

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

Joseph Henrich

University of Michigan Business School

701 Tappan Dr.

Ann Arbor, MI 48019-1234

henrich@umich.edu

(734) 763-0370

9 October 1999

Abstract

In challenging the pervasive model of individual actors as cost-benefit analysts who adapt their behavior by learning from the environments, this paper analyzes the temporal dynamics of both environmental (individual) learning and biased cultural transmission processes by comparing these dynamics with the robust “S-shaped” curves emerged from the diffusion of innovations research. The analysis shows three things: 1) that environmental learning alone never produces the *S*-shaped adoption dynamics typically observed in the spread of novel practices, ideas and technologies; 2) that biased cultural transmission always produces the *S*-shaped temporal dynamics; and 3) that a combination of environmental learning and biased cultural transmission can generate *S*-dynamics, but only when biased cultural transmission is the predominate force in the spread of new behaviors. These findings suggest that biased cultural transmission processes are much more important to understanding the diffusion of innovations and sociocultural evolution than is often assumed by most theorists. [diffusion of innovations, cultural transmission, learning, cultural evolution]

Efforts to understand human behavioral change have produced a multiplicity of different approaches. Most of these, whether they come from anthropologists, economists, sociologists or political scientists, share one common core element: individuals select among alternative behaviors by performing benefit/cost analyses using payoff-relevant information (i.e. data about the costs and benefits). Laying aside the hyper-rational, omniscient beings of classical economics formations,¹ more plausible approaches model individuals as goal seekers with limited computational abilities and incomplete information, who rely on trial & error learning, experimentation, and long experience in similar environments to achieve locally-effective solutions (Earle 1997; Netting 1993; Harris 1979; Camerer & Tek 1997; Young 1998; Erev & Roth in press). Using data from the vast diffusion of innovations literature, I argue that human behavioral change does not result primarily from individual-level, trial & error learning or cost/benefit analysis. Instead, I show that the dynamics of diffusion demand a primary reliance on some form of biased cultural transmission.

The intuitively persuasive model of human behavior that pervades the social sciences proposes that individuals acquire and evaluate *payoff-relevant* information about alternative behavioral options by action and interaction in their local social, economic and ecological environments. The adjective “*payoff-relevant*” emphasizes that the information analyzed by individuals is directly applicable to evaluating behavioral alternatives, according to some set of prescribed goals. Such goals may involve concepts like self-interest, reproductive fitness, social prestige, income or group-benefits, etc. Here, I argue against this standard model by showing three things: (1) environmental learning models alone, without substantial contributions from biased cultural transmission, do not generally produce the empirical “S-shaped” cumulative adoption curves that dominate the diffusion of innovations literature; (2) biased transmission models alone, and especially those with a *conformist transmission* component, consistently produce the particular S-dynamics found throughout the literature; and (3) a combined model, with both environmental learning and biased cultural transmission, allows us to predict the conditions that produce the different kinds of empirically-observed diffusion dynamics, and only generates S-dynamics when

biased transmission predominates.

S-shaped adoption curves

One of the most robust findings from over 3,000 studies in the diffusion of innovation literature is the *S*-shaped cumulative adoption curve (Rogers 1995: 23). This vast literature contains data for the spread of an enormous variety of practices, technologies and ideas in communities and countries throughout the world. These cases include the adoption of innovations such as hybrid corn among Iowa farmers, bottle-feeding practices among impoverished third-worlders, new governance practices among Fortune 500 companies, chemical fertilizers in peasant communities, novel approaches to teaching mathematics (the “new math”) in secondary schools, and the practice of not smoking among Americans. Typically, the cumulative adoption curve for the spread of these practices has an *S*-shape. For example, Figure 1 shows the *S*-curve that emerged from Ryan & Gross’ classic study (1943) of the spread of hybrid corn in two Iowa farming communities—this general shape captures the temporal dynamics encountered in a wide range of diffusion studies.

However, not all adoption curves are *S*-shaped. Of the small fraction of curves that are not *S*-shaped, most display a single alternative shape, which I will call an *R*-curve. *R*-curves lack the slow growth during the initial portion of the spread, which characterizes *S*-curves (the bottom-left part of the curve in Figure 1). Instead, *R*-curves begin at their maximum rate of growth (at $t = 0$), and then slowly taper off towards equilibrium (see Figure 3). Coleman et. al. (1966), for example, found that *R*-curves describe the cumulative adoption dynamics for the spread of the practice of prescribing Tetracycline, among both “interconnected” and “isolated” doctors. *R*-curves are also characteristic of a variety of non-social learning processes in which individuals acquire increasing proficiency in some skill or ability through practice (see Jovanovic & Nyarko 1995).² As I will discuss, the ‘combined model’ of environmental learning and biased transmission produces both *S*- and *R*-curves, depending on the parameter values.

To explore the relative importance of environmental (individual) learning vs. biased cultural

transmission, I have analyzed the dynamics of three models: a generalized individual-learning model, a biased cultural transmission model, and a combined model. This combined model integrates the first two models and allows us to compute the relative contribution of each to the *S*- and *R*-adoption dynamics.

The environmental learning model

Figure 2 graphically depicts a simple, though quite general, model of environmental learning for two behavioral traits.³ In the typical diffusion of innovation's situation, tracking only two traits is sufficient to capture the essential process. In this model, *Trait 1* represents the presence of the novel trait (the 'innovation'), while *trait 2* indicates the absence of the trait. If we are, for example, studying the spread of a new nitrogen fertilizer, an individual possesses *trait 1* if he uses the fertilizer, and possesses *trait 2* if he does not use the fertilizer. The symbol q gives the frequency of individuals with *trait 1* in the population, while $(1-q)$ gives the frequency of individuals with *trait 2*. The normal curve in Figure 2, with mean μ and variance σ^2 , shows the distribution of relative payoff information provided by the environment. Individuals may acquire this information through observation, experience, interaction and/or experimentation in the environment. During each time cycle (a fixed time period), individuals receive one draw from this normal distribution. This single draw provides a measure of the difference in payoffs (X) between the two alternative behaviors. However, people just don't switch to a novel behavior based on one piece of information, unless the suggested difference (i.e. X) is sufficiently large. How large this value of X needs to be depends on the quality of environmental information available, which is captured by μ and σ^2 , and on the individual's 'threshold of evidence'—which is parameterized by d . If, for example, the X drawn during a given cycle falls between $-d$ and $+d$, the individual stays with their previous behavior (from the previous time cycle). However, if the X drawn exceeds d , then the individual switches to *trait 1*. If they already possess *trait 1*, they stick with it. If X falls below $-d$, then the individual switches, mistakenly, to behavioral *trait 2*, or retains it if they already have it. This is a 'mistake' because the situation depicted in Figure 2, by the fact that $\mu > 0$, indicates that behavioral *trait 1* is superior in the current environment. Superior means that *trait 1* brings higher payoffs, on average,

relative to whatever individuals want, strive for, or hope to maximize.

To illustrate this phenomena, suppose a farmer, who currently plants wheat variety *A*, decides to plant a small patch of his land with a novel wheat seed, variety *B*, as an experiment.⁴ This experimental patch provides our farmer with a single measure of average yield (in kilograms of wheat harvested per hectare, for example), which he can compare against his average yield for variety *A*. The difference between the yield per hectare for the experimental patch and the average yield per hectare for variety *A* provides a value of X —an observed difference in payoffs between the two varieties. If the yield from variety *B* is about the same as, or less than that from variety *A* (implying $X < d$), then our farmer does not change from variety *A*. However, if the yield from variety *B* is sufficiently greater than the yield from *A* ($X > d$), then our farmer switches and sows only variety *B* in the following year.

Now, let's derive the population dynamics for the spread of *trait 1* into a group in which everyone currently possesses behavioral *trait 2*. As mentioned earlier, q represents the frequency of individuals in the population who have adopted the novel trait (*trait 1*). Initially, $q = 0$, but with each time cycle we update the value of q . The new value of q , in the next time cycle, is represented as q' (which reads “ q prime”). Applying the individual learning model described above and depicted in Figure 2, we arrive at the following recursion:

$$q' = P_1 + Lq \quad (1)$$

The new frequency of individuals with *trait 1*, q' , depends on P_1 , L and q . P_1 is the probability of learning the new trait from environmental information obtained during this time cycle. Restated, it is the probability that the payoff differences observed between the two behaviors exceeds the threshold of evidence (d)—it's also the gray area under the curve on the right side of Figure 2. L , is the probability that the environmental information is inconclusive, represents the area between $-d$ and $+d$ under the curve (in Figure 2). Individuals who receive inconclusive information will stick with their current behavior. Both P_1 , and L are derived from d via the normal distribution shown in Figure 2—Appendix A outlines this derivation. By iterating equation (1) recursively through successive time cycles, we can plot its temporal

dynamics and the cumulative adoption curves that it generates (see Figure 3).

For those readers who, like me, are interested in longer-term cultural evolutionary processes—in which the frequency of different ideas, beliefs, values and practices may change over many generations—we can interpret equation (1) in a slightly different way. During each time cycle, or perhaps each generation, naïve individuals (those who do not currently possess a particular behavior) acquire environmental information about the relative payoffs of alternative behaviors. If the difference in payoffs is clear (that is, if X is greater than $+d$ or less than $-d$), then individuals adopt the behavior indicated by their environmental information. However, if X falls between $+d$ and $-d$, individuals rely on *unbiased* cultural transmission or simple imitation. This means that individuals either copy their parents (also termed “vertical transmission”), or someone at random from the population. At the population-level, *unbiased* transmission simply replicates the distribution of behaviors found in the preceding generation. Boyd and Richerson (1985, 1988) call this cultural evolutionary model, which combines *unbiased* transmission and individual learning, *guided variation*.

Just as the above environmental learning model formalizes the cost/benefit model held by many social scientists, *guided variation* captures the fundamental processes that many economically-oriented anthropologists believe underlies much of sociocultural evolution. For example, while Harris maintains that, “As a species we have been selected for our ability to acquire elaborate repertoires of socially learned responses...” (1979: 62), he believes that sociocultural evolution is driven by individuals opportunistically selecting among cultural/behavioral variants according to their benefit/cost ratios. Obviously, the second assertion about benefit/cost ratios can only be true if the apparent social learning abilities of humans do not substantially bias the intergeneration transmission of cultural/behavioral variants. Consequently, Harris’ position, and that of many anthropologists, that sociocultural evolution results from unbiased social learning plus opportunistic benefit/cost analysis (environmental learning) is exactly what *guided variation* attempts to formalize. Having formalized this idea, we can now better analyze its evolutionary dynamics. If the empirical data shows that cultural transmission biases do substantially affect the frequency of alternative cultural/behavioral variants from one generation to the

next,⁵ then Harris' approach fails to capture an important component of sociocultural evolution.

Using data from the diffusion of innovation literature, we can address the applicability of the environmental learning and guided variation models. If the evaluation of costs and benefits, based on environmental information, is the dominant force in the spread of novel practices, then empirically-observed cumulative adoption curves should reveal the basic *R*-shape generated by equation (1) and shown in Figure 3. Interestingly however, most adoption curves constructed from empirical data have the *S*-shape shown in Figure 1, not the *R*-shape seen in Figure 3. An examination empirical *S*-curves tell us that the change in *q* over each time cycle must first increase to a maximum point somewhere in the middle of the *S*, and then begin decreasing toward zero. Computing Δq , the change in *q* over each time cycle, we get:

$$\Delta q = q' - q = P_1 - q(1-L) \quad (2)$$

Note that both P_1 and $(1-L)$ are positive constants, so Δq *must* decrease as *q* increases. Consequently, equation (1) will *never* produce an *S*-shape. Although *R*-curves do occasionally pop up in the diffusion literature, *S*-curves are, by far, the dominant shape of the temporal dynamics. Therefore, either this general environmental learning model somehow fails to capture the logic of humans as cost-benefit analysts (maintained by many social scientists), or that humans are *not* primarily individual learners doing cost-benefit analysis. Later, after I have completed presenting the basic *biased* cultural transmission model and the combined model, I will modify this environmental learning model and add the assumption that individuals vary in their degree of 'innovativeness.' As you will see, this modification does not change the basic results just derived.

The Biased Cultural Transmission Model

Instead of assuming that individuals acquire novel traits by figuring things out on the basis of pay-off relevant information, a substantial amount of empirical work from throughout the social sciences suggests that humans rely on social learning or cultural transmission to acquire the majority of their behaviors (Tarde 1903; Miller & Dollard 1941; Bandura 1977; Boyd & Richerson 1985; Cavalli-Sforza &

Feldman 1981; see Henrich 1999 for summary). However, people don't simply imitate random things from random people. Here, I'll describe three categories of biased transmission: *direct bias*, *prestige-bias*, and *conformist bias*. Under *direct bias*, people copy ideas or practices with specific qualities, regardless of who possesses them. The practice of purchasing and using cooking oil, for example, spreads rapidly even through remote villages—far from the reach of advertising—because there is something about the behavior or idea that appeals to people (Boyd & Richerson 1985). Other times, people copy ideas or practices from *individuals* with specific qualities or attributes, regardless of the characteristics of the behaviors or ideas that are copied. Gil-White and I (1999) have demonstrated that people will copy a wide range of traits from prestigious or successful people, even when the behaviors, ideas or opinions have nothing to do with the person's prestige or success. We call this process *prestige-biased transmission*. Americans, for example, will use a certain type of cologne, or even shave their heads, if Michael Jordan does (or they believe he does), despite that fact that Jordan's scent and hairstyle are probably not connected to his basketball prowess, prestige and overall success. Finally, under *conformist transmission*, humans preferentially imitate ideas and behaviors that are expressed by a majority of the group, over traits expressed by the minority, even when their personal opinions or behavior will *not* be known by the other group members (Baron et. al. 1996; Insko et. al. 1985; see Henrich & Boyd 1998 for theoretical treatment).

Equation (3) formalizes *biased* cultural transmission and was derived using basic replicator dynamics (Weibull 1995; Boyd & Richerson 1985). As in equation (1), q represents the frequency of individuals with the novel behavioral trait (*trait 1*), and q' is the frequency of individuals with *trait 1* in the next time cycle.

$$q' = q + (1 - q)q(r_1 - r_2) = q + q(1 - q)B \quad (3)$$

The term $(r_1 - r_2)$, or simply B , ranges from -1 to 1 and represents the overall difference in the replicatory propensities of *traits 1* and *2*. These r 's can each include the influence of *prestige-bias*, *conformist-bias* and *direct-bias*. For now, we will leave *conformist transmission* out, and assume that B

aggregates only the effects of prestige and direct bias. We will also assume that these biases are not a function of either q or time.⁶ Later, I'll incorporate terms for *conformist transmission* (which is frequency dependent) and examine its influence on diffusion. The reader should be aware that equation (3) are not some special case of cultural transmission, but a general form for any replicator process. It has been independently re-derived for a variety of purposes in a number of different fields—including economics (Gintis 2000, Weibull 1995), genetics (Hartl & Clark 1989), epidemiology (Waltman 1974), and cultural transmission (Boyd & Richerson 1985; Cavalli-Sforza & Feldman 1981; Bowles 1998).

Figure 4 presents four cumulative adoption curves generated using equation (3) for different values of B . Note the similarity between the empirical curve in Figure 1 and the curves in Figure 4. In fact, the different S -shapes captured in Figure 4 resemble a wide range of the empirical adoption curves found in the diffusion of innovations literature. This similarity suggests that cultural transmission models may capture an important component of human behavioral change.

Combined Model: learning + biased transmission

So far, I have contrasted two quite different models of human cognition and information processing. However, it seems both intuitively and empirically true that humans do both cultural transmission and environmental learning. That is, we do some imitating and some figuring things out on our own. Theoretically, I have developed this idea using computer simulations that modeled the biological evolution of the parameter d —which, as I discussed earlier, determines an individual's degree of reliance on environmental learning vs. cultural transmission (Henrich & Boyd 1998). Under a wide range of conditions, in both spatially- and temporally-varying environments, this theoretical work suggests that our reliance on environmental learning is a small, but important, component of human adaptive behavioral plasticity. Consequently, the question becomes: how much biased cultural transmission must be added to the environmental learning (or guided variation model) to generate the empirically-observed S -curves? Or, what is the predominate force in human behavioral change?

To address this, we combine equations (1) and (3). However, because simply substituting

equation (1) into (3) gives a slightly different answer than substituting (3) into (1),⁷ an additional step is required. We assume that, during each time cycle, not everyone attempts individual learning and/or biased transmission. Instead, only a fraction of the population updates their behavior based on one of these two sources of information. For environmental learning, the symbol ξ represents the fraction of individuals in the population that consider updating their behavior via environmental learning per unit time. This can be thought of as the update rate for environmental learning, or as the probability of using environmental learning in each time cycle. Similarly, the symbol γ represents the fraction of individuals in the population that update with biased cultural transmission per unit time. In both cases, Δt represents one unit of time, or one time cycle. Therefore, $\xi\Delta t$ provides the fraction of individuals who consider updating with environmental learning in each time cycle, while $\gamma\Delta t$ gives the fraction who deploy biased transmission in a given time cycle. Applying this additional step to equation (1) yields equation (5):

$$q'_L = q(1 - \xi\Delta t) + (P_1 + Lq)\xi\Delta t = q + \xi\Delta t(P_1 + (L-1)q) \quad (5)$$

Applying the same process to equation (3) yields equation (6):

$$q'_T = q(1 - \gamma\Delta t) + \gamma\Delta t(q + q(1 - q)B) = q + \gamma\Delta t(q + q(1 - q)B) \quad (6)$$

Since we want to arrive at the derivative of q with respect to time, so we substitute equation (5) into (6) [or (6) into (5)], solve for $\Delta q/\Delta t$, and then take the limit of $\Delta q/\Delta t$ as Δt approaches zero. This gives us:

$$\frac{dq}{dt} = \xi(P_1 - (1 - L)q) + \gamma Bq(1 - q) \quad (7)$$

An examination of the typical S -curve suggests the general form of dq/dt . The rate of change of the frequency of the novel trait must first ascend to a peak, somewhere in the middle of the S , and then decline to zero. Figure 5 plots dq/dt for both the combined model (CM) and for the environmental learning model (ENLR). For the environmental learning model, the maximum value of dq/dt occurs at $q = 0$, so no S -curve is generated. For the biased cultural transmission model (BCT), the maximum value of dq/dt (the middle of the S) occurs at $q = 0.50$ (curve not shown). For the combined model (without

conformist transmission), the maximum value of dq/dt occurs below $q = 0.50$. When this maximum value occurs between $q = 0$ and 0.50 , some form of S is produced. When this maximum value occurs below $q = 0$, the part of the dq/dt curve between $q = 0$ and 1 looks just like the environmental learning curves. Consequently, an R -curve, not an S -curve, is produced. To visualize this, imagine sliding the CM curve, shown in Figure 5, to the left until the maximum value of dq/dt drops below zero.

If we take the derivative of (7) with respect to q , set it equal to zero, and solve for q , we get an expression for the value of q when dq/dt is maximized.

$$q_{\max} = \frac{1}{2} - \frac{(1-L)\zeta}{2B\gamma} \quad (8)$$

In order to produce the S -shape, mathematically speaking, q_{\max} must be greater than zero. So, solving equation (8) for $q_{\max} > 0$ yields:

$$B > \frac{\zeta(1-L)}{\gamma} = \phi(1-L) \quad (9)$$

Because ζ and γ are both update rates, we can simplify (9) by defining $\phi = \zeta/\gamma$, where ϕ represents a ratio of the fraction of the population that updates via environmental learning to the fraction that updates via biased cultural transmission. If the update rates are equal, then $\phi = 1$; if people update their behavior more frequently using environmental information, then $\phi > 1$; if people use cultural transmission more frequently, then $\phi < 1$. Figure 6 graphs the S and non- S regions (R -regions) of B and L . This plot shows that in order to consistently produce S -curves either B , the replicatory bias created by the trait (or the individual(s) possessing the trait) must be big, or L , the degree to which humans rely on cultural transmission over individual learning, must be big.⁸ I discuss this in greater detail at the end of the paper.

Figure 7 shows the effect of moving the value of ϕ away from one. Increasing the update rate of environmental learning relative to cultural updating (i.e. changing ϕ to 1.2) moves the B -intercept (at $L = 0$) up to 1.2 , which shrinks the S -region—which is the plot area to the right of the curve. Conversely, increasing the cultural update rate relative to environmental learning (decreases ϕ to 0.8) moves the B -intercept down to 0.8 and expands the S -region. Depending on the details of a particular diffusion

situation, one might argue that is greater than or less than one, but its actual value will be difficult to measure empirically, because it depends partly on human psychology and partly on environmental constraints. For much of the coming discussion I will assume that $\phi = 1$. For empirical purposes, it's best to incorporate γ and ζ in B and $(1-L)$, respectively.

When biased transmission opposes environmental learning

So far, we've considered only the situation in which both biased transmission and individual learning favor the spread of the novel trait. In this section, we explore the temporal dynamics of the diffusion of a novel trait when individual learning successfully spreads the novel trait against the force of biased transmission. That is, the bias favors the other, initially common, trait when B is negative. In the next section, we analyze the opposing case, in which biased transmission spreads a novel trait in the face of individual learning. This occurs when environmental information indicates that a trait is not beneficial, but transmission biases spread it anyway. Exploring these two situations, we can ask which set of dynamics more closely matches the empirically-observed temporal dynamics of trait diffusion.

Figure 8 shows the adoption curves for five different sets of parameters (B , L , and P_1). As either L or B increases (B is negative), the equilibrium value of q rises, and the curves ascend more quickly. However, nothing remotely resembling an S -curve emerges. Equation (9) tells the same story. The right-side of equation (9) is always positive (or zero) and B is always negative in this case, so condition (9) is never satisfied and S -curves do not emerge. Given that S -curves are empirically rampant in diffusion contexts, the situations in which individual learning overpowers biased cultural transmission to spread a beneficial trait seem relatively rare. Two possible explanations present themselves. One suggests that our database is somehow biased against these kinds of diffusions, so they only seem rare. The second is that L is large—meaning that biased transmission is the predominate component of human cognition. In the future, researchers should look for diffusion cases in which trial and error learning clearly favors one trait, but transmission bias favors another (e.g., only low status people initially adopt).

When environmental-learning opposes biased transmission

What do the curves look like when biased transmission swims upstream against environmental learning? Examining equation (9) alone suggests that *S*-curves may or may not be generated, depending on the values of *L* and *B*. This condition holds even when environmental information does not favor the novel trait. Because the bias (*B*) must overcome environmental learning, we can use equation (9) to set *B* to its maximum value that still produces *R*-curves, $B = \phi(1-L)$. If we substitute this into equation (7), set it equal to zero, and solve for *q*, we get equation (10):

$$q_{eq} = \sqrt{\frac{P_1}{P_1 + P_2}} \quad (10)$$

This is the equilibrium frequency of the novel trait when an *R*-curve is produced by biased transmission flowing upstream against individual learning. Remember, in this case P_1 is less than P_2 because environmental information opposes the spread of the trait. Consequently, the frequency produced by equation (10) is small; it's the probability of selecting the "wrong" behavior (i.e. the one not favored by environmental information). Typically, this equilibrium value is so small that it would never 'count' as a diffusion. Figure 9 shows eight curves for differing values of *B*, *L*, and P_1 that illustrate the basic point. As a consequence, all substantial diffusions driven by biased transmission diffusions generate *S*-curves, not *R*-curves. Curve 8, the only *R*-curve on Figure 9, shows the case when $B = \phi(1-L)$. The equilibrium frequency of Curve 8 is 0.32.

Conformist transmission and long tails

Thus far, we have ignored conformist transmission. However, Figure 1 displays an interesting feature that suggests another form of biased cultural transmission—*conformist transmission*—may also be at work. Note the slow growth of *q* during the initial stages of the diffusion process—I call this slow growth a 'long tail' (see notation on Figure 1). It took nine years for the frequency of hybrid planters to reach 0.20, but only six more years for it to reach fixation at 0.99. In an effort to account for this recurrent phenomenon of long tails, we can incorporate a simple conformist component into the existing model,

and then examine its effects on the temporal dynamics of adoption.

So far, we have dealt with B , the replicatory or transmissive bias on the novel trait, as a constant in any particular situation, not as a function of time or frequency. Now B has two components, a constant part and a frequency-dependent part, which are shown in equation (12):

$$B = b(1 - \alpha) + \alpha(2q - 1) \quad (12)$$

The second term in (12), $\alpha(2q-1)$, is the component of the overall bias contributed by conformist transmission. The symbol α , which varies between zero and one, gives the relative strength of conformist transmission in human cognition—it scales the cognitive weight given to the frequency of a behavior *relative to* other biases. Generally, it's best to consider α small because when α is large few, if any, traits can spread—for example when $\alpha > 0.5$ nothing rare ever spreads. The term $(2q-1)$ varies between -1 and 1 . When the frequency of the novel trait is low (less than 50%) this conformist component is negative, which reduces the value of the total bias (B), and may actually make it negative (depending on the relative sizes of the other components). When $q > 0.50$, this conformist term increases the overall size of the bias. The other term, $b(1-\alpha)$, is the contribution to the overall bias made by direct bias and non-frequency-dependent prestige biases. The symbol b is the constant bias, while its complement $(1-\alpha)$ gives the weight accorded to the non-conformist component of the transmission bias. Substituting (12) into (7) yields:

$$\frac{dq}{dt} = \xi(P_1 + (L-1)q) + \gamma q(1-q)\{b(1-\alpha) + \alpha(2q-1)\} \quad (13)$$

Using this expression we can follow the same procedure as before to derive the conditions when (13) generates S -curves. Note the similarity between equations (14) and (9).

$$b > \frac{\zeta(1-L) + \gamma\alpha}{\gamma(1-\alpha)} = \frac{\phi(1-L) + \alpha}{(1-\alpha)} \quad (14)$$

Figure 10 illustrates the curves for equation (14) when $\phi = 1$. As the strength of conformist transmission increases, the region of b and L values that generates S -curves shrinks—remember, S -curves begin appearing as one moves to the right of, or above, the curve. Consequently, if conformist transmission is even a small component of human psychology, we should expect either: 1) that all the

various values of b represented throughout the diffusion of innovations literature are quite high; or 2) that the value of L in human psychology is substantial—otherwise R -curves would be more common.

Back to the question I asked at the beginning of this section: can conformist transmission account for the long tail observed in Figure 1? Figure 11 shows the temporal dynamics for a series of α values, ranging from zero to 0.27. By comparing Figure 1 and Figure 11, we observe that conformist transmission does generate the long tails observed in some empirical data. Assuming α is fairly small, such tails occur when the biases generated by the non-conformist components of our cultural capacities are relatively weak ($b \cong \alpha$). When these biases are large ($b \gg \alpha$), the effect of α almost disappears.

More generally, this slow growth period is a common feature of many adoption curves. Rogers (1995: 259-260) explains that potential adopters initially seem resistant to new ideas until a “critical mass” is achieved and the diffusion process “takes-off.” This intuitive explanation supports the idea, formalized in conformist transmission (Boyd & Richerson 1985), that individuals use the frequency of a trait as an indirect indicator of its worth. Hence, the frequency of a trait inhibits its spread when rare, but encourages the trait’s diffusion once it becomes common.

Conformist transmission can also help in predicting the take-off points described by applied diffusion researchers. In attempting to actively spread novel innovations, governments, states and organizations will sometimes provide “pump-priming” incentives to adopters, often in the form of direct cash payments, until the innovation spreads past some critical frequency, often thought to lie between 20% and 30%. Once this threshold is reached, the innovation is considered self-sustaining, which means that it will continue to spread on its own. If we assume that the contribution of environmental learning to diffusion is negligible (i.e. L is big), then we can derive a simple expression for the take-off frequency using equation (12). If we don’t assume L is big, then environmental learning will always spread beneficial traits, and take-off points should not exist.⁹ Note, the empirical existence of take-off points supports both the claim that L is big and that conformist transmission is real, but small. With this assumption, the diffusion process becomes self-sustaining when:

$$B = b(1 - \alpha) + \alpha(2q - 1) > 0 \quad (17)$$

A diffusion process is not self-sustaining when the magnitude of the conformist component of B , $| \alpha(2q - 1) |$, exceeds $b(1 - \alpha)$, thereby making the overall bias less than zero. Remember, the conformist component is negative when $q < 0.50$. Solving equation (17) for the take-off frequency, q_p , requires setting $B = 0$, and solving for q . At this point B crosses over from negative to positive values. This yields:

$$q_p = \frac{1}{2} - \frac{b(1 - \alpha)}{2\alpha} \quad (18)$$

Equation (18) tells us two things. First, if it exists, the take-off frequency lies between zero and 0.50. And second, if $b(1 - \alpha)/2\alpha > 0.5$, then the process will never be self-sustaining. Empirical data indicates that pump-priming incentives do often work (but not always), and critical points always seem to lie between zero and 0.5 (Rogers 1995).

Modifying the environmental learning model still won't produce S-shapes

A great deal of diffusion research has adopted the intuition that the diffusion dynamics, including the *S*-shape, result from differences among individuals in their degree of “innovativeness” or their fear of uncertainty. For example, Rogers (1995: 258) writes:

Many human traits are normally distributed, whether the trait is a physical characteristic, such as weight or height, or a behavioral trait, such as intelligence or the learning of information. Hence, a variable such as the degree of innovativeness is expected also to be normally distributed.

The idea is that a few individuals with a high degree of innovativeness adopt early, most people adopt somewhere in the middle, and a few stragglers, with low innovativeness, adopt late. Although it may be true that individuals vary in their degree of innovativeness,¹⁰ building this into the environmental learning model does not produce the anticipated *S*-dynamics, as I will demonstrate. Furthermore, I have already shown that *S*-dynamics can be produced without assuming people are different (also see Cavalli-Sforza & Feldman 1981). All the models so far have assumed people are psychologically and socially identical, yet they still produce cumulative logistic curves—i.e. *S*-curves under a wide range of conditions.

We can construct environmental learning models that incorporate individual variation in two ways: 1) assume individuals do environmental learning first, and then, if they remain uncertain, they rely on *unbiased* transmission (copy someone at random)—this provides a transgenerational model; or 2) assume that individuals do repeated trials and that the dynamics of learning are fast relative to an individual's lifetime (or that individual's live forever). In the afore-described environmental learning model, the parameter d (see Figure 2) represents an individual's threshold of evidence or their willingness to proceed under uncertainty. Innovative individuals are those willing to adopt a new trait based on limited evidence, under uncertainty. Thus, this parameter captures what many researchers mean by innovativeness. Following the standard diffusion approach—to classify people into adopter categories—I will define five types of individuals: Innovators, Early Adopters, Early Majority, Later Majority and Laggards (Rogers 1995:262). The subscript i indexes these categories 1 to n ($n = 5$ in this case). Each category i is characterized by its own value, d_i . Innovators have the smallest value of d and Laggards have the largest value of d . Each value of d_i generates, via the cumulative normal distribution shown in Figure 2, corresponding values of P_i and L_i . For the first version of the model (with unbiased transmission), the frequency of the novel trait among members of category i (e.g. Early Adopters) in the next time cycle is shown in (15).

$$q'_i = P_i + qL_i \quad (15)$$

Further, assume the symbol F_i represents the proportion of the total population that adopter category i comprises. For example, if 10% of the population are Laggards, then $F_5 = 0.10$. To find the new frequency of the novel trait in the overall population, we compute the expected value of equation (15).

$$q' = E(q'_i) = \sum_i F_i(P_i) + \sum_i F_i(L_i q) = \bar{P}_1 + q\bar{L} \quad (16)$$

Equation (16) demonstrates that, when individuals vary in their innovativeness, the cumulative adoption curves depend only on the average values of P_i and L_i (regardless of their distribution). This means that equation (16) behaves just like equation (1), and therefore, does not produce S -curves.

In the second version of individual variation, instead of unbiased transmission, I assume the dynamics of learning are fast relative to the lifetime of individuals. Thus, we get equation (17):

$$q'_i = P_{1i} + q_i L_i \quad (17)$$

Note, the only difference between equations (17) and (15) is the subscript i on q . This occurs because, instead of copying someone at random from the population every time period, these long-lived individuals simply stick with their current behavior—that is, the environment does not provide sufficiently convincing data to justify a change. Taking the expectation of q'_i , to get q' , we arrive at equation (18):

$$q' = E(q'_i) = E(P_{1i}) + E(L_i q_i) = \underbrace{\bar{P}_1 + q\bar{L}}_{NS} + COV(q_i L_i) \quad (18)$$

Does equation (18) produce S -curves? We have already seen that the terms labeled ‘NS’ in equation (18) will never produce an S -curve. The final term, $COV(q_i L_i)$, is the covariation between q_i and L_i , which varies for different values of q (or over time). At $t = 0$ (and $q = 0$), $COV = 0$. In Figure 12, although the frequency of *trait 1* rises for each of the subgroups in the population (q_i), the subgroup with the smallest value of d_i learns the novel trait most quickly—note that the different values of q_i can be observed at the points where the vertical line crosses the different curves (which have different values of d_i). The initially rapid adoption of the trait by more innovative individuals (those with lower d values) generates a negative covariation between L_i and q_i . This negative association remains until the curves cross one another in the middle of Figure 12. After this crossover, the $COV(q_i L_i)$ crosses through zero and stabilizes at a positive equilibrium value. These dynamics for covariation remain robust because more innovative individuals adopt novel behaviors more rapidly, but achieve lower equilibrium values of q_i than less innovative individuals. Lower equilibrium values occur because more innovative individuals are subject to more erroneous switchbacks, as their standard of evidence for changing behaviors is lower—it’s the price of innovativeness.¹¹ The dynamics of $COV(q_i L_i)$, when added to the standard R -shaped curves produced by the NS terms in equation (18), will never produce an S -curve.

Many diffusion researchers believe the S -shaped cumulative adoption curve to be the product of

an underlying, normally-distributed “time-to-adoption” curve that captures the varying degrees of innovativeness distributed throughout the population. In this view, “time-to-adoption” acts as the inverse of innovativeness. When researchers test their empirically-derived $\Delta q/\Delta t$ (dq/dt) curves for deviations from normality, sometimes they pass (and cannot be distinguished from a normal distribution), and sometimes they do not. When such curves do not pass the normality test, researchers claim that they “approach normality.” For example, using the Iowa farmer data (Figure 1), researchers went to great lengths to show the data were normally distributed—their efforts even included getting time on a supercomputer. Yet, they still failed to show normality, because of the distribution’s long tail. However, from the perspective I have presented here, there’s no *reason* to expect underlying normality. Often equation (13) does produce time-to-adoption distributions that look approximately normal,¹² but knowing if they are approximately normal or not does not tell us anything more about the underlying social-decision processes. For example, the time derivative of a logistic curve (its probability density function) looks quite normal, and would certainly appear normal if one sampled from it. More importantly, equation (13) can also produce underlying, non-normal, time-to-adoption distributions that are much more similar to that produced by the diffusion of hybrid corn or of Tetracycline, than any normal distribution.

Many efforts to fit the *S*-dynamics of the diffusion literature have been made, especially in the marketing and new product literature. For a long time, researchers have recognized that logistic curves in various forms can fit many of the *S*-curves fairly well. Unfortunately, the parameters in these functional forms have little meaning because such curve-fits lack any *a priori* theoretical foundation in human psychology or decision-making (Bass 1969). However, some researchers have managed to construct environmental learning models (similar to the one in this paper), in which individuals vary in their degree of risk aversion, that under some conditions will generate logistic *S*-curves (Jensen 1982; Kalish 1985; Oren & Schwartz 1988). Although these models can produce *S*-curves, based on individual differences in risk aversion and Bayesian learning processes, the circumstances that produce the *S*-dynamics depend critically on the initial distribution of beliefs in the populations, the specific shape of the utility curves, and the details of the information-gathering processes. In Oren and Schwartz’s model (1988), for

example, deriving the logistic form depends on assuming both constant proportional risk aversion ($xU''(x)/U'(x) = \text{constant}$), and that risk aversion is exponentially distributed across the population. No empirical justification for either of these rather narrow assumptions is provided. Without any empirical support, it is difficult to believe that these assumptions are as robust across the world's populations as are the S -dynamics of diffusion. Similarly, under some conditions, the environmental learning model in this paper will produce S -curves if innovative individuals are assumed to acquire or process information better than less innovative people. However, getting an S depends on exactly how innovativeness and information processing abilities are distributed across the population.

Some readers may criticize this analysis because they realize that a wide variety of mathematical formulations of environmental learning or rational calculation could generate S -curves, and I have not begun to exhaust the possible formulations. This is true, however, merely having equations with the symbols arranged in a particular fashion is not a sufficient riposte. In my view, the trick is to formulate a learning model that's evolutionarily plausible, empirically grounded, tied directly to individual psychology and produces S -curves under a wide range of general conditions. I hope skeptics who favor environmental learning will endeavor to generate and test such competing models.

Discussion and Summary

Many scholars have the intuition that cultural transmission is, at best, a minor force in human behavior and behavioral change (Tooby & Cosmides 1992; Stigler & Becker 1978; Pinker 1997; Harris 1979; Buss 1999). However, if cultural transmission is merely a weak component of the psychological processes that generate human behavior—meaning L is fairly small—then we would expect the real world, and the diffusion of innovations literature, to contain a large proportion of R -curves relative to the proportion of S -curves. If people have small L values, S -curves should result only when the replicatory bias (B) is quite high. Remember, B is generated by the qualities of the trait itself (e.g. eating high fat foods or believing in a good god), or by the qualities of the trait's possessors (i.e. their local prestige or success). So, the rest of the time, when B is medium or low, environmental learning should generate only

R-curves. However, in real world (or at least in the available empirical data), *R*-curves are relatively rare, while *S*-curves are rampant. This suggests that biased cultural transmission dominates the diffusion process, and that *L* must be pretty big—or somehow hundreds of researchers studying everything from the spread of insecticides among Colombian peasants to the diffusion of “poison pills” among Fortune 500 companies must have systematically biased the database and selected only traits with very high biases values (*B*).

Further evidence for a substantial reliance on cultural transmission comes from the spread of maladaptive or costly behavioral traits. My analysis indicates that maladaptive traits may spread against the force of individual learning—to produce an *S*-curve—as long as *L* and *B* are sufficiently large. For example, the practice of bottle-feeding infants spread throughout the third world despite the fact that this inappropriate practice produces higher rates of sickness, infection and death in infants under third-world conditions than does breast-feeding (Rogers 1995). Such costly, maladaptive practices abound in the anthropological literature (Edgerton 1992). In many societies, food taboos restrict the consumption of nutritionally-valuable foods (Descola 1994; Wilbert 1993; Baksh 1984). Even in places where protein and dietary fat are limited, people still refuse to eat valuable nutritional resources. The Machiguenga of the Peruvian Amazon, for example, would not consider eating snake meat, even when the dead snake is known to be non-venomous. Similarly, the Warao, who inhabit the extremely marginal environs of the Orinoco river delta, refuse to hunt large mammals (which include some of the most valuable animal resources in South America) because they “have blood like people” (Wilbert 1993: 18). Furthermore, nearly half of all cultures surveyed throw out the valuable colostrum that precedes mother’s milk, and which helps infants develop their immune systems while providing essential minerals (Morse et. al. 1990). Without the predominance of biased transmission, it would be difficult to explain the prevalence of costly, maladaptive traits in populations throughout the world. Remember, environmental learning models, like Oren & Schwartz’s model, predict that only beneficial, utility-maximizing and/or adaptive traits will spread through populations, but that is not all we observe.

On the flip side, if our reliance on biased transmission were weak (if *L* were small), then

environmental learning would frequently spread beneficial traits against the tide of negatively biased cultural transmission. However, my analysis indicates that we should record an *R*-curve every time our cost-benefit analysis overcomes our social learning tendencies. Yet, *R*-curves are rare, so biased transmission is most likely a substantial component of human behavioral plasticity.

Finally, how can environmental, cost-benefit learning account for the empirical phenomena of long-tails and take-off points? Why do diffusion processes sometimes begin so slowly and finish up so rapidly? Why doesn't this occur other times? Why do some behaviors have threshold adoption frequencies at which they begin spreading on their own (without paying people for adoption), even when the behavior later turns out to be a bad idea? As I've described, the simple models of biased transmission presented in this paper can account for all these phenomena, but it remains quite unclear whether cost-benefit learning approach can be modified to account for them as well.

What kind of information flows through social networks?

Many social scientists believe that by diffusing 'information,' social networks generate the classical diffusion dynamics. Rogers writes, "they [diffusion networks] convey information to decrease uncertainty about a new idea" (1995:281). By using the term "innovation-evaluation information" Rogers captures what I described earlier as 'pay-off relevant' information, which is the essential ingredient in the cost-benefit model. Although the biased cultural transmission processes I've modeled here do involve the transfer of information among individuals, this imitation process does *not* directly involve the transmission of innovation-evaluation information—that is, information used by individuals to evaluate the costs and benefits of alternative practices. Biased imitation involves copying an idea or practice for reasons not directly related to its costs and benefits. Despite the intuitions of many people, the available empirical data supports the kinds of imitation processes I have described, and not innovation-evaluation hypothesis.

For example, in *prestige-biased transmission*, individuals copy traits possessed by prestigious individuals, regardless of how these traits affect the success of the prestigious model or the copier (Gilwhite & Henrich 1999). Generally, the enormous importance of what diffusion researchers call "opinion

leadership” confirms the theoretical predictions of prestige-biased transmission. For example, the same farming practice will spread rapidly in places where the locally high prestige individuals favor the novel idea, but entirely fail to spread in other places where the prestigious individuals dislike the novel practice. Similarly, Van den Ban (1963, from Rogers 1995) effectively demonstrates the importance of prestige-biased transmission over the evaluative information processing in his study of farmers in the Netherlands. He shows that small-scale farmers copied the farming practices of prestigious, large-scale farmers *even* when such practices were clearly inappropriate for their particular situation.

Like prestige-biased transmission, conformist transmission does not depend directly on the costs and benefits of alternative behaviors, but still seems to be an important component of adoption dynamics. Besides the long tails and take-off points observed in many diffusion curves, *conformist transmission* can also account for the spatial or socio-spatial clustering of traits frequently observed in the diffusion literature. For example, in studying the spread of contraceptive methods in rural Korean villages, Rogers & Kincaid (1981) found that choices clustered by village. There were “pill villages,” IUD villages,” and even “vasectomy villages.” All these contraceptive methods were being promoted equally by the government campaign, and each village contained individuals with differing degrees of wealth and social standing. Cost-benefit analyses, environmental learning, and most kinds of direct biases can neither generate nor maintain such patterns. Eventually, given any social connection between villages (which Rogers & Kincaid did clearly observe), the contraceptive method with the highest bias or greatest benefits/ratio should spread to all villages. Or, if all methods were somehow exactly equal in benefits and costs or direct biases, then we would expect these methods to scatter across the social landscape, and not cluster in village networks. In contrast, *conformist transmission* predicts socio-spatial clusters of similar traits any time the differences between the costs and benefits or the biases of alternative practices are relatively small. Similar patterns of innovation clusters were observed by Whyte (1954) in his study of the spread of air-conditioning units in Philadelphia.¹³

Before concluding, I’d like to point out that a great deal of empirical work has been done on the characteristics of “innovators” and “early adopters”—those who adopt early in the diffusion processes. At

first glance, these patterns are convincing. According to this work, early adopters tend to have larger social networks, higher status, more money, more cosmopolitan contacts, and more exposure to mass media outlets. The assumption seems to be that these characteristics (causally) increase an individual's likelihood of adopting an innovation early on in the diffusion process. Unfortunately, the literature's focus on successful diffusions produces an extremely biased database. The only situations included in the database involve those in which the trait actually did spread; in contrast, all those times when the trait did not spread are *not* included. So, the more accurate empirical claim would be: early adopters tend to have larger networks, higher status, etc. *given that the trait eventually spreads to high frequency*. It's quite possible that all individuals, regardless of their economic position, media exposure, etc., are *equally* likely to adopt an innovation early, but that the subsequent diffusion of an innovation depends on the characteristics of the initial adopters. Things like large social networks and high status may have nothing to do with an individual's chances of innovating, but they may be critical to the subsequent transmission of these traits. When poor, low-status individuals innovate nobody copies them, so the trait never diffuses, and they never get into the database as 'innovators.'

Future work could turn the diffusion problem on its head and explain why certain societies, particularly peasant groups, seem slow or resistant to the spread of novel behavioral traits, ideas and 'innovations.' As well, this work could address how cultural transmission mechanisms, under certain circumstances, can produce upper-middle-class conservatism or the 'Cancian-dip' (Cancian 1979). Finally, such work could use diffusion data from a wide variety of sources and numerical computer simulations to estimate parameter distributions for L , P , b , and α .

Endnotes

¹ Recently, a vast amount of work in cognitive psychology and experimental economics has severely criticized the extreme, hyper-rational models of classical economics (e.g. Gigerenzer & Goldstein 1996; Rabin 1998; Kagel & Roth 1995; Henrich 1999; Kahneman et. al. 1982); consequently, many economists and other students of human behavior are increasingly turning to cognitively more-realistic models of human learning and decision-making.

² These curves also describe the spread of milk-bottle opening behaviors among pigeons (Lefebvre & Giraldeau 1994).

³ Throughout this paper, I use 'behavioral trait' or simply 'trait' to stand for a whole range of things that could be termed 'innovations,' 'cultural traits,' 'practices,' 'beliefs,' 'ideas' and/or 'values'.

⁴ This kind of experimentation is common in both traditional and modern agricultural systems—see Johnson 1972 and Rogers 1995, respectively.

⁵ By 'substantially,' I mean an effect of the same order of magnitude or larger than the cost/benefit effect.

⁶ Prestige-biases may be either constant or frequency-dependent (dependent on q), depending on whether the frequency of the transmitted trait in the population significantly affects the success, payoffs or prestige of the trait's possessors. In this paper, I do not incorporate frequency-dependent payoffs, but many transmission models have built this in using evolutionary game theory (Henrich & Boyd 1999; Bowles 1998; Gintis 1999).

⁷ This occurs because whichever equation is second in the life cycle (meaning whichever one gets substituted into) exerts a small bias on the final result. It's a sampling bias that favors the most recent recursion.

⁸ Figure 6 and equation (9) provide the minimum mathematical conditions to produce an *S*-curve. However, for humans to discern an *S*-shape in the curve, q_{\max} should be set at 0.1 or more. This shrinks the region of B and L that generates *S*-curves, thus making the argument stronger.

⁹ Furthermore, if we don't assume L is big, then equation (12) yields a cubic equation in q , which can be solved, but does not yield any useful insights. We also assume α small, because otherwise nothing spreads when L is big.

¹⁰ Dewees and Hawkes (1988) found that particular commercial fishermen could not be generally characterized as 'Innovators' or 'Laggards' in their study of six different fishing-related innovations. Their work shows that same individuals were not consistently early adopters, or consistently adopters at all.

¹¹ The mean (or variance) of the normal distribution shown in Figure 2 represents the quality of environmental information that is available to every individual—this variable tells us how difficult the problem is. In this model, I assume that everyone receives the same quality of information and has the same abilities to process this information.

¹² The time derivative of a logistic curve (its probability density function) looks quite normal, and would certainly appear normal if one sampled from it.

¹³ Admittedly, there are other explanations for this kind of clustering besides *conformist transmission*, including combinations of other types of cultural transmission mechanisms (see Boyd & Richerson 1985). Another possibility is that if the costs or benefits of an innovation were frequency-dependent, then once one method attains high frequency by whatever stochastic processes, it stays at high frequency. In some situations this hypothesis seems mildly plausible, but in other situations, like the spread of contraceptive methods, it's difficult to see the frequency dependence.

Appendix A

Derivation of P_1 , L and P_2 from Figure 2

P_1 , L and P_2 (from Figure 2) can all be related through the cumulative normal distribution using μ , σ^2 and d . If $F(\mu, \sigma^2, x)$ represents the cumulative normal distribution evaluated at x , then

$$P_2 = F(\mu, \sigma^2, -d - \mu)$$

$$P_1 = 1 - F(\mu, \sigma^2, d - \mu)$$

$$L = 1 - P_1 - P_2 = F(\mu, \sigma^2, d - \mu) - F(\mu, \sigma^2, -d - \mu)$$

Derivation of equation (3)

Equation (3) is a robust result of a variety of approaches to formalizing biased cultural transmission and replicator dynamics. Here I only outline the simplest derivation. More extensive treatments can be found in Gintis (2000), Boyd and Richerson (1985) and Weibull (1995).

This is a two-trait model. The symbol q tracks the frequency of individuals with *trait 1*, while $(1 - q)$ tracks the frequency of individuals with *trait 2*. Naïve individuals enter the world and acquire the trait of their parents. Later in life, as adolescents, they pick an individual at random from the population and compare the r -value of this individual's trait with the r -value of the trait they possess (which they acquired from their parents). The probabilities of switching traits or keeping the current trait are shown in Table 1 below.

Table 1. Probabilities of switching traits

Naïve's current trait	Model's trait	Probability of trait 1	Probability of trait 2
1	1	1	0
1	2	$\frac{1}{2}\{1+(r_1-r_2)\}$	$\frac{1}{2}\{1-(r_1-r_2)\}$
2	1	$\frac{1}{2}\{1+(r_1-r_2)\}$	$\frac{1}{2}\{1-(r_1-r_2)\}$
2	2	0	1

Remember, r -values—the replicatory propensities for each of the traits—contain two parts. The first part depends on the qualities of the trait itself, while the second part depends on the frequency of the trait in the current population. When a naïve individual encounters someone with a trait different from his

own, he quickly samples the population and uses this frequency assessment in his imitation decision. The r -values are described as follows:

$$r_1 = b_1(1 - \alpha) + \alpha(q - \frac{1}{2})$$

$$r_2 = b_2(1 - \alpha) + \alpha(1 - q - \frac{1}{2})$$

$$r_1 - r_2 = b_{12}(1 - \alpha) + \alpha(2q - 1)$$

Table 2. Frequency of Possible Pairings

Possible Pairings	Frequency of Pairings
trait 1- trait 1	q^2
trait 1- trait 2	$q(1-q)$
trait 2- trait 1	$(1-q)q$
trait 2- trait 2	$(1-q)(1-q)$

Using the frequency of each possible pairing (shown above), we can calculate the frequency of *trait 1* after this imitation process by multiplying the frequency (or probability) of each pairing by the probability of ending up with *trait 1*. We get the following recursion:

$$q' = q^2(1) + q(1-q)\frac{1}{2}(1+(r_1-r_2)) + (1-q)q\frac{1}{2}(1+(r_1-r_2)) + (1-q)(1-q)(0)$$

If we simplify this, we get equation (3).

$$q' = q + (1-q)q(r_1 - r_2) = q + q(1-q)B = q + q(1-q)\{b_{12}(1-\alpha) + \alpha(2q-1)\}$$

References Cited

1. Baksh, Michael G. 1984 Cultural Ecology and Change of the Machiguenga Indians of the Peruvian Amazon. University of California at Los Angeles.
2. Bandura, Albert. Social learning theory. 1977. Englewood Cliffs, N.J, Prentice Hall.
3. Baron, Robert, Vandello, Joseph, and Brunsman, Bethany. The forgotten variable in conformity research: impact of task importance on social influence. *Journal of Personality & Social Psychology* 71(5), 915-927. 96.
4. Bass, F. M. 1969. A new product growth model for consumer durables. *Management Science* 15:215-27.
5. Bowles, Samuel. Cultural Group Selection and Human Social Structure: The effects of segmentation, egalitarianism, and conformism. 1998 University of Massachusetts at Amherst.
6. Boyd, Robert and Peter J. Richerson
1985 *Culture and the Evolutionary Process*. edition. Chicago, IL: University of Chicago Press.
1988 An evolutionary model of social learning: the effects of spatial and temporal variation.
In Social Learning: Psychological and Biological Perspectives. Thomas R. Zentall and Bennett G. Galef, eds. Pp. 29-48. Hillsdale, NJ : Lawrence Erlbaum Associates.
8. Buss, David 1999 *Evolutionary Psychology: The New Science of the Mind*. Boston: Allyn & Bacon.
9. Camerer, Colin and Teck-Hua Ho forthcoming Experience-Weighted Attraction Learning in Normal Form Games. *Econometrica*

10. Cancian, Frank 1979 *The Innovator's Situation: Upper Middle Class Conservatism in Agricultural Communities.* edition. Stanford: Stanford University Press.
11. Cavailli-Sforza, Luca L. and Marcus Feldman 1981 *Cultural Transmission and Evolution.* edition. Princeton: Princeton University Press.
12. Coleman, James S., Elihu Katz, and Herbert Menzel 1966 *Medical Innovation: A Diffusion Study.* edition. New York: Bobbs-Merrill.
13. Descola, Philippe 1994 *In the Society of Nature: a native ecology in Amazonia.* New York: Cambridge University Press.
14. Dewees, Christopher M. and Glenn R. Hawkes 1988 *Technical Innovation in the Pacific Coast Trawl Fishery: The Effects of Fishermen's Characteristics and Perceptions on Adoption Behavior.* *Human Organization* 47(3):
15. Earle, Timothy 1997 *How Chiefs Come to Power.* edition. Stanford : Stanford University Press.
16. Edgerton, Robert B. *Sick societies : challenging the myth of primitive harmony.* 1992. New York, Free Press.
17. Erev, Ido and Alvin Roth. In press *Learning, Reciprocation and the Value of Bounded Rationality.* *In Bounded Rationality: The Adaptive Toolbox.* Gerd Gigerenzer and Reinhard Selten, eds. MIT Press.
18. Gigerenzer, Gerd and Goldstein, Daniel G. *Reasoning the fast and frugal way: models of bounded rationality.* *Psychological Review* 103(4), 650-669. 96.

19. Gintis, Herbert 2000 *Game Theory Evolving*. Princeton : Princeton University Press.
20. Harris, Marvin. *Cultural materialism : the struggle for a science of culture*. 1979. New York, Random House.
21. Hartl, Daniel L. and Andrew G. Clark 1989 *Principles of Population Genetics*. edition. Sunderland, Mass.: Sinauer Associates.
22. Henrich, Joe and Robert Boyd 1998 The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior* 19:215-242.
23. Henrich, Joseph. *Rationality, Cultural Transmission and Adaptation: the problem of culture and decision-making in Anthropology*. webuser.bus.umich.edu/henrich. 99.
24. Insko, Chester A., Smith, Richard H., Alicke, Mark D., Wade, Joel, and Taylor, Sylvester. *Conformity and group size: the concern with being right and the concern with being liked*. *Personality & Social Psychology Bulletin* 11(1), 41-50. 85.
25. Jensen, R. 1982 Adoption and diffusion of an innovation of uncertain profitability. *Journal of Economic Theory* 27:182-193.
26. Johnson, Allen 1972 *Individuality and Experimentation in Traditional Agriculture*. *Human Ecology* 1(2):149-159.
27. Jovanovic, B. and Y. Nyarko 1995 A bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity* 1247-305.

28. Kagel, John H and Roth, Alvin E. The handbook of experimental economics. 95. Princeton, N.J, Princeton University Press.
29. Kahneman, Daniel, Paul Slovic, and Amos Tversky 1982 Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
30. Kalish, S. 1985 New product diffusion model with price, advertising and uncertainty. Management Science 31:1569-85.
31. Lefebvre, Louis and Giraldeau, Luc-Alain. Cultural transmission in pigeons is affected by the number of tutors and bystanders present. Animal Behaviour 47(2), 331-337. 94.
32. Miller, N. E. and Dollard J. 1941 Social learning and imitation. edition. New Haven: Yale University Press.
33. Morse, J. M., C Jehle, and D. Gamble 1990 Initiating breastfeeding: a world survey of the timing of postpartum breastfeeding. International Journal of Nursing Studies 27(3):303-313.
34. Netting, Robert M. Smallholders, householders : farm families and the ecology of intensive, sustainable agriculture. 93. Stanford, Stanford University Press.
35. Pinker, Steven 1997 How the Mind Works. edition. New York: W. W. Norton & Company.
36. Rabin, Matthew 1998 Psychology and Economics. Journal of Economic Literature XXXVII:1-46.
37. Roger, Everett M. and D. L. Kincaid 1981 Communication Networks: Toward a new Paradigm for Research. edition. New York: The Free Press.

38. Rogers, Everett M. *Diffusion of innovations*. 95. New York, Free Press.
39. Ryan, Bryce and Neal C. Gross 1943 *The Diffusion of Hybrid Seed corn in Two Iowa Communities*. *Rural Sociology* 8:15-24.
40. Stigler, George J. and Gary S. Becker 1977 *De Gustibus Non Est Disputandum*. *American Economic Review* 67(1):76-90.
41. Tarde, Gabriel 1903 *The Laws of Imitation*. edition. New York: University of Chicago Press.
42. Tooby, John and Leda Cosmides 1992 *The psychological foundations of culture*. *In The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Pp. 19-136-xii, 666.
43. Waltman, P. 1974 *Deterministic Threshold Models in the Theory of Epidemics*. edition. Berlin: Springer-Verlag.
44. Weibull, Jorgen W. *Evolutionary game theory*. 95. Cambridge, Mass, MIT Press.
45. Whyte, William H. 1954 *The Web of Word of Mouth*. *Fortune* 50:140-143, 204-212.
46. Wilbert, Johannes 1993 *Mystic Endowment: Religious Ethnography of the Warao Indians*. edition. Cambridge, Mass: The President and Fellows of Harvard College.
47. Young, H. Peyton. *Individual strategy and social structure : an evolutionary theory of institutions*. 98. Princeton, N.J, Princeton University Press.



Frequency of Iowa Farmers adopting Hybrid Corn Seed

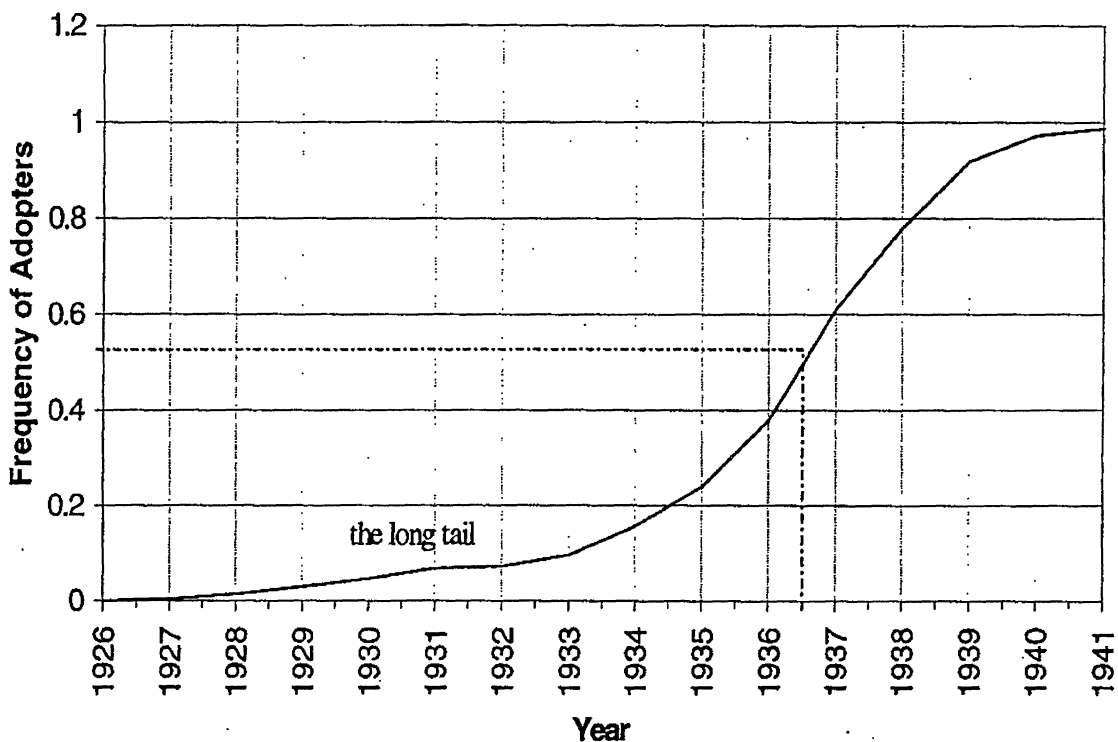


Figure 1. Diffusion of hybrid corn among two Iowa farming communities.

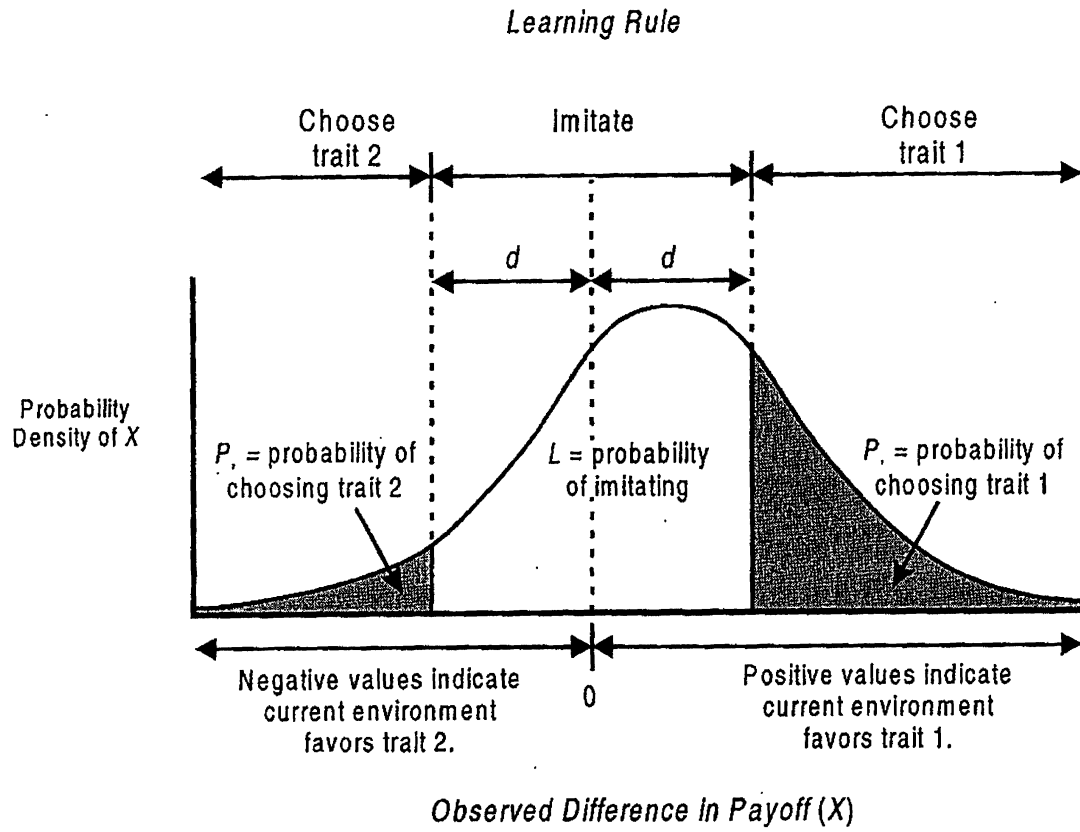


Figure 2. The Individual Learning Model.

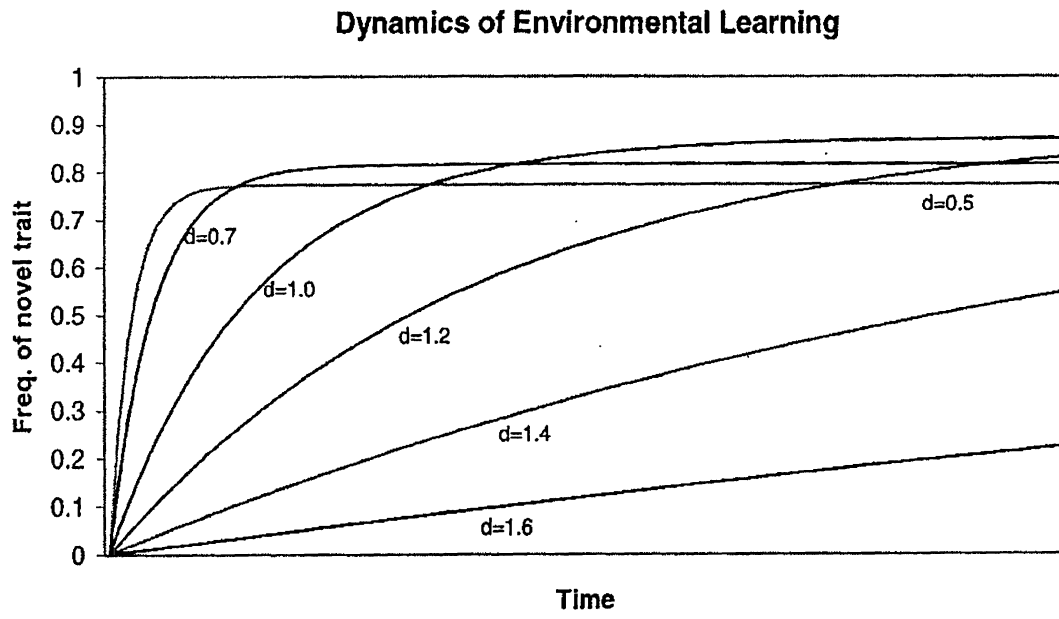


Figure 3. Environmental Learning R -curves for different values of d .

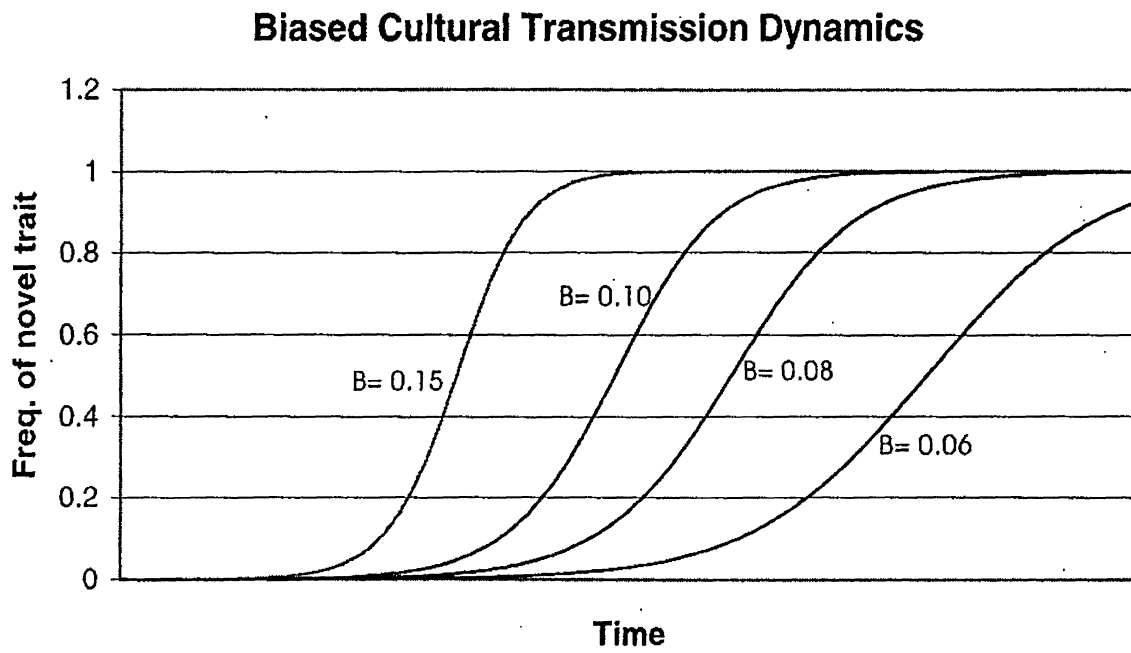


Figure 4. Biased Cultural Transmission Dynamics using four values of B

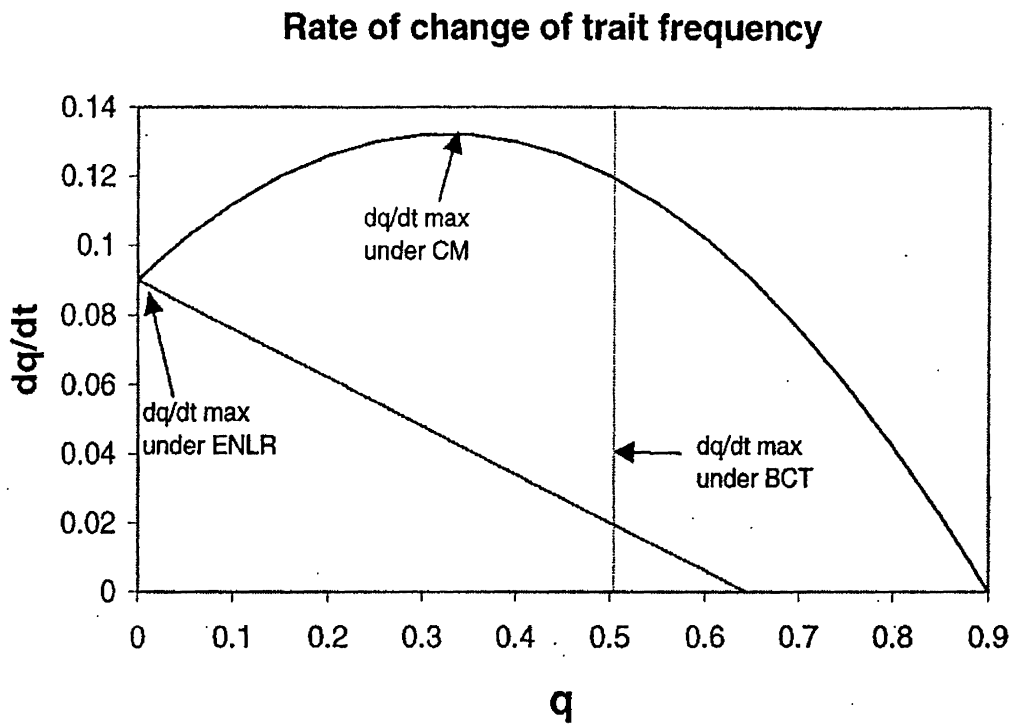


Figure 5. The rate of change of the frequency of the novel trait under the environmental learning model (ENLR) and under the combined model (CM).

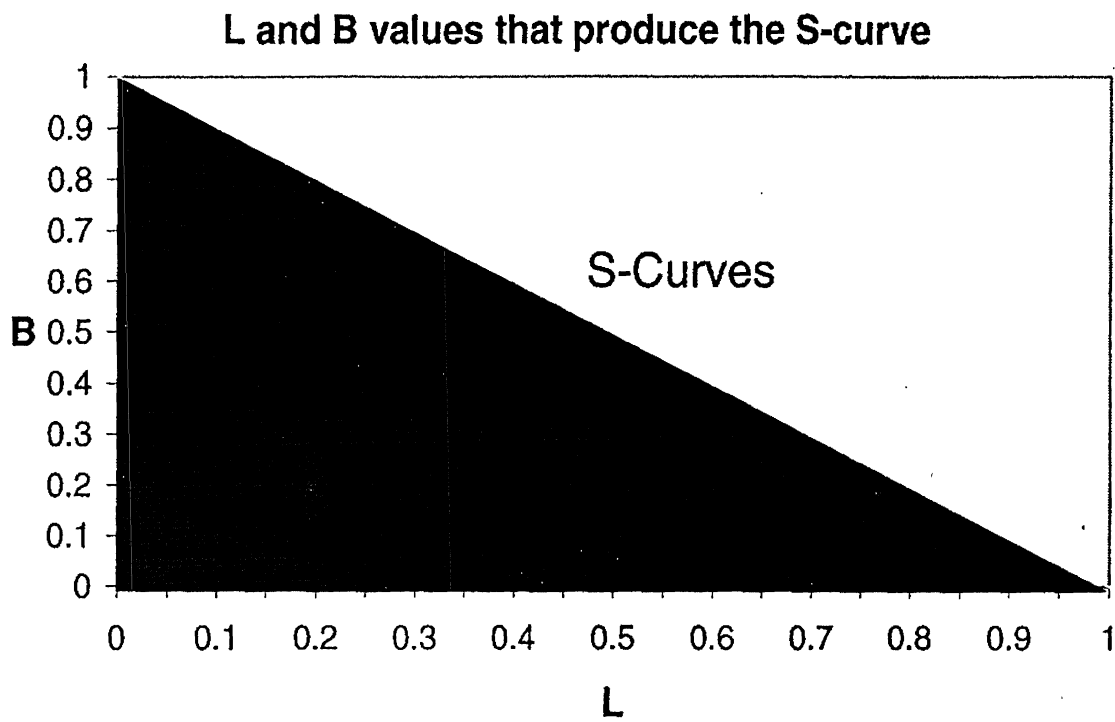


Figure 6. The regions of B and L that produce S -dynamics

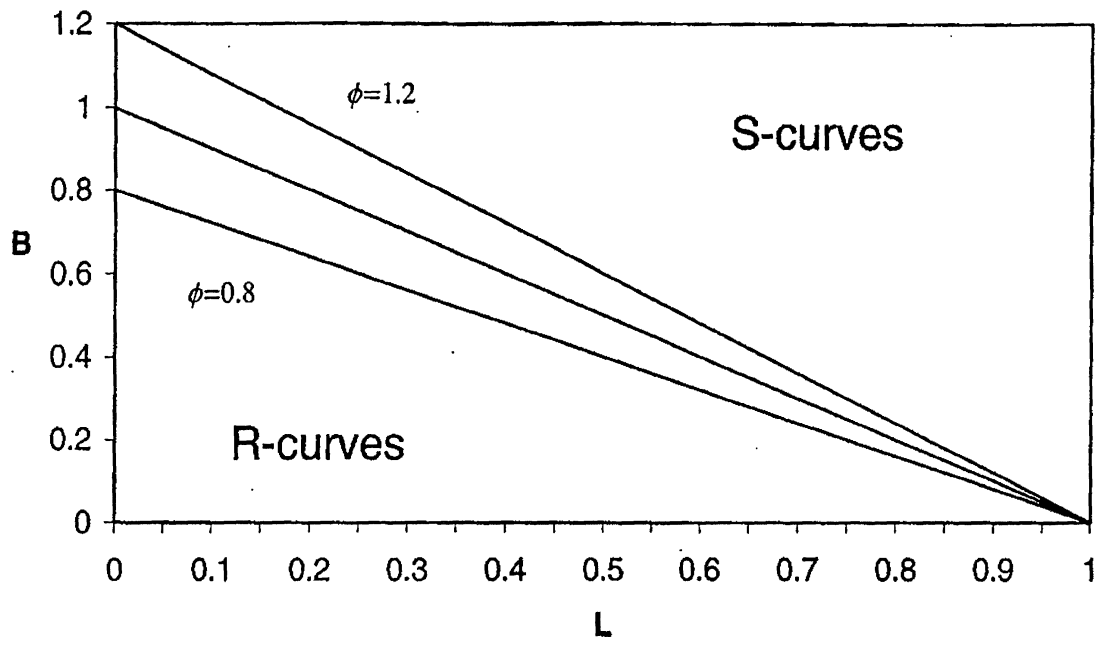


Figure 7. S-curve regions for different values of ϕ

Environmental learning overpowers biased transmission

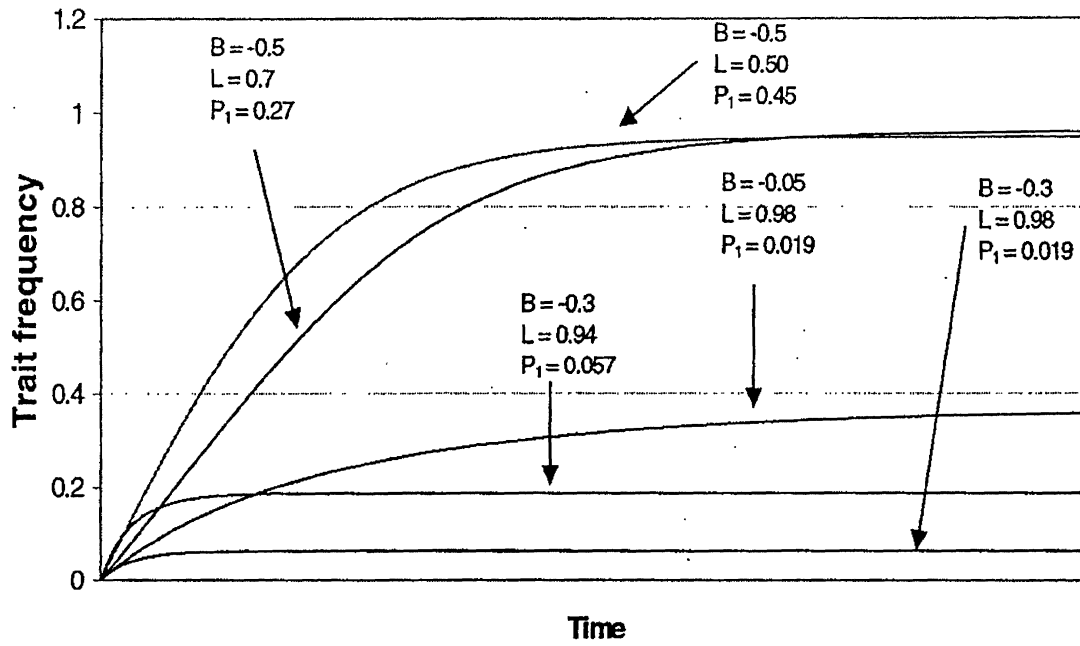


Figure 8. Temporal dynamics when individual learning spreads a novel trait against biased transmission for five sets of parameters

Biased Transmission spreads against Environmental Learning

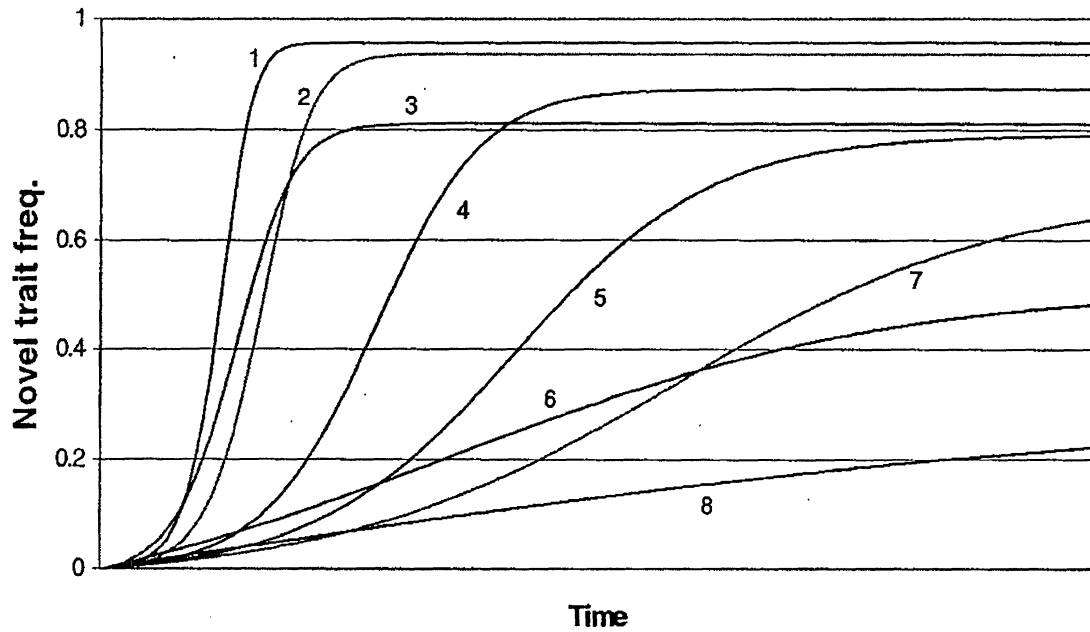


Figure 9. Temporal dynamics when biased transmission spread a novel trait against individual learning for 8 sets of parameters. Values of B , L and P_1 (respectively) for Curve 1: 0.45, 0.98, 0.001; Curve 2: 0.30, 0.98, 0.001 Curve 3: 0.30, 0.94, 0.003; Curve 4: 0.15, 0.98, 0.001; Curve 5: 0.90, 0.80, 0.01; Curve 5: 0.90, 0.80, 0.01; Curve 6: 0.5, 0.7, 0.03; Curve 7: 0.6, 0.8, 0.01; Curve 8: 0.2, 0.8, 0.02.

S-curve Regions with Conformist Transmission for $\phi=1$

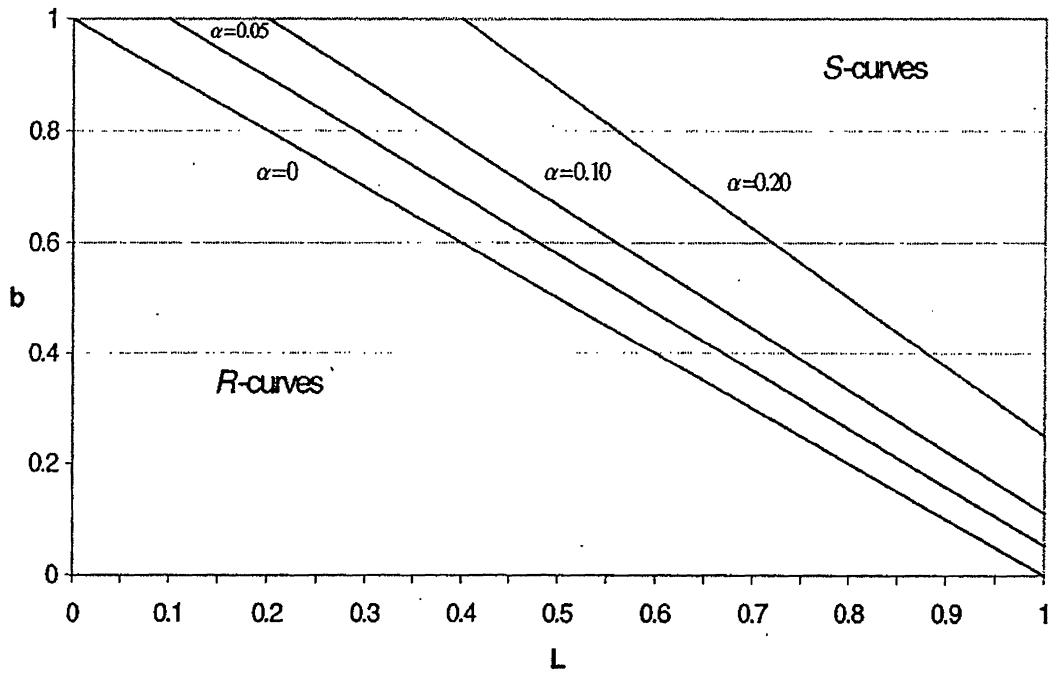


Figure 10. S-curve regions for different strengths of conformist transmission.

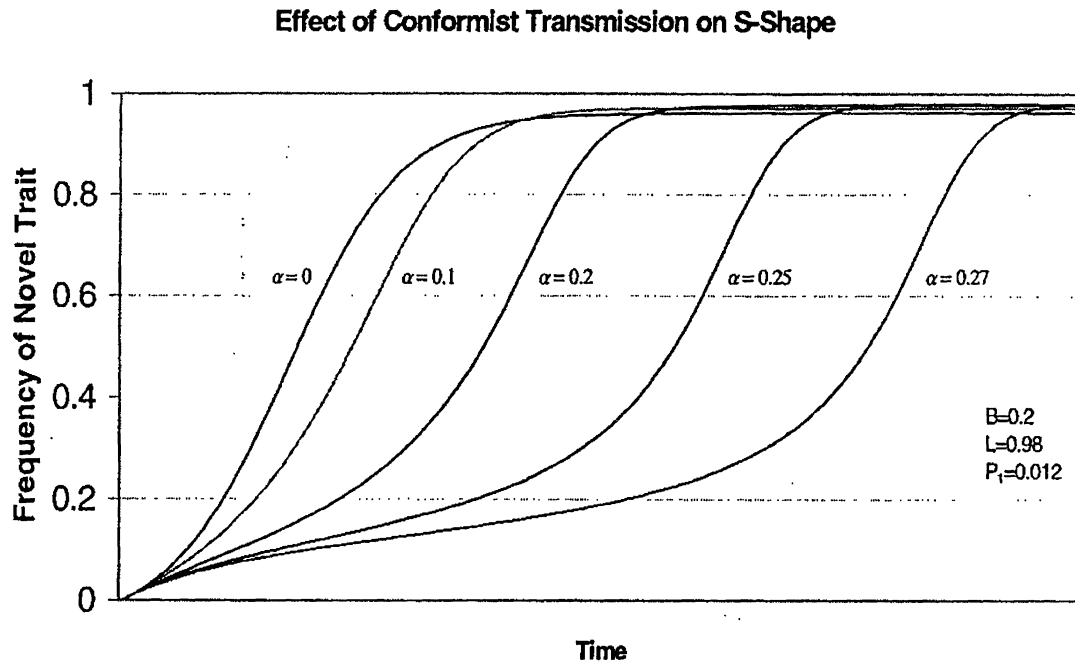


Figure 11. Examples of adoption curves for different values of α

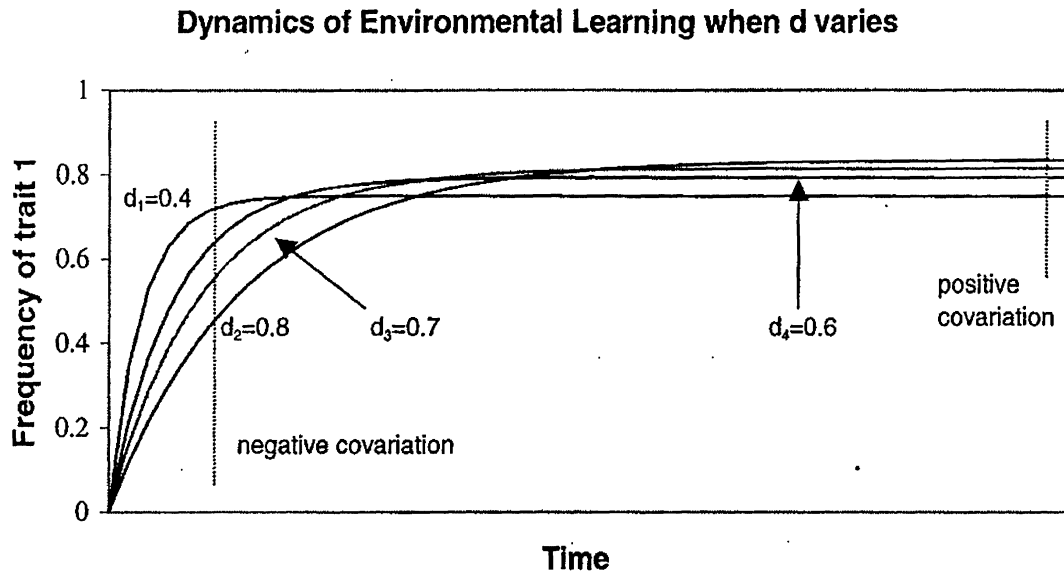


Figure 12. Environmental learning dynamics for four subgroups with differing degrees of innovativeness (d).