Cultural group selection and coevolutionary processes explain human prosociality and large-scale cooperation[*]

Joe Henrich
University of Michigan Business School
701 Tappan Road, D3276
Ann Arbor, MI 48109-1234
(734) 763-0370
henrich@umich.edu

Abstract

Standard evolutionary approaches to cooperation/altruism fail to explain both the degree of prosociality observed in experimental settings, and the actual levels of cooperation and punishment found in many human societies. Using a coevolutionary approach to human prosociality, I show how the details of our evolved cultural learning capacities (i.e. imitative abilities) generate both prosocial, behavioral equilibria not available to genetic evolution alone, and a mechanism of equilibrium selection. Further, in the novel social environments left in the wake of these cultural evolutionary processes, natural selection is likely to favor prosocial genes that would not be expected without this culture-gene coevolutionary approach.

Key words: cooperation, altruism, group selection, coevolution, dual inheritance theory

JEL code: D00, C72

Standard approaches to the evolution of cooperation/altruism (i.e., kinship and reciprocal altruism) are insufficient to explain both the levels of prosociality (cooperation and punishment) observed in many human societies, and the degree of variation in prosociality recorded across human societies and behavioral domains. Evidence from one-shot, anonymous, experimental games (among non-relatives) shows both highly prosocial behavior and tremendous cross-cultural variability, neither of which can be easily explained by kin selection or reciprocal altruism. In Dictator games for example, proposers often bestow significant portions of their endowment on anonymous receivers (Forsythe et. al. 1994; Hoffman et. al. 1994), even when large sums are involved (Burks et. al. 2000). Dictator offers also vary substantially across societies, and in some groups the modal offer in the Dictator game is a 50/50 split (Henrich et. al. 2000; Burks et. al. 2000; Smith 2000).[1] Using Ultimatum game data, Henrich et. al. (2000) show that proposers from many societies offer too much (assuming income maximization) to their anonymous partners given the actual distribution of rejections across offers. On the receiving end, responders across a wide range of societies (but not all) reject low Ultimatum offers (Roth et. al. 1991; Roth 1995; Henrich et. al. 2000). Similarly, in one-shot, anonymous, Public Goods games and the first round of repeated games, many players from a variety of societies contribute non-zero sums to the group interest (Ledyard 1995), and the secondary modal contribution is often 'give it all to the group' (Fehr & Gachter 2000; Henrich & Smith 2000; Henrich et. al. 2000). However, depending on the society, mean contributions range from 20% to 70% (Henrich et. al. 2000). Real life corroborates these experimental measures as people from many societies vote (Mueller 1989), pay taxes (Skinner & Slemrod 1985), recycle, share game (Kaplan and Hill 1985), stop to help stranded motorists, threaten defectors who butt into line (Frank 1994), fight in wars, conserve electricity, etc. All of this indicates levels of prosociality and variability not observed in other animals, and well beyond the explanatory range of standard evolutionary approaches.

---

[1] Among non-student populations in the U.S., we find Dictator results that are indistinguishable from (or even more "fair" than) Ultimatum game offers from the same groups (Burks *et. al.* 2000; Smith 2000).

2

In addressing the puzzle of human prosociality, I will show how culture-gene coevolutionary processes can explain prosocial, group-beneficial, behavioral traits that cannot be explained by standard evolutionary approaches. In pursuit of this, I'll first clarify the debate about 'group selection' by showing how the force of natural selection acting on genetic variation can be divided into within-group ('individual selection') and between-group ('group selection') components. Then, in light of this clarification, I show how all explanations for the evolution of cooperation/altruism rely on exploiting some underlying informational regularity that is maintained by the imposition of some kind of constraint. In most altruism models, including typical genetic 'group selection' models, the plausibility of a solution depends on the plausibility of the underlying constraint(s) that makes the model work. Through this demonstration, I also show why kinship and reciprocity-based explanations are insufficient in the human case. Having laid out the logic of group selection, and exposed the underlying structure of cooperation, I'll show why the between group component of cultural evolution (i.e., 'cultural group selection') is much more likely to lead to high levels of prosocial behavior then is the case in genetic evolution. Finally, I'll discuss how, in the wake of cultural group selection, individual-level natural selection can favor prosocial/altruistic genes that would not otherwise be favored.

The logic of selection within- and between-groups

Debates about the potential relevance of 'group selection' (or 'multi-level selection'), particularly in its application to human behavior and cooperation, continue despite a fair amount of agreement among the biologists, economists and anthropologists who have seriously explored the question (Wade 1985; Frank 1995; Maynard Smith 1998; Bowles 1998, 2000; Gintis 2000; Soltis et. al. 1995). Since the publication of George Williams' *Adaptation and Natural Selection* in 1966—which quite appropriately laid waste to a particularly naïve form of group- or species-functionalism that had been prominent in biology and anthropology during the preceding decades—a whole generation of biologists and anthropologists learned to scorn any explanation

that involves selection among groups, or proposes 'group functional' traits. However, only a few years after Williams' book, George Price (1970, 1972) provided an elegant formalization that showed, among other things, how the force of natural selection acting on genes can be partitioned into 'group-level' and 'individual-level' components. Unfortunately, the insight derived from Price's simple demonstration did not spread very far outside of theoretical evolutionary biology, and failed to impede the spread of the belief that group-selectionist-thinking is somehow logically flawed, wrong-headed, or wishful thinking. This *a priori* dismissal of group selection, without understanding the relevant details, has slowed progress in our understanding of both genetic and (by faulty analogy) cultural evolution.

Unfortunately, the way evolutionary scholars often talk about 'group selection' leads people to think (mistakenly) of it as a separate process, somehow fundamentally different from 'individual selection' or 'natural selection' (e.g. Lowe 2000; Ridley 1993; Dawkins 1976). Genetic evolution, at least from one perspective, is about changes in the frequency of alternative alleles, not about the frequencies of individual organisms or groups of organisms. Having said that, it's both possible and sometimes useful to write down accounting systems that track the frequencies of these alleles by examining their effects on fitness from different points of reference. Useful approaches might involve tracking the fitness of alleles, individuals, families, social groups, genomes, chromosomes, etc. So, to begin, I'll derive Price's Covariance Equation of evolutionary change, which partitions the forces of selection-driven evolutionary change into within- and between-group components. With this analytical tool, we can clarify the conditions under which the between-group component may overpower the within-group component and favor group-beneficial traits, like altruism. It's important to realize that Price's equation is a very general formulation applicable to any evolutionary system. However, for the sake of clarity, I'll derive it from the perspective of evolutionary genetics, in more concrete terms than the formulation demands. Later, in applying the Price Equation to cultural evolution, and then to

4

gene-culture co-evolution, I'll show how it can be generalized. In this section I've drawn from

Hamilton (1975), Frank (1995, 1997, 1998) and Price (1970).

We start with a population subdivided into groups indexed by i. There are no restrictions

on how the groups are composed, except that all groups must contain at least one individual. The

variable x_i gives the frequency of the trait in subpopulation i, x_i' represents the same frequency in

the next time period (or generation), and $\Delta\bar{x}$ expresses the average change in the frequency of the

trait under investigation. By definition we have,

$$\Delta\bar{x} = E(\Delta x_i) = E(x_i') - E(x_i)$$
(1).

Incorporating q_i, the proportion of total population accounted for in group i, and q_i', the
same proportion in the time step, yields,

$$\Delta\bar{x} = \sum_i q_i' x_i' - \sum_i q_i x_i .$$

Noting that $\Delta x_i = x_i' - x_i$ gives us,

$$\Delta\bar{x} = \sum_i q_i'(x_i + \Delta x_i) - \sum_i q_i x_i .$$

We can relate q_i and q_i' by comparing the fitness of group i, w_i, to the mean fitness

across all groups as follows

$$q_i' = \frac{q_i w_i}{\bar{w}} .$$

Substituting this into (1) gives us,

$$\Delta\bar{x} = \sum_i \frac{q_i w_i}{\bar{w}}(x_i + \Delta x_i) - \sum_i q_i x_i = \sum_i q_i x_i(w_i/\bar{w} - 1) + \sum_i q_i(w_i/\bar{w})\Delta x_i .$$

Then, using the standard definitions of *Covariance* and *Expectation*, we arrive at the

Price Equation:

$$\bar{w}\Delta\bar{x} = Cov(w_i, x_i) + E(w_i \Delta x_i) .$$
(2)

The covariance term in the above equation neatly summarizes the effect of selection

between groups on $\Delta\bar{x}$. If w_i, the fitness group i, positively covaries with the frequency of the

5

trait in group i (x_i), then the covariance term will favor an increase in the trait. The

term $E(w_i\Delta x_i)$, however, conceals the effect of natural selection within groups among the

effects of transmission (i.e., mutation, non-random mating, etc.) We can separate out the

influence of selection within groups by re-applying the above technique to the term $w_i\Delta x_i$, which

is itself the product of two expectations—the average fitness of group i (averaged across the

individuals in group i) and the average change in the frequency of the trait in group i. To do this,

individuals within group i are indexed by j. For our purposes in this derivation, x_{ij} is the

frequency of an allele (trait) in individual j, and takes on the values of 1 (presence) or 0 (absence).

Substituting this into (2) yields:

$$\overline{w}\Delta\overline{x} = \underbrace{Cov(w_i, x_i)}_{\substack{Selection \\ between-groups}} + \underbrace{E_i(Cov(w_{ij}, x_{ji}) + E_j(w_{ij}\Delta x_{ij}))}_{\substack{Selection\&\ transmission \\ within-groups}} \qquad (3)$$

If we ignore any effects arising from the transmission process between parents and

offspring (e.g. recombination, mutation, etc.), then $\Delta x_{ij} = 0$. With this, and a standard definition

of *Covariance*, we can rewrite the equation as follows:

$$\overline{w}\Delta\overline{x} = \underbrace{\beta_{w_i, x_i} Var(x_i)}_{\substack{Selection \\ between-groups}} + \underbrace{E(\beta_{w_{ij}x_{ij}} Var(x_{ij}))}_{\substack{Selection \\ within-groups}} \qquad (4)$$

Equation (4) tells us that the change in the frequency of an allele created by natural

selection acting on individuals can be partitioned into between-group and within-group

components. The magnitude of the between-group component depends on the amount of variation

between groups and the size of the partial regression coefficient of the frequency of the allele (or

trait) within groups on group fitness. The sign of the between-group component depends on the

sign of the partial regression coefficient. If having a higher frequency of the trait predicts higher

group fitness (e.g., cooperation), then the between-group component is positive. Similarly, the

magnitude of the within-group component depends on the variation within groups and on the

partial regression coefficient of individual allele frequency on individual fitness within groups. The sign of the within-group component depends on the sign of the regression coefficient. Often the within- and between-group components have the same sign, so the partitioning provides little insight. However, if x tracks the frequency of an altruistic allele, then the within-group regression coefficients will be negative (by definition, altruists are exploited by non-altruists), while the between group regression coefficient is positive (groups with more altruists do better than groups with fewer). 'Genetic group selection' refers to the between-group component of natural selection acting on genes.[2]

The between-group component is often small relative to the within-group component

Under a wide variety of conditions, the between-group term in (4) is often small (or rapidly becomes small) relative to the within-group term. This occurs because migration, random group formation and other kinds of genetic mixing among groups deplete the variation among groups ($Var(x_i)$), while increasing (or at least maintaining) the variation within groups, $Var(x_{ij})$. To see this, consider a population consisting of two equally-sized groups, one initially composed entirely of kindly altruists, who bestow fitness benefits on all other members of their group, and the other of selfish egoists, who do not bestow benefits. This condition makes the between-group component that favors altruism as big as possible, $Var(x_i) = 0.25$, while making the within-group variation zero, $Var(x_{ij}) = 0$. According to equation (4), altruistic genes will initially spread, no matter what the relative difference (in magnitude) is between the regression coefficients because the within-group term is zero. If we keep the groups isolated for a long time and assume no mutation, the population will someday be dominated entirely by altruists. However, if we allow a small amount of migration between the populations, say 5% exchanged per generation, the variation between groups, $Var(x_i)$, will begin to decline and rapidly approach zero. At the same

---

[2] I specify 'acting on genes' because natural selection can act on any kind of heritable phenotypic variation. Both cultural transmission and the transmission of acquired immunities provide examples of non-genetic,

time, migration will begin driving the variation within-groups, $Var(x_{ij})$, towards its maximum value of 0.25. This migration rate, ignoring the fitness bounty reaped by immigrant egoists into the altruistic population, will reverse the initial values of the between- and within-group variations in about 40 generations. In general, a great deal of theoretical work shows that genetic group selection will only lead to substantial levels of altruism when groups are very small, migration rates are low, and the intensity of selection among group is high compared to the intensity of selection within groups (Rogers 1990; Crow & Aoki 1982; Aoki 1982; Boorman & Levitt 1980).[3] In short, these genetic group selection models work fine under the right constraints, however most researchers don't expect these constraints to be satisfied very often. Further, evidence from paleoanthropology and extant small-scale societies does not support such stringent constraints. To the contrary, the mating systems in many small-scale societies favor substantial genetic mixing between groups (Hartl & Clark 1989:300-301; Lee 1979; Richerson & Boyd 1998).

Multiple stable equilibria can resist the force of migration & maintain variation between groups

Unlike the particular problem of altruism/cooperation, it's important to realize that many different forms of social interaction (i.e. different payoff structures) have multiple stable (Nash) equilibria—locally optimal solutions. When the nature of local interaction produces multiple stable equilibria, the between-group component of selection can strongly influence the long-term equilibrium of the system. Simple examples include coordination interactions, interactions that combine coordination and conflict (e.g. battle-of-the-sexes games), mutualistic interactions (e.g. the 'stag hunt' game, Hirshleifer 1982; Connor 1995) and some models of reciprocity (Axelrod & Hamilton 1981; Boyd 1988). Unlike 'group-selected' solutions in cooperative interactions,

---

heritable variation that could be subject to natural selection.
[3] Kelly (1992) has shown that some models of the evolution of altruism, which yield very restrictive conditions for genetic group selection, contain an implicit form of density-dependent regulation that works

8

neither small group size nor low migration rates are necessary for between-group selection to have an important effect because *within-group selection* counteracts the mixing generated by migration. That is, within-group selection acts to reduce the variation within groups, while simultaneously maintaining the variation between groups. This allows between-group selection to favor groups at equilibria that produce the highest mean group fitness. Therefore, when populations are structured and social interactions produce multiple stable equilibria, the between-group component is likely to have an important influence on the final distribution of behaviors/strategies (Boyd & Richerson 1990). Thus, contrary to popularized claims about the general unimportance of group selection for understanding social behavior (laying aside cooperative social interactions), understanding the between-group component may be essential to explaining a wide range of behavior, especially in highly-social species.[4] Selection between human organizations, such as firms, may also result from between-group variation being stabilized by such mechanisms (Nelson and Winter, 1985: Chapter 5).

To illustrate, imagine two proto-human groups. Due to random variation, one group possesses a small number of mutants capable of using gestures (hand signals, body positions and facial expressions) to communicate with other such mutants, while the other group contains a similar small number of mutants capable of using verbalizations to communicate with other mutant-verbalizers. But, gestural communicators cannot communicate with verbalizers, and vice versa. If increased communicative abilities, in any form, are favored by natural selection in both groups, then the relative strength of selection on a particular form of communication (gestural vs. verbal) depends, in part, on the frequency of other individuals in the group who are capable of the same form of communication. Natural selection acting within the first group favors all gestural

---

against the between-group component. Decoupling the effect of migration from population regulation makes the conditions less restrictive, although still difficult to satisfy.

[4] Male vs. female exogamy provides an example of a social behavior with multiple equilibria. Assuming that there is a big cost to inbreeding and a polygynous mating system that produces lots of unknown paternal half sibs, selection should favor one sex leaving the natal group. It is plausible that in many cases there are two stable equilibria: 1) males leave and 2) females leave.

communicators, while in the second group it favors all verbalizers. Within-group forces will drive the first group towards all-gesturers and the second group towards all-verbalizers. Unlike cooperative dilemmas, migration between groups, which would tend to deplete the variation between groups, will be opposed by within-group frequency-dependent selection. If this within-group selection is strong enough, the variation between groups will remain high. Although verbalizers cannot invade the gestural group because of the coordination problem, groups of verbalizers may have a number of advantages over gestural communicators, such as: 1) communication at night, through walls, in dense forest, and around topographical features is substantially easier; 2) communication is possible while using your hands for something else; and 3) communicating while injured is easier. As a consequence, the mean fitness of the verbalizer-group may be higher than the first group. If so, in the long run verbalizer genes will be favored over gestural genes because of the genetic group selection (i.e. the between-group component). As we'll see later, cultural processes are even more likely to generate multiple stable equilibria, so the between-group component of cultural evolution is even more likely to be an important force of selection among culturally-evolved equilibrium.

Solutions to the problem of altruism are all based on constraints

The core of the issue

Let's return to the evolution of cooperation and altruism. Solutions to the genetic evolution of altruism work by taking advantage of some informational structure in the environment. Usually, this informational structure is maintained by some constraint that is imposed, either implicitly or explicitly, in setting up the problem. In generating altruism, natural selection 'looks' to exploit statistically reliable patterns of association, which usually result from the conditions of the world, or of the model. It has long been recognized that kin-based, reciprocity-based, and group-selection-based models of the evolution of altruism all work according to the degree in which 'being an altruist' (in this case, having an altruistic gene)

predicts that one's partners are altruistic. We can derive this using (2). Let i index individuals (rather than groups) in the population, so that x_i gives the frequency of altruistic alleles (0 or 1) in individual i. We are only interested in the effects of selection in one unstructured population, so we assume $\Delta x_i = 0$. Assuming $Var(x_i)$ does not equal zero, the frequency of altruists in the population will increase when

$$\beta_{w_i x_i} > 0.$$

If we express w_i using linear regression as (following Frank 1998),

$$w_i = \alpha + \beta_{w_i x_i \bullet \bar{x}} x_i + \beta_{w_i \bar{x} \bullet x_i} \bar{x} + \varepsilon,$$

where $\square$ is the baseline fitness, $\beta_{w_i x_i \bullet \bar{x}_i}$ is the partial linear regression coefficient of x_i on w_i (holding $\bar{x}$ constant), $\beta_{w_i \bar{x} \bullet x_i}$ is the partial linear regression coefficient of $\bar{x}$ on w_i (holding x_i constant), and ε is the uncorrelated error term. Taking the partial derivative of this fitness with respect to x_i and substituting it into (5), we arrive at a very general form of the conditions for the evolution of altruism, which we might also call "Hamilton's rule" (Queller 1992).

$$\beta_{w_i x_i \bullet \bar{x}} + \beta_{w_i \bar{x} \bullet x_i} \beta_{\bar{x} x_i} > 0$$

If we rewrite this using conventional notation (Hamilton 1975; Frank 1998), in which B represents the fitness bestowed on the rest of the group by an altruist (i.e. $\beta_{w_i \bar{x} \bullet x_i}$), and C (which is positive by convention) is the fitness cost to the altruist (i.e. $\beta_{w_i x_i \bullet \bar{x}_i}$), we arrive at

$$\beta B - C > 0 \quad (5)$$

where β (i.e., $\beta_{\bar{x} x_i}$) represents the partial regression coefficient of 'being an altruist' on the expected frequency of altruists in the population. It tells of the degree to which being an altruist predicts the frequency of other altruists in the group. β is *not*, in general, a measure of relatedness by descent from a recent common ancestor—although, as we'll see, that is *one way* to get a positive value of β. Note that, Hamilton (1975) re-derived his famous expression (equation (5),

11

Hamilton 1964) using the logic of selection within and between groups (i.e., equation 4) a quarter of a century ago, but the theoretical implications of his demonstration are still not widely appreciated by students of evolution and human behavior.

Equation (4) indicates that the trick to solving the problem of altruism relies entirely on evolving, or at least maintaining, a positive statistical relationship between being an altruist and bestowing benefits on other altruists (i.e. being in a 'group' of other altruists). All models for the evolution of cooperation somehow create, or at least sustain, a positive value of β. To understand models of altruism, it is important to identify and evaluate (if possible) the particular constraint or constraints that sustain the non-random association captured in β. Remember, however, that the greater the value of β, the greater the amount of altruism that can evolve, and the greater selective pressures for mutant genes that can 'beat the system' by exploiting β.

Many students of evolution and human behavior regard kinship- and reciprocity-based models of altruism as more acceptable and legitimate solutions to the problem of altruism than 'group selection' hypotheses. In my view, the acceptability or legitimacy of a theoretical solution depends on an evaluation of the constraints that give rise to the statistical association (β), which is what allows for the evolution of cooperation. The applicability of particular theories of altruism (i.e. particular constraints) depend on the details of particular species. The details of an organism's social structure, migration patterns, genome, cognitive abilities, migration patterns or imitative abilities may support some hypotheses and undermine others. Of course, by emphasizing 'constraints' I do not mean to suggest that each 'solution' is equally likely to be observed in a randomly selected species. Surely, some constraints are more frequently satisfied in nature than other constraints, however, the rarely satisfied constraints may provide the most interesting forms of altruism. Below, I'll briefly discuss some of the commonly deployed approaches to the problem of altruism and highlight the constraint(s) involved. In doing this, I

hope to show that explanations other than those based on kin and reciprocity have, at least in some specific cases, should an equal claim to legitimacy and acceptability.

Greenbeards

I'll begin with the classic example of a solution to the altruism dilemma based on an 'illegitimate' constraint: the "greenbeard" problem (Dawkins 1976). Imagine a gene that causes its bearer to have a greenbeard and to cooperate with other greenbearded individuals. Greenbearded cooperators will merrily focus their benefits only on other cooperators. Thus, even when rare, natural selection will strongly favor this altruistic gene because $\beta = 1$ (unless they are so rare that fellow greenbeards never find one another). More generally, if altruists can recognize other altruists (with a better than random chance, $\beta > 0$) and preferentially bestow benefits on them, then altruism can be favored by natural selection, depending on (5). In this solution, the informational regularity exploited by natural selection is the perfect correlation between greenbeards and altruism. The problem with this solution is that this informational regularity arises from a genetic (or physiological) constraint on the power of mutation to create selfish egoists with greenbeards (greenbearded defectors). Such a non-altruistic gene will exploit the greenbeard-cooperator regularity (β), and destroy it in the process. Egoists with greenbeards will invade until the regularity is gone—β is driven to 0. Consequently, this model works only if we allow the mutation constraint. Is this constraint acceptable? Popular writers on evolution often scoff at the absurdity of such a constraint, although the available empirical evidence is somewhat mixed. Experimental work suggests that people have some ability to distinguish cooperators from defectors (Frank et. al. 1993), while other work says they cannot (Ockenfels & Selton 1998; Ekman 1992; Ekman & O'Sullivan 1991; Henrich & Smith 2000).

The greenbeard problem can be addressed as a problem in costly signaling (Frank 1988). If it costs less for cooperators to cue their strategy than for non-cooperators to falsely cue, then cooperation can be favored. This occurs because such signals allow the benefits of cooperation to

be preferentially delivered to other cooperators. From the perspective of equation (4), cooperation can evolve because this signaling allows for the non-random formation of groups, which reduces the variation within groups (mostly cooperators, or mostly defectors) and increases the variation between groups (Wilson & Dugatkin 1997). Unfortunately, many models of cooperation simply assume that individuals know the strategies of other individuals, without providing any justification (theoretical or empirical) as to why this should be so. This is tantamount to assuming the answer (β).

There are four problems with the costly signaling approach to greenbeard altruism. First, it relies on the constraint that it's cheaper for cooperators to send the cues of cooperation, than for defectors. Mutations that allow defectors to send these signals for the same cost as altruists cannot be permitted—this is a key constraint. Second, as noted above, empirical evidence for an ability to distinguish 'cooperators' from 'defectors' without extensive interaction is contradictory and ambiguous. The ability of good actors (e.g., in the movies) to evoke powerful emotions in us by pretending to be either a valiant 'cooperator' or a malevolent 'defector' suggests there is some variation out there in the ability to send false signals (why aren't we all good actors?). Third, we also want to explain cooperation and trust among anonymous people (who cannot see one another) in one-shot encounters, such as that which we observe in experimental economic games (Kagel & Roth 1995; Davis & Holt 1993). In most of these situations, there's no chance for signaling of any kind. So, at best, this approach solves only one kind of altruism problem. Fourth, if such costly signaling (involving emotional commitments) explains the degree and kind of cooperation in humans, why hasn't this mechanism generated much more cooperation in non-human animals, such as chimpanzees, elephants and dolphins?

Kin-based Selection

Kin selection models also make use of constraints. In some kin-based approaches, the nature of a species' social (or family) structure produces a whole host of possible informational

14

regularities that can be exploited by natural selection. For example, if sisters tend to hang around their mom, genes that cause their bearer to cooperate with others 'hanging around mom' may spread because—if they are really your sisters—they have at least a 50% chance of having the same altruistic gene as you. 'Hanging around mom' merely provides one possible predictor of kinship, and indirectly, of altruism. Consequently, social structures in which individuals interact with relatives more frequently than non-relatives provide a potentially valuable informational regularity. If 80% of those receiving benefits from an altruist are full siblings of the altruist and the rest are non-relatives, then $\square$ is approximately 0.40 (assuming the frequency of altruists in the rest of the population is very small).

This social structure constraint is only exploitable if, by whatever means, the underlying non-random association can be maintained. If non-altruistic mutants, posing as fake-family members, can frequently slip into nests or family groupings and obtain the benefits, then kinship solutions will suffer the same problem as the greenbeard-models. Among animals lacking the requisite social structure (e.g. snakes), we should not expect the corresponding forms of altruism. Furthermore, the social structure—which is what often creates the requisite informational regularity—is usually taken as a given, rather than being a product of evolution with its own costs and benefits. Models rarely endogenize the costs of maintaining a particular family or social structure in analyzing the evolution and maintenance of altruism.

Other kin-based solutions are founded on kin-recognition mechanisms (Fletcher & Michener 1987) that allow individuals to probabilistically spot their kin and (among other things) direct altruism towards them. Qualitatively, this solution is the same as the greenbeard solution, except that having many of the same genes, being reared together, and the potential existence of other evolutionary pressures favoring kin recognition (e.g. inbreeding·avoidance) makes the explanation much more plausible—though not fundamentally different. Combinations of kin-cues may make it quite costly for natural selection to produce faker-defectors.

Finally, the rules of inheritance, which make kinship a valuable predictor of altruism, are often taken as given despite the fact that certain alleles can violate these rules (e.g. meiotic drive). A non-altruistic gene that is transmitted from parent to offspring with a greater than 50% chance (contrary to Mendelian rules) could spread and gradually corrode the informational regularity upon which some forms of kin altruism are based.

Because kin-based mechanisms should be designed to focus benefits only on close relatives, kin-selection does not help us solve the problem of cooperation among large groups of unrelated individuals, unless our kin-psychology is making a lot of big mistakes by confusing large numbers of non-relatives with relatives. This version of the "big mistake hypothesis" (Boyd & Richerson, In press) proposes that, because our psychology evolved in small groups with high degrees of interrelatedness, kin-selection favored a psychology in humans that is designed to generously bestow benefits on members of their groups. However, in the supposedly novel world of large-scale, complex, societies, this once adaptive mechanism misfires, and we get large-scale cooperation (Tooby & Cosmides 1989).

There are a number of problems with this explanation. First, even in small-scale societies, there are plenty of distant relatives that altruists need to distinguish from close kin—β (relatedness by descent for alleles not on the Y-chromosome) decreases geometrically as the circle of kin expands from sibs, to half-sibs, to first cousins, etc. Second, although large-scale cooperation is prevalent in many societies, people everywhere favor their kin over non-kin— showing that we can, and do, distinguish these behaviorally (Daly & Wilson 1988; Westermarck 1894; Sepher 1983; Wolf 1970). And third, lots of non-human primates also live in small-scale societies, but show no generalized tendency to cooperate with all members of their group. When non-human primates are placed in larger (evolutionarily-novel) groups, cooperation does not expand to the enlarged group (Boyd & Richerson, In press, summarizes the empirical evidence against this). It's hard to believe non-human primates do not make these big mistakes, but humans do.

16

Reciprocity

Reciprocity provides another well-studied solution to the altruism dilemma. In reciprocity-based models with repeated interactions (usually modeled as a repeated prisoner's dilemma), it has been frequently shown that strategies that both bestow benefits on individuals who have bestowed benefits on them in the past, and that withhold future benefits from strategies that fail to reciprocate, can be favored by selection if they are sufficiently common (Axelrod 1984; Axelrod & Hamilton 1981; Trivers 1971). Repeated interaction using these kinds of reciprocating strategies produces the requisite informational regularity (β). The longer the repeated game (the more rounds), the higher β can be (and in this case β depends on the frequency of different reciprocal strategies).

However, making such models work requires a couple of assumptions, conditions, or constraints that often go unstated. First, most models analyze only two strategies at a time (e.g. Axelrod 1984), and don't consider the presence of additional strategies (even one additional strategy) maintained at low frequency by mutation, immigration or non-heritable environmental variation. It turns out that if the full range of other strategies is allowed to occasionally mutate into existence at low frequencies, then *no pure strategy,* such as "tit-for-tat" (TFT) and "always defect" (ALLD),[5] is evolutionarily stable (Boyd & Lorberbaum 1987). For example, if the game is repeated a sufficient number of times, tit-for-two-tats (TF2T) can invade a population of mostly TFT if "suspicious tit-for-tat" (STFT) is present at low frequency. TF2T is like TFT but requires two consecutive defections before defecting, while STFT defects on the first round and then plays TFT. Similarly, TF2T can be maintained at high frequency with low frequencies of STFT and TFT, unless ALLD enters the fray at low frequency, in which case it's possible for STFT to invade and become common. Notably, in a population of mostly STFT, we will observe mostly defection. The point here is, whether or not reciprocating strategies can remain stable depends on

---

[5] TFT cooperates on the first round, then plays whatever its partner played on the previous turn. ALLD always defects on every turn.

exactly how the mutational spectrum is restricted (errors can take the place of mutations, but then one needs to know that probability distribution of errors). Interestingly, proponents of reciprocity criticize greenbeard explanations because they require restricting the mutational spectrum. [6]

Second, it's also important to realize that most theoretical models of reciprocity explore only 2-person interactions. Despite this, many scholars have falsely assumed that the qualitative aspects of the 2-person result can be generalized to n-person situations (e.g. Patton 2000). However, Boyd & Richerson's (1988) analysis of an n-person repeated prisoner's dilemma shows that the results do not generalize for groups larger than about ten individuals. Among other discouraging results, they show that the threshold frequency of reciprocators necessary to maintain stable cooperation in an n-person group increases with group size by the power $1/n$. They further show that combining kinship (relatedness by common descent) with reciprocity does not substantially improve the result, as it does in the 2-person case. Reciprocity, on its own, is unlikely to solve the problem of cooperation in large groups (also see Bendor & Mookherjee 1987; Joshi 1987).

The final implicit constraint in reciprocity models lies in the fact that individuals must be able to accurately recognize their partner(s) in repeated games. If there are mutants capable of posing as the individual(s) who bestowed benefits in the last turn, then the benefits of reciprocity evaporate (as β 0). Some might think that, while this is a problem for cooperation in cognitively-limited animals, it's not a serious problem for big-brained humans. Perhaps it's simply too costly for natural selection to produce doppelgangers capable of fooling human cognition—a cognition which is presumably maintained for other evolutionary reasons (which puts the costs of that cognition outside the problem). However, to the contrary, this is a serious problem for the kind and degree of cooperation we observe in many societies. Consider the

---

[6] However, if pure strategies sometimes make mistakes by cooperating when their strategy indicated defection, or defecting when they wanted to cooperate, some pure strategies can be evolutionarily stable. Stable pure strategies include both reciprocating and non-reciprocating strategies, but not TFT (Boyd 1989).

effects of increasing the group size from 2 to 2000, or of increasing the delay between the time you receive benefits and the time when your next chance to bestow them occurs. In large-scale societies, reciprocators would potentially need to keep track of the acts of hundreds or thousands of individuals for years. Recently, a guy that I have never seen before (i.e. don't remember seeing) volunteered to help me jump-start my stalled car. I thought this was a very nice gesture, but I've already forgotten what he looks like. A wide range of overweight, brown-haired, white guys could easily pose as my benefactor, if they were so motivated.[7]

Using the same logic as for kin-selection, some evolutionary psychologists have argued that, although reciprocity is not adaptive in large-scale cooperative societies (including our societies), it was adaptive in the small-scale, face-to-face, societies of our evolutionary history, so it remains part of our psychology (Alexander 1974, 1987; Hamilton 1975; Tooby & Cosmides 1989). In large-scale complex societies, this reciprocal psychology causes us to mistakenly cooperate when we should defect. This explanation suffers the same problem as the kin-selection version of the "big mistake hypothesis." Boyd & Richerson (In press) summarizes the evidence against this explanation.

Indirect reciprocity

By assuming individuals know something of their partner's previous behavior, models of indirect reciprocity attempt to explain generalized cooperation by expanding on the idea of reciprocity (Alexander 1987). In a recent model of this (Nowak & Sigmund 1998), individuals decide whether or not to bestow benefits on a partner based on that partner's reputation for bestowing benefits on previous partners in past interactions. They show that the probability of knowing the reputation (accurately) of one's partner must exceed the cost-to-benefit-ratio of the altruistic act. That is, they re-derived a simplified version of equation (5), for a specific case.

---

[7] I've had a lot of car problems over the years, so that just the latest in a string of similar examples.

The success of this model rests entirely on two constraints (1) that individuals can acquire accurate reputational information, and (2) that the predictive value of reputational information is stable (resists systematic manipulation). It is thus another form greenbeard model. If a mutant defecting-strategy can generate an inflated reputational signal, perhaps by paying other individuals to lie about the mutant's cooperative tendencies, then the predictive value of reputation will corrode and cooperation will collapse.

Accepting this constraint, indirect reciprocity still fails to explain large-scale cooperation and altruism in one-shot, anonymous, situations. The amount of cooperation supported by indirect reciprocity declines exponentially with increasing group size (Nowak & Sigmund 1998),[8] so large-scale cooperation is not explained. And, if the reputation information is removed from the model, indirect reciprocity can only sustain cooperation in very small groups (Boyd & Richerson 1989). In contrast, empirical evidence shows that people cooperate in both experimental and real settings without any reputational information, and increasing group size does not reduce levels of cooperation (Ledyard 1995; Henrich & Smith 2000).

Punishment

Many scholars have attempted to solve the problem of cooperation in large groups by incorporating punishment (e.g. McAdams, 1998; Hirshleifer & Rasmussen, 1989; Fundenberg & Maskin, 1986; Axelrod 1986). If cooperators punish defectors, then cooperation can be favored. However, if punishment is costly for the punisher (which it must certainly be), then cooperators who don't punish can invade because they avoid both the costs of being punished (for not cooperating) and the costs of punishing defectors (they are 2nd order free riders). If the private benefits derived from punishing are greater than the costs of administering it, punishment may initially increase, but cannot exceed a modest frequency (Boyd & Richerson 1992). Using

---

[8] This claim may not be clear from Nowak & Sigmund (1998), but see the caption of Figure 3. The authors write, "the averages over time of the frequency of cooperative strategies (defined by k ≤ 0) are 90%, 47%, and 18% for, respectively, n = 20, 50 and 100" (Nowak & Sigmund 1998: 574).

punishment, the problem of cooperation becomes one of how to maintain punishers in the population. One way to do this is to limit the mutational spectrum by eliminating strategies that cooperate, but don't punish (e.g. Axelrod 1986; Axelrod and Dion 1988; Gintis 2000). How we evaluate this depends on how likely such a mutational constraint is in nature. Another way to solve the problem of punishment is to incorporate a recursive punishing strategy in which punishers punish individuals who don't cooperate and those who fail to punish individuals who don't punish (Boyd & Richerson 1992). This solution is a mathematical trick that eliminates the cost of punishment by spreading it out over an infinite space. Do people, or can people, track defections through a nearly infinite set of stages? Later, I'll discuss a culture-gene coevolutionary solution, based on punishment, for the evolution of one-shot, n-person, cooperation (Henrich & Boyd 2000) that builds on a third, more-plausible, constraint.

Viewing kinship and reciprocity as group selection

It's my view that 'kin selection', 'reciprocity' and 'group selection' are historically derived labels for different types of constraints—constraints that generate opportunities for natural selection to solve the problem of altruism. Often the labels seem to depend mostly on how the problem was initially set up. For example, why not consider explanations of altruism based on kin-recognition as plausible forms of 'greenbeard-selection'?[9] Many students of evolution and human behavior do not realize that 'inclusive fitness' (which was first used to derive (5) using kinship; Hamilton 1964), 'individual fitness' (which was used to derived (5) for 'reciprocal altruism', Trivers 1971), and 'group selection' (i.e. the partitioned Price Equation, equation (4), which was used to derive (5) in Hamilton 1975) are simply three systems of gene-tracking and fitness accounting from three different perspectives. Any solution can be reformulated from each perspective to yield the identical answer. Hamilton (1975), for example, reformed kin-selection using the Price Equation, instead of inclusive fitness.

---

[9] Note that sexual recombination does not work against kin recognition in the same way that it does against non-kin Greenbeard genes.

To demonstrate the relationship between 'group-selection' solutions with those based on reciprocity, I'll analyze a standard, 2-person, reciprocity model using the partitioned Price Equation (4). In doing this, we'll observe that reciprocity favors cooperation when the between-group component overpowers the individual component. Of course, it's equally possible to re-derive typical 'group-selection' solutions (selective emigration, assortative interaction, etc.) using inclusive fitness, or even by tracking individual fitness. None of the accounting techniques can claim general superiority over the others (cf. Sober & Wilson 1998). However, despite the fact that it is possible to do this, it's neither easy nor desirable to do so (Queller 1992; Frank 1998).

In this prisoner's dilemma game, individuals are paired into groups and play for m rounds. During each round, each individual plays either 'cooperate' (C) or 'defect' (D). Table 1 gives the payoffs received by the row player. The index i labels three groups 1, 2, 3. Each group consists of 2 individuals labeled $j = 1$ or 2. Group $i = 1$ contains two individuals who both play the strategy tit-for-tat. Group $i = 2$ contains one tit-for-tat'er and one defector (who always plays D). Group $i = 3$ contains two defectors. The variable x_{ij} equals 1 for tit-for-tat'ers and 0 for defectors. The variable x_j represents the frequency of tit-for-tat'ers in group i. Equation 10 will provide the change in frequency of tit-for-tat'ers. Let's calculate the components of (10).

$$E(\beta_{w_{ij}x_{ij}} Var(x_{ij})) = \frac{1}{3}\left[(0)(0) + (-4)(.25) + (0)(0)\right] = -0.33$$
$$Var(x_i) = 0.167$$
$$\beta_{w_i,x_i} = \frac{2}{3}(m+1)$$

Substituting these into equation (10) gives us,

$$\overline{w}\Delta\overline{x} = \underbrace{0.11(1+m)}_{BetweenGroup} - \underbrace{0.33}_{WithinGroup}$$

This equation shows that it is the between-group component of natural selection that favors reciprocity (in this case tit-for-tat), not the individual selection. In fact, the within-group component is always negative—meaning 'individual selection' always selects *against* reciprocity.

Within groups, tit-for-tat never does better than a fitness tie (with another tit-for-tat'er), and always loses against defectors—thus the within-group regression coefficients, $\square_{ij}$, are always zero or negative. The magnitude of the positive between-group component (i.e. group selection) depends on m, the number of rounds of play. The variable m essentially controls the migration rate between groups. As m increases, groups become more stable because they re-mix at lower rate. When $m = 1$, groups completely remix after every round and the between-group component is minimized. Reciprocity is favored when m is sufficiently large that teams of reciprocators can run-up their combined fitness total sufficiently to overcome the relative fitness losses they suffer at the hands of defectors, within groups (for more on this, see Sober & Wilson 1998). The fact that the between-group component is what drives reciprocity is often lost because it's substantially easier to solve this problem using the 'individual fitness-accounting' approach—in which a strategy's relative fitness is calculated by averaging across all possible groups.

[Table 1 about here]

The same logic applies to reciprocity models in which the groups are larger than 2. Increasing the group size reduces the variation between groups, which reduces the size of the between-group component, and makes reciprocity less likely to evolve. This also increases the within-group advantage of invading defectors because they can reap benefits from more than one individual. This is why, as I've mentioned, reciprocity models are unlikely to explain cooperation in large groups.

Kinds of Cooperation and Altruism

We are not limited to picking only one solution to the altruism dilemma. Our psychology may embody the effects of natural selection having taken advantage of different stable informational regularities over our evolutionary history. For example, the nature of the human family—with its division of labor, stability, and monogamy (relative to our closest primate cousins; Boyd & Silk 1997; Klein 1989)—suggests that human ancestors were likely to have

repeatedly found themselves in small sub-groups (called 'families') in which other individuals were likely to have the same genes by descent from a recent ancestor. Consequently, like other non-human primates, we are likely to have some cognitive machinery dedicated to bestowing benefits on kin. Similarly, life in small-scale, fairly-stable, groups may provide for lots of repeated interaction with unrelated or distantly related individuals. Thus, assuming our cognition or emotional system was complex enough that mutants could not easily pose as friends, then we should expect a psychology that favors helping well-known, reliable reciprocators (i.e. friends). Unfortunately, kin-selection can only explain cooperation among close kin, and reciprocity (including indirect reciprocity) is limited to small groups with lots of repeated interaction (putting aside the "big-mistake hypothesis"). Neither of these solutions seems capable of explaining the large-scale cooperation among non-relatives that we observe in both modern and pre-modern societies, including foraging societies.[10] Experimental findings from many small- and large-scale societies show that people will trust, cooperate and not exploit anonymous individuals in simple one-shot games (Henrich et. al. 2000; Roth et. al. 1991; Bateson & Shaw 1991)—a finding further confirmed by lots of ethnographic evidence (Richerson & Boyd 1998, 1999).

Furthermore, none of the other explanations I've discussed above seem up to the task of explaining large-scale cooperation (also see Boyd & Richerson 2000). Laying asides its empirical problems, solutions based on our ability to spot other "cooperators" based on signals, without repeated interaction (Frank 1988), cannot explain large-scale public goods cooperation like warfare, waiting in line, conservation, recycling, and voting. Furthermore, adding punishment to models of cooperation can solve the cooperation dilemma on large scales, in an intuitively

---

[10] There are two things worthy of note with regard to foraging societies. First, the nature of sharing, especially of sharing game, cannot be explained by either reciprocity- or kin-based theories (Bliege Bird & Bird 1997), even in the smallest small-scale societies. Many of the simplest societies such as the !Kung, have quite strong institutions for maintaining cooperation on a scale much larger than the family or the band (Wiessner, 1983). Second, contrary to the view of foraging created by anthropologists studying extant groups, lots of archaeological and ethnohistorical data indicates that foraging societies can be politically, economically and socially complex, with large-scale cooperation, social stratification and a substantial division of labor (Arnold 1996).

pleasing fashion (i.e. our experience tells us that defectors do get punished). However, no one has explained the maintenance of widespread punishment without some kind of unpleasant (and in my view, unlikely) constraint.

Above, I've argued that there's no *a priori* reason to dismiss 'group selection' on theoretical or logical grounds, and that the success of each solution depends on the plausibility of the constraints that it invokes to maintain the required informational regularity. However, given what we know about human evolutionary history, genetic 'group selection' models that rely on low or heavily biased migration between groups probably don't apply to humans, as there probably was plenty of migration and intermixing in prehistory, as there is in extant small-scale societies.

In the final part of this paper, I'll show why cultural evolution and culture-gene co-evolution are quite plausible explanations for the large-scale cooperation observed in many human societies. By applying the Price Equation to cultural evolution, I'll explain why, unlike in genetic evolution, the between-group component is likely to be larger than the within-group component in a wide range of circumstances. I'll show how a simple combination of cultural transmission mechanisms (imitation biases) and culturally-transmitted punishment can produce stable, large-scale, cooperation in one-shot public-goods games. Finally, using a coevolutionary version of the Price Equation, I'll demonstrate that prosocial genes (those favoring cooperation and punishment) can spread into a population in the wake of cultural evolution.

Cultural group selection, cultural transmission & large-scale cooperation

Humans (at least in some societies) cooperate on a larger scale than any other species, with the possible exception of eusocial insects. Interestingly, humans are also the most proficient at, and most reliant on, social learning to acquire behavioral practices and strategies. Our cognitive abilities to acquire information via imitation and other forms of direct social learning far exceed that of any other species (Tomasello 2000). Many operational details of these learning

mechanisms appear to have been 'designed' by natural selection to extract useful information from the social world—that is, from the minds of our conspecifics (Boyd & Richerson 1985; Henrich & Gil-White 2000; Henrich & Boyd 1998). Perhaps it's merely an interesting coincidence that humans are both the most cooperative species, and also the most reliant on, and proficient at, social learning (Tomasello 1994, 2000). However, I argue that the nature of our cultural transmission capacities, and of human psychology more-generally, creates stable behavioral equilibria, consisting of combinations of cooperation and punishment, that are not available to genetic evolutionary processes in acultural species. The existence of these additional, culturally-evolved, behavioral equilibria make the group selection component of cultural evolutionary processes a much more powerful, relative to the within-group components of cultural transmission, than can occur in genetic evolution. Further, I also show how, once cultural processes have stabilized certain behavioral patterns, the within-group component of natural selection (acting on genes) can favor altruistic or other prosocial traits. In the context of the discussion above, cultural group selection creates stable informational regularities that natural selection (acting on genes) can exploit. By producing, as an incidental byproduct, reliable informational regularity in the human environment, cultural transmission creates the conditions for natural selection to favor prosocial genes that could not otherwise be favored in mammalian social species because non-humans lack the requisite social learning capacities (e.g. high fidelity imitation) that give rise to the coevolutionary process.[11]

We have empirical reasons to believe that culturally-transmitted ideas, beliefs and values (i.e. information) are important for understanding human cooperation. First, unlike other animals, the domains of cooperative behavior in humans vary from place to place and from group to

---

[11]Two notes. First, for a model and discussion of why more animals don't have the requisite social learning abilities see Boyd & Richerson (1996). Second, many people have the idea that human 'intelligence' or 'high-level' cognitive abilities can account for our level of cooperation. These explanations generally fail because such capacities facilitate deception, deviousness and defection, just as much—if not more—than they facilitate cooperation. The problem of cooperation is just as acute for rational actors as for the mindless biological automaton of evolutionary theory.

group. In some societies, people may cooperate in fishing and house-building, but not warfare. In neighboring groups that inhabit the same physical environment, folks may cooperate in warfare and fishing, but not house-building. So, unlike cooperation in eusocial insects and kin-based altruism in non-human primates, there is a tremendous amount of variation in the cooperative domains among human groups that is independent of differences in physical environments or local ecologies (Kelly 1985; Henrich & Boyd 1998; Henrich et. al. 2000). Second, unlike non-human animals, the scale of human cooperation varies from no cooperation outside the nuclear family (Johnson 2000; Johnson & Earle 1987) to massive cooperation on the level of nation states containing millions of individuals. The scale of cooperation in non-human social groups does not vary much among groups, only among species. Taking into account both this variation in cooperative domains and in cooperative scales, cultural evolutionary processes seem much more likely to generate and explain these patterns of variation than genetic evolution acting alone—especially given the modest genetic heterogeneity found among humans as a species, as compared to other species (such as chimpanzees), and the relatively recent (last 5,000 years) rapid emergence of very large scale cooperation.

Cultural Group Selection

Interestingly, 'group-functional' (group selective) explanations for cooperation and other forms of group-beneficial behavioral patterns have long been part of anthropology and sociology (Spencer 1891; Rappaport 1994; Harris 1977; Vayda 1971; Turner and Maryanski 1979). However, by the end of the 1970's the anti-group-selection movement had penetrated cultural anthropology, and, by analogy with biological evolution, was used to argue that the individual was the relevant level of analysis in cultural evolution (Harris 1979: 60-61). The problem was that anthropology lacked a sufficiently clear understanding of the differences between genetic and cultural evolution to understand why between-group processes that were unlikely to account for

seemingly group-beneficial behavioral traits in genetic evolution could still operate effectively

(even rapidly; Boyd & Richerson, manuscript) in cultural evolution. Below, I explain why.

The Price Equation derived at the outset of this paper, with genetic evolution in mind,

turns out to be a very general statement about any evolutionary system, which we'll use to frame

our thinking about cultural evolution. The 'altruist gene' (x) that we focused on in the earlier

derivation could be any characteristic of an evolving system, including the frequency of hydrogen

atoms in a cluster of galaxies or a quantitative phenotypic measure like I.Q., 'cooperativeness in

group fishing ventures', managerial success, malaria resistance, or the variance in height of

brothers. To discuss cultural evolution using the Price Equation, I'll rename the variables to avoid

confusion. ϕ will represent a quantitative phenotypic trait that can be influenced by cultural

transmission. As a behavior or strategy, ϕ_{ij} could measure individuals' willingness to die for their

country (or tribe) in war, the amount of time or money an individual contributes to charity, or

how much of a hunter's total 'catch' he brings back to camp to share (as opposed to that portion

he eats alone at the kill site). Replacing w_{ij} with f_{ij} gives us the cultural fitness for particular

values of ϕ_{ij}, and equation (6).

$$\bar{f}\Delta\bar{\phi} = \underbrace{\beta_{f_i,\phi_i}Var(\phi_i)}_{\substack{Selection \\ between-groups}} + \underbrace{E(\beta_{f_{ij},\phi_{ij}}Var(\phi_{ij}))}_{Selection\,within-groups} \qquad (6)$$

Cultural fitness measures the degree to which a particular value of ϕ, which represents stuff

stored in the head of individual j in group i at time t, affects its proportional representation in the

population at time $t + 1$. It may be thought of as replicability, transmissibility or simply

influence—note that no where in this formulation is there a need for cultural stuff to 'replicate' or

be discrete (for a discussion of the confusion created by cultural 'replicators', see Boyd &

Richerson In Press). The cultural fitness of ϕ_{ij} is jointly determined by the operational details of

our social learning psychology and by how that psychology interfaces with the environment (see

Boyd & Richerson 1985 for an extended treatment of this).

Cultural transmission, human psychology & between-group variation

As I discussed earlier, the main problem with 'group-selection' is the maintenance of between-group variation in the face of migration or other forms of genetic mixing. Unlike genetic evolution, several different mechanisms will act to reduce the within-group term in equation (6), $Var(\phi_{ij})$, while maintaining the between-group term, $Var(\phi_i)$, even in large groups and in the face of substantial migration. I will discuss four of these mechanisms. The first two are rooted in the details of how our cognition exploits the distribution of behaviors and ideas among members of a social group in order to 'decide' which of these traits to acquire. As I'll show, both evolutionary modeling and empirical data support the existence of these cultural transmission mechanisms. The third and fourth mechanisms are not cultural transmission capacities, but rather psychological 'tastes' or preferences for (1) avoiding behaviors that deviate from the common pattern, and for (2) punishing those who do not conform to the expected pattern. In my view, these two 'tastes' are probably either the products of purely cultural evolution (driven by cultural group selection), or coevolved products of genes responding to the novel social environments created by cultural group selection. Thus, they cannot be responsible for initiating the coevolutionary process that led to high levels of cooperation in humans. Nevertheless, once brought into existence by the first two mechanisms, they further catalyze the cultural group selection process by bolstering the forces that create and maintain differences between groups. From the perspective of social interaction occurring in single groups, these four mechanisms create many novel equilibria that do not exist in genetic systems, including equilibria with high-levels cooperation and punishment in one-shot, *n*-person games. Cultural Group Selection supplies a process that selects among alternative stable cultural equilibria (Boyd & Richerson 1990).

The first mechanism, *conformist transmission,* is a psychological propensity to preferentially copy high frequency behaviors. By biasing individuals in favor of copying common behaviors or behavioral strategies, this transmission bias tends to homogenize social groups. There are both theoretical and empirical reasons to believe that humans possess a tendency to

preferentially copy the most common behavior. Theoretically, Henrich & Boyd (1998) have shown that genes favoring a heavy reliance on social learning and *conformist transmission* (copying the majority) can outcompete genes favoring individual learning in both spatially and temporally varying environments. This model predicts two important things 1) that individuals should increase their reliance on social learning when individual (or environmental) information becomes less certain or as the difficulty of the problem increases, and 2) that individuals should rely on copying the majority (conformist transmission) under a wide range of conditions (also see Boyd & Richerson 1985, and Ellison & Fundenberg 1993).

Independent experimental work in psychology supports both predictions, as well as a number of other predictions arising from this model. Psychologists studying conformity have shown that, as a task's difficulty and financial incentives rise, individuals *increase* their reliance on imitation (vs. individual analysis) regardless of whether others will know how they behave (reducing any fear of social sanctions; see Baron et. al. 1996; Insko et. al. 1985). Furthermore, with real money on the line, other experiments show that individuals rely on copying the majority in *social dilemmas*, both when self-interest conflicts with the group-interest, and when self-interested choices correspond to group-interested choices (Smith & Bell 1994; Wit 1999). Finally, Henrich (2000b) shows that the slow take-offs, and 'critical mass tipping-point' observed in many empirical studies of the diffusion of innovations are quite consistent with the effects of conformist transmission.

The presence of this adaptive bias on our social learning cognition means that, in the absence of unambiguous information from the environment, or other decisive social learning stimuli (such as a prestigious individual, see the next mechanism), individuals will preferentially copy the most common ideas, beliefs, values and practices. Because new immigrants and the offspring of immigrants will preferentially adopt the common practices, conformist transmission can maintain group differences in a way that genetic transmission cannot—because offspring

30

acquire their genes from their parents, not from the group.[12] Consequently, as a by-product of its

evolved design, conformist transmission decreases the phenotypic variation among individuals

within groups, thereby depleting the strength of both within-group cultural and within-group

genetic forces (both of which operate on phenotype). As stochastic forces like cultural drift

(sampling errors transmission), biological shocks (e.g. plagues) and environmental disasters

introduce random variation between groups, conformist transmission will act to maintain this

variation—variation that would otherwise be depleted by migration between groups, natural

selection and pay-off-biased forms of cultural transmission. Thus, by reducing within-group

variation, and increasing between group variation, conformist transmission provides the raw

materials for cultural group selection.[13]

Other cultural transmission mechanisms can create the same effect through other means.

If individuals possess a psychological bias to preferentially copy people who are both more

successful (get higher payoffs) and similar to themselves in some marker trait like language or

dress, then, under a wide variety of conditions (even with substantial migration rates), these

cultural transmission mechanisms push the variation within groups towards zero (opposing the

force of migration), while sustaining substantial amounts of variation among groups (McElreath

et. al. 2000; Boyd & Richerson 1987).

There are both theoretical and empirical reasons to believe that people preferentially copy

successful individuals and that people possess some preference to copy people like themselves.

Recently, Gil-White and I have argued that, with the rise of imitation in the human lineage,

natural selection favored cognitive abilities to rank potential models according to their payoffs,

and preferentially imitate highly ranked models. Among the cues individuals use to rank potential

---

[12] This is consistent with a substantial amount of work in psychology and behavioral genetics showing that children do not acquire much, via social learning, from their parents (Harris 1998).

[13] This process explains why humans have different 'cultures' and other animals don't. That is, conformist transmission provides one important reason why people in the same social group tend to believe the same things and why these beliefs persist over long periods. Without a conformist component to create 'cultural

models is the amount of prestige-deference an individual receives from other people. This deference acts as an honest signal of who other individuals believe is highly successful or skilled because deference is 'paid' to such individuals in exchange for copying opportunities. This rank-based copying bias, which we call *prestige-biased transmission*, allows individuals to shortcut environmental or trial and error learning process, and leap directly to better-than-average skills by imitating successful or skilled cultural models. Further, because the world is a noisy, uncertain place, and it's often not entirely clear why a particular individual acquires great prestige or success, humans have evolved the propensity to copy a wide-range of cultural traits from prestigious individuals, only some of which may actually relate to the individuals' success (Henrich & Gil-White 2000).

A substantial amount of psychological, economic and ethnographic literature confirms many of these predictions (see Henrich & Gil-White 2000), but most importantly for our purposes, it confirms that people preferentially imitate prestigious individuals and over weight their opinions in making judgments. In a synthesis of the diffusion of innovations literature, Rogers (1995) has shown that the rate of spread of novel technologies and new economic practices into different cultural groups depends on how quickly prestigious, local "opinion leaders" adopt these innovations. In the laboratory, using a multi-round, market game with substantial incentives in which the results of each players' decisions were posted between sessions, experimental economists unexpectedly found that MBA students tended to mimic the decisions of successful players, even though rewards were distributed on a competitive basis. Allowing imitation also moved the group average substantially closer to the optimal decision predicted by Portfolio Theory (Kroll & Levy 1992). In a different experiment, Offerman & Sonnemans (1998) show that subjects making investment decisions tended to copy the beliefs of successful individuals (about the current environment), even when players clearly knew that these

---

clumps', social learning models predict (incorrectly) that populations should be a smear of ideas, beliefs, values and behaviors, and that group differences should only reflect local environmental differences.

individuals had the same information as they do about the current situation (also see Pingle 1995 and Pingle & Day 1996).

A self-similarity transmission bias, especially when combined with prestige-biased transmission, makes good adaptive sense. Individuals in social groups throughout human history needed to coordinate their beliefs, norms and expectations in order to make economic exchanges, marry and raise children. In theoretical work examining the interaction of genes and culture in solving these coordination problems, McElreath et. al. (2000) show that natural selection will favor the evolution of a bias to cue off salient symbolic markers when they covary with individuals' underlying norms of interaction—so by using symbolic markers individuals avoid acquiring norms that will produce costly un-coordinated interactions. Although no one has precisely examined imitation and ethnic markers, evidence for social learning biases towards self-similar comes more generally from the diffusion of innovations literature (Rogers 1995: 286), laboratory psychology (Rosekrans 1967; Stotland & Dunn 1962, 1963) and studies of child development (Harris 1998).

A third mechanism, *punishment* of non-conformists or norm violators, will act to homogenize social groups. A great deal of ethnographic and experimental research suggests that people in many societies are willing to inflict punishment on individuals who violate group norms of behavior (Sober and Wilson 1998, Roth et. al. 1991; Henrich et. al. 2000). If violators of group norms receive costly punishments, then either prestige-biased transmission, or simply trial & error learning, will reduce the variation between groups. Costly punishments will mean lower payoffs for norm violators, which, under prestige-biased transmission, means that the behaviors of norm-violators are less likely to spread. The problem with this norm-based punishment, as I discussed earlier, is the difficulty in explaining how it could evolve in a purely genetic system of inheritance. If punishing norm-violators is costly to the punisher, then punishing strategies are unlikely to evolve under natural selection. In the next section, I'll describe how stable punishing

33

behaviors can be maintained if humans have an arbitrarily small (non-zero) amount of conformist transmission bias in their social learning psychology.

A fourth mechanism, *normative conformity*, arises because individuals want their behavior to match the common behavior in their social group. Normative conformity differs from conformist transmission in that individuals aren't using the frequency of a behavior, belief or idea as an indirect indicator of its worth. Instead, in normative conformity, individuals alter their socially-displayed behavior (without necessarily changing their minds) because they want their behavior to match the majority, not because they 'believe' the majority is probably doing the smart thing. A vast amount of experimental work beginning with Asch's famous studies (e.g. Asch 1951), and including cross-cultural work (Furnham 1984), show that normative conformity is a robust part of human psychology (Neto 1995), at least in complex societies (nobody has done work among small-scale societies). Baron et. al.'s (1996) experimental work attempted to differentiate normative conformity from the effects of conformist transmission. Their results strongly suggest that people (i.e. university students) have both conformist transmission and normative conformity components to their psychology. These two psychological processes probably evolved for separate reasons. While conformist transmission probably evolved as a short-cut means of acquiring useful information, normative conformity, which may be a product of cultural evolution or culture-gene coevolution, may have evolved in response to the spread punishing strategies, and/or to provide a means to manipulatively cue one's 'similarity' to other group members (and obtain the advantages of in-group membership). Also, note that normative conformity *may* result in conformist transmission. Under some circumstances, imitators may mistakenly infer the underlying preferences or goals in observing the compliant behavior of models (by assuming a models 'likes' to perform certain behaviors), and thereby acquire both the outward behavior and the underlying supporting preferences (which were not possessed by the models).

Individually, and in combination, these mechanisms increase the importance of cultural group selection by creating a myriad of additional stable equilibria for all kinds of cultural traits, ideas, beliefs, values and practices, including those that govern social—and in particular cooperative—interactions. As I discussed earlier, the between-group component becomes important in circumstances with multiple stable equilibria because within-group selective processes, which act to push the system back to the locally stable equilibria, oppose the effects of migration between groups that reduce the variation between groups (Var (ϕ_i) in equation (6)). The nature of the mechanisms described above means that most of the time these stable equilibria will be monomorphic (everyone is doing the same thing), thus placing $Var(\phi_{ij})$ near zero. This means the cultural group selection component is likely to be quite important on longer time scales.

In thinking about how the difference between cultural and genetic evolutionary processes might affect the spread of n-person cooperation, it's also important to keep two other differences between cultural and genetic transmission in mind. First, cultural evolution is likely to proceed much more rapidly than genetic evolution because mechanisms like prestige- and conformist-biased cultural transmission, as well as other kinds of direct biases, favor rapid horizontal and oblique transmission, which can spread novel behaviors, ideas and practices among populations within a single generation (Boyd & Richerson 1985; Boyd & Richerson, manuscript)—it's worth noting that lots of evidence shows that most cultural transmission is not vertical (i.e. not parent to offspring; Harris 1998). Second, cultural transmission is likely to be more subject to drift and random variation than genetic transmission. The imitative skills of humans are good, and qualitatively better than any other animals, but certainly worse than genetic replication. The combination of these two differences means that human groups, under cultural evolution, will drift into the domains of attraction of alternative equilibria more often than in genetic evolution. This means that cultural group selection will "see" a greater variety of equilibria than genetic evolution over the same amount of time (Young 1998).

Conformist transmission can stabilize cooperation by stabilizing punishment

Before discussing cultural group selection, it's important to see how cultural transmission mechanisms can combine to create alternative stable prosocial equilibria. Henrich & Boyd (2000) have shown that if our social learning psychology contains both a transmission bias to copy successful individuals ('pay-off' or prestige-biased transmission) and a bias to copy high frequency behaviors (conformist transmission), and an arbitrary number of punishing 'levels', then highly cooperative equilibria can exist even if conformist transmission is only a weak component of human cultural transmission.

A tendency to copy high frequency behaviors can stabilize costly cooperative strategies without punishment, but only if this conformist transmission is quite strong compared to pay-off biased transmission. All other things being equal, pay-off biased transmission causes higher payoff variants to increase in frequency, and thus cooperation is not evolutionarily stable under plausible conditions—because not-cooperating leads to higher individual level payoffs than cooperating. Thus, on its own, pay-off biased transmission suffers the same problem as natural selection in genetic evolution. However, if our social learning psychology contains a combination of conformist and prestige-biased transmission (as I've argued above), then, if cooperation becomes common, conformist transmission will oppose payoff-biased transmission and favor cooperation. When cooperation is not too costly, conformist transmission will maintain cooperative strategies in the population at high frequency. However, because both theory and evidence (Henrich 2000) suggests that conformist transmission is relatively weak compared to payoff-biased transmission (and the costs of cooperation are probably substantial), it is seems unlikely that conformist transmission will be able to maintain cooperation.

A quite different logic applies to the maintenance of punishment. Suppose that culturally-transmitted punishing and cooperating strategies are both common, and that being punished is sufficiently costly that cooperators have higher payoffs than defectors. Rare invading 2nd order free riders who cooperate but do not punish will achieve higher payoffs than punishers because

they avoid the costs of punishing. However, because defection doesn't pay, the only defections will be due to rare mistakes, and thus the *difference* between the payoffs of punishers and 2nd order free riders will be relatively small. Hence, conformist transmission is more likely to stabilize the punishment of noncooperators than cooperation itself. As we ascend to higher order punishing, the difference between the payoffs to punishing vs. non-punishing decreases geometrically towards zero because the occasions that require the administration of punishment become increasingly rare. Second order punishing is required only if someone erroneously fails to cooperate, and then someone else erroneously fails to punish that mistake. For third order punishment to be necessary, yet another failure to punish must occur. As the number of punishing stages (i) increases, conformist transmission, no matter how weak, will at some stage overpower payoff-biased imitation and stabilize common i-th order punishment. Once punishment is stable at the i-th stage, payoffs will favor strategies that punish at the $i-1$ order, because common punishers at the i-th order will punish non-punishers at stage $i-1$. Stable punishment at stage $i-1$ order means payoffs at stage $i-2$ will favor punishing strategies, and so on down the cascade of punishment. Eventually, common 1st order punishers will stabilize cooperation at stage 0.

It is important to see that the stabilization of punishment is, from the gene's point of view, a maladaptive side-effect of conformist transmission. If there were genetic variability in the strength of conformist transmission (α) and cooperative dilemmas were the *only* problem humans faced, then conformist transmission might never have evolved. However, human social learning mechanisms were selected for their capability to efficiently acquire adaptive behaviors over a wide range of behavioral domains and environmental circumstances—from figuring out what foods to eat, to deciding what kind of person to marry—precisely because it is costly for individuals to determine the best behavior. Hence, we should expect conformist transmission to be important in cooperation as long as distinguishing cooperative dilemmas from other kinds of problems is difficult, costly or error prone. Looking across human societies we find that cooperative dilemmas come in an immense variety of forms, including harvest rituals among

agriculturalists, barbasco fishing among Amazonian peoples, warfare, irrigation projects, taxes, voting, meat sharing and anti-smoking pressure in public places. It's difficult to imagine a cognitive mechanism capable of distinguishing cooperative circumstances from the myriad of other problems and social interactions that people encounter.

As I've mentioned, natural selection favors the evolution of conformist transmission because it solves individual-level problems (Henrich & Boyd 1998), as well as many forms of social interaction (e.g., coordination problems). Consequently, in order for natural selection to favor an ability to "switch-off" the conformist effect upon encountering a cooperation problem, individuals must be able to accurately distinguish cooperative dilemmas from all other dilemmas. To accomplish this, individuals must be able to acquire sufficient information about the relative payoffs received by individuals behaving in alternative ways. But, the amount of payoff information necessary to distinguish a cooperative dilemma from other problems (e.g. coordination interactions) is also sufficient to determine the optimal solution to the problem without any social learning. If people could acquire and process sufficient information to 'know' when to switch-off conformist transmission, then they would be able to determine the optimal choice in the situation, and we wouldn't expect to observe any social learning in social interactions. Boyd and Richerson (1985) call this the "costly information hypothesis" for the origin of culture and of cultural evolution's interesting departures from strict genetic fitness optimizing expectations. Whenever experimentalists have provided opportunities for social learning in economic interactions, subjects always rely heavily on it, as if we have evolved to treat the individual acquisition of information as generally more costly than social learning or imitation (Kroll & Levy 1992, Ball et. al. 1999; Smith & Bell 1994; Pingle 1995; Pingle & Day 1996; Sonnemon & Offerman 1998; Wit 1999). Smith & Bell (1994) have tested this explicitly by comparing subjects' strategic use of social information with their imitative use of this information, and found that subjects imitate others in *both* public goods dilemmas and

coordination problems. No difference is observed in the relative strengths of people's imitative tendencies.[14]

Different Processes of Cultural Group Selection

Once stable culturally transmitted differences arise between groups, at least two different forms of cultural group selection may influence the evolution of practices, beliefs, ideas and values. I label these *inter-group competition* and *prestige-biased group selection.* One type of inter-group competition, which I term, *demographic swamping*, produces changes in the frequency of cultural traits because one cultural group, as a consequence of some set of ideas or practices that are relatively stable in that group, simply reproduces new individuals faster than other groups. Demographic swamping is consistent with the spread of early agriculturalists into regions once dominated entirely by hunter-gatherers. Agriculturalists gradually replace foragers, increasingly compressing them into tracks of inarable land (Cavalli-Sforza et. al. 1994; Young & Bettinger 1992; Diamond 1997). Actual cases of demographic swamping suggest that this is probably the slowest kind of cultural group selection, operating on time scales of millennia.

In intergroup competition, different cultural groups may also compete directly for access to resources through warfare and raiding. Cultural practices and beliefs that provide a competitive edge to groups in warfare will proliferate at the expense of traits that make groups less effective in competition (and more likely to be defeated and dispersed). Such cultural traits might relate to beliefs about patri-locality, heroism, patriotism, economic cooperation (leading to surplus production), the villainy of foreigners and the proper forms of social or political organization. In exploring cultural group selection resulting from intergroup competition, Soltis et. al. (1995) calculated evolutionary rates using a model based on group "extinctions" ("extinction" only implies that the group members must be disbanded and scattered, not necessarily killed). Using

---

[14] Further, the availability of the payoff information required to distinguish between cooperative and other kinds of social interactions is limited. If an individual enters the world and everyone is playing strategy "A", then he has no idea what will happen if he plays "B". And, if the underlying game is a PG game, and

data from New Guinean horticultural groups, Soltis et. al. estimated that a group beneficial trait could spread to fixation on time scales of 500 to 1000 years.

One of the best-documented cases of cultural group selection occurred during the 18th century among the anthropologically famous ethnic groups of the Nuer and the Dinka. Before 1820, the Nuer and Dinka (Kelly 1985) occupied adjacent regions in the southern Sudan. Despite inhabiting similar environments and possessing identical technology, the two groups differed in significant ways. Economically, both the Dinka and the Nuer raised cattle, but the Dinka maintained smaller herds of approximately nine cows per bull, while the Nuer maintained larger herds with two cows per bull. The Nuer ate mostly milk, corn and millet and rarely slaughtered cows. The Dinka, however, frequently ate beef. Politically, the Dinka lived in small groups, which corresponded to their group's wet season encampment. In contrast, the Nuer organized according to a patrilineal kin system that structured tribal membership across much larger geographic areas. Consequently, the size of a Dinka social group was limited by geography, whereas the Nuer system could organize much larger numbers of people over greater expanses of territory. Despite the similarity of their environments, these two groups showed substantial economic and political differences. Over about 100 years, starting in about 1820, the Nuer drastically expanded their territory at the expense of the Dinka, who were driven off, killed or captured and assimilated. As a result, Nuer beliefs and practices spread, fairly rapidly, across the landscape relative to Dinka beliefs and practices—despite the fact that the "Nuer" were soon living in the once "Dinka environment."

Another, quite subtler, form of cultural group selection is likely to operate on even shorter time scales than inter-group competition: prestige-biased group selection. Prestige-biased transmission means people will preferentially copy individuals who get higher payoffs. The higher an individual's payoff, the more likely that individual is to be imitated. If individuals

---

A is "cooperate", then everyone will get the same payoffs, and it will *look* like (be indistinguishable from) a coordination game.

occasionally have opportunities to copy people in neighboring groups, people from groups (or populations) at cooperative equilibria will be preferentially imitated by individuals in non-cooperative groups because the average payoff to individuals from cooperative groups is much higher than the average payoff of individuals in non-cooperative groups. Boyd & Richerson (manuscript) have shown that, under a wide range of conditions, this form of cultural group selection will deterministically spread group-beneficial behaviors from a single group (at a group-beneficial equilibrium) through a meta-population of other groups, which were previously stuck at a more individualistic equilibrium. This process can probably occur on time-scales of decades.

Culture-Gene Coevolution: a coevolutionary form of the Price Equation

Above, I have shown the conditions under which the between-group component of culture evolution is likely to be important relative to the within-group component, and I've explained why, given what we know about human psychology and social learning, we expect the between-group component to play a more important role in cultural evolution than it does in genetic evolution. In this section, we'll briefly explore the coevolution of prosocial traits.

Suppose z_{ij} is a phenotypic measure of some group beneficial, prosocial trait—like altruism—that is affected by both cultural and genetic transmission.[15] As before, i indexes the groups, while j indexes individuals within group i. Using linear regression, z_{ij} can be expressed as the weighted sum of all the alleles and cultural traits that influence an individual's prosocial phenotype.

$$z_{ij} = \underbrace{\sum_{All\,k} b_{ijk} X_{ijk}}_{the\,breeding\,value,\,x_{ij}} + \underbrace{\sum_{All\,v} d_{ijv} \Phi_{ijv}}_{the\,"cultural\,value",\,\phi_{ij}} + \varepsilon \qquad (7)$$

Here, X_{ijk} denotes the presence or absence of the k different alleles (at k loci) that effect phenotype z_{ij}; b_{ijk} specifies the relative contribution (as a partial regression coefficient) for each of these k

---

[15] This approach can be further generalized to include non-genetic, non-cultural, transmissible components of the environment, like road systems, monuments, and legal systems ("niche construction;" Laland et. al. 2000).

alleles. Similarly, Φ_{ijv} denotes the values of v cultural traits—either dichotomous or continuous—

that influence an individual's prosocial phenotype. Φ_{ijv} could represent such things as a belief in

an almighty god that punishes non-prosocial behavior, an unconscious habitualized practice, a

familiarity with historical myths about great patriots, a belief in an afterlife for prosocial people,

etc. In fact, cultural systems seem to have 'devised' all sorts of way to exploit our existing

psychological propensities for altruism towards kin and close friends in small groups to increase

our overall, phenotypic, 'prosocialness' (Richerson and Boyd 1998, 1999). Partial regression

coefficients, d_{ijv}, specify the relative contribution of these traits to the overall phenotype. The

uncorrelated error is ε. To keep things manageable and avoid clouding the issue, equation (7)

leaves out all kinds of interactive genetic and cultural terms (e.g., dominance, epistasis), and

gene-culture interactive terms (interaction between certain alleles and particular cultural traits,

$\sum TX\Phi$).

Before substituting into the Price Equation for the change in the mean value of phenotype

z, I'll simplify the notation by combining all the additive genetic effects into one variable, x_{ij}. And

similarly, ϕ_{ij} will capture the additive effects of culturally-learned traits on z_{ij}. These two give us,

$z_{ij} = x_{ij} + \phi_{ij}$. Putting this into the Price formulation we get,

$$\Delta \bar{z} = \overbrace{\underbrace{\beta_{w_i x_i} Var(x_i)}_{} + \underbrace{E(\beta_{w_{ij}x_{ij}} Var(x_{ij}))}_{GeneticComponents}}^{\text{Between-Group} \qquad \text{Within-Group}} + \overbrace{\underbrace{\beta_{\mu_i \phi_i} Var(\phi_i)}_{} + \underbrace{E(\beta_{\mu_j \phi_{ij}} Var(\phi_{ij}))}_{CulturalComponents}}^{\text{Between-Group} \qquad \text{Within-Group}} \qquad (8)$$

Equation (8) partitions the forces that contribute to phenotypic change in a cultural species into

the within-group and between-group components of both natural selection acting on genes, and

social learning processes acting on cultural variation. To simplify things, I have submerged the

mean genetic and cultural fitnesses, $\bar{w}$ and $\bar{\mu}$, into their respective β coefficients. And, in

deriving (8) I have also ignored phenotypic changes due to mutation (for both genetic and cultural

components) as well as individual learning—I have dealt with adding individual learning

elsewhere (Henrich 2000).[16]

Above, I discussed how, under certain circumstances, both the between-group genetic

component and the within-group cultural component are likely to be relatively small. If we

assume our groups are linked by a substantial degree of migration, this will drive the genetic

variation among groups, Var(x_i), towards zero—thereby substantially reducing the between-group

genetic component. Similarly, the forces of cultural transmission described earlier, as well as

normative conformity and punishment, will drive populations to monomorphic cultural equilibria.

When examined on the longer time scales appropriate for genetic evolution and between-group

cultural processes, the expected value of the variation within-groups, $E[Var(\phi_{ij})]$, will also

approach zero. With these assumptions, (9) reduces to:

$$\Delta \bar{z} = \underbrace{\overbrace{E(\beta_{w_{ij}x_{ij}}Var(x_{ij}))}^{Within-Group}}_{Genetic\,Components} + \underbrace{\overbrace{\left(\beta_{\mu_i,\phi_i}Var(\phi_i)\right)}^{Between-Group}}_{Cultural\,Component}. \tag{9}$$

Equation (9) shows that phenotype $\bar{z}$ will increase when the cultural group selection

component (which is positive by definition for a prosocial trait) is larger than the within-group

component of natural selection acting on genes (which is negative, by the same definition). Re-

arranging, we arrive at the conditions for the spread of prosocial behavior:

$$\beta_{\mu\phi} > \frac{C\bar{V}(x_{ij})}{Var(\phi_i)}. \tag{10}$$

As before, C gives the partial regression coefficient of x_{ij} on w_{ij} (Hamilton 1975) and

$\bar{V}(x_{ij})$ is the average amount of variation within groups. This tells us that "prosocialness", $\bar{z}$,

will increase as long as the benefits (in terms of the effect on the spread of the cultural traits)

---

[16] The idea of "cultural mutation" is not some weird concept arising from a faulty analogy with genetics. If social learned ideas, stories, beliefs and values change, in the absence of additional external input, while they sit in people's brains (e.g., some configurations might be easier to store), or if they change in the

exceeds the individual level fitness cost multiplied by the ratio of the average within-group

additive genetic variance to the cultural variation among groups. This ratio of variances is likely

to be quite small because the processes that create cultural variation between-groups (cultural

drift, transmission errors, group decisions, etc.) appear to operate at greater rates than the

mutational forces in genetic evolution.

In addition to knowing the conditions for the spread of a prosocial phenotype in a

genetically well-mixed population, we also want to know the conditions for the spread of a

prosocial gene in such a population. Using the Price formulation to calculate the conditions for

the spread of a prosocial gene in a structured population, with x_{ij} tracking the frequency of a

prosocial gene in individual j in group i, we start with equation (3):

$$\overline{w}\Delta\overline{x} = Cov(w_i, x_i) + E(Cov(w_{ij}\Delta x_{ij}) + E(w_{ij}\Delta x_{ij}))\,.$$

As before, using linear regression, we write down the fitness of a gene as the sum of a

constant term ($\square$) plus the effect of individual ij's phenotype on her fitness (holding the group's

mean phenotype constant) plus the effect of being in group i on ij's fitness (holding individual ij's

phenotype constant), plus an uncorrelated error term.

$$w_{ij} = \alpha + \beta_{w_{ij}z_{ij} \bullet z_i} z_i + \beta_{w_{ij}z_i \bullet z_{ij}} \overline{z} + \varepsilon \tag{10}$$

If we assume the population is genetically well-mixed (i.e., there is sufficient migration such that

Var(x_i) is quite small), and ignore the effects of mutation, substituting (10) into equation (3) gives

$$\overline{w}\Delta\overline{x} = E(\beta_{w_{ij}z_{ij}} Cov(z_{ij}, x_{ij}) + \beta_{w_{ij}z_i} Cov(z_i, x_{ij})) \tag{11}$$

Using the standard notation for the regression coefficients and solving for the conditions under

which an altruistic gene will spread, we arrive at

$$B(\partial z_i / \partial z_{ij}) > C\,. \tag{12}$$

---

processes of retrieval and output to phenotypic expression, then cultural mutation may be an important
component in some kinds of cultural evolution.

$\partial z_i / \partial z_{ij}$ represents $\square$ from equation (5), but is now described in terms of phenotypes rather than genotypes.

This tells us that, besides depending on the costs and benefits, the spread of a prosocial gene depends on the degree to which *having a prosocial phenotype predicts being in a prosocial group*. This is not a novel finding (Frank 1998). In the first part of the paper, I explained why generating and (more importantly) sustaining a high value of β is the essential element in all solutions to cooperation. In the second part of the paper, I've explained why biased cultural transmission (conformist transmission, self-similar biases and prestige-biases), conformism, and punishment, as well as the fast, error-prone dynamics of cultural transmission, will land social groups at a multiplicity of different quasi-stable *monomorphic* equilibria. This means that whatever an individual's behavior (his phenotype z_{ij}), it's likely to be an excellent predictor of the mean phenotype of his group (z_i). Our cultural capacities mean $\partial z_i / \partial z_{ij}$ (β) is likely to be large.

From the perspective on cooperation that I constructed earlier in the paper, the nature of our evolved social learning mechanisms produces (quite accidentally) informational regularities in phenotypic behavior that natural selection can exploit by spreading prosocial genes. Just as the family structure of some animals creates informational regularities to allow for the evolution of kin-based altruism, the nature of our social learning capacities create similar reliable statistical structures that form the foundation for the evolution of prosocial genes.

As I explained for kinship and reciprocal altruism, it's always possible to imagine mutants capable of invading and destroying the information regularity that sustains cooperation. In this case, if mutant social learners were capable of distinguishing cooperative dilemmas from the myriad of social interactions and individual problems people face (and thereby avoid acquiring common prosocial cultural traits), and were then also able to avoid detection and punishment with sufficient frequency, then such mutants could corrode this culturally-maintained $\square$. It remains quite unclear how the task of distinguishing cooperative dilemmas from all other

problems could be accomplished (see previous section). And, it may be quite possible to avoid detection and punishment in modern society, but it is more difficult to see how this could occur in small-scale societies—although some meager ability to distinguish highly profitable opportunities to defect should create cultural equilibria with lots of cooperation and some defection.

Having argued that cultural transmission creates the required informational regularities for the subsequent genetic evolution of prosocial genes, I must now briefly summarize some of the reasons why the high levels of cooperation we observe across human societies might be entirely a product of cultural group selection—independent of any sequence of genetic responses. First, the time since the evolution of the requisite cultural capacities in humans may be too short. If we mark the starting point for full-blown cultural capacities in the Upper Paleolithic (the first clear flowering of human culture occurs 50,000 years ago), then the span of time to the rise of high-levels of cooperation (i.e. cooperation involving large numbers of people emerges 5,000-10,000 years ago) seems rather short for genetic evolutionary processes.

Second, even if we post a start date for full-blown human cultural capacities at 75,000 years ago (giving a generous 20,000 or 30,000 years for cultural evolution to ramp-up before the big flowering) and assume strong selection, the massive environmental fluctuations that characterized the period from 75,000 to 12,000 (as compared to the last 12,000 years) may have periodically smashed β by repeatedly fragmenting and remixing cultural groups (Richerson *et. al.* 2000, but cf. Yengoyan 1968). Perhaps the last 10,000 years are the first period since the evolution of our cultural capacities that the environment has remained stable long enough for cultural group selection to have cobbled-up prosocial behavior and societal complexity. Of course, it's equally sensible to argue that cultural capacities arose much earlier (say 400,000 years ago at the end of the Acheulian tool industry), and the 'flowering' at 50,000 years ago resulted from some interaction between existing cognitive capacities and novel environmental conditions.

Third, cooperation may not be a dispositional trait of individuals, but rather a specific behavior or value tied only to certain cultural domains. Some cultural groups, for example, may

cooperate in fishing, and house-building, but not warfare. Other groups may cooperate in warfare, and fishing, but not house-building. Such culturally-transmitted traits would have the form 'cooperate in fishing,' 'cooperate in house-building,' and 'don't cooperate in warfare,' rather than the more dispositional approach of simply 'cooperate' vs. 'don't cooperate.' If this is the case, then the spread of prosocial genes among social groups becomes more difficult. As prosocial genes spread among groups with different stable cooperative domains, individuals with such genes would be more likely to mistakenly cooperate in noncooperative cultural domains. For example, in cultures where people cooperate in fishing, but not warfare, individuals with prosocial genes may be more likely to mistakenly cooperate in warfare (and pay the cost), as well as less likely to mistakenly defect in cooperative fishing.

Conclusion

This co-evolutionary approach solves two problems in understanding large-scale (n-person) cooperation in humans. First, if large-scale cooperation is a product of purely natural selection acting on genes to favor something like indirect reciprocity, why isn't large-scale cooperation more wide-spread in nature? If humans can establish massive amounts of cooperation among unrelated individuals based on kinship, reciprocity, indirect reciprocity, etc., why don't other social species like baboons, sea lions, chimpanzees and dolphins cooperate at human levels? Zoos and research centers create evolutionarily-novel, large social groups for many primate species, so why don't kin- and reciprocity-based altruism produce large-scale cooperation as in human societies? By rooting the coevolution of cooperation in the details of human social learning, my account address this challenge. Other animals don't cooperate to the degree humans do (except for eusocial insects), because they lack the social learning abilities that produce cultural evolution—which generates the informational regularities that favor prosociality. Of course, this only pushes the question back to, why don't more animals have human-style cultural capacities. In answering this question, Boyd and Richerson (1996) have shown that there is an

adaptive valley in the evolution of cultural abilities that inhibits the spread of these abilities when rare. Thus, such cultural abilities should be rare in nature, but once a species crosses the 'cultural threshold', whole new evolutionary vistas open up (also see Henrich & Gil-White 2000).

Second, unlike other animals, human cooperation varies in both scale and behavioral domains across social groups. As I described, there are many cases in which social groups inhabit the same physical environment and possess the same technology, but cooperate to differing degrees and in different domains (e.g. the Nuer and Dinka). Furthermore, recent cross-cultural experimental research reveals a substantial amount of between group variation in how people from small-scale societies behave in bargaining and public goods games (Henrich *et. al.* 2000). Standard evolutionary approaches struggle with these uniquely human observations—in other animals, cooperation and sociality does not vary much from group to group, especially the non-cultural components of the environments are identical. However, the co-evolutionary approach suggests that rapid within-group cultural forces will drive different social groups to one of a myriad of different quasi-stable equilibria, where they may remain for long periods until the between-group component of cultural evolution sorts among the various equilibria. This process has a time scale on the order of millennia, not even taking account of the possibility that some sets of equilibria may be adaptively equivalent. The rate of cultural evolution under group selection has about the right time scale to explain human history, whereas genetic explanations are too slow and rational choice ones too fast. Prosocial genes, because they co-evolved with culture, in a world with substantial culturally-produced phenotypic variation, will create psychologies designed to cue off cultural norms and will thereby not homogenize human sociality and cooperation across groups.

Reference List

Alexander, R. (1987). The Biology of Moral Systems. New York: Aldine De Gruyter.

Alexander, R. (1974). The Evolution of Social Behavior. Annual Review of Ecology and Systematics, 5, 325-383.

Aoki, K. (1982). A Condition for Group Selection to Prevail Over Counteracting Individual Selection. Evolution, 36(4), 832-842.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. H. Guetzkow (Ed), Groups, Leadership and Men (pp. 177-190). Pittsburgh: Carnegie.

Axelrod, R. (1984). The Evolution of Cooperation. U.S.: Basic Books.

Axelrod, R. (1986). An Evolutionary Approach to Norms. American Political Science Review, 80(4), 1095-1111.

Axelrod, R., & Dion, D. (1988). The Further Evolution of Cooperation. Science, 242, 1385-1390.

Axelrod, R., & Dion, D. (1988). The Further Evolution of Cooperation. Science, 242, 1385-1389.

Axelrod, R., & Hamilton, W. (1981). The Evolution of Cooperation. Science, 211, 1390-1396.

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

Batson, D. a. S. L. (1991). Evidence for Alturism: Toward a Pluralism of Prosocial Motives. Psycholigical Inquiry, 2(2), 107-122.

Bendor, J., & Mookherjee, D. (1987). The American Political Science Review, 81(1), 129-154.

Boorman, S., & Levitt, P. (1980). The Genetics of Altruism. New York: Academic Press.

Bowles, S. (1998). Cultural Group Selection and Human Social Structure: The effects of segmentation, egalitarianism, and conformism.

Bowles, S. (2000). Individual Interactions, Group Conflicts and the Evolution of Preferences. S. Durlauf, & P. Young (editors), Social Dynamics . Washinton D. C.: Brookings Institution.

Boyd, R. (1989). Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game. J. Theor. Biol., 136, 47-56.

Boyd, R., & Lorderbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. Nature, 32(6117), 58-59.

Boyd, R., & Richerson, P.Group Beneficial Norm Can Spread Rapidly in a Structured Population.

Boyd, R., & Richerson, P. (In press). Solving the Puzzle of Human Cooperation. S. Levinson (Editor). Cambridge MA: MIT Press.

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

Boyd, R., & Richerson, P. J. (1987). The Evolution of Ethnic Markers. Cultural Anthropology, 2(1), 27-38.

Boyd, R., & Richerson, P. J. (1989). The evolution of indirect reciprocity. Social Networks, 11(3), 213-236.

Boyd, R., & Richerson, P. J. (1988). The Evolution of Reciprocity in Sizable Groups. J. Theor. Biol., 132, 337-356.

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

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology & Sociobiology, 13(3), 171-195.

Boyd, R., & Silk, J. (1997). How Humans Evolved. New York: W. W. Norton & Company.

Boyd, R. a. R. P. (1996). Why Culture is Common, but Cultural Evolution is Rare. Proceedings of the British Academy, 88, 77-93.

Burks, S., Gary, C., Carpenter, J., & Verhoogen, E. (2000). Bargaining with Real Stakes and Real People.

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). The History and Geography of Human Genes. Princeton : Princeton University Press.

Crow, J. a. A. K. (1982). Group Selection for a Polygenic Behavioral Trait: A Differential Proliferation Model. Proceedings of the National Academy of Sciences of the United States of America, 79(8), 2628-2631.

Daly, M., & Wilson, M. (1988). Homicide. New York: Aldine de Gruyter.

Davis, D. D., & Holt, C. A. (1993). Experimental economics. Princeton, N.J: Princeton University Press.

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

Diamond, J. M. (1997). Guns, germs, and steel : the fates of human societies. (pp. 480 p., [32] p. of plates). New York: W.W. Norton & Co.

Ekman, P. (1992). Telling Lies. New York: Norton.

Ekman, P., & O'Sullivan, M. (1991). American Psychologist, 46, 913-920.

Fehr, E., & Gachter, S. (2000). Cooperation and Punishment in Public Goods Experiments. American Economic Review, 90(4), 980-995.

Fletcher, D. J. C., & Michener, C. D. (1987). Kin Recognition in Animals. New York: John

Wiley & Sons.

Forsythe, R., Horowitz, J., Savin, N., & Sefton, M. (1994). Fairness in Simple Bargaining Experiments. Games and Economic Behavior, 6, 347-69.

Frank, R. (1994). Microeconomics and Behavior. New York: McGraw Hill.

Frank, R. (1988). Passions within Reason: The strategic rle of the emotions. New York: W. W. Norton & Company.

Frank, R. H. (1994). Group selection and "genuine" altruism. Behavioral and Brain Sciences, 17(4), 620-621.

Frank, R. H., Gilovich, T., & Regan, D. T. (1993). The Evolution of One-Shot Cooperation: An Experiment. Ethology Aand Sociobiology, 14(247-256).

Frank, S. (1997). The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis. The Society for the Study of Evolution, 51(6), 1712-1729.

Frank, S. A. (1998). Foundations of Social Evolution. Princeton: Princeton University Press.

Frank, S. A. (1995). George Price's Contributions to Evolutionary Genetics. Journal of Theoretical Biology, 175(3), 373-388.

Fudenberg, D., & Maskin, E. (1990). Evolution and Cooperation in Noisy Repeated Games. New Developments in Economic Theory, 80 (2), 275-279.

Furnham, A. (1984). Studies of cross-cultural conformity: a brief and critical review. Psychologia: an International Journal of Psychology in the Orient, 27(1), 65-72.

Gintis, H. (forthcoming). Strong Reciprocity and Sociality. Journal of Theoretical Biology.

Hamilton, W. D. (1975). Innate Social Aptitudes of Man: an Approach from Evolutionary Genetics. R. Fox (editor), Biosocial Anthropology (pp. 133-156). London: Malaby Press.

Harbaugh, W. T., & Krause, K. (2000). Children's Altruism in Public Goods and Dictator Experiments. Economic Inquiry, 38(1), 95-109.

Harris, J. R. (1998). The Nuture Assumption: Why children turn out the way theydo. New York: Touchstone.

Harris, M. (1977). Cannibals and Kings: The origins of cultures. New York: Random House.

Hartl, D. L., & Clark, A. G. (1989). Principles of Population Genetics. Sunderland, Mass.: Sinauer Associates.

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

Henrich, J., & Smith, N. (2000). Comparative experimental evidence from Peru, Chile and the U.S. shows substantial variation among social groups. webuser.bus.umich.edu/henrich.

Henrich, J. (2000). 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. http://webuser.bus.umich.edu/henrich/.

Henrich, J., Boyd, R., Bowles, S., Gintis, H., & Fehr, E. (2000). Recriprocity, Cooperation and Punishment: Experiments in 15 small-scale societies.

Henrich, J., & Gil-White, F. (forthcoming). The Evolution of Prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior.

Hirshleifer, J. (1982). Evolutionary models in economics and law: cooperation vs. conflict strategies. Research in Law and Economics, 4, 1-60.

Hirshleifer, R., & Rasmusen, E. (1989). Cooperation in a Repeated Prisoner's Dilemma with Ostracism. Journal of Economic Behavior and Organization, 12, 87-106.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games. Game and Economic Behavior, 7, 346-380.

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

Johnson, A. (2000). The Matsigenka of the Peruvian Amazon: A Psychoecological Study. Stanford: Stanford University Press.

Johnson, A., & Earle, T. (1987). The Evolution of Human Societies: From Foraging Group to Agrarian State. Stanford: Stanford University Press.

Joshi, N. V. (1987). Evolution of cooperation by reciprocation within structured demes. Journal of Genetics, 66(1), 69-84.

Kagel, J. H., & Roth, A. E. (1995). The handbook of experimental economics. Princeton, N.J: Princeton University Press.

Kaplan, H., & Hill, K. (1985). Food Sharing Among Ache Foragers: tests of explanatory hypotheses. Current Anthropology, 26, 223-245.

Kelly, R. C. (1985). The Nuer Conquest. Ann Arbor: University of Michigan Press.

Kelly, R. L. (1995). The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways. Washington D. C.: Smithsonian Press.

Klein, R. G. (1989). The Human Career: Human Biological and Cultural Origins. Chicago: University of Chicago.

Kroll, Y., & Levy, H. (1992). Further Tests of the Separation Theorem and the Capital Asset Pricing Model. American Economic Review, 82(3), 664-670.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. The Handbook of Experimental Economics (pp. 111-194). Princeton, New Jersey: Princeton University

Press.

Lee, R. B. (1979). The !Kung San: Men, Women, and Work in a Foraging Society. Cambridge: Cambridge University Press.

Low, B. S. (2000). Why Sex Matters: a Darwinian look at Human Behavior. Princeton: Princeton University Press.

Maynard Smith, J. (1998). The origin of altruism. Nature, 393(18), 639-640.

McAdams, R. H. (1998). The origion, development and regulation of norms. Michigan Law Review, 96, 338.

McElreath, R., Boyd, R., & Richerson, P. (2000). Shared norms can lead to the evolution of ethic markers. http://www.sscnet.ucla.edu/anthro/faculty/boyd/EthnicMarkers1.8.pdf.

Mueller, D. (1989). Public Choice II. Cambridge: Cambridge University Press.

Nelson, R. R., & Winter, S. G. (1985). Evolutionary Theory of Economic Change. Cambridge, MA: Harvard University Press.

Neto, F. (1995). Conformity and independence revisited. Social Behavior & Personality, 23(3), 217-222.

Nowak, M. a. S. K. (1994). The Alternating Prisoner's Dilemma. J. Theor. Biol., 168, 219-226.

Nowak, M. a. S. K. (1998). Evolution of indirect reciprocity by image scoring. Nature, 393, 573-577.

Ockenfels, A., & Selten, R. (1998). An Experiment on the Hypothesis of Involuntary Truth-Signaling in Bargaining. Working Paper.

Offerman, T., & Sonnemans, J. (1998). Learning by experience and learning by imitating others. Journal of Economic Behavior and Organization, 34(4), 559-575.

Patton, J. Q. (2000). Reciprocal Altruism and Warfare: A case from the Ecuadorian Amazon. L. Cronk, N. Chagnon, & W. Irons (Editors), Adaption and Human Behavior: An anthropological perspective (pp. 417-436). New York: Aldine de Gruyter.

Pingle, M. (1995). Imitation vs. rationality: An experimental perspective on decision-making. Journal of Socio-Economics, 24, 281-315.

Pingle, M., & Day, R. H. (1996). Modes of economizing behavior: Experimental evidence. Journal of Economic Behavior & Organization, 29, 191-209.

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

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

Queller, D. (1992). Quantitative Genetics, Inclusive Fitness, and Group Selection. The American Naturalist, 139(3), 541-558.

Queller, D. C. (1992). A General Model For Kin Selection. Evolution, 46(2), 376-380.

Rappaport, R. A. (1984). Pigs for the ancestors. New Haven: Yale University Press.

Richerson, P., & Boyd, R. (2000). Complex Societies: the evolutionary dynamics of a crude
    superorganism. Human Nature, 10, 253-289.

Richerson, P., & Boyd, R. (1998). The Evolution of Ultrasociality. I. Eibl-Eibesfeldt, & F. K.
    Salter (editors), Indoctrinability, Ideology and Warfare (pp. 71-96). New York: Berghahn
    Books.

Ridley, M. (1993). The Red Queen: Sex and the Evolution of Human Nature. New York: Penguin
    Books.

Rogers, A. (1990). Group Selection by Selective Emigration: The Effects of Migration and Kin
    Structure. The American Naturalist, 135(3), 339-413.

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

Rosekrans, M. A. (1967). Imitation in Children as a Funtion of perceived similarity to a social
    model and vicarious reinforcement. Journal of Personality and Social Psychology, 7(3),
    307-315.

Roth, A. E. (1995). Bargaining Experiments. J. H. Kagel, & A. E. Roth (editors), The handbook
    of experimental economics (pp. 253-248). Princeton: Princeton University Press.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and Market
    Behavior in Jerusalem, Ljubljana, Pittburgh and Tokyo: An Experimental Study.
    American Economic Review, 81(5), 1068-95.

Sepher, J. (1983). Incest, the biosocial view. New York: Academic Press.

Skinner, J., & Slemrod, J. (1985). An Economic Perspective on Tax Evasion. National Tax
    Journal, 38, 345-53.

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

Smith, M. J. (1979). Game Theory and the Evolution of Behaviour. Proceedings of the Royal
    Society of London. Series B, Biological Sciences, 205(1161), 475-488.

Smith, N. (2000). "Ultimatum and Dictator Games among the Chaldeans of Metro Detroit." Paper
    given at the Ethno-Experiment Research Group meetings, Los Angeles.

Sober, E., & Wilson, D. S. (1998). Unto Others: The Evolution and Psychology of Unselfish
    Behavior. 1998: Harvard University Press.

Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural
    group selection? An empirical test. Current Anthropology, 36(3), 473-494 .

Spencer, H. (1891). Essays: Scientific, political, and speculative. London: Williams and Norgate.

Stotland, E. a. D. R. (1963). Empathy, Self-Esteem and Birth Order. Journal of Abnormal and Social Psychology, 66(6), 532-540.

Stotland, E. a. D. R. (1962). Identification, "Oppositeness," Authoritarianism, Self-Esteem, and Birth Order. Psycholigical Monigraphs: General and Applied, 76(9), 1-30 & 1-19.

Tomasello, M. (2000). The Origins of Human Cognition. Cambridge: Harvard University Press.

Tomasello, M. (1994). The question of chimpanzee culture. Chimpanzee Cultures (pp. 301-317-xxiii, 424).

Tooby, J., & Cosmides, L. (1989). Evolutionary psychology and the generation of culture: i. Theoretical considerations. Ethology & Sociobiology, 10(1-3), 29-49.

Trivers, R. (1971). The Evolution of Reciprocal Alturism. The Quarterly Review of Biology, 46, 34-57.

Turner, J., & Maryanski, A. (1979). Functionalism. Menlo Park: Benjamin/Cummings.

Vayda, A. P. (1971). Phases of the process of war and peace among the Marings of new Guinea. Oceania, 42, 1-24.

Wade, M. (1985). Soft Selection, Hard Selection, Kin Selection, and Group Selection. The American Naturalist, 125(1), 61-73.

Westermarck, E. (1894). The history of human marriage. London: Macmillan.

Wiessner, P. (1983). Style and social information in Kalahari San Projectile Points. American Antiquity, 48(2), 253-275.

Williams, G. (1966). Adaption and Natural Selection: A Critique of Some Current Evolutionary Theory. Princeton: Princeton University Press.

Wilson, D. S. a. D. L. (1997). Group Selection and Assortative Interactions. The American Naturalist, 149(2), 336-351.

Wilson, D. S. a. S. E. (1998). Reintroducing group selection to the human behavioral sciences. Behavioral and Brain Sciences, 21(2), 304-306.

Wit, J. (1999). Social Learning in a Common Interest Voting Game. Games and Economic Behavior, 26, 131-156.

Wolf, A. P. (1970). Childhood association and sexual attraction: A further test of the Westermarck hypothesis. American Anthropologists, 72, 503-515.

Young, D., & Bettinger, R. L. (1992). The Numic Spread: A computer simulation. American Antiquity, 57(1), 85-99.

Young, H. P. (1998). Individual strategy and social structure : an evolutionary theory of institutions. Princeton, N.J: Princeton University Press.

Table 1. Two-person Prisoner's Dilemma Payoff Matrix. Payoffs are for the row player (Player 1). Players 1 and 2 are interchangeable. 'C' indicates cooperation and 'D' indicates defection.

Player 1 plays	Player 2 plays	
	C	D
C	2	-1
D	3	0