations imply $w_2 = w_3$. However, above we chose variants 2 and 3 so that the latter has superior diffusion potential, whereas the retention potentials are equal, and hence the outcome of a competition between these 2 variants must be that variant 3 takes over. Thus, any success index w must also satisfy $w_2 < w_3$. But no index can simultaneously satisfy $w_2 = w_3$ and $w_2 < w_3$, so a success index does not exist.

In conclusion, memetics in its simplest form, where the success of a trait is determined by its ability to "infect" individuals, works only when individuals have a single learning opportunity. With several learning opportunities, the individual and its previous experiences and choices become important, so individual-based models will be more useful than meme-based models.

Discussion

We will now discuss some implications. Research on the diffusion of innovations has found that it tends to follow a standard S-shaped curve (19). Fig. 3 illustrates that this comes out of our model as well. Our model also predicts that the speed by which the new variant takes over will depend on the particular values of the diffusion and retention potentials among variants: Increasing d and/or decreasing r will result in faster evolutionary change. For instance, evolution among competing traits that are initially attractive but fail to satisfy individuals, like different schemes for losing weight, will display fast diffusion and replacement whereas satisfying traits may stay in the population for a long time, even if they initially spread slowly. Of particular interest is, as illustrated in Fig. 2C, that a variant that starts spreading more quickly than other variants is not necessarily the winner in long-term evolution. A variant with a high diffusion potential but sufficiently low retention potential will enjoy temporary success, but in the end it will only have delayed the ultimate success of the competing variant. Although similar scenarios of temporary successes can occur in evolutionary games when evolution is determined by frequency-dependent payoffs (20), this is a different phenomenon:
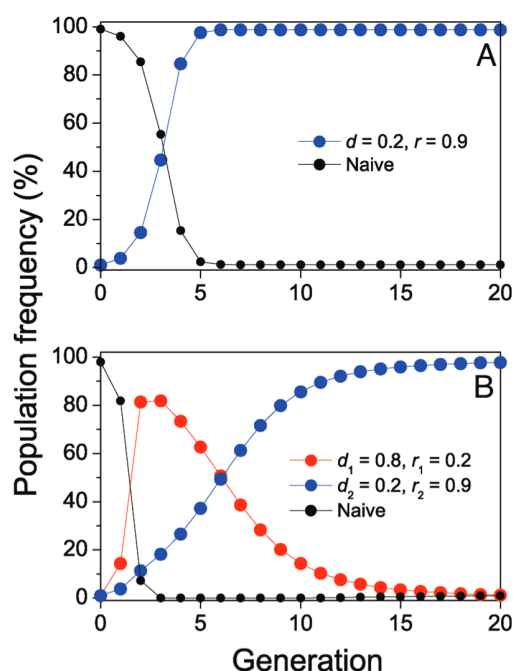
Fig. 3. Patterns of cultural evolution (oblique inheritance). (*A*) A new trait takes over the population in a number of generations, consistent with the S-curve typically found in empirical studies of diffusion of innovations. (*B*) If 2 variants are introduced simultaneously, the variant that eventually loses may initially be more successful and delay the establishment of the winning variant (cf. *A*).

The dynamics of our model are not based on payoffs but only on repeated transmission.

Our theory can also account for some of the mixture of rational and irrational elements that occur in human culture. It seems reasonable to assume that retention is more related to individual rationality than is diffusion because experiences allow individuals to make more informed choices. With this assumption, our model suggests that both rational and irrational factors determine cultural evolution and that the outcome cannot be predicted solely on the basis of function to the individual. The results also suggest that rational/functional traits are more likely to be favored when the number of learning trials increases. Our model can also explain why advertisements work. As long as traits are roughly equal in quality or function, like different brands of coffee, attractive but irrelevant factors that increase d will easily tip the balance toward one brand. In addition, advertisements may have an impact by increasing the number of learning trials.

Critics of memetics have said that the analogy between memes and genes is pointless without a rigorous theoretical framework for the ecology of memes (21). Although the main purpose of our model is to develop the theory of cultural evolution, it can also serve as a starting point for empirical studies of "meme ecology". Properties of cultural traits, such as their potentials for diffusion and retention, should be possible to measure in the laboratory. For instance, cognitive research on religion has shown that it is possible to experimentally measure differences between various supernatural notions by how easily they are transmitted and recalled in memory (22, 23). Our results suggest that it is also important to measure the impact a currently held cultural trait has on the cultural traits an individual will adopt in the future. Measurements of this kind can then be combined with mathematical models, such as the one presented here, to create a theory of cultural evolution that offers more substantial explanations and specific predictions.

In order to develop more powerful theories of cultural evolution, we need to abandon the idea that such evolution follows the exact same laws as biological evolution. Introducing a richer characterization of cultural traits and exploring in detail the consequences of repeated learning is one necessary step in this direction.

Materials and Methods

We will consider the dynamics of a fixed set of mutually exclusive cultural variants spreading across generations in a large population of individuals. New individuals learn through repeated interaction with a set of cultural role models uniformly drawn from the previous generation. Let x_v^{parents} denote the proportion of the parent population that holds variant v. We will only consider variants that are not already extinct, so all $x_v^{\text{parents}} > 0$.

Each cultural variant v is characterized by the value pair (d_v, r_v). These values define what happens in social-learning interactions: If a social learner has variant v and the cultural model has variant u, the probability that the social learner switches traits from v to u is $d_u(1 - r_v)$. Intuitively, d_u is the probability that a social learner will be able to learn u in this interaction, whereas r_v is the probability that she will keep variant v anyway.

Before they start learning, new individuals are assumed to be "naive" in the sense that they do not have any variant at all. We model this nonvariant by the value pair $(0, 0)$. This means that these nonvariants cannot spread, and are never kept when it is possible to learn another variant.

The distribution of cultural variants in the children generation at a given moment can be represented by a column vector

$$x^{\text{children}} = \begin{pmatrix} x_{\text{naive}} \\ x_1 \\ x_2 \\ \vdots \end{pmatrix}.$$

Let $x^{\text{children}}(i)$ denote the distribution after i learning opportunities. Then

$$x^{\text{children}}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

and for any $i \geq 0$ (assuming an infinite population, so that there are no stochastic effects) we can compute $x^{\text{children}}(i+1)$ recursively:

$$x^{\text{children}}(i+1) = Mx^{\text{children}}(i),$$

where M is the transition matrix of switching probabilities, which is easily set up. For instance, for the case of only 2 cultural variants, this matrix looks as follows:

$$M = \begin{pmatrix} 1 - P(\text{naive} \to 1) - P(\text{naive} \to 2) & 0 & 0 \\ P(\text{naive} \to 1) & 1 - P(1 \to 2) & P(2 \to 1) \\ P(\text{naive} \to 2) & P(1 \to 2) & 1 - P(2 \to 1) \end{pmatrix}.$$

Here $P(u \to v)$ denotes the probability that someone with variant u switches to trait v. It follows from our assumptions that this probability is $x_u^{\text{parents}} d_v (1 - r_u)$ if the individual learns from a role model drawn uniformly at random from the parent generation ("oblique transmission"). If instead learning is only derived from a smaller set of role models, like 2 parents ("vertical transmission"), then the factor x_u^{parents} must be replaced with the proportion of variant u in that set of role models.

We will now use this recursion to investigate 4 important cases. These cases are given by different combinations of the number of learning opportunities (either 2 or infinitely many) and the number of cultural role models (either learning from just 2 role models or from the whole parent generation).

Oblique Learning, 2 Learning Opportunities. Our object here is to determine how the distribution of variants will change between generations under oblique learning with 2 learning opportunities: Which variants will increase in frequency? Mathematically, this means that for a given cultural variant v we want to determine the conditions under which we obtain the inequality

$$x_v^{\text{children}}(2) > x_v^{\text{parents}}. \qquad [3]$$

For convenience, we introduce the notation

$$\widetilde{d} = \sum_v x_v^{\text{parents}} d_v, \quad \widetilde{d(1-r)} = \sum_v x_v^{\text{parents}} d_v (1 - r_v)$$

for the parent generation's weighted means of d values and $d(1 - r)$ values, respectively. Now, after 2 learning opportunities, our recursion gives the distribution

$$x^{\text{children}}(2) = M^2 x^{\text{children}}(0) = M^2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

It is straightforward but tedious to compute this expression and see that inequality shown in Eq. **3** translates into

$$d_v(2 + \widetilde{d(1-r)} - \widetilde{d}(2 - r_v)) > 1.$$

This inequality implies that whether or not a variant increases in frequency will depend on the weighted average values of d and $d(1 - r)$ in the parent generation. It is this frequency dependence that makes mixed equilibria possible (as illustrated in the main paper).

Oblique Learning, Infinite Number of Learning Opportunities. An infinite number of learning opportunities can be treated in the same way. For simplicity, we will assume just 2 cultural variants (1 and 2), but the results generalize to any number of variants. We are interested in studying under which conditions we have the inequality

$$x_v^{\text{children}}(\infty) > x_v^{\text{parents}},$$

and using the expressions obtained from the recursion

$$x^{\text{children}}(\infty) = M^\infty x^{\text{children}}(0),$$

we obtain the condition

$$\left(1 - x_v^{\text{parents}}\right) \frac{d_v}{1 - r_v} > x_u^{\text{parents}} \frac{d_u}{1 - r_u},$$

where u denotes the other variant. With an infinite number of learning opportunities, there will eventually be no naive children left, so in the coming parent generations there will be no naive parents. This means that $1 - x_v^{\text{parents}} = x_u^{\text{parents}}$, which simlipifies the inequality so that

$$\frac{d_v}{1 - r_v} > \frac{d_u}{1 - r_u}$$

is the condition under which the frequency of variant v increases over generations. Thus, the eventual survivor is the variant that maximizes $\frac{d}{1-r}$.

Vertical Learning, 2 Learning Opportunities. We now move on to the case in which each child has 2 cultural role models (drawn uniformly at random from the parent generation at the birth of the child). We need to keep track of what cultural role models a particular child has, so that we can calculate the probability that the role model she picks in a particular round has a particular variant.

Again, to simplify, we study the case for 2 variants. Let $M_{a,b,c}$ denote the transition matrix for a child who has a cultural role models who are naive, b role models with variant 1, and c role models with variant 2. In general, the probability that the cultural role model picked in a particular round has trait 1 is $\frac{b}{a+b+c}$; specializing to the case where everyone has 2 parents, we obtain the probability $\frac{b}{2}$.

There are children with different combinations of role models. For each combination, we must weight the corresponding matrix with the proportion of children that have that particular combination of role models. To make the equation less cumbersome, we will here keep track only of the role-model couples that affect trait 1.

$$x_1^{\text{children}}(2) = x_1^{\text{parents}} \left(2 x_{\text{naive}}^{\text{parents}} M_{1,1,0}^2 + x_1^{\text{parents}} M_{0,2,0}^2 \right.$$
$$\left. + 2 x_2^{\text{parents}} M_{0,1,1}^2 \right) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (0, 1, 0).$$

The inequality

$$x_1^{\text{children}}(2) > x_1^{\text{parents}}$$

now boils down to the condition

$$d_1(4 - d_1 + \widetilde{d(1-r)} - \widetilde{d}(2 - r_1)) > 2,$$

which, just as in the oblique learning case, opens up for mixed equilibria.

Vertical Learning, Infinite Number of Learning Opportunities. Finally, we come to the case wherein each individual has 2 cultural role models and an infinite number of opportunities to learn. Following the previous case, we have

$$x_1^{\text{children}}(\infty) = x_1^{\text{parents}} \left(2 x_{\text{naive}}^{\text{parents}} M_{1,1,0}^\infty + x_1^{\text{parents}} M_{0,2,0}^\infty \right.$$
$$\left. + 2 x_2^{\text{parents}} M_{0,1,1}^\infty \right) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (0, 1, 0).$$

The inequality

$$x_1^{\text{children}}(\infty) > x_1^{\text{parents}}$$

translates into the condition

$$\left(1 - x_1^{\text{parents}}\right) \frac{d_1}{1 - r_1} > \left(2 x_2^{\text{parents}} - \left(1 - x_1^{\text{parents}}\right)\right) \frac{d_2}{1 - r_2}. \qquad [4]$$

As in the oblique case, there will eventually be no naive individuals left, whence we obtain the same optimization condition:

$$\frac{d_1}{1 - r_1} > \frac{d_2}{1 - r_2}.$$

Horizontal Learning, Infinite Number of Learning Opportunities. We will now show that the results for infinite learning opportunities also hold if the model is changed to reflect learning within the same generation. We still have the same basic recursion

$$x(i+1) = Mx(i),$$

where M is the transition matrix of switching probabilities, which only differs from the oblique case in that there is no distinction between the learning generation and the parent generation, so $P(u \to v) = x_u d_v (1 - r_u)$. From this recursion, it is straightforward to compute that the inequality

$$x_v(i+1) > x_v(i)$$

translates into the condition

$$\sum_u x_u x_v (d_v(1-r_u) - d_u(1-r_v)) > 0.$$

Hence, the variant with the highest value of $d/(1-r)$ will steadily increase in frequency until it takes over the entire population.

1. Dawkins R (1975) *The Selfish Gene* (Oxford Univ Press, London).
2. Cavalli-Sforza LL, Feldman MW (1981) *Cultural Transmission and Evolution* (Princeton Univ Press, Princeton).
3. Basalla G (1988) *The Evolution of Technology* (Cambridge Univ Press, New York).
4. Dennett DC (1995) *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (Simon & Schuster, New York).
5. Laland K, Brown G (2002) *Sense and Nonsense* (Oxford Univ Press, New York).
6. Carneiro RL (2003) *Evolutionism in Cultural Anthropology* (Westview Press, Boulder, CO).
7. Jablonka E, Lamb MJ (2005) *Evolution in Four Dimensions* (MIT Press, Cambridge, MA).
8. Mesoudi A, Whiten A, Laland K (2006) Towards a unified science of cultural evolution. *Behav Brain Sci* 29:329–347.
9. Blackmore SJ (2000) *The Meme Machine* (Oxford Univ Press, New York).
10. Stark R, Bainbridge WS (1987) *A Theory of Religion* (Peter Lang, New York).
11. Boyd R, Richerson PJ (1985) *Culture and the Evolutionary Process* (Univ Chicago Press, Chicago).
12. Rogers A (1988) Does biology constrain culture? *Am Anthropol* 90:819–831.
13. Kendal JR, Laland KN (2000) Mathematical models of memetics. *J Memetics* 4.
14. Enquist M, Eriksson K, Ghirlanda S (2007) Critical social learning: A solution to Rogers'paradox of non-adaptive culture. *Am Anthropol* 109:727–734.
15. Aunger R (2000) *The Electric Meme: A New Theory of How We Think* (Free Press, New York).
16. Aunger R, ed (2000) *Darwinizing Culture* (Oxford Univ Press, New York).
17. Holbrook MB, Hirschman EC (1982) The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *J Consumer Res* 9:132–140.
18. Williams GC (1992) *Natural Selection: Domains, Levels and Challenges* (Oxford Univ Press, New York).
19. Rogers ME (2003) *Diffusion of Innovations* (Free Press, New York), 5th Ed.
20. Weibull J (1995) *Evolutionary Game Theory* (MIT Press, Cambridge, MA).
21. Pigliucci M (2008) Memes concept too broad and imprecise. *Skeptical Inquirer* 32:65–66.
22. Barrett JL, Nyhof MA (2001) Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *J Cognit Cult* 1:69–100.
23. Boyer P, Ramble C (2001) Cognitive templates for religious concepts: Cross cultural evidence for recall of counter-intuitive representations. *Cognit Sci* 25:535–564.