Rapid cultural adaptation can facilitate the evolution of large-scale cooperation

Robert Boyd[1], Peter J. Richerson[2], and Joseph Henrich[3]

[1]Department of Anthropology UCLA; [2]Department Environmental Science and Policy and Center for Population Biology, UC Davis; [3]Departments of Psychology and Economics, University of British Columbia.

Abstract: Over the past several decades we have argued that cultural transmission facilitates the evolution of cooperation because it leads to more rapid adaptation, and therefore to more variation among groups, compared to genetic adaptation. Recently, Lehmann, Feldman and colleagues have published several papers questioning our argument. They analyze models showing that cultural evolution can actually reduce the range of conditions under which cooperation can evolve, and interpret these models as indicating that we were wrong to conclude that culture facilitates the evolution of cooperation. Throughout, Lehmann and colleagues assume that rates of adaption are so weak that they do not substantially affect variation among groups. Lehmann et al reach different conclusions because they have made different assumptions, and therefore their work has little bearing on our conclusions unless their assumptions their assumptions are more appropriate for cultural evolution than ours. We argue that Lehmann, et al.'s assumptions, apparently made for analytical convenience, are not consistent with the empirical data on cultural learning and cultural variation in humans, and thus do not provide a useful analysis of the cultural evolution of human cooperation.

Most human societies cooperate on larger scales than do other mammalian social groups. In other mammals, cooperation is limited to very small groups, there is little division of labor, no trade, no large scale conflict, no social support for sick or disabled, no moral systems enforced by third parties. In stark contrast, even in simple foraging societies people regularly cooperate in large groups. Division of labor, trade, and large scale conflict are important nearly everywhere. The sick and disabled are often cared for, and social life is regulated by shared moral systems that specify the rights and duties of individuals enforced, albeit imperfectly, by third party sanctions.

In a series of publications, we have argued that the scale of human cooperation, along with its other unique features, is a consequence of adaptive cultural learning. This hypothesis rests on two claims:

1. Over the last million years or so, our lineage has evolved the psychological machinery capable of producing cumulative, cultural adaptation. This process permits human populations to develop highly refined adaptations, including social adaptations (e.g., norms and institutions), to local environments much more rapidly that is possible by genetic adaptation and individual learning.
2. Rapid cultural adaptation vastly increases heritable variation between groups. Systems of reciprocity and reputation can stabilize a vast range of heritable behaviors ranging from ruthless spite to prosocial cooperation. Conformist social learning rules can also stabilize many different behaviors. Rapid cultural adaptation can then lead to persistent differences between local social groups, and then competition between groups can favor the spread of behavior that enhance the competitive ability of groups. As a result cultural evolution can search multiple stable equilibria faster than genetic evolution.

We have developed a number of theoretical models that show that this argument is cogent, presented empirical data that suggests that the assumptions of the models are realistic, and reviewed empirical examples of cultural changes that result from competition between groups. Summaries and further references can be found in Richerson and Boyd (2005), Boyd and Richerson (2009), Henrich (2004, 2008) and Henrich and Henrich (2007).

In four recent papers, Lehmann and a series of collaborators have challenged the logical foundations of our hypothesis about the evolution of human cooperation. In two papers, (Lehmann et al 2008a, 2008b) they present models showing that one form of cultural transmission can actually reduce the range of parameters that allow the evolution of altruistic traits; based on these results, they argue that we have been premature in concluding that culture facilitates the evolution of cooperation. In two other papers, they argue against specific components of our hypothesis. Lehmann et al (2006) show that the punishment of noncooperators cannot evolve unless the transmitted variants giving rise to punishment and to

cooperation are linked. Lehmann and Feldman (2008) analyze models showing that a conformist social learning psychology does not enhance the evolution of altruism.

In all of their modeling efforts, Lehmann and colleagues assume that rates of cultural adaption are so weak that cultural adaptation does not affect variation among groups. This assumption radically changes the nature of the forces that generate relatedness among interacting individuals. Otherwise, their models are structurally similar to ours. They reach different conclusions because they made different assumptions. Their work provides a competing explanation for human cooperation. It bears on our conclusions only if their assumptions are better supported empirically than ours. In what follows we delineate the differences between our work and that of Lehmann and colleagues, and explain why these differences in assumptions lead to different conclusions. We then present data indicating that the Lehmann et al models are inconsistent with the observed scale of cultural variation, and thus do not provide a plausible alternative explanation for the evolution of large scale of human cooperation.

Does culture facilitate the evolution of cooperation?

The way that Lehmann and et al. (2008b) frame the relationship of their research to ours is misleading. They claim that we believe that culture *always* makes it easier for cooperation to evolve. They then refute this alleged claim by analyzing models in which altruism evolves under a narrower range of conditions with cultural transmission than with genetic transmission. However, we have never argued that culture *always* makes it easier for cooperation to evolve. Given the vast diversity of possible cultural transmission mechanisms, and the poor state of our empirical knowledge about such mechanisms, it would be foolish to claim that *any* imaginable form of cultural transmission facilitates the evolution of cooperation. Instead we have made a much more specific argument: human cooperation depends on systems of norms maintained by punishment, reputation and conformist cultural transmission, and that because cultural adaptation can be much more rapid than genetic adaptation, cultural evolution generates more stable behavioral variation among large groups, even very large groups.

In constructing their models, Lehmann and colleagues assume that adaptive forces are very weak, and as a result their models closely resemble those used in population genetics (e.g. Lehman and Keller 2006). Lehmann and colleagues do not mention this crucial difference between their models and ours in any of their papers. (In fact, with the exception of Lehmann et al 2006, the weak forces assumption is not discussed in the introductions or conclusions, and is only introduced deep in the mathematics) Instead they assert that their models capture the Boyd/Richerson cultural group selection models, and in one case state that their model "exactly" captures ours (Lehmann et al 2008b, p. 22).

To see why this difference in assumptions leads to quite different evolutionary dynamics, consider the models presented in Lehmann et al (2008b). In this paper, they compare the

evolution of altruism under cultural and genetic transmission. The genetic model is similar to models of the evolution of altruism in viscous populations (e.g. Taylor 1992). The population is structured into a large number of groups. Local population regulation maintains groups at a fixed, finite size and each in generation groups exchange migrants with all other groups. There are two haploid genetic variants, an altruistic variant that produces a benefit to all group members and a selfish variant that does not. Benefits and costs affect individual survival so that altruists have lower survival than selfish individuals within their own group, but groups with more altruists have higher average fitness and produce more emigrants.

The cultural evolutionary model is meant to be a generalization of the model in Boyd and Richerson (2002). Again the population is structured into a large number of fixed, finite sized groups. Lehmann et al consider two different payoff structures: In the body of the paper they present a model in which there is an altruistic variant and a selfish variant just as in the genetic model. This is an important modification of Boyd and Richerson (2002) because the altruistic variant is not favored in any group, and adaptive processes cannot maintain variation among groups. In the online supplementary materials they analyze a second model in which the payoff structure is given by a Stag Hunt game, as in Boyd and Richerson (2002). There are two variants. Each has higher payoff than the other when it is common, but one type increases the average payoff of all in the group while the other type does not. Cultural transmission is payoff-biased: individuals meet another individual, the "model", and adopt the model's behavioral variant with a probability proportional to the difference in payoffs. With some probability, the model is drawn from another randomly chosen group; otherwise the model is drawn from the individual's own group this leads to the exchange of cultural variant among groups. Models drawn from groups with a higher frequency of group beneficial behavior are more likely to be copied.

To analyze these models, Lehmann et al restrict parameter values so that the adaptive forces, natural selection in the genetic model and payoff-biased imitation in the cultural models, are much weaker than changes caused by the flow of genes or cultural variants among groups. As a result, the relatedness within groups (or, equivalently, the variation among groups) rapidly comes to a "quasiequilibrium" determined only by interplay of random sampling variation and the flow of heritable variants among groups—adaptive forces (natural selection and payoff-biased transmission) are ignored. This assumption greatly simplifies the analysis for several reasons: First, because relatedness equilibrates much more rapidly than changes in gene or cultural variant frequencies due to adaptive forces, the equilibrium relatedness can be taken as a fixed parameter in calculating the inclusive fitness of each variant. Second, the equilibrium relatedness can be calculated using a simple linear recursion, and third the adaptive forces are slow enough that changes in frequency result from the average relatedness over all groups. Lehmann and colleagues use this approach to analytically derive the conditions under which an altruistic variant can increase, and show that the conditions are less restrictive under genetic inheritance than cultural transmission because emigrants leave the group and therefore there is less competition among descendants in the genetic case.

In contrast, we assume that adaptive processes are strong compared to migration. This assumption is supported by empirical evidence from many sources. Recent studies of cultural learning suggests that learning mechanisms generate potent effects (see reviews in Henrich and Gil-White 2001; Henrich and Henrich 2007: Chapter 2; McElreath et. al. 2008, and Chudek et. al. n.d.), and the literatures on the diffusion of innovations, public health, business, history, and anthropology provide much evidence of rapid cultural change (reviewed in Richerson and Boyd 2005; Henrich 2001). Often novel cultural traits, including new norms and practices, spread to fixation in less than one generation. Since trait frequencies evolve rapidly, local conditions are important. We also focus only on payoff structures typical for repeated interaction and contingent behavior in which the cultural adaptive forces lead to more than one stable equilibrium. This means that adaptive cultural forces like payoff-biased transmission can favor different variants in different groups. Such systems cannot be analyzed using the weak-selection, quasiequilibrium approach used by Lehmann et al, and instead must be analyzed as high dimension dynamic systems, something that is not easy to solve analytically, particularly if groups are finite and the system is stochastic. However, simulating the behavior of such systems is straightforward.

To illustrate the impact of the different assumptions about the strength of adaptive forces, we simulated a model very similar to that presented in Lehmann et al (2008b) with weak payoff bias and strong payoff bias. A population of size N is divided into groups with n individuals. There are two cultural variants, labeled 0 and 1. Let x_i be the frequency of variant 0 in group i. The life-cycle has three steps: First, there is a mutation-like process. With probability μ, each variant spontaneously transforms into the alternative variant. In all simulations $\mu = 10^{-4}$. Second, individuals interact socially. The payoffs in group i are:

$$w_0 = 1 + gx_i \tag{1}$$

$$w_1 = 1 + s(x_i - \tilde{x}) + gx_i \tag{2}$$

Thus, variant 1 produces a benefit to every member of the group proportional to g, and has higher payoff than variant 0 if $x_i > \tilde{x}$ where $s > 0$ and $0 < \tilde{x} < 1$. Thus, both variant 0 and variant 1 are favored when common, giving a version of the Stag Hunt payoff structure described above. The parameter s controls the magnitude of this effect. Third, after social interaction individuals meet a model and observe its payoff. With probability m the model is drawn from another, randomly chosen, group, and with probability $1 - m$ from the individual's own group. The learner adopts the model's variant with probability

$$\frac{1}{2}\left(1 + \beta\left(w_m - w_f\right)\right) \tag{3}$$

Where w_m and w_f are the payoffs of the model and focal respectively, which models individuals' tendency to switch to the model's behavior if the model has a higher payoff. The parameter β

controls the strength of biased transmission, with larger values of β create more rapid adaptive change. The MatLab code used in the simulations is available from the first author on request.

Increasing the strength of biased transmission changes the nature of the forces that shape relatedness within groups. To see why consider the special symmetric case in which $\tilde{x} = 0.5$ and $g = 0$. There are 500 groups each with 100 individuals. Initially the frequency of variants 0 and 1 is one half, and groups are either all one type or all the other type. This means the relatedness within groups (approximately the fraction of variance among groups for groups of this size) in the population is initially equal to one.

Figure 1(a) shows the results for a weak transmission bias ($s = 0.1$, $\beta = 0.01$). Relatedness declines rapidly to steady state value of around 0.2, about the value predicted by the weak bias approximation given in Lehmann et al (2008b), and as a result, adaptive processes like selection and payoff-biased transmission can lead to the spread of low levels of individually costly group beneficial behaviors. Compare these results to those shown in Figure 1(b) where bias is strong ($s = 0.1$, $\beta = 0.5$). Now the relatedness within groups stabilizes at a much higher value, around 0.8. Note that we are conforming to the contemporary definition of relatedness as a measure of the extent to which an individual's cultural variant predicts the variants of others in the group. This is not Wright's coefficient of relatedness as originally used by Hamilton and it does not, necessarily measure the extent to which individuals are similar by common descent. We think that this usage is confusing because it conflicts with ordinary English usage. (There can be high relatedness between individuals who are not related). However, we adopt it as it has become standard.

To see why relatedness within groups is greater in the strong bias case, consider the distribution of frequencies of the group beneficial trait across groups in the two cases. When bias is weak, the distribution of frequencies across groups is unimodal. The variance across groups exceeds the level that would result from random group formation, so an individual's own type predicts the types of others in its group. When bias is strong, the cultural analog of disruptive selection creates a bimodal distribution of frequencies across groups. Because most of the groups are composed of mostly one variant or mostly the alternative variant, an individual's own variant is a much better predictor of the variants of others in its group than in the weak bias case.

As is shown in Figure 2, the effect of group size on relatedness is very different in weak bias and strong bias. When payoff-biased transmission is weak ($\beta = 0.01$) relatedness gets smaller as group size increases because relatedness derives from common descent, and the probability that two individuals have the same cultural parent declines as groups get bigger. In contrast, when the payoff-biased transmission is strong ($\beta = 0.5$), variation among groups and (therefore relatedness within groups) is mainly created and maintained by biased transmission. Because the strength of bias does not depend on group size, the within group relatedness remains high even when groups are very large, and as a consequence very few individuals are similar due to shared descent.

Both the strong and weak biases can generate group beneficial behavior but in very different ways. Lehmann et al (2008b) derive a condition for the group beneficial variant to increase when rare. In the notation of the present model this condition can be rearranged to become:

$$s\tilde{x} < \frac{R(s + mg)}{\left(1 - R(1 - m)\right)} \tag{4}$$

Where $R = 1/\left(nm(2 - m)\right)$ is the equilibrium relatedness if bias is weak and groups are large. When individuals with the group beneficial variant are rare, they suffer a payoff disadvantage relative to the common variant. The magnitude of this disadvantage is proportional to $s\tilde{x}$, so the left hand side is the cost associated with the group beneficial variant when it is rare. The right hand side is proportional to the relatedness and gives the inclusive fitness benefit associated with the group beneficial variant. Because R is proportional to $1/n$ this condition becomes hard to satisfy when groups are large unless migration rates are very low. The dynamics in the weak bias case are given in Figure 3. Relatedness quickly converges to the predicted equilibrium value, and the group beneficial trait increases because the inclusive fitness benefits exceed the cost. Relatedness here is due to common descent.

Now suppose that payoff bias is strong enough that once either trait is common within a group, it will remain common because payoff-biased transmission maintains the trait at high frequency against the force of migration. In this case nothing happens even though (4) is satisfied. The group beneficial trait still raises payoffs, and individuals in the group in which it is common are still disproportionately imitated by individuals in other groups, and relatedness is high. However, the group beneficial trait does not spread to groups in which it is not common, because payoff bias acts strongly against the trait in such groups. Thus, the group beneficial trait remains common in the initial populations, but cannot spread. Notice that this provides another example that supports the claim that cultural transmission need not facilitate the evolution of cooperation.

The group beneficial trait can spread, even in very large groups, if the model is modified in one of two ways. First suppose that groups with a higher frequency of the group beneficial trait are less likely to suffer extinctions, and that empty habitats are recolonized by individuals drawn from a single randomly selected group (Boyd and Richerson 1990). This assumption is consistent with ethnographic, historical, and archaeological research (Soltis et. al. 1995, Keeley 1997, Bowles 2009). Figure 4 shows that the group beneficial trait increases, and relatedness remains high even though the groups are an order of magnitude larger than in the weak bias case. When n = 1000 and payoff-bias is weak, R is very low, and the group beneficial trait is unlikely to spread. Notice that the dynamics of the distribution of frequencies is very different than in the weak bias case—throughout the process strong bias maintains quite different frequencies of the group beneficial trait among groups, and adaptation occurs because groups with a low frequency of the group beneficial variant are more likely to go extinct than groups with a high frequency of the variant.

The second way to modify the strong bias model is to use a stepping stone population structure so that individuals only imitate models in a small number of neighboring groups (Boyd and Richerson 2002). There is only a small difference between island and stepping stone models in the weak bias case because traits diffuse rapidly throughout the population. However, when bias is strong, the difference is crucial, and cannot be ignored. If the group beneficial trait becomes common in one group, the high payoff causes individuals in neighboring groups to adopt the group beneficial variant, which can tip the neighbors into the basin of attraction of the group beneficial trait. This results in a cascade which spreads the group beneficial trait throughout the population.

Both of these modifications require that the group beneficial trait become common in one group—a sort of group level "mutation". There are at least three plausible processes that can lead to the necessary initial shifts from one basin of attraction to another. First,sampling variation leads to cultural drift in a manner analogous to genetic drift (Cavalli-Sforza and Feldman 1981, Neiman 1996, Shennan 2001). Such cultural drift can lead to "peak shifts" for the same reasons as genetic drift. Moreover, if as some authors have argued (Claidre and Sperber 2007, Henrich 2009), the cultural analog of mutation rates are much higher than genetic mutation rates, the equilibrium frequency of deleterious traits resulting from the balance of adaptive bias and cultural mutation will be higher, and therefore waiting times for peak shifts should be much shorter. Note that this mechanism depends on sampling variation and should be less effective in large groups. There are, however, two mechanisms in which waiting times for peak shifts do not depend on population size. Second, in random environments, linkage leading to "pseudo drift" (Gillespie 2000) can also lead to peak shifts, and the rate at which this occurs does not depend on population size. Linkage in cultural transmission means that you acquire two traits from the same person, either because that person is a particularly salient model or because acquiring one trait one trait increases receptivity to a second trait. This leads to correlations between traits analogous to linkage disequilibrium. Then payoff biases that increase the frequency of one trait also tend to increase the frequency of correlated traits, and in a fluctuating environment this leads to random, nonadaptive temporal variation in frequencies that can cause the shift from one basin of attraction to another. Finally, individual learning and biased transmission depend on environmental cues. One cue will cause an individual to preferentially adopt one variant, while a different cue will cause her to adopt the alternative variant. The cues observed by members of a group may often be highly correlated (think floods, famines, epidemics, etc.). For example, the disastrous loss of WWII seems to have shifted Japanese from militaristic moral system to a more pacifistic one (Dower 2000). However, such cues often have a strong random component, and as a result will lead to random fluctuations in the frequency of different behaviors. For example, the Battles of the Coral Sea Midway the might easily have gone differently, and if they had and the US sued for peace, the Japanese might have "learned" that militarism pays.

Punishment can evolve even when not linked to cooperation

In Boyd et al (2003) we presented a simulation model showing how competition among groups could enhance the evolution of costly punishment when adaptive forces are strong. The model assumed that there is a population structured into groups of size n. Each generation individuals can contribute to a collective good at a cost c. Then individuals can punish any other individual reducing their fitness an amount p at a cost k. Next, with probability ε, each group enters into a conflict with another randomly chosen group. It loses the conflict with a probability proportional to the difference in the frequency of cooperators between the two groups. Losing groups go extinct and are replaced by a clone of the winning group. Traits "mutate" with probability μ. Then, individuals choose another randomly chosen individual and acquire their trait from a model with a probability proportional to the difference in payoffs according to the rule given in (3) above. Finally, a fraction m of models are drawn from at random from the population as a whole, and $1 - m$ from the individuals own group. We considered competition between three strategies, cooperators who contribute to the collective good but do not punish, punishers who contribute and punish non-contributors, and defectors who neither contribute nor punish. Our simulations indicate that plausible amounts of intergroup conflict can maintain cooperation and punishment at high levels as long as payoff biases are strong compared to migration.

Lehmann et al (2007) analyze a model of the evolution of cooperation and punishment very similar to that in Boyd et al. (2003). The main differences are that in Lehmann et al. 2007 collective action did not reduce extinction rates, and instead increased the average fitness of local groups, and that the two traits are transmitted genetically. They conclude that selection can only lead to the evolution of punishment when it is tightly linked to cooperation, so that increasing the frequency of cooperators can also increase punishment. While acknowledging that their model is not directly comparable to ours, they note that in our model punishment and cooperation are linked, and conjecture that selection would not lead to the evolution of punishment if this were not the case.

This conjecture is incorrect. We have modified the simulation used in Boyd et al (2003) so that individuals acquire the variant of the cooperation trait (contribute or defect) and the variant of the punishment trait (punish or don't punish) from two different, randomly chosen models. This corresponds to a recombination rate equal to one, and means that in each time period immediately after transmission, the correlation due to linkage between the cooperation trait and the punishment trait within groups is zero.

When transmission biases are weak, Lehmann et al are correct—punishment does not evolve. However, when bias is strong compared to migration between groups, they are incorrect. The results are qualitatively similar to the original results which assumed that punishment and cooperation were linked. To see why, consider the results shown in Figure 5. Both simulations assume that groups consist of 128 individuals and that migration rate is one percent, and both

assume that initially one population has high frequencies of punishment and cooperation and the rest have no punishment and no cooperation. The figure shows the distribution of frequencies across groups after 1000 time periods. In (a) bias is weak ($\beta = 0.05$). At steady state, the frequency of cooperation is about 0.3, the frequency of punishment is close to zero, and there is no correlation across groups. This makes sense. Finite group size and limited migration lead to an equilibrium relatedness of around 0.3, in this case due to common descent. Thus the lower extinction rates that are generated by cooperation lead to an inclusive fitness benefit to cooperators. Punishment does not substantially increase the frequency of cooperation within groups because the transmission bias is weak, so punishment cannot create an inclusive fitness benefit, and does not increase in frequency.

In contrast, in (b), the bias is strong ($\beta = 0.5$). Now at steady state, the population average frequencies of punishment and cooperation are about 0.4 and 0.9 respectively, and there is a substantial positive correlation between cooperation and punishment across groups—groups with more punishers have more cooperators. In groups in which punishers are common, defectors are heavily punished and thus have a lower payoff than cooperators who are not punished, and this tends to maintain cooperation in such at a higher frequency than in groups in which punishers are rare. Groups with a low frequency of cooperators go extinct at a higher rate than those with a high frequency of cooperators, and because there is a positive correlation across groups this means that groups with high frequency of punishers have a lower extinction rate than groups with a low frequency of punishers. Of course, punishers always have lower payoffs than non-punishers in their group, and thus the frequency of punishment within groups tends to decrease. However, in groups in which defectors are rare, there is little cost to being a punisher, and thus frequency of punishment declines slowly. As long as the increase in frequency due to differential extinction is greater than the decrease within groups, punishment and cooperation, are sustained at high frequencies.

Conformism can facilitate the evolution of cooperation

In a series of studies we (e.g. Boyd and Richerson 1985: ch 7; Henrich and Boyd 2001) we have argued that a "conformist" bias in social learning may facilitate the evolution of altruism. Conformist bias occurs when individuals are disproportionately likely to acquire the more common variant from among the variants that they observe. So, for example, if three of four models have one variant and the fourth model has a second variant, the probability of acquiring the common variant is greater than three quarters. Such a bias creates an evolutionary force that increases the more common variant in the population. Theoretical work predicts that natural selection will favor conformist biased social learning in some kinds of spatially and temporally varying environments (Henrich and Boyd 1998, Nakahashi 2007, Wakano and Aoki 2007, Kendal et al 2009). In such environments, many models predict that more than one cultural variant will coexist, but the most adaptive variant will tend to be more common. Thus, all other

things being equal, a predisposition to adopt the more common variant increases the chance of acquiring the most adaptive variant. Recent experimental work indicates that human social learning is subject to conformist bias (Efferson et al 2008, McElreath et al 2008).

Conformist bias can facilitate the evolution of altruism because it creates multiple stable equilibria which in turn can create and sustain variation among groups. To see how this works, consider what happens when a potentially altruistic trait evolves under the influence of payoff-biased transmission alone. Altruists produce a benefit to the group at a cost to themselves. This means that payoff bias decreases the frequency of altruists in every group; altruism can increase only if groups are small enough or migration rates low enough that there is enough relatedness within groups to create a compensating inclusive fitness benefit (or equivalently, maintain sufficient variation among groups so that there is sufficient between group selection in favor of altruists). Now, suppose that in addition to payoff bias, there is also a conformist bias. Remember that conformist bias tends to increase the common type. So if conformist bias is strong enough, it will maintain altruists at high frequency even though they achieve a lower payoff. Thus, conformist bias creates multiple equilibria in situations with an altruistic payoff structure, and if the bias is strong compared to migration in a structured population this allows altruism to evolve in very large groups for the same reasons describe above.

We have also argued that this mechanism is probably of particular importance for the maintenance of punishment (Henrich and Boyd 2001). Because punishment suffers only a small disadvantage when cooperation is common, weak conformism can stabilize punishment within groups, and this allows selection among groups to maintain punishment and cooperation in much larger groups than is possible without conformism (Guzman et. al. 2007). Interestingly, these authors also show that conformism is favored in this model even without any other form of variation in payoffs.

Lehman and Feldman (2008) analyze the effect of conformism on the evolution of altruism using the assumption that transmission biases are very small relative to migration. This means that conformism cannot maintain variation among groups, and that the differential extinction and stepping stone processes discussed above cannot lead to the spread of group beneficial altruistic behavior. In their model the main effect of conformism is to make it hard for rare traits to increase. So, once again, they reach quite different conclusions because they have made different assumptions.

Evidence that cultural evolution is subject to strong adaptive forces

Our models and those of Lehmann et al are based on different assumptions about the processes that maintain cultural variation among groups. The fact that Lehmann et al do not mention this difference in the introductions or discussions of any of their papers misleads their readers about the relevance of their work to ours. Their work does not refute ours. Rather, properly interpreted,

their efforts yield an alternative hypothesis about the processes that govern the cultural evolution of large scale cooperation in human populations. Lehmann et al do not give an empirical defense of their alternate assumptions. Indeed, given their very limited discussion of the difference, one might assume that they made the weak forces assumption for analytical convenience in the mistaken assumption that it is a harmless simplification. We have shown that it is not harmless, and now turn to showing that it is empirically dubious.

Lehmann et. al.'s approach cannot be a sufficient explanation of human cooperation because it is inconsistent with what is known about the scale of human cultural variation. One of the striking puzzles about human cooperation is that people frequently cooperate in large groups. This is obviously true in the agricultural societies of the last 10,000 years in which thousands of individuals are mobilized for military activity and the construction of large capital facilities like roads, fortifications, and ceremonial centers. However, it is also true for small-scale human societies. For example, hunter-gatherers recruit war parties numbering up to two hundred individuals. Such groups are too big to have high average genetic relatedness in a primate species with substantial migration between groups (See Richerson and Boyd 2005 and Henrich and Henrich 2007 for references).

In the Lehmann et al models, cultural relatedness arises from common descent as in genetic models, thus naively, one might expect that there would be low cultural relatedness in large groups. However, Lehmann et al point out that prestigious individuals may often play a disproportionate role in cultural evolution, and that this could lead to substantial cultural variation among larger groups. They illustrate this idea with the "teacher" model (Cavalli-Sforza and Feldman 1981). With probability t, each individual in the group acquires his or her cultural variant from a single focal individual, the teacher, and, with probability $1 - t$, imitates a randomly chosen individual. This means that when groups are very large, the cultural relatedness within groups converges to t^2, a result they believe provides an explanation for observed cultural variation among large groups.

This account is empirically implausible. Remember that that the goal is to explain why cooperative social norms spread. In the modern world, there is substantial variation in beliefs and norms among ethnic groups and nation states that number millions of individuals. For example, Bell et al (2009) show that a lower bound on cultural F_{ST} is more than 10 times the genetic F_{ST} for neighboring nation states (F_{ST} is the fraction of total heritable variation that is among groups. For large groups it is approximately equal to the relatedness within groups). They also show that lower bounds on the cultural F_{ST} values for four large East African ethnolinguistic units are quite high, even when ecological variation is controlled for. It is not plausible that four million Kamba, (one of the East African groups) share a language and many beliefs because a substantial fraction of them acquired their beliefs by imitating the same person. But that is what is required by the Lehman et al model.

Nor is the Lehman et al account believable for smaller scale foraging societies because the scale of cultural variation in such societies is typically much larger than the scale of everyday interaction. For example, among Australian aboriginal foragers, ethnolinguistic units that shared a common language and culture typically numbered between 500 and 5000 (Keen 2004), and migration rates between ethnic groups were probably substantial. If we assume that bands numbered between 10 and 100 people, and that everybody in a band imitates a single individual, then the formula used by Lehman and colleagues predict that only a small fraction of cultural variation will be between ethnolinguistic units.

Moreover, this is a best case because it assumes that a single individual dominates cultural transmission. Empirical evidence indicates that while some individuals are more important in cultural transmission than others, probabilities of common descent are lower than in the teacher model. Assuming discrete traits, and accurate social learning, the probability that two individuals who share n models acquire the same variant by common descent is $\sum a_i^2$ where a_i is the probability that model i is imitated by each individual in the next generation. Henrich and colleagues (in preparation) have measured these a_i values in a small Fijian village with 210 residents. A sample of 146 subjects were asked to indicate which other individuals they would go to if they were seeking information in three different domains: (1) knowledge about fishing, (2) yam horticulture, and (3) medicinal plants. Assuming that this is a measure of the importance of individuals in cultural transmission, these data can be used to estimate $\sum a_i^2$. The values are 0.043 for fishing, 0.053 for yam horticulture, and 0.053 for medicinal plants. These data indicates that some individuals have more influence than others, but the probability that two individuals acquire their beliefs by common descent is still quite small. For comparison the average genetic relatedness computed from a complete genealogy is 0.018.

Nor can the teacher model explain the persistence of differences between large neighboring groups over hundreds of generations unless it is assumed that migration rates are unrealistically low. For example, the Romance/Germanic linguistic boundary is roughly where the Roman advance came to rest two millennia ago despite massive flows of people across the boundary. This boundary also separates peoples with different norms that lead to measurably different behavior in important economic contexts (Brügger et al 2009). Lehmann et al would seem to imply that Germans keep speaking German, and keep adhering to German social norms and the French do the same because there is a significant probability that they acquire their linguistic and social norms from a small number of people. A much more plausible hypothesis is that strong biased transmission maintains cultural boundaries (McElreath et. al. 2003; Boyd and Richerson 1987). When people move from one culture to another they, and especially their children, modify their language and social behavior in response to local conditions, so that they will be understood, and approved of. If this process is sufficiently rapid compared to the rate of migration, the boundary will be maintained. This is exactly the process assumed in the Boyd/Richerson/Henrich models, but not in the models of Lehmann and colleagues.

Mister Natural sez, Use the right tool for the job

Lehmann and colleagues seem to believe that cultural and genetic evolution are very similar. Traits have discrete values, transmission is sufficiently accurate that errors can be ignored, and adaptive forces are weak compared to migration. The only substantive difference between genetic and cultural transmission in their models is that cultural traits can be acquired horizontally, and that this process biased so that successful, wealthy, or prestigious people are more likely to be imitated. The latter process leads to dynamics that are similar to natural selection, but as Lehmann et al show, different in interesting ways. Accordingly, Lehmann et al apply the weak selection/inclusive fitness formalism that has proven to be very useful for understanding the genetic evolution of social behavior. However, cultural evolution is very different from genetic transmission in a number of ways. Here, we have focused on the fact that cultural metapopulation dynamics can be very different than genetic dynamics if the relative strength of adaptive forces and migration are different. Elsewhere we have argued that the adaptive advantage of having a cultural system of inheritance is that biased social learning allows humans to evolve complex adaptations much more quickly than genes adapting by natural section (Richerson and Boyd 2005). If this argument is correct, culture exists *because* it violates the weak forces assumption. For social traits with multiple equilibria, this causes much more cultural variation among groups, but also reduces the ability of immigrants to affect the cultural composition of their new homes. If we are correct that strongly biased social learning maintains cultural variation among large groups, then the quasiequilbrium assumption cannot be used to compute relatedness when modeling the cultural evolution of cooperation.

As our understanding of cultural transmission increases, we suspect that many more important differences between cultural and genetic evolution will emerge. For example, much evidence supports the idea that culture is heritable at the group level. There are important persistent differences between large groups that are not due to the environmental or genetic differences. In genetic transmission, heritability at the group level results from accurate transmission at the individual level. It is not clear that the same is true of culture. Based on ideas about the cognitive psychology of social learning, Claidre and Sperber (2007) have forcefully argued that cultural transmission is very inaccurate. If they are correct, there may be little cultural heritability at the individual level, and group heritability is based on some kind of strongly biased frequency dependent cultural transmission (Henrich and Boyd 2002). The Boyd/Richerson/Henrich strong biased models won't be much affected if this turns out to be correct because they don't depend critically on the transmission mechanisms the within groups (Henrich and Boyd 2002). In contrast, the quasiequilibrium approach that Lehmann and his colleagues have used in their models will be substantially affected.

Human cultural evolution is best conceptualized as a population process, and as a result theoretical tools from population biology can be very useful in understanding cultural evolution. However, it is important to resist the temptation to think that cultural transmission is just like

genetic transmission. The genetical theory of the evolution of social behavior is highly developed, and replete with subtle, powerful insights, and well-worked out mathematical tools. In genetic evolution, selection is often weak and transmission accurate. This means that relatedness through common descent is often sufficient to predict patterns of social interaction, providing a powerful tool for understanding the evolution of social behavior. The evidence suggests that cultural variation is affected by strong biased transmission. If so, despite its elegance and power, the quasiequilibrium approach must be avoided when it is not the right tool for the job.

References cited

Bell, A. V., P. J. Richerson, & R. McElreath. 2009. Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences*, 106, 17671-17674.

Bowles, S. 2009. Did warfare among ancestral hunter-gatherer groups affect the evolution of human social behaviors. *Science*, 324, 1293-98.

Boyd, R., R. and P. J. Richerson, 1985. *Culture and the Evolutionary Process*, University of Chicago Press, Chicago, IL.

Boyd, R. and P. J. Richerson, 1990. Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology,* 145, 331-342.

Boyd, R. and P. J. Richerson, 2002. Group beneficial norms spread rapidly in a structured population, *Journal of Theoretical Biology,* 215, 287–296, 2002.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences (USA)* 100, 3531–3535.

Boyd, R. and P. J. Richerson. 2009. Culture and the evolution of cooperation, *Philosophical Transactions of the Royal Society, B*. 364, 3281-3288.

Brügger, B., R. Lalive, and J. Zweimüller. 2009. Doe culture affect unemployment? Evidence from the *Röstigraben*, Institute for the Study of Labor, University of Zurich, Discussion Paper No. 4283.

Cavalli-Sforza, L.L. and Feldman, M. W. 1981. *Cultural Transmission and Evolution*, Princeton University Press, Princeton.Claidière, N. and D. Sperber. 2007. The role of attraction in cultural evolution. *Journal of Cognition and Culture, 7, 89-111*

Chudek, M., Heller, S., Birch, S., Henrich, J. n.d. Prestige-biased learning in children: Attention from others as a cue for cultural transmission.

Dower, J. 2000. *Embracing Defeat: Japan in the Wake of World War* II**,** W.W. Norton, Co. NY.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. 2008. Conformists and mavericks: The empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, 29, 56-64.

Gillespie J.H. 2000. Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics*. 155, 909-919

Andrés Guzmán, R., C. Rodríguez-Sickert, R. Rowthorn, 2007. When in Rome, do as the Romans do: the coevolution of altruistic punishment, conformist learning, and cooperation, *Evolution and Human Behavior*, 28, 112-117

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

Henrich, J. & F. Gil-White 2001. The Evolution of Prestige: freely conferred status as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22, 1-32.

Henrich, J. 2001. 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*, 103, 992-101

Henrich, J. and R. Boyd, 2001. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208, 79–89.

Henrich, J. and R. Boyd. 2002. Culture and cognition: Why cultural evolution does not require replication of representations, *Culture and Cognition*, 2, 87–112,

Henrich, J., 2004. Cultural group selection, coevolutionary processes and large-scale cooperation, *Journal of Economic Behavior and Organization*, 53, 3-35.

Henrich, J., and Henrich, N., 2007. *Why Humans Cooperate*, Oxford University Press, Oxford.

Henrich, Joseph 2009. The evolution of costly displays, cooperation, and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behaviour, 30,* 244-260.

Keeley, L. 1997. *War Before Civilization*, Oxford University Press, Oxford.

Keen, I. 2004. *Aboriginal Economy and Society: Australia at the Threshold of Colonisation*, Oxford U. Press, Oxford.

Kendal, J. Luc-Alain Giraldeau, and Kevin Laland, 2009. The evolution of social learning rules: Payoff-biased and frequency-dependent biased transmission, *Journal of Theoretical Biology*, 260, 210-219.

Lehmann L. and L. Keller L.2006. The evolution of cooperation and altruism - A general framework and a classification of models. *Journal of Evolutionary Biology* 19,1365-1376.

Lehmann L., F. Rousset, D. Roze and L. Keller. 2007. Strong-reciprocity or strong-ferocity? A population genetic view of the evolution of altruistic punishment. *The American Naturalist* 170,21-36

Lehmann L. and M. W. Feldman 2008. The coevolution of culturally inherited altruistic helping and cultural transmission under random group formation. *Theoretical Population Biology* 73,506-516.

Lehmann L., K. Foster K., E. Borenstein, and M.W. Feldman 2008a. Social and individual learning of helping in humans and other species. *Trends in Ecology and Evolution* 23:664-671

Lehmann L., M.W., Feldman and K. Foster 2008b. Cultural transmission can inhibit the evolution of altruistic helping. *The American Naturalist* 172:12-24

Neiman, F. D. 1995. Stylistic variation in evolutionary perspective: Inferences from decorative diversity and interassemblage distance in Illinois Woodland ceramic assemblages, *American Antiquity* 60, 7-36.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. 2008. Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and payoff-biased social learning strategies. *Philosophical Transactions of the Royal Society B*, 363, 3515-3528.

P. J . Richerson and R. Boyd, , 2005, *Not by Genes Alone: How culture transformed human evolution*, University of Chicago Press, Chicago, IL.

Shennan, S. 2001. Demography and cultural innovation: a model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal.* 11, 5-16.

J. Soltis, R. Boyd, and P.J. Richerson, Can group functional behaviors evolve by cultural group selection? An empirical test. *Current Anthropology.* 36, 473–494, 1995.

Taylor , P.D. 1992, Inclusive Fitness in a Homogeneous Environment, *Proceedings of the Royal Society*, *B.* 249, 299-302

Nakahashi, W. 2007, The evolution of conformist transmission in social learning when the environment changes periodically, *Theoretical Population Biology*, 72, 52-66

Wakano, J. H. and K. Aoki 2007. Do social learning and conformist bias coevolve? Henrich and Boyd revisited, *Theoretical Population Biology*, 72, 504-512.
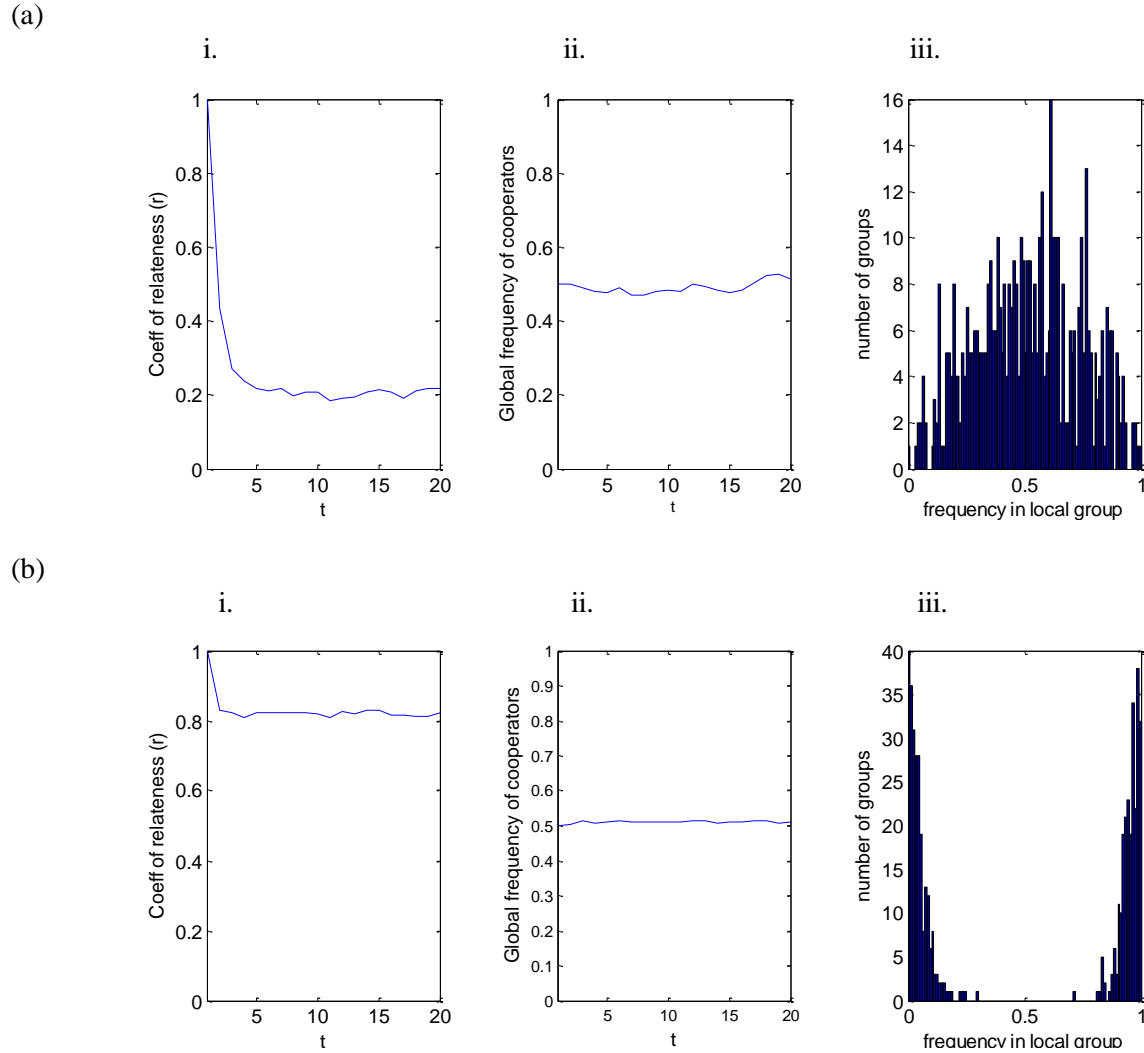
Figure 1. (a) Model behavior when biased cultural transmission is weak and both variants have a higher payoff when common. The payoffs are symmetric so that the basins of attraction of both equilibria are the same, and there is no group benefit. Population size is 10,000. There are 500 groups each with 100 individuals. The probability of choosing a model from outside the group is 0.02. Initially half of the groups are all one variant and other half are the other variant. (i) The relatedness within groups converges rapidly to a value predicted by the analytical treatment given in Lehmann et al (2008). ii) The overall frequency does not change due to the symmetry of the model. (ii) The distribution of frequencies across groups in the final time period. The distribution is modal, but because relatedness is approximately equal to 0.2, the variance of this distributions is much greater than that would be predicted if groups were sampled at random with probability 0.5. (b) Model behavior when biased cultural transmission is strong. Parameter values as in (a), except that the strength of payoff-biased transmission in increased by a factor of 20 ($\beta = 0.5$). (i) The relatedness within groups converges rapidly to a value that is much higher than in the weak bias case. Relatedness here is a measure of the extent to which one individual's type predicts the types of others in its group. (iii) The distribution of frequencies across groups in

the final time period. This shows why relatedness is so high—the cultural analog of disruptive selection creates a bimodal distribution of frequencies across groups. Because most of the groups are either mostly one variant or mostly the alternative variant, an individual's own variant is a good predictor of the variants of others in its group.
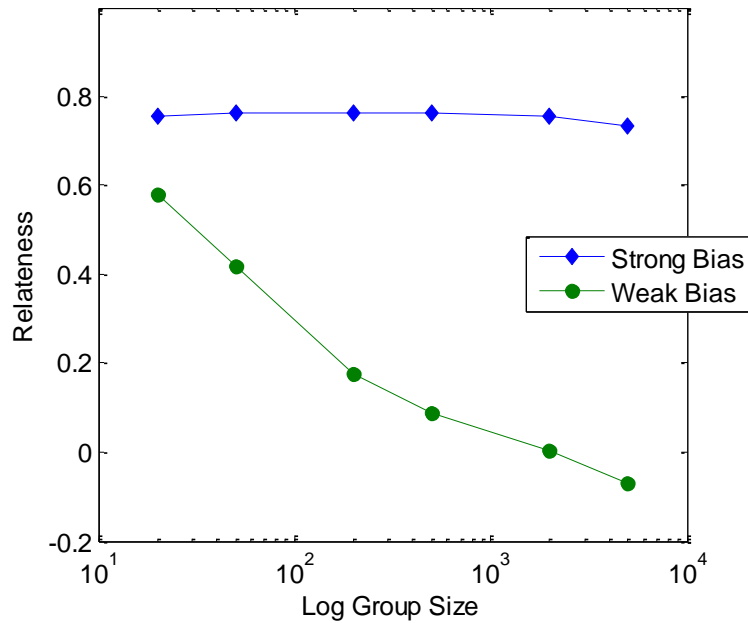
Figure 2. Mean relatedness within groups during the last period of the simulation as a function of group size. When payoff-biased transmission is weak ($\beta = 0.01$) relatedness among group members declines as group size increases because relatedness derives from common descent, and probability that two individuals have the same cultural parent declines as groups get bigger. In contrast, when the payoff-biased transmission is strong ($\beta = 0.5$), variation among groups and (therefore relatedness within groups) is mainly created and maintained by bias and common descent plays a minor role. Since the strength of bias does not depend on group size, the within group relatedness does not depend on group size.
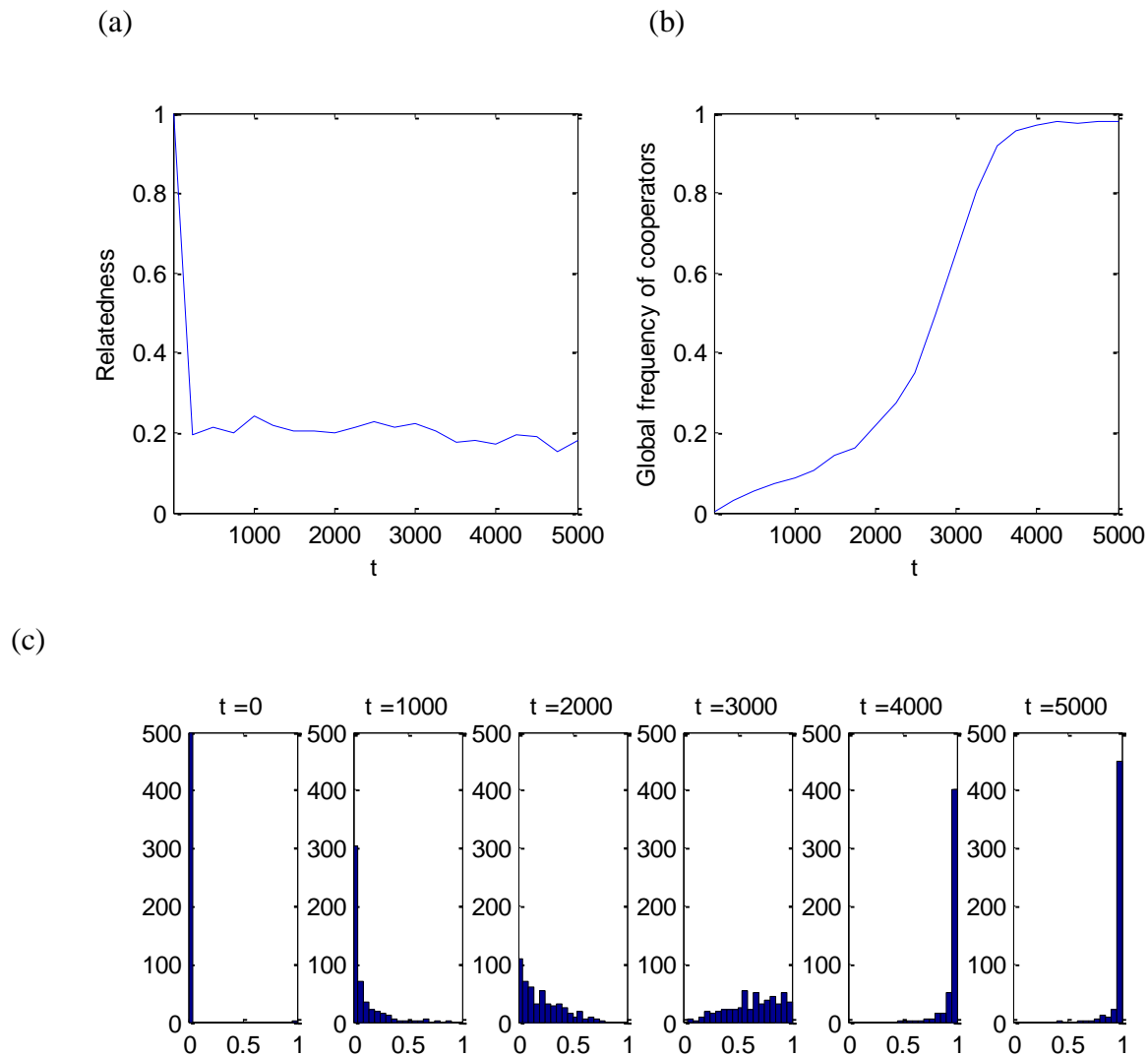
Figure 3. The dynamics of the group beneficial trait when payoff bias is weak ($n = 100$, $m = 0.02$, $s = 0.2$, $g = 1$, $\tilde{x} = 0.25$, $b = 0.05$). Initially, there is one group in which the frequency of the group beneficial variant is one in one group, and zero in all other groups. (a) shows that relatedness quickly attains the predicted equilibrium value (≈ 0.2 for these parameter values), and (b) shows that because the inclusive fitness benefits exceed the costs, the group beneficial trait increases in frequency. (c) Gives the distribution of frequencies across groups as the group beneficial trait increases. Individuals with the group beneficial variant rapidly diffuse throughout the population, and then the distribution of frequencies results from the interplay of migration and common descent. Adaptation can be understood as responding to the average over the entire distribution.
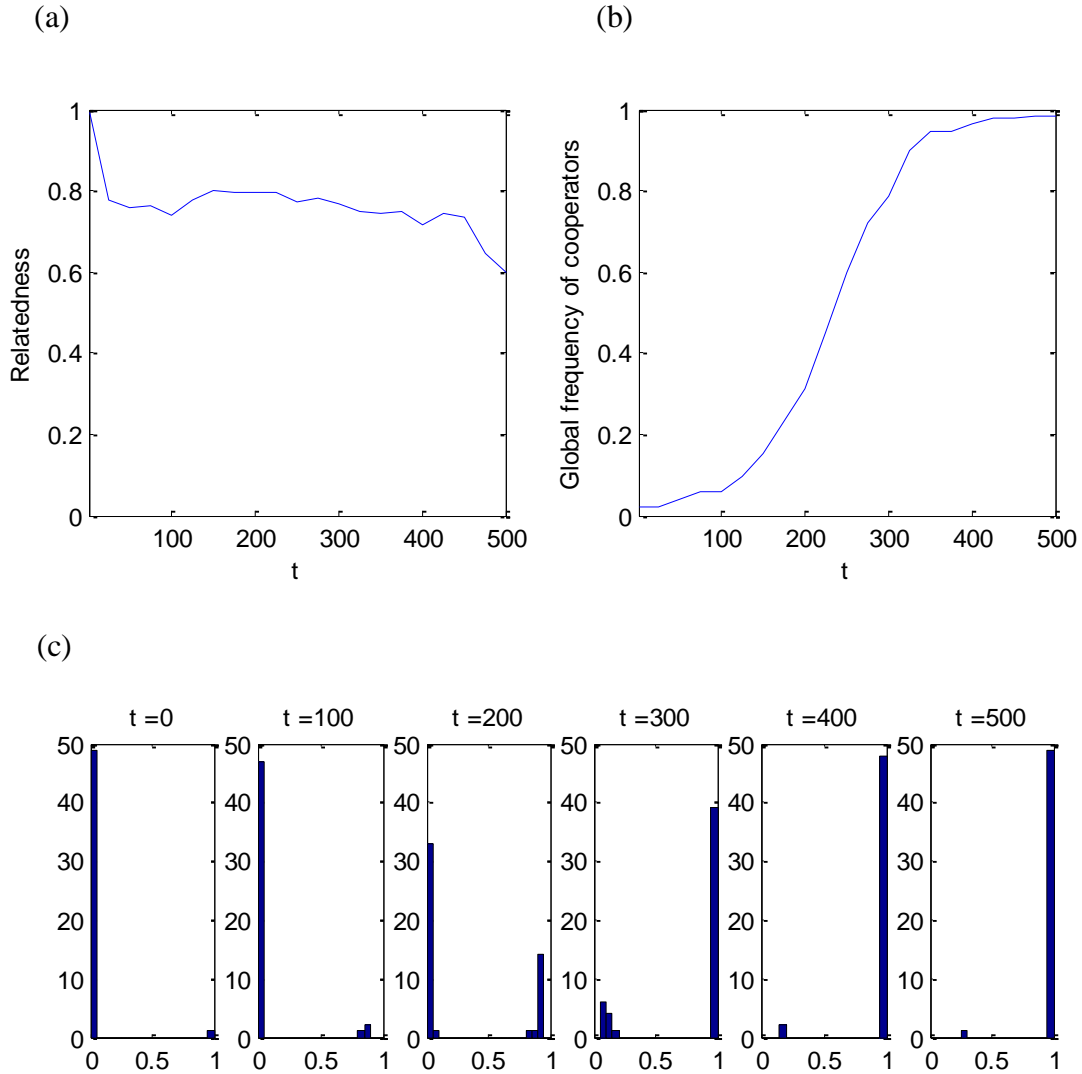
Figure 4. The dynamics of the group beneficial trait when bias is strong and the group beneficial trait lowers extinction rates ($n = 1000$, $m = 0.02$, $s = 0.2$, $g = 0$, $\tilde{x} = 0.5$, $\beta = 1.0$). The probability of extinction in group i each time period is $\varepsilon(1 - x_i)$ where x_i is the frequency of the group beneficial trait in group i. The rate of group conflict $\varepsilon = 0.015$, a value that when combined with the distribution of frequencies yields extinction rates roughly consistent with those observed in tribal societies (Soltis et al. 1995) assuming simulation time periods of one year. Empty habitats are recolonized by immigrants from a single surviving group. Initially, there is one group in which the frequency of the group beneficial variant is one, and zero in all other groups. (a) Relatedness quickly reaches an equilibrium value of about 0.8 even though groups are quite large because strong bias maintains the group beneficial norm at either a high or low frequency in every group. Here relatedness is mainly not the result of common descent. (b) The group beneficial trait spreads because such groups are much less likely to become extinct. (c) Gives the distribution of frequencies across groups as the group beneficial trait increases. Throughout the process, strong bias maintains groups at strongly different frequencies, and

adaptation occurs because groups with a low frequency of the group beneficial variant are more likely to go extinct than groups with a high frequency of the variant.
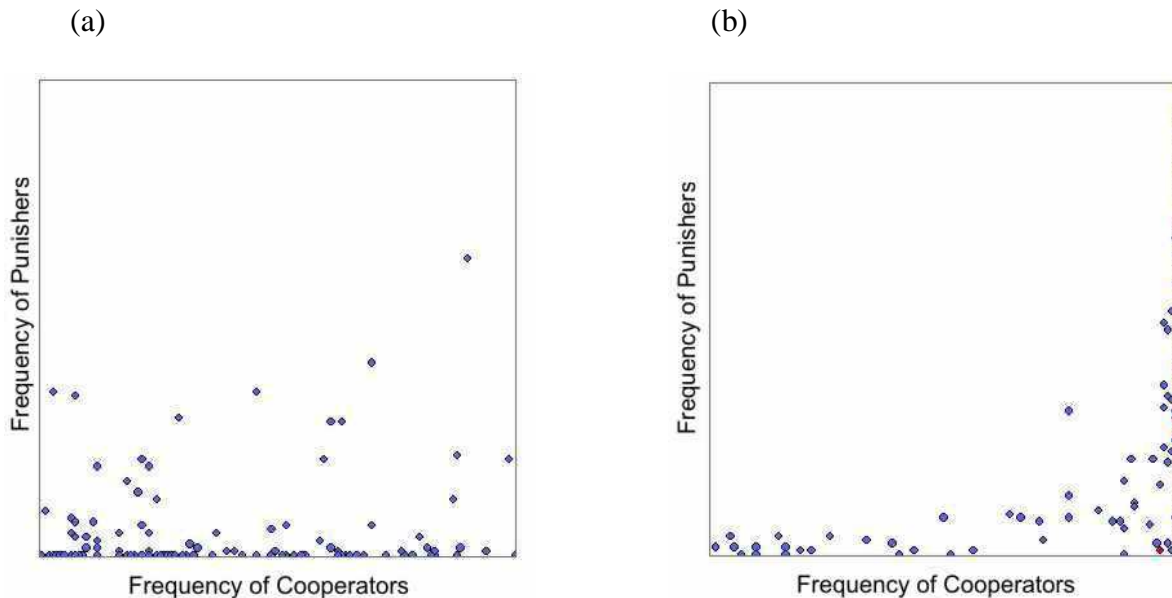
Figure 5. Each dot represents the frequencies of punishment and cooperation in one of 128 groups during the last period (2000) of the simulation. In both simulations $c = 0.2$, $p = k = 0.8$, $m = 0.01$, and $n = 128$. There is complete recombination each time period. Individuals acquire their punishment variant and their cooperation variant from different randomly selected models so that no linkage exists. (a) Weak forces ($\beta = 0.05$, $\varepsilon = 0.0015$, $\mu = 0.0001$). Relatedness (based on common descent builds) up to substantial levels, and since extinction rates are proportional to the frequency of cooperators in groups, selection increases the frequency of cooperation to a modest level (averaging about 0.3) at steady state. However, punishment and cooperation are uncorrelated both within and across groups, so punishment is selected against. It is maintained in the population by the cultural analog of mutation. (b) Strong forces ($\beta = 0.5$, $\varepsilon = 0.015$, $\mu = 0.001$). As in the weak bias case, relatedness builds up due to finite populations and limited migration. However, in groups with a high frequency of punishers, defectors are selected against thus maintaining a high frequency of cooperators. Complete recombination means that there is no correlation between cooperation and punishment within groups, but there is a strong correlation *across* groups that is generated by the fact that punishment lowers the payoff of defectors. Thus the extinction of groups with few cooperators increases the frequency of punishers, and since punishment has low cost in groups in which cooperators are common, punishment is maintained at a substantial frequency (about 0.4) and cooperation at a higher frequency (about 0.9) than in the weak forces case.