On modeling cognition and culture

How formal models of social learning can inform our understanding of cultural evolution

Joseph Henrich
University of Michigan Business School
701 Tappan Road, D3276
Ann Arbor, MI 48109
j.henrich@wiko-berlin.de

Robert Boyd
Department of Anthropology
University of California
Los Angeles, CA 90095
boyd@wiko-berlin.de

Draft: November 8, 2000. Do not quote or cite without authors' permission.

Abstract

Formal models of cultural evolution analyze how cognitive and affective processes combine with patterns of social interaction to generate the distributions and dynamics of ‘representations’—ideas, beliefs, schemas, or mental models. Recently, cognitive anthropologists have made three criticisms of such models: First, mental representations are non-discrete. Second, cultural transmission is highly inaccurate because public representations provide only incomplete information to social learners. And third, mental representations are not replicated, but rather are ‘reconstructed’ through an inferential process that is strongly affected by cognitive ‘attractors,’ which shape the kinds of representations that are likely to be acquired. Based on these three claims, critics have concluded that: 1) models that assume use replication or replicators are inappropriate, 2) selective cultural learning cannot account for ‘cultural inertia’ (stable traditions), and 3) selective cultural learning also cannot generate cumulative adaptive evolution.

Here we analyze three formal models to show that even if the three premises of this critique are correct, the deductions that have been drawn from them are false. In the first model, we assume continuously varying representations under the influence of weak selective transmission and strong attractors. We show that if the attractors are sufficiently strong relative to selective forces, the continuous representation model reduces to the standard discrete-trait replicator model, and the weak selective component determines the final equilibrium of the system. In the second model, we assume representations are discrete, but that replication is very inaccurate. We show that very low fidelity replication of representations at the individual level does not preclude accurate replication at the population level, and therefore, accurate individual-level replication of representations is not necessary for selective forces to generate either cultural inertia or cumulative cultural adaptation. In the third model, we assume continuous (non-discrete) cultural representations, incomplete transmission and substantial inferential transformations. We derive the conditions for cumulative adaptive evolution.

Introduction

Formal models of cultural evolution (e.g. Boyd and Richerson 1985) provide an important tool for understanding cultural phenomena. Culture is shaped by both psychological processes that determine how people think and feel, and social processes that determine how people interact, whether they succeed or fail, and if they live or die. Formal cultural evolutionary models are useful because they can combine cognitive and affective processes with patterns of social interaction to generate explicit predictions about the distributions and dynamics of 'representations'—ideas, beliefs, schemas, or mental models. With such predictions, empirical data on behavior and beliefs from populations can be used to select among alternative models, or among the cognitive and social components of these models, in a manner that informs our understanding of cognition and connects laboratory experiments with real life (e.g. Henrich 2001).

Recently Sperber (1996), Atran (2000), and Boyer (1994, 1999) have argued that these "Neo-Darwinian" models[1] are inappropriate for studying cultural evolution because they are "based on a serious distortion of the relevant facts" (Sperber 1996: 118). Sperber, Atran and Boyer ('SAB' for short) suggest three interrelated problems with such approaches: First, cultural transmission processes are usually incomplete and imperfect, so, unlike genetic systems, accurate replication rarely occurs. Replication is the exception, rather than the rule. Second, unlike DNA replication, inferential processes "transform" these representations during their transmission and reconstruction. This suggests that *mutation*-like processes are much more important than *selection-like* processes in shaping cultural variation. Third, unlike genes, cultural representations are rarely discrete units, suggesting that the idea of a 'replicator' (or meme) makes little sense for most types of cultural representations.

Building on these three points, SAB (e.g., Sperber 1996: 102-103; Atran 2000: 3; Boyer 1994) have claimed the following: First, because representations are non-discrete, using replicators and replicator dynamics is entirely inappropriate for culture systems. Second, because of the high levels of inaccuracy and the incompleteness of cultural transmission, selective processes cannot explain 'cultural inertia'—the existence of social groups in which individuals share a set of relatively stable representations (traditions). From this, Sperber (1996: Chapter 5) further asserts that inferential transformations create 'strong cognitive attractors' that swamp any influence from the selective components of cultural evolutionary processes,[2] leading to the claim that the selective components of cultural transmission cannot generate adaptation cultural evolution.[3]

[1] SAB use the term "Neo-Darwinian" for such models even when they are explicitly derived from disease epidemiology (Cavalli-Sforza and Feldman 1981: 46-53), learning theory (see Boyd and Richerson 1985, chapter 4), or cognitive psychology (Lumsden and Wilson 1981).

[2] Boyer calls these 'attractors' 'triggered representations' or cognitive tracks (1999), while Atran (2000) calls them 'cognitive elicitors'

[3] While SAB's characterization applies to the informal theorizing of memeticists (e.g. Dawkins 1982, Dennett 1995 or Blackmore 1999), it is wide of the mark for much formal cultural theory. Boyd and Richerson (BR 1985: 75) explain that there is no need to assume particulate "units" in order to build evolutionary models (1985: 70-74) and

Whether SAB are correct that social learning is low fidelity, non-discrete and strongly affected by cognitive transformations is an empirical matter, although it is plausible that they are correct in wide range of cultural domains. Nevertheless, the deductions they make from these assumptions are matters of logic, and here we use three mathematical models to show that these deductions are wrong. In the first model, we assume continuously varying representations under the influence of weak selective transmission and strong attractors. We show that if the attractors are sufficiently strong relative to selective forces, the continuous representation model reduces to the standard discrete-trait replicator model, and the weak selective component determines the final equilibrium of the system. In the second model, we assume representations are discrete, but replication is very inaccurate. We show that very low fidelity replication of representations at the individual level does not preclude accurate replication at the population level, and therefore, accurate individual-level replication of representations is not necessary for selective forces to generate either cultural inertia or cumulative cultural adaptation. In the third model, we assume continuous (non-discrete) cultural representations, incomplete transmission and substantial inferential transformations. Despite these assumptions, the model shows that adaptive cultural evolution is likely in empirically plausible conditions—predictions derived from the models have been tested elsewhere using ethnographic data (Henrich, in prep.).

Model 1: Strong attractors give rise to discrete replicators

Formal models of cultural evolution have made use of both continuous and discrete representations of cultural variation, depending on the situation. SAB categorically reject the use of discrete-representation models. At the same time, they emphasize the role of intra-individual cognitive transformations of representations (‘strong attractors’ or’ cognitive tracks’), and argue that, by comparison, selective cultural processes have little or no effect on the epidemiology of representations (e.g. Sperber 1996: Chapter 5). In this section, we analyze a simple model that assumes continuous representations, strong attractors and weak selective forces. We show that this model reduces to discrete-representation replicator-dynamics in which the weak selective forces determine the ultimate outcome. The stronger the cognitive attractors are relative to the selective transmission forces, the better is the discrete-replicator approximation. Therefore, assuming that human cognition gives rise to more than one attractor per domain, it is inconsistent to simultaneously advance the idea of strong attractors (or ‘cognitive tracks’ or ‘triggered representation) and to criticize the use of discrete-representations in models.

19 of the 38 models presented are continuous (non-discrete) trait models, which allow for an arbitrary amount of transmission error. Similarly, Cavalli-Sforza and Feldman (CSF 1981) devote one of the five chapters to continuous trait models. These continuous models all allow for substantial error and other forms of non-replication. BR also explicitly distinguish public representations from mental representations (though using different terminology) throughout the book, and repeatedly specify the inferential transformation between observed behavior and representation formed. For example, in describing one model, they write, “The offspring uses each model’s actual behavior [public representation] to estimate [make inferences about] his or her cultural variant [mental representation (1985: 75-79).” They also make explict reference to much research in psychology on the nature of social learning. After a review of the psychological literature, especially Bandura’s work on social learning, they specify the following pathway: Modeled events → Attention Processes → Retention Processes → Motor Reproduction → Motivation Processes → Matching. Of the 396 references, 101 were by psychologists, or published in psychology journals. Chapters 4 and 5 discuss how what Sperber (1996) would later call “attractors” biase cultural change so that some outcomes are more likely than others, and even use some of the same examples as Boyer (1999).

Here we study the evolution of the distribution of mental representations in a population. We assume that each individual's mental representation in this domain is a single real number, x, between zero and 1.[4] Individuals holding different representations (i.e. different values of x) behave differently on-average, or in SAB's terminology, generate different "public representations". During each time period, people in the population observe the behavior of another individual, infer the mental representation of the model from his or her behavior, and adopt this inference as their own representation.[5] Following Sperber, cognitive 'attractors' strongly bias the inferential process that underlies social learning—individuals whose model has a representation x, on average, infer a representation that is nearer to one of two attractors. In particular, we assume that the attractors are $x = 0$ and $x = 1$, and that an individual who observes a model with representation x infers a representation $x + \Delta x$, where $\Delta x < 0$ for $x < m$ and $\Delta x > 0$ for $x > m$. We assume that Δx has the form shown in Figure 1.

[Figure 1 here]

The following example illustrates how domain specific cognition might create multiple attractors. Suppose individuals at Attractor-0 ($x = 0$) perceive the Moon is a self-aware, conscious, entity with goals, emotions, and motivations, thus the Moon's behavior can be understood using folk psychology (Theory of Mind; Leslie 1994). Thus, $(1 - x)$ tells us the degree to which an individual attributes the Moon's shape, color, position and movements to underlying goals, emotions or motivations, or to what degree the Moon's goal-driven actions influence events and individuals, such as weather, tides, personal moods, werewolves, etc. In contrast, individuals with $x = 1$ (those at Attractor-1) see the Moon as simply a big rock, lacking goals, consciousness, and emotions. These individuals attribute the Moon's color, shape and movement to the effects of non-agentic interactions of light, and other mindless bodies, governed by physical laws that operate throughout the Universe (even if they, themselves, don't understand those laws). Any effects the Moon has on things such as tides, moods and calendars, are merely unintentional consequences of the moon's mass and movement. Now, it is possible to imagine Moon-concepts that mix these poles ($1 > x > 0$). One could believe, for example, that the moon's movement and shape are out of its control (governed by physical laws), while its color or hue expresses its mood, which in turn influences the weather. Or, perhaps the Moon's color is 23% controlled by its emotions and 77% controlled by the law of light refraction. One might also believe that on Tuesdays and Thursdays the Moon is a goal-oriented agent, on Monday, Wednesdays and Fridays the Moon is a big rock, and on the weekends these two alternate minute by minute. From this perspective, such beliefs might seem odd to us because they violate intuitive expectations, and would consequently be transformed by cognitive attractors. In contrast, $x = 1$ or 0 are "easier to think." To capture this example in a formal model the underlying mental representations, x, would likely need to be a multi-dimensional vector, and the analysis represented below could be easily adapted to multi-dimensional variables—none of the qualitative finding would change. However, to keep the presentation tractable, we will limit our analysis to one-dimensional x variables.

[4] We restrict the x to [0,1] only for simplicity of exposition. Representations that varied over other ranges, or multi-dimensional would produce the same basic result.

[5] For the moment we do not allow individuals to observe a sample models and make use of the pattern of observed in the sample. We return to this question below.

A second example illustrates how an attractor might be represented by a single quantitative dimension. Young and Burke (2001) have shown that an overwhelming majority of the share cropping contracts in Illinois are one of two types: In some communities the farmers and the landowners split the crop yield, while in other communities the landowners receive two shares for every one received by the farmer. Interestingly, despite wide variation in the quality of the land within communities, there is little variation in share crop contracts within communities. Let us suppose that we were studying a population of farmers in this region, and that x_i is individual *i*'s mental representation of the fair or appropriate share for the land-owner in a share-cropping contract. We scale these representations so that Attractor-0 ($x = 0$) is a 1:1 share, while Attractor 1 ($x = 1$) represents a 2:1 share. We assume that the integer ratios, using numbers (1, 2, 3) that human understand intuitively, are easier to represent and to remember (e.g. easier than 1.38:3.13). When a person is observed to use another contracts, he or she is often thought to be using (and thinking) about one of the two focal contracts.

We also assume in that individuals are selective in picking their cultural models—that is, in deciding on whom to focus their inferential processes. In particular, we assume individual focus on individuals with higher values of x. Individuals with higher values of x might have more success, interact more, or be more salient for other reason. This assumption about human psychology is well founded: A substantial amount of evidence from both laboratory and field research demonstrates than humans pay particular attention to, preferentially interact with, and tend to imitate successful/prestigious individuals (Henrich & Gil-White 2001 summarize this evidence and provide lay a evolutionary foundation). Using our above example, it could be that individuals who believe in a 2:1 contract have more success, on average, than individuals who have other mental representations, and that their relative success will make them preferred models. We let w_i be the likelihood that individual i is selected as a model, and assume that there is a positive correlation between x_i and w_i.

We analyze this model using the Price Equation (Price 1970). This equation (1) says that the mean change in the mean of *any* property of a population of things, genes, mental representations, or the hydrogen atoms in the Andromeda galaxy, can be decomposed into two parts: the extent to which the properties of things covary with their rate of replication,[6] and the rate at which the things themselves change with time. There are no assumptions about replication fidelity, discreteness or underlying distributions. This form is particularly useful here because it partitions the change in a population of mental representations into the effect of selecting particular cultural models ("selective transmission") from the effect of inferential processes ("incomplete inference").

$$\overline{w}\Delta\overline{x} = \underbrace{\mathrm{Cov}(w_i, x_i)}_{\text{Selective Transmisson}} + \underbrace{E(w_i \Delta x_i)}_{\text{Incomplete Inference}} \qquad (1)$$

$\Delta\overline{x}$ is the change in the average value of x_i per time step, Δx_i captures the effect of the cognitive attractors, and $\overline{w}$ is the average replicability of the x_i values carried in the minds of the n individuals in the group.

[6] In the form expressed here we have assumed that the x's proportional representation from one time step to the next is a function of it's relative replicability (or fitness) and its current representation in the population.

Another useful property of the Price equation is that it makes it easy to deal with populations that are structured into groups (Price 1972). Here, we divide the population into two groups, one for each domain of attraction, and express the dynamics for the change in the average value of *x* in each domain of attraction. Group 0 is composed of individuals whose mental representation is less than *m*, and group 1 individuals whose mental representation is greater than *m*. Let $\overline{x}_j$ and $\overline{w}_j$ be the mean value of *x* and the mean replication rate for group *j*, and let x_{ji} and w_{ji} be the mental representations of the *i*th individual in group *j*. Then we can rewrite (1) as follows:

$$\overline{w}\Delta\overline{x} = \text{Cov}(\overline{w}_j, \overline{x}_j) + E(\overline{w}_j\Delta\overline{x}_j) \tag{2}$$

That is, the change in the mean value over the whole population can be decomposed into the covariance between group means and group mean replication rate, and the average change within each group. But, notice that the average change within each group has the same form as the left hand side of the Price equation. Thus, we can apply the Price equation again, this time to the dynamics *within* each group—substituting the expressions given in (3) into the expectation term in (2).

$$\overline{w}_j\Delta\overline{x}_j = \text{Cov}(w_{ji}, x_{ji}) + E(w_{ji}\Delta x_{ji}) \tag{3}$$

The covariance term gives the change within the groups due to psychology of model selection, and the second term gives the change within each group due to inferential transformation by attractors.

Now we make use of the specific form of the inferential transformation shown in Figure 1. In group 0, $\Delta x_{0i} = -\beta_0 x_{0i}$, and in group 1, $\Delta x_{1i} = \beta_1(1 - x_{1i})$. Further, assume that an individual's likelihood of being selected as a model depends linearly on their value of x_{ji} so that $w_{ji} = 1 + sx_{ji}$. With these assumptions

$$\overline{w}_0\Delta\overline{x}_0 = -\beta_0\overline{x}_0 + s\left(\text{Var}(x_{0i}) - \beta_0 E(x_{0i}^2)\right) \approx -\beta_0\overline{x}_0 \tag{4}$$

and

$$\overline{w}_1\Delta\overline{x}_1 = \beta_1(1 - \overline{x}_{1i}) + s\left[Var(x_{1i}) - \beta_1(\overline{x}_{1i} - E(x_{1i}^2)\right] \approx \beta_1(1 - \overline{x}_{1i}) \tag{5}$$

If intra-psychic transformation are strong relative to the selective forces (β_1, β_0 >> *s*), then the first terms on the right-hand sides of (4) and (5) will dominate the dynamics and determine the equilibrium values of the mean values of *x* in each group.

We can now write down the approximate epidemiological dynamics for the entire population. First, let *p* be the fraction of individuals in group 1 and 1 – *p* the fraction of the population in group 0. Then, substituting (4) and (5) into (2) yields:

$$\overline{w}\Delta\overline{x} \approx \underbrace{\left[p(\overline{w}_1\overline{x}_1) + (1-p)(\overline{w}_0\overline{x}_0) - \overline{x}\overline{w}\right]}_{\text{Covariance Term from (2)}} + \underbrace{\left[p\beta_1(1-\overline{x}_1)) - (1-p)\beta_0\overline{x}_0\right]}_{\text{Expectation Term from (2)}} \tag{6}$$

Because the β's are so much larger than s, the dynamics of the system will be initial governed by the attractors—$\bar{x}_0$ will rapidly evolve toward zero, and $\bar{x}_1$ toward 1, and thus $\bar{x} \approx p$. Substituting these values into (6) gives the following equation for the approximate dynamics after this initial period:

$$\Delta\bar{x} \approx \Delta p \approx \frac{p(\bar{w}_1 - \bar{w})}{\bar{w}} \tag{7}$$

Equation (12) is the common formulation for discrete-trait replicator dynamics. With this formulation, we see that the longer-term dynamics and the final equilibrium, $\bar{x} = 1$, are determined by selective cultural transmission and approximated with standard replicator dynamics. The stronger cognitive attractors are relative to selective cultural transmission, the *better* the discrete replicator approximation.

These findings are very general. Elsewhere, we (in prep) show this finding applies to more than two attractors, for a wide range different intra-psychic transformations (Δx_{ji}), and to representations in any number of dimensions. Here we also discuss the effect of more complex forms for selective cultural transmission.

Numerical simulations of the model indicate that the approximate analysis derived above is quite accurate. In our simulation, we assumed an initially uniform distribution of x_i in a population in a finite population. For parameters, n = 200, $\beta_0 = \beta_1$ = 0.50, s = 0.05, and m = 0.60, Figure 2 shows the dynamics of the average value of x in group 0 ($\bar{x}_0$), the average value of x in group 1 ($\bar{x}_1$), the mean of x (overall: $\bar{x}$) for one run of the simulation, the mean value of x for 10 simulations, and the predictions of replicator dynamics. As anticipated above, $\bar{x}_0$ goes rapidly to 0 and $\bar{x}_1$ goes rapidly to 1. The $\bar{x}$ curve from one run of the simulation shows the random effects of drift produced by finite populations, but is still clearly tracked by the replicator approximation. Across 10 simulations, the effects of drift are averaged out and replicator dynamics tracks $\bar{x}$ quite closely. A series of simulations with a variety of parameter combinations confirms the findings derived above. In addition, it shows how cultural drift is affected by s, n and m. As n decreases the force cultural drift increases, making it more likely for x to drift into Domain 0. If s gets too small, drift is also more likely to drive x to zero. As m get large, x values are more likely to drift into Domain 0 and be driven to zero.

[Figure 2 here]

Model 2: Cultural inertia and cumulative cultural adaptation can occur even when inference is inaccurate.

In this second model, we explore the effect of inaccurate replication. We assume that cognitive processes generate strong attractors, but that inferences, based on the available public representations, are highly inaccurate. Building on our findings in Model 1, we usediscrete-representations to show that, even when transmission fidelity is low (and very low), cultural transmission can still create cultural inertia and adaptive cultural evolution.

Much of the confusion about the whether selective cultural transmission can produce cultural inertia, shared traditions and adaptive evolution arises from a failure understand how human social learning distinguishes cultural from genetic evolution. In genetic evolution both evolutionary inertia and the possibility of gradual cumulative cultural change have the same cause, accurate, unbiased genetic replication. In sexual species like humans, each individual has two parents and this has important consequences for the distribution of genotypes in populations. However, except for extremely rare mutations, each gene is a faithful copy of a single gene carried by a member of the previous generation—genetic replication is very accurate. Moreover, each gene that an individual carries, again with a few exceptions, is equally likely to be included in the individuals' gametes—thus genetic reproduction is unbiased. These properties of genetic reproductions result in Mendel's laws for the distribution of genotypes.

To see how accurate, unbiased replication gives rise to genetic inertia and makes cumulative adaptive change possible consider a simple genetic system with only two alleles (or genes) at a single locus segregating in a large population. We label them *A* and *a*. Let *q* be the frequency of *A* before genetic transmission, and *q'* the frequency after genetic transmission. Finally, suppose that each allele mutates to the other with probability *u.* This means that,

$$q' = (1 - 2u)q + u\,. \tag{8}$$

A plot of *q'* as a function of *q* is a straight line with a slope slightly smaller than one as shown in Figure 3. If the slope were exactly one (i.e., no mutation), then transmission would lead to no change in the population ($q'= q$) from across the generations. A population that had 80% *A* alleles before transmission would have 80% after transmission, and a population that had 20% before transmission would have 20% afterward. However, any amount of mutation tends to reduce the frequency of the more common allele, and thus if no other evolutionary processes affect the population it will move toward a state in which both alleles have the same frequency. Because mutations are very rare ($u \approx 10^{-6}$), this will occur very, very slowly. Thus, in the absence of other forces, populations will change very slowly, and this explains much genetic inertia. Similarly, as long as natural selection is stronger than mutation, it can lead to cumulative adaptation.

[Figure 3 here]

By analogy from the process just described, SAB (Sperber 1996: 103-118; Atran 1999: 3-4) argue that cultural replication is highly inaccurate, and therefore, that the social processes of cultural transmission cannot give rise to cultural inertia or cumulative cultural adaptation. Unlike genes, mental representations are not replicated during cultural transmission. Instead, mental representations give rise to behaviors (or "public representations" in the SAB's terminology) that are observed by others, who must then *infer* the underlying mental representations that gave rise to the behavior. Because individuals differ and public representations provide incomplete information, this inferential processes is, SAB assert, highly inaccurate. If they are correct, it does follow that cultural transmission can not give rise to cultural inertia for the *same* reasons as genetic transmission. Too see this, suppose there are only two possible mental representations in some domain, labeled *A* and *B*. Each generates different but overlapping distributions of public representations. When cultural learning occurs, naïve individuals, perhaps children, observe a sample of individuals from these distributions, make inferences, and then adopt their ownmental representation. Following SAB, we suppose this process is very inaccurate; it is governed by equation (8) but now u = 0.2 (5 orders of magnitude larger than the rate genetic mutation). To get a feeling for the magnitude of u, remember that u = 0.5 means that there is no transmission at all—naïve individuals adopt mental representations at random, independent of who they observe. Thus, we are assuming that the acquisition of mental representations is 40% error and 60% transmission. As is shown in figure 4, very high error rates lead to very different dynamics than accurate replication. Under such high error rates, populations very rapidly converge to a random distribution of mental representations. There is no inertia, and extremely strong selective forces would be required to generate cumulative adaptation.

However, SAB have carried the gene analogy too far. Even if cultural transmission is inaccurate, it does not follow there can be no cultural inertia or cumulative evolution of adaptations. Any transmission process that leads to accurate replication *at the level of the population* will lead to cultural inertia and allow cumulative, gradual adaptation. Put another way, any transmission process that produces a plot like that shown in figure 3 will do the job. SAB's error is to assume that the only process that can give rise accurate replication at the level of the population is accurate replication at the level of individuals.

As before, we assume that there are two possible mental representations, A and B, and that a fraction q of the experienced population has A. Every naïve individual observes the public representations of n models and makes inferences about the mental representations that gave rise to these public representations. In each case, this inferential process is subject to frequent errors. The probability of inferring the correct mental representation is $1 - u$ and the incorrect representation is u. Thus, the probability, p, that a social learner infers that any one of the models has mental representation A is:

$$p = q(1-u) + (1-q)u$$

Finally, suppose that individuals adopt, as their own, the mental representation that they believe is most common among their models. This means that the probability that a naïve individual acquires A, and therefore, the frequency of A after transmission, q', is:

$$q' = \sum_{i>n/2}^{n} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}$$

Because individuals tend to acquire the more common mental representation, this model is one way to represent a "conformist" bias in social learning.

Here, in a model constructed on well-established evolutionary and psychological foundations, we show that even highly inaccurate cultural transmission can give rise to inertia—both data and theory indicate that human social learning has a conformist bias. Theoretically, evolutionary modeling of the cognitive capacity for conformist transmission suggests that genes leading to a conformist bias in social learning will be supported in virtually any environment that also favors social learning (Henrich and Boyd 1998). Empirically, psychological evidence for conformity from pre-1984 studies is summarized in Boyd & Richerson (1985). Much of recent evidence can be found in Baron et. al. (1996), Insko et. al. (1985), Smith and Bell, Bond & Smith (1996), and Campbell & Fairey (1989). These studies analyze everything from economic decisions and strategy choice in strategic games to perceptual tasks and food choices. Henrich (2001) produces evidence of conformist transmission from field data on the diffusion of innovations.

A conformist bias at the individual level leads to more accurate replication at the population level. Figure 5a plots *q'* as a function of *q* for $n = 5$ and $u = 0.2$. The shows that social learning processes come much closer to accurately replicating the population frequency of the mental representations, even though the transmission process is error prone at the individual level. The *maximum* error rate at the population level (at $q = 0, 1$) is approximately 0.02, one tenth of the individual rate. The reason for this is simple: errors have a bigger effect on populations in which one mental representation is common compared to populations in which both mental representations have similar frequencies. As a result, the conformist bias in transmission corrects for the effect of errors. Furthermore, even higher error rates can be compensated for by larger samples of cultural parents: Figure 5b plots the effect of transmission when $u = 0.3$ and $n = 10$. Recall that $u = 0.30$ means errors occur 60% of time.

This conformist learning mechanism can lead cultural inertia because it tends to increases the frequency of common mental representations, and reduces the frequency of rare ones (Figure 6a). In models with low fidelity learning this creates equilibria that balance the effect of low fidelity against the force of conformist transmission. Two initially different populations can remain different even in the same environment.

Conformist social learning also allows selective forces to generate more accurate adaptations to the environment. Suppose that individuals with mental representation *A* are twice as successful as individuals with representation *B* on average, and that the probability that individuals are chosen to be models is proportional to the ratio of an individual's success to the average success in the population ('prestige-biased transmission,' see below). Figure 6b shows that when we combine prestige bias, conformist bias and high error rates the more successful representation readily spreads.

Clearly, the combination of high error rates and a conformist bias does not result in the same kind of "frictionless" adaptation as genetic replication. Highly accurate genetic replication allows minute selective forces to generate and preserve adaptations over millions of years. Error prone cultural replication, even when "corrected" by a conformist bias imposes modest, but still significant forces on the cultural composition of the population. This means that only selective forces of similar magnitude will lead to cumulative adaptation. We do not think this is a problem because the selective forces acting on cultural variation are probably much stronger than those that shape genetic variation because they work on social learning time scales, and are driven by psychological not demographic events—e.g. the diffusion of innovations literature shows time scales of decades, not millennia (Rogers 1995).

Model 3: Adaptive evolution in a model without accurate replication or discrete representations

Evolutionary theory predicts, and both field and laboratory data empirical data confirm, that individuals have a psychological propensity to copy particularly skillful, successful and prestigious individuals. We call this *prestige-biased transmission*; Henrich & Gil-White 2001 lay out the theory and summarize the empirical data. This tendency creates a selective force that can, under the right circumstances, generate cumulative adaptation. It is easy to see how this would work if cultural replication were accurate. Suppose populations vary in their ability to perform some important task, say prey tracking or arrow manufacturing techniques. Further, suppose that in choosing models, people preferentially imitate the techniques used by the best trackers in the group. If the mental representations that underlie the success of the best trackers can be accurately copied, the mean tracking success of the population will increase through time (Boyd and Richerson 1985, Ch 8) because good trackers will leave more cultural 'descendants'. However, SAB have argued that, because cultural representations are not discrete, and cultural replication is inaccurate, selective forces, such as prestige-biased transmission, are unlikely to be important.

Here we show that prestige biased transmission can lead to cumulative adaptation, even when cultural transmission is inaccurate and representations are not discrete. Consider a population of N individuals who are numbered $i = 1,\ldots,N$. Individual i has a z-value (z_i). This value measures the individual skill in some domain like canoe making, arrow manufacture or medicinal plant selection. Later we will show that z may measure either an individual's mental representation (e.g. how long they think arrows should be), or some phenotypic measure that aggregates many skills or measures success (like lifetime tapir kills). However, for now we assume that it is a mental representation. Each individual is also characterized by a variable f that specifies the relative likelihood that an individual will be chosen as a cultural model. If people's social learning attention is drawn to skilled hunters (and larger values of z lead to hunting skill), then f and z will have a positive partial regression coefficient (other things can affect f as well). The pool of individuals could represent the same group of individuals (same farmers) from one year to the next, or it could different generations or cohorts.

Once again, we use the Price Equation to study the combined effects of selective transmission and inaccurate inference on continuous representations.

$$\Delta\bar{z} = \underbrace{\mathrm{Cov}\left(f_i, z_i\right)}_{\text{Selective Transmisso n}} + \underbrace{E\left(f_i \Delta z_i\right)}_{\text{Incomplete Inference}}$$

$\Delta\bar{z}$ is the change in the average value of z_i per time step. If $\Delta\bar{z}$ is positive, then adaptive evolution is taking place because, for example, people are becoming better trackers or arrow makers. Cov(f_i, z_i) is the covariation between f_i and z_i, and gives the effect of selective cultural forces on $\Delta\bar{z}$. In this case, it captures our psychological tendency to copy successful/skillful people—here we will assume that all learners attempt to copy the most skilled individual. $E(f_i \Delta\bar{z}_i)$ is the replicability-weighted average of all the individual intra-psychic transformations on cultural representation.

To capture the idea that inferential processes are incomplete and inaccurate, we assume that the inferential processes that underpin social learning are inaccurate in two senses: First, they are noisy, so that copiers never accurately replicate the z value of their model, and second, they are biased so that the behavior acquired by copiers is, on average, less skilled than the behavior of their model. More formally (as illustrated in Figure 7), individuals who attempt to copy a model with z-value, z_i, end up with a z-value drawn from a Gumbel probability distribution[7] with mode $z_i - \alpha$ and dispersion β. Typically, copiers construct representations that are on average worse than their model's *z*-values by an amount *α*, but occasionally—through lucky guesses or errors—individuals construct representations that yield z-values higher than their model. The probability of that occurring for an individual is the area under the distribution to the right of the dashed line (the model's z-value). It's also worth noting that the probability of an exact copy is zero, and the probability of two people arriving at exactly the same copy is zero—no replicas in this model.

[Figure 7 here]

With these assumptions, equation (1) becomes:

$$\Delta\bar{z} = -\alpha + \beta(\varepsilon + Ln(N)) \qquad (9)$$

The first term represents the effect of systematic errors and is always negative—it operates against adaptive evolution. In the second term, *Ln*(*N*) is the natural logarithm of N, the number of social learners in the population. The parameter ε is the Euler-gamma constant, which equals 0.577. β, *Ln*(N) and ε are all always positive, so the second term is always positive—and favor adaptive evolution. This means that adaptive cultural evolution depends on the relative sizes of the two terms. Interestingly, the two components of inference, systematic bias (measured by α) and random noise (measured by β), have opposite effects on adaptive evolution. Systematic bias operates against adaptive evolution, while noise—the tendency of individuals to make *different* inferences from observing the operation of the same underlying representation—*favors* adaptive evolution. The more individual tend to make different inferences, the faster cultural evolution goes (or the more likely it is to be adaptive). Similarly, the larger the population of social learners, *N*, the faster adaptive evolution proceeds (or the more likely selective forces will favor adaptive processes).

Setting $\Delta\bar{z} > 0$ and solving (9) yields the conditions under which selective cultural transmission will drive adaptive cultural evolution:

$$N^* > e^{\frac{\alpha}{\beta} - \varepsilon}$$

*N** is the critical number of social learners necessary to produce cumulative adaptive cultural evolution for a specified set of inferential processes (α, β). Figure 8 plots this inequality and shows graphically shows the conditions under which selective transmission will drive adaptive evolution.

[Figure 8 here]

This result provides several insights about cultural adaptation in a noisy environment. First, cumulative adaptive evolution is facilitated by both a large pool of social learners and an inferential tendency to make different "mistakes" (large β, a tendency for individuals to form

[7] The details of this distribution do not qualitatively impact our results. The details of the construction of this model, and the choice of inference distribution, are discussed in Henrich (in prep).

quite different representations). This finding directly contradicts Sperber (1996) and Atran (1999), who have claimed that because cultural transmission leads to a different representations in each individuals head, cumulative cultural evolution is not possible. Second, no matter how poor individuals are at imitating, there are some values of β and N that generate cumulative evolution—elsewhere Henrich (in prep) has shown how a sudden drop in population size (N) initiated a process of technological devolution in Tasmania over the last 8,000 years. So, both the social environment and the nature of cognition influence the conditions for cumulative adaptation. Looking only the inferential processes and the conditions for adaptation, what matters is not α, but α/β. So, if the typical inaccuracy of copies doubles (α), while the variation in different representations triples, the cumulative adaptation will still occur. Finally, having incorporated SAB's claims about the nature of culture into simple formal models, we have shown that such models bear no resemblance to genetic models (other than both being about evolutionary processes) can deal with these assumptions and derive, precise, non-intuitive insights.

SAB argued that evolutionary models can only work if there identified 'units' of transmission. This is not the case, and suggests a failure to apprehend the more general nature of evolutionary processes (Boyd and Richerson 2000). Neither genetic, nor cultural evolutionary, models require 'units of transmission'—genetic evolutionary models were, e.g., providing insights into the nature of evolution long before anyone knew anything about DNA discreteness or replication (not to mention that genetic models of quantitative characters, like height, still ignore these "units"). The most general form of the Price Equation, for example, is exact for anything one can measure about the constituent parts of any evolutionary system. This feature allows us to study the evolution of, and interrelation between, phenotypically expressed behavior and "public representations," and the epidemiology of the underlying mental representation. Suppose now, instead of a mental representation of a skill, z measures success in some domain such as hunting (perhaps quantified in lifetime tapir kills), combat (in 'heads-taken'), canoe making, or farming (in sacks harvested per hectare of wheat sown). Equation (2) would still govern the evolution of z. However, if we wanted to get at the underlying mental representations that produce particular values of z, we would have to specify how each mental representation contributes to an individuals' behavioral expression (to their success). For illustrative purposes, suppose z is hunting returns (a phenotypic measure) and y and ϕ are mental representations related to prey pursuit time and arrow length—presumably there are many more relevant representations for hunting. Using a linear regression equation, we can express the causal relationship of between mental representations y and ϕ on success, z, as follows.[8]

$$z_i = \mu + \lambda_1 y_i + \lambda_2 \phi_i + \varepsilon$$

The λ's give the relative contribution of an individual's y and ϕ mental representations to the observed success, z_i. ε gives uncorrelated random error, and μ specifies the constant term. With this, we can express the change in the average value of representation y as:

[8] In general, we can do this for any number of mental representations, and study the interaction of different mental representations.

$$\Delta\bar{y} = \underbrace{\mathrm{Cov}\left(f_i, y_i\right)}_{\text{Selective Transmisso n}} + \underbrace{E\left(f_i \Delta y_i\right)}_{\text{Incomplete Inference}}$$

But, because an individual's likelihood of being selected as a cultural model depends on *f*, and *f* depends on individuals observing z (e.g., hunting returns), then this can be rewritten as,

$$\Delta\bar{y} = \underbrace{\lambda_1 \rho Var\left(y_i\right)}_{\text{Selective Transmisson}} + \underbrace{E\left(f_i \Delta y_i\right)}_{\text{Incomplete Inference}} \quad ,$$

where ρ is the partial regression coefficient on *f* on z. The introduction of λ_1 shows us that effective transmission of *y* depends of how much it affects z. The Δ*y* term would have the same form as above; it depends of how difficult it was to infer the underlying y_i by observing the models' behavior. The reminder of the derivation proceeds as above.

Conclusion

Formal models are a necessary component to developing a more complete of understanding of cognition and culture. People are not well equipped to understand the population-level outcomes that arise from numerous minor interactions, weak cognitive biases, random errors, migration rates and micro-level decisions. The history of ecology, disease epidemiology and evolutionary biology over the last 40 years provides strong evidence that simple mathematical models provide powerful tools for understanding such problems.

The simple models described in this paper yield several finding that challenge the untutored intuitions of many:

1. Strong attractors generate replicator dynamics, such that weak selective forces determine the final state of the system. The stronger the attractors, the better the assumption of discrete-trait replicator dynamics.

2. Conformist transmission can compensate for even very high error rates in social inference in manner that fundamentally violates the intuitions that SAB import from genetics. Prestige biased transmission and combines with conformist transmission to spread adaptive representations; together they predict the conditions for inertia and change.

3. Adaptive cultural evolution can occur even when representations are continuous and inferences biased against adaptation. Poor inferences can be compensated for by larger pools of social learners, or greater amount of random error (by the fact that people don't make the same inferential mistakes).

4. Discrete units of transmission are not necessary for adaptive evolution. Dawkins (1976, 1982) proclamations about the requirements for adaptive evolution (replicators) are flat wrong.

The crux of SAB's position is that the transmission of culture requires innate, domain specific cognitive mechanisms, and that therefore that population dynamic models of culture are necessarily wrong. We accept that social learning, like all other forms of learning, requires innate expectations about objects in the environment and the nature of relationships among them. How these innate structures shape the human mind is obviously of great importance for understanding human culture. SAB's mistake is to see these ideas as incompatible with making population dynamic models of cultural change. It will never be enough to focus on the mind and ignore the interactions between different minds. To keep track of such interactions some kind of population dynamic models will be necessary. What is needed is both more effort by coevolutionary theorists to incorporate rich cognition into formal models of social learning, and more effort by cognitive scientists to consider how innate cognitive structure interacts with social processes and the cognition of social learning to influence the epidemiology of representations and its associated behavioral products.

References

Atran, S. (2002). The Religious Landscape. Cambridge University Press.

Atran, S. (unpublished). The Trouble with Memes: Inference versus Imitation in Cultural Creation.

Baron, R., Vandello, J., & Brunsman, B. (1996). The forgotten variable in conformity research: impact of task importance on social influence. Journal of Personality & Social Psychology, 71(5), 915-927.

Blackmore, S. (1999). The Meme Machine. Oxford: Oxford University Press.

Bond, R., & Smith, P. B. (1996). Culture and conformity: a meta-analysis of studies using asch's (1952b, 1956) line judgment task. Psychological Bulletin, 119(1), 111-137.

Boyd, R., & Richerson, P. J. (1985). Culture and the Evolutionary Process. Chicago, IL: University of Chicago Press.

Boyd, R., & Richerson, P. J. (2000). Memes: Universal Acid or a Better Mouse Trap. R. Aunger (editor), Darwinizing Culture: The Status of Memetics as a Science (pp. 143-162). Oxford: Oxford University Press.

Boyer, P. (1994). The Naturalness of Religious Ideas. Berkeley: University of California.

Campbell, J. D., & Fairey, P. J. (1989). Informational and normative routes to conformity: the effect of faction size as a function of norm extremity and attention to the stimulus. Journal of Personality & Social Psychology, 57(3), 457-468.

Cavailli-Sforza, L. L., & Feldman, M. (1981). Cultural Transmission and Evolution. Princeton: Princeton University Press.

Dawkins, R. (1982). The Extended Phenotype. Oxford: Oxford University Press.

Dawkins, R. (1976). The Selfish Gene. Oxford: Oxford Unversity Press.

Dennett, D. (1995). Darwin's Dangerous Idea . London: Penguin Press.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. Evolution and Human Behavior, 19, 215-242.

Henrich, J. (forthcoming). Cultural Transmission and the Diffusion of Innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change and much of sociocultural evolution. American Anthropologist.

Insko, C. A., Smith, R. H., Alicke, M. D., Wade, J., & Taylor, S. (1985). Conformity and group size: the concern with being right and the concern with being liked. Personality & Social Psychology Bulletin, 11(1), 41-50.

Leslie, A. M. (1994). ToMM, ToBY, and Agency: Core architecture and domain specificity. L. A. Hirschfeld, & S. A. Gelman (editors), Mapping the Mind: Domain specificity in cognition and culture (pp. 119-148). Cambridge: Cambridge University Press.

Lumsden, C., & Wilson, E. O. (1981). Genes, Mind and Culture. Cambridge: Harvard University Press.

Price, G. (1972). Extensions of Covariance Selection Mathematics. Annals of Human Genetics, 35, 485-490.

Price, G. R. (1970). Selection and Covariance. Nature, 520-521.

Rogers, E. M. (1995). Diffusion of innovations. New York: Free Press.

Smith, J. M., & Bell, P. a. (1994). Conformity as a determinant of behavior in a resource dilemma. Journal of Social Psychology, 134(2), 191-200.

Sperber, D. (1996). Explaining culture : a naturalistic approach. (pp. vii, 175 p). Oxford, UK. Cambridge, Mass: Blackwell.

Young, H. P., & Burke, M. A. (1998). Competition and Custom in Economic Contracts: A Case Study of Illnois Agriculture.

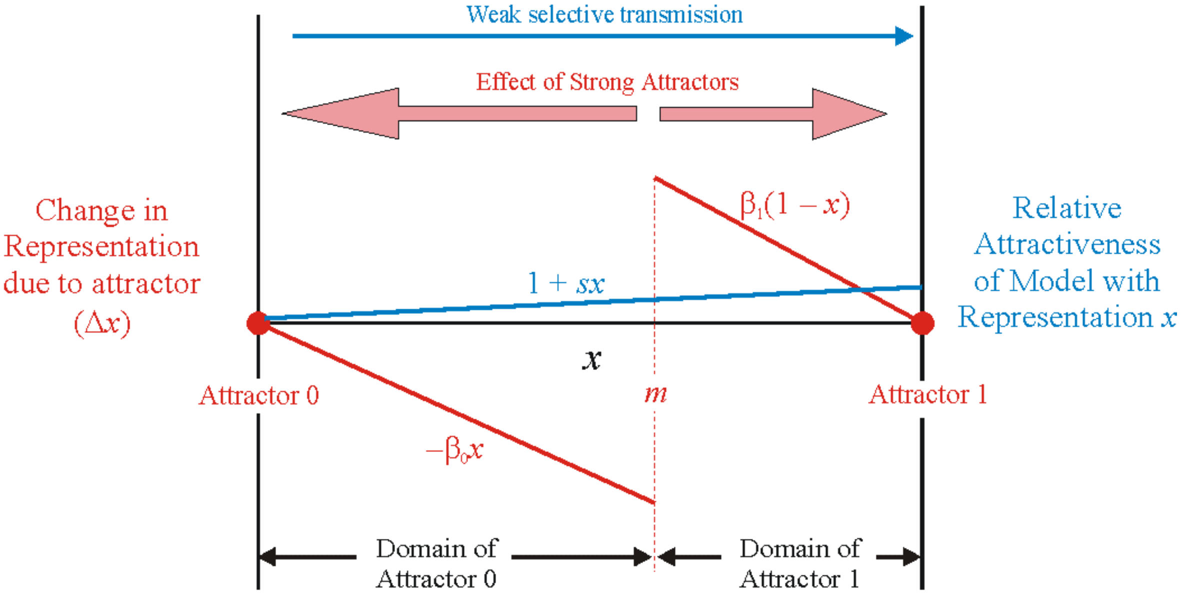
Weak selective transmission
Effect of Strong Attractors
Change in Representation due to attractor (Δx)
$\beta_1(1-x)$
$1 + sx$
Relative Attractiveness of Model with Representation x
x
Attractor 0
m
Attractor 1
$-\beta_0 x$
Domain of Attractor 0
Domain of Attractor 1

Figure 1 Graphical representation of the assumptions in model 1. A learner who observes a model with representation, x, infers a representation $x + \Delta x$, where $\Delta x = -\beta_0 x$ if $x < m$ and $\Delta x = \beta_1(1 - x)$ if $x > m$. This creates a strong force that shift representations in the population toward attractors at 0 and 1. The probability that an individual is chosen as a model is $1 + sx$. This selective transmission process creates a weak force that increases the frequency of larger values of x in the population.

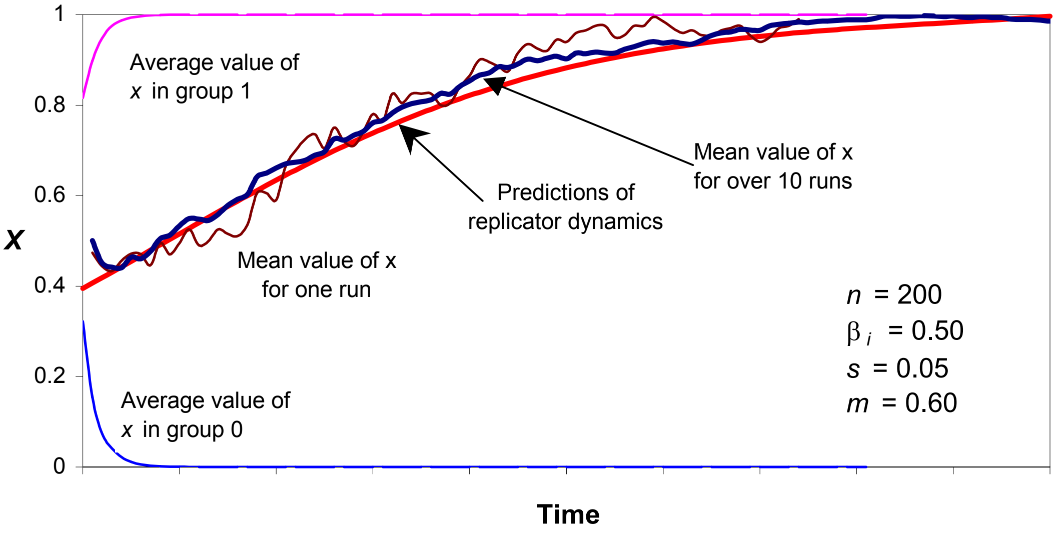

Figure 2 Results from simulating model described in text. The blue and magenta lines show that the mean representation in each domain rapidly approaches the attractors 0 and 1. The overall evolution of the population is very well approximated by a discrete model in which only weak selective forces are present.

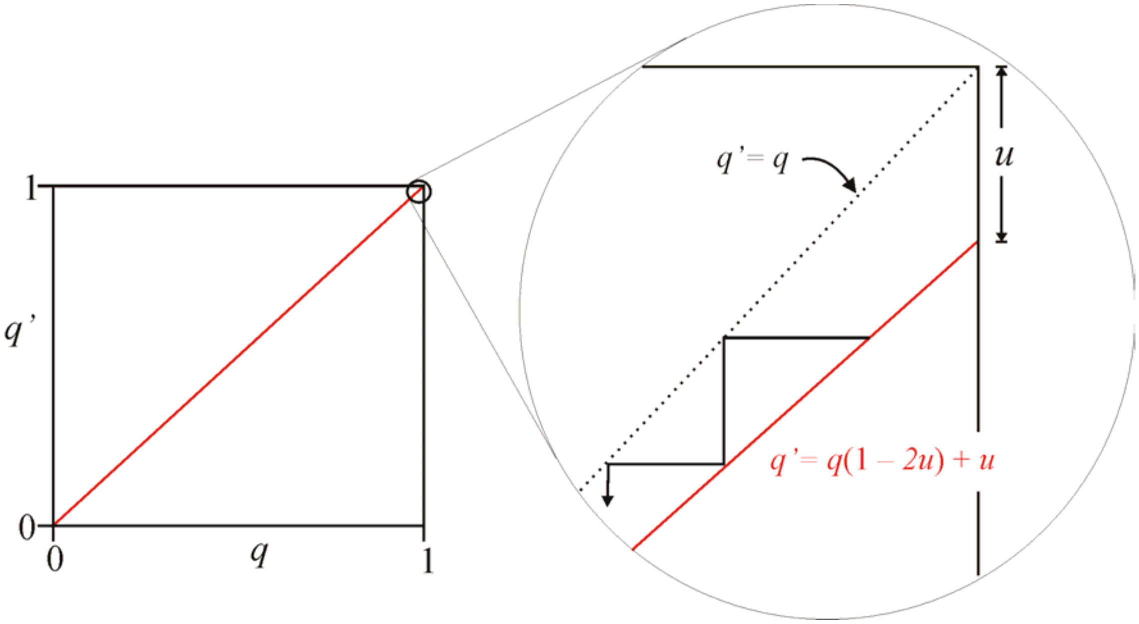

Figure 3. The left hand figure plots the frequency of gene after reproduction, *q',* as a function of the frequency before reproduction *q.* As is shown in the right hand part of the figure, mutation tends to drive gene frequencies toward 0.5, but because the mutation rate, *u,* is very low, this happens very slowly.

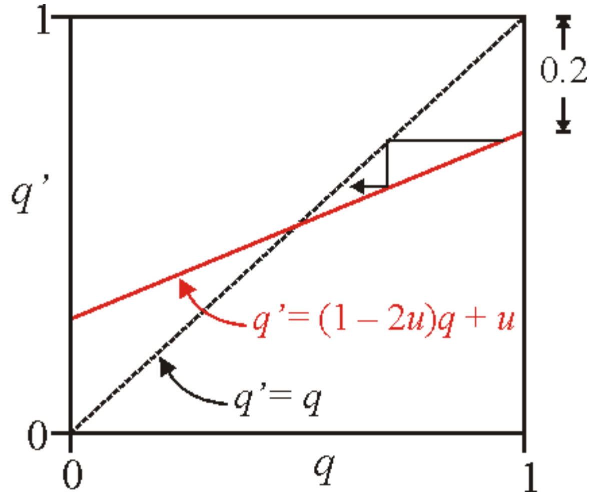

Figure 4 The frequency of representations after reproduction, *q'*, as a function of the frequency before reproduction *q* when the error rate, *u,* is high. As before, errors tend to drive representation frequencies toward 0.5, but now very rapidly because the error rate is large.

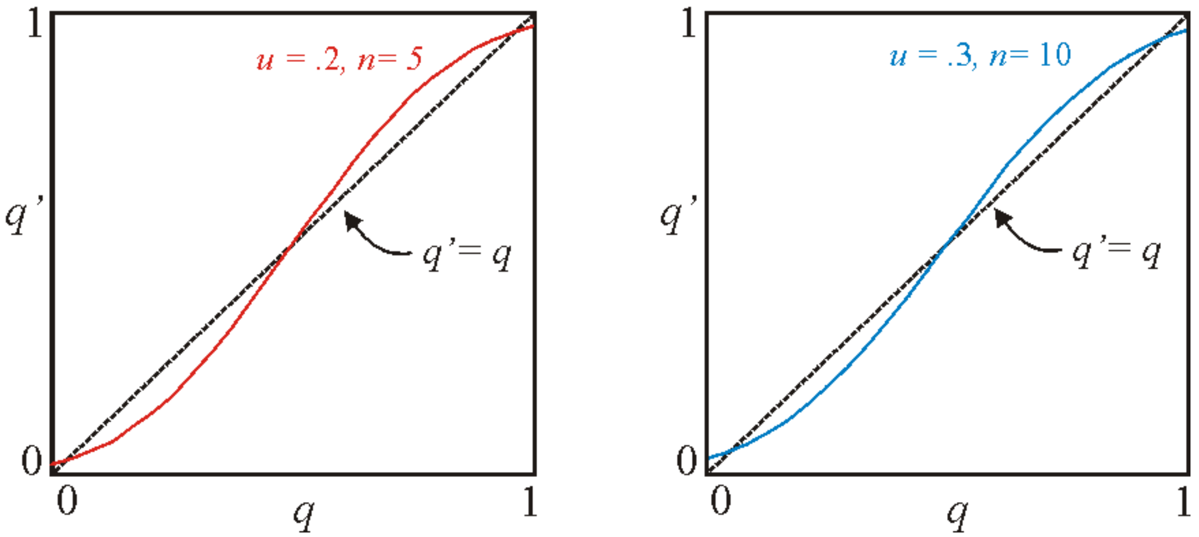

Figure 5. The post transmission frequency of a cultural representation as a function of the pre-transmission frequency when the error rate is high but there is a conformist bias in the the cultural transmission rule. In the left panel (a) the error rate is 0.2 and then number of models is 5, while in the right panel (b) the error rate is 0.3 and the number of models is 10.

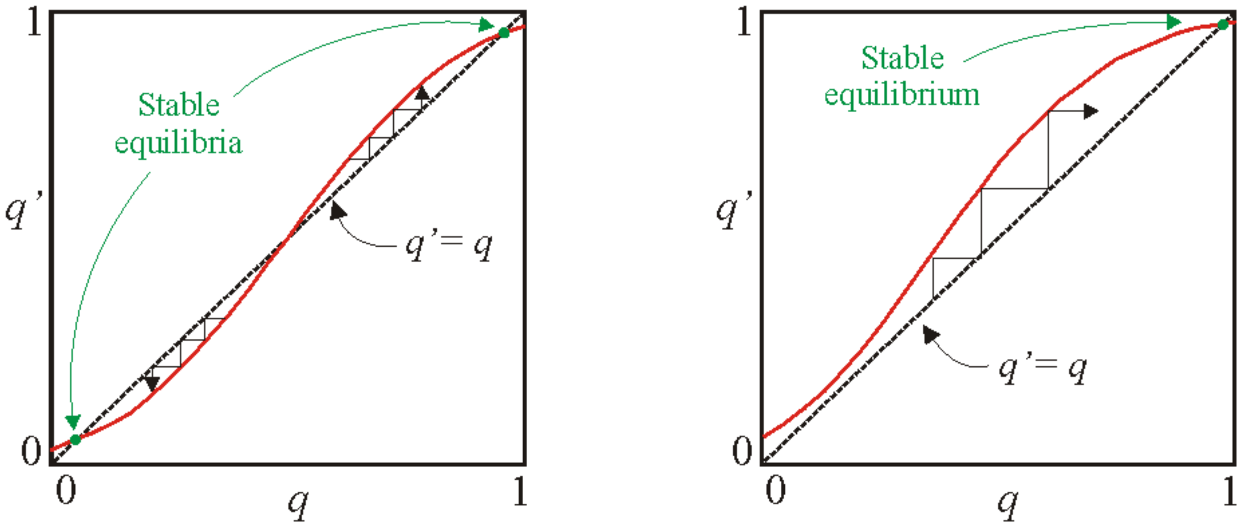

Figure 6 (a) The combination of frequency dependent bias and very high error rate ($u = 0.3$) acts to preserve variation between groups because common representations tend to increase in frequency. (b) When more successful individuals are more likely to be imitated, the more successful variant spreads.

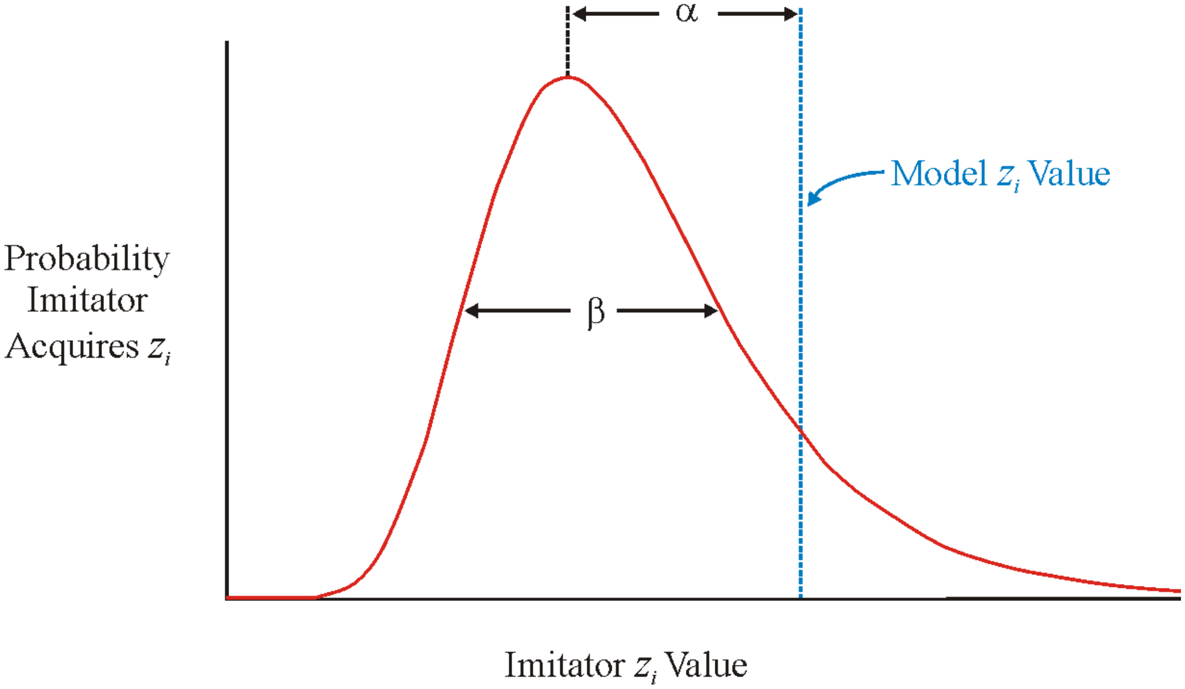

Figure 7. A Graphical representation of the assumptions of model 3. Learners observe a model with a behavior *zi*, given by the blue line. They infer a value drawn from the probability density function plotted in red. Their most likely value is *a* less than the models value, and since larger values of *z* are better, this represents systematic error. However, there is some probability that individuals acquire larger values.

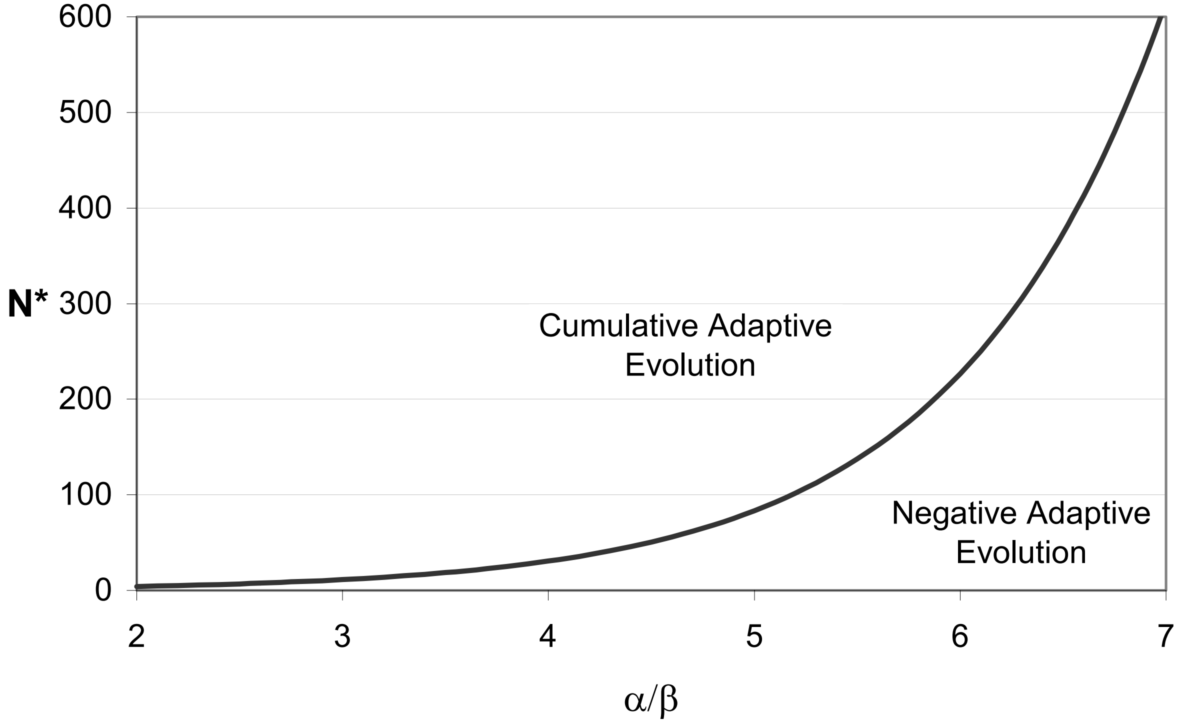

Figure 8 Cumulative adaptation occurs when population sizes are large, and the ratio of the systematic error to the unsystematic error in the inferential process is small.