Strong Reciprocity and the Roots of Human Morality

**Herbert Gintis · Joseph Henrich · Samuel Bowles ·
Robert Boyd · Ernst Fehr**

Abstract   Human morality is a key evolutionary adaptation on which human social behavior has been based since the Pleistocene era. Ethical behavior is constitutive of human nature, we argue, and human morality is as important an adaptation as human cognition and speech. Ethical behavior, we assert, need not be a means toward personal gain. Because of our nature as moral beings, humans take pleasure in acting ethically and are pained when acting unethically. From an evolutionary viewpoint, we argue that ethical behavior was fitness-enhancing in the years marking the emergence of Homo sapiens because human groups with many altruists fared better than groups of selfish individuals, and the fitness losses sustained by altruists were more than compensated by the superior performance of the groups in which they congregated.

Keywords   Morality · Human nature · Evolution · Reciprocity · Sociobiology · Altruism

H. Gintis (✉)
Santa Fe Institute, Santa Fe, New Mexico and, The Central European University, Budapest, Hungary
e-mail: hgintis@comcast.net

J. Henrich
University of British Columbia, Vancouver, Canada

S. Bowles
Santa Fe Institute, Santa Fe, NM, USA

R. Boyd
University of California, Los Angeles, CA, USA

E. Fehr
University of Zurich, Zurich, Switzerland

The Nature of Human Morality

The two volumes under review summarize a body of research suggesting that human morality is a key evolutionary adaptation on which human social behavior has been based since the Pleistocene era. Ethical behavior is constitutive of human nature, we argue, and human morality is as important an adaptation as human cognition and speech. Ethical behavior, we assert, need not be a means toward personal gain. Because of our nature as moral beings, humans take pleasure in acting ethically and are pained when acting unethically. From an evolutionary viewpoint, we argue that ethical behavior was fitness-enhancing in the years marking the emergence of *Homo sapiens* because human groups with many altruists fared better than groups of selfish individuals, and the fitness losses sustained by altruists were more than compensated by the superior performance of the groups in which they congregated.

Price (2008) acknowledges the quality of the research on which these books are based, he affirms the importance of strong reciprocity as a human behavioral pattern documented in the laboratory, and he does not deny its importance in everyday human affairs.[1] However, he suggests that moral behavior in humans is not an adaptation, but rather a maladaptive response to modern social conditions: we are charitable to strangers because individuals encountered few "strangers" in our hunter-gatherer past, and hence our Pleistocene brains confuse strangers with close acquaintances or kin. Moreover, he argues, people behave ethically under conditions of anonymity because anonymity was very rare in our hunter-gatherer past, so our stone-age brains never really believe we are not being closely watched and socially evaluated.

Price's (2008) interpretation of morality as enlightened self-interest has a venerable history in evolutionary psychology. Hamilton (1964), Williams (1966), Trivers (1971), and others convinced a generation of researchers that personal sacrifice on behalf of others is really just long-run self-interest. Richard Dawkins, for instance, struck a responsive chord when, in *The Selfish Gene* (1989[1976]), he asserted "We are survival machines—robot vehicles blindly programmed to preserve the selfish molecules known as genes. ...Let us try to *teach* generosity and altruism," he advises, "because we are born selfish." Yet, even social morality, according to R. D. Alexander, the most influential ethicist working in the Williams–Hamilton tradition, can only superficially transcend selfishness. In *The Biology of Moral Systems* (1987), Alexander asserts, "ethics, morality, human conduct, and the human psyche are to be understood only if societies are seen as collections of individuals seeking their own self-interest" (p. 3). In a similar vein, Ghiselin (1974) wrote "No hint of genuine charity ameliorates our vision of society, once sentimentalism has been laid aside. What passes for cooperation turns out to be a mixture of opportunism and exploitation...Scratch an altruist, and watch a hypocrite bleed" (p. 247). Our research brings seriously into question this view of human morality.

---

[1] Readers not acquainted with the range of evidence can consult Henrich et al. (2006), Gintis, Bowles, Boyd, and Fehr (2003), and Fehr, Fischbacher, and Gächter (2002).

We will outline the implications of our findings for modeling human moral behavior, and then comment on Price's (2008) alternative explanation of strong reciprocity as a maladaptation.

Strong Reciprocity

By *strong reciprocity* we mean a propensity, in the context of a shared social task, to cooperate with others similarly disposed, even at personal cost, and a willingness to punish those who violate cooperative norms, even when punishing is personally costly. We deem this "reciprocity" because it embraces an ethic of treating others as they treat us, bestowing favors on those who cooperate with us, and punishing those who take advantage of our largesse. We call this reciprocity "strong" to distinguish it from forms of reciprocity, such as tit-for-tat (Axelrod & Hamilton, 1981) and reciprocal altruism (Trivers, 1971), that are the forms of long-run enlightened self-interest.

Strong reciprocity is a universal structure of human morality, but it acquires concrete content only in the context of specific cultural values concerning the legitimate rights and obligations of individuals. This is why, for instance, our study of fairness in *Foundations of Human Sociality* produced considerably more varied behavior than previously found in the studies of advanced market economies. However, there is considerable agreement among individuals in advanced industrial societies concerning the content of moral behavior (Fong, Bowles, & Gintis, 2005). Finally, strong reciprocity is but one of a number of human moral constructs that have been studied in the experimental laboratory, others including character virtues such as honesty and trustworthiness, and other-regarding emotions such as shame, envy, empathy, and the taste for retribution.

Strong Reciprocity and the Social Welfare System

Our most important finding, evident from the diversity of behaviors in *Foundations of Human Sociality*, is that most individuals treat moral values as ends in themselves, not merely means toward maintaining a valuable social reputation or otherwise advancing their self-interested goals. This conclusion follows from observing that even in one-shot, anonymous interactions of the sort studied in our experimental work, individuals behave in ways that reflect the moral standards of their particular social group. This insight has helped us understand the social welfare systems of the advanced welfare states.

We argue that altruistic punishment is critically important to both the health of egalitarian systems, as well as to their demise. On the one hand, a small fraction of altruistic punishers can induce self-interested individuals to cooperate, on threat of being punished for defecting. On the other hand, when the frequency of free riding is too high, altruistic punishers will withdraw their participation, thereby exacerbating the problem of low participation rates, leading to the complete unraveling of social cooperation.

A case in point is the increased opposition to income redistribution witnessed in the United States and some European countries in recent years. According to the norms associated with strong reciprocity, we would expect egalitarian policies that reward people independent of whether and how much they contribute to society to be considered unfair, even if the intended recipients are otherwise considered worthy of support, and even if the incidence of non-contribution in the target population is rather low. This would explain the shift from support for to opposition to welfare measures for the poor, since such measures are thought to have promoted various social pathologies that have become serious in recent decades. At the same time it explains the continuing support for social security and medicare in the United States, since the public perception is that the recipients are 'deserving' and the policies do not support what are considered anti-social behaviors.

A striking fact about the decline in the support for the former Aid to Families with Dependent Children, Food Stamps, and other means-tested social support programs in the United States, however, is that overwhelming majorities came to oppose these programs, whatever their income, race, or personal history with such programs. This pattern of public sentiment, we think, can be accounted for in terms of strong reciprocity.

We rely mainly on two studies. The first (Farkas & Robinson, 1996) analyzes data collected in late 1995 by Public Agenda, a non-profit, non-partisan research organization. The authors conducted eight focus groups around the country, then did a national survey, involving half-hour interviews, of 1,000 randomly selected Americans, plus a national oversample of 200 African-Americans. The second, political scientist Martin Gilens' *Why Americans Hate Welfare*, is an analysis and review of several polls executed during the 1990s and earlier by various news organizations. A third study by Weaver, Shapiro, and Jacobs (1995), drawing in addition on NORC and General Social Survey data, comes to broadly similar conclusions.

In the Public Agenda survey 63% of respondents thought the welfare system should be eliminated or "fundamentally overhauled" while another 34% thought it should be "adjusted somewhat." Only 3% approved of the system as is. Even among respondents from households receiving welfare only 9% expressed basic approval of the system, while 42% wanted a fundamental overhaul and an additional 46% wanted some adjustments.

The cost of welfare programs cannot explain this opposition. While people generally overstate the share of the Federal budget devoted to welfare, this cannot account for the observed opposition. Farkas and Robinson note that

> By more than four to one (65% to 14%), Americans say the most upsetting thing about welfare is that "it encourages people to adopt the wrong lifestyle and values," not that "it costs too much tax money." ...Of nine possible reforms presented to respondents—ranging from requiring job training to paying surprise visits to make sure recipients deserve benefits—reducing benefits ranked last in popularity (Table 4).

The cost, apparently, is not the problem. In focus groups:

Participants invariably dismissed arguments about the limited financial costs of welfare in almost derisive terms as irrelevant and beside the point.

The perception of fraud cannot account for this opposition. While 64% of respondents, and 66% of respondents on welfare, believe welfare fraud is a serious problem, most do not consider it more serious than in other government programs, and only 35% of survey respondents would be more "comfortable with welfare" if fraud were eliminated.

In the Public Agenda study, respondents overwhelmingly consider welfare to be unfair to working people and addictive to recipients. Indeed, 70% (71% of welfare recipients) say welfare makes it "financially better for people to stay on welfare than to get a job," while 57% (62% of welfare recipients) think welfare encourages "people to be lazy" and 60% (64% of welfare recipients) say the welfare system "encourages people to have kids out of wedlock." The truth of such assertions is beside the point. Whether or not, for example, welfare *causes* out of wedlock births, for example, or fosters an unwillingness to work, citizens object that the system provides financial support for those who undertake these socially disapproved behaviors. Their desire is to bear witness against the behavior and to disassociate themselves from it, whether or not their actions can change it.

This interpretation is supported by a careful study by Luttmer (2001), who matched U. S. General Social Survey (GSS) data with census tract information on the number and characteristics of the GSS subjects' neighbors who were on welfare. He found that the number of people in the surrounding area who were receiving public assistance predicted opposition to welfare spending if those on welfare were predominantly not working (or working very little) and if many of those on welfare were unmarried mothers *and* the subject voiced disapproval of premarital sexual relations. The relevant fact for our interpretation is that opposition was conditioned on the non-working and unmarried mother status of the recipients and the moral beliefs of the subjects.

On the other hand, surveys support 'basic needs generosity,' a virtually unconditional willingness to share with others to assure them of some minimal standard, especially, as the survey data show, when this is implemented through provision of food, basic medical care, housing, and other essential goods. The interplay of basic needs generosity and strong reciprocity, we think, accounts for the salient facts about public opinion concerning the welfare state.

Our analysis supports the notion that declining voter support for the welfare state, where it has occurred, is due not to the selfishness of the electorate, but rather the failure of social welfare programs to tap powerful commitments to fairness and reciprocity. There is substantial support for generosity toward the less well off as long as they are "deserving" poor, who have provided or tried to provide a *quid pro quo* and are in good standing. Poverty is often the result of low returns to such socially admired behaviors as hard unskilled work, independent small-scale entrepreneurship, and studious behavior in poor educational environments. Policies designed to raise the returns to these activities might garner widespread support. Strong reciprocity sentiments might also support policies that insure individuals against the vagaries of the weather or the market without compensating them for losses to laziness or poor judgment.

We urge social justice researchers to review the behavioral game theoretical evidence concerning fairness and reciprocity, including not only our own work but the valuable contributions of Smith (1982), Yamagishi (1986), and Frohlich and Oppenheimer (1992), and their associates, as well as the many others who have contributed to our understanding of normative behavior through laboratory experiments over the past two decades.

Michael Price's Critique: Stone-age Minds in Modern Brains

Our interpretation of strong reciprocity as an adaptation is reported in both books under review in this issue of *Social Justice Research*, in Richerson and Boyd (2004), and in several journal articles, including Gintis (2000), Henrich and Boyd (2001), Bowles, Choi, and Hopfensitz (2003), Gintis (2003), and Boyd, Gintis, Bowles, and Richerson (2003). The alternative view expressed by Price (2008) extends an interpretation proposed by Tooby and Leda (1997), who see "inappropriate" contemporary human behavior (e.g., obesity, substance abuse, weakness of will, unsafe sex) as stemming from the fact that

> our modern skulls house a stone age mind. The key to understanding how the modern mind works is to realize that its circuits were not designed to solve the day-to-day problems of a modern American—they were designed to solve the day-to-day problems of our hunter-gatherer ancestors. ...[O]ur minds are ...very sophisticated computers, whose circuits are elegantly designed to solve the kinds of problems our ancestors routinely faced.

In line with this general approach, Price asserts that contemporary altruistic behavior of the sort that we have identified evolved under conditions in which it advanced the self-interest of the actor, but now is a fitness-reducing maladaptation to novel environments for which the human brain has not evolved the appropriate fitness-enhancing modules.

In support of his position, Price argues that our models of morality as an adaptation are based on "group selection," which he contends was long ago shown to be implausible by eminent biologists, while evolutionary psychologists such as himself support the more traditional "individual selection" view, according to which true altruism can never extend beyond sacrificing for close kin. In fact, group selection in the case of humans is quite plausible in the context of what we term "gene-culture coevolution," an evolutionary dynamic not considered in the classical objections to group selection. More broadly, far from being rejected by biologists, recent years have seen a resurgence of support for group selection models in non-human social species (Odling-Smee, Laland, & Feldman, 2003). Indeed, the argument that multicellularity itself is a product of group selection is now widely accepted (Maynard Smith & Szathmary, 1997). Also, it is estimated that eusociality evolved independently about 15 times in insects, and eusocial insects make up about half the biomass of all insects. Moreover, sociality in insects cannot be accounted for in terms of kinship, so group selection is the most plausible mechanism accounting for the preeminence of social insects (Gadagkar, 1991; Frank, 1995,

1998; Wilson & Holldobler, 2005). Group selection is probably not common in the world of biota, but it does occur, and when it does, it can produce highly fit organisms. Our species is but one example of the power of group selection.

Our ability to infer moral values from observed behavior is based on our use of what we term the *beliefs, preferences, and constraints* (BPC) model of human choice, an extension of the economist's rational actor model. Price argues that our use of the BPC model is illegitimate because, "minds are executors of adaptations, rather than maximizers." However, the BPC model does not presume or imply that individuals consciously "maximize." Rather, the BPC model presumes only that individuals have consistent preferences. As is well known from decision theory, this assumption allows us to represent the individual's choice behavior as though it were the product of conscious maximization, whatever the underlying decision mechanism (Kreps, 1988). This usage of a maximizing model is no different from being able to predict the behavior of an animal based on an optimal foraging model, whatever the mechanism whereby the animal turns sensory information into physical movements. In short, the BPC model is perfectly compatible with the notion of the mind as an "executor of adaptations" (Gintis, 2007).

Gene-Culture Coevolution

In species that produce complex environments (e.g., beaver dams, bee hives), these environments themselves alter the fitness of individual genes and gene complexes, so are best analyzed at the level of the social group, as suggested in niche construction theory (Odling-Smee et al., 2003). Niche construction models use *multi-level selection*, in which the structure at some higher level (in this case, that of the niche-constructing unit) affects evolution at a lower level (in this case, the individual). Gene-culture coevolutionary theory, which applies almost exclusively to our species, is a form of niche construction theory in which cultural rules create a group structure that affects individual fitness. While gene-culture coevolution facilitates group selection, it is not itself a type of group selection, because it is not based on competition among groups. A dynamic pattern of genetic adaptation, followed and preceded by cultural evolution within a single group, is sufficient for gene-culture coevolution to occur.

Price contends that human morality is the effect of genetic fitness maximization at the individual level, culture playing an intermediary role that can be "factored out" in the long run. This view is incompatible with gene-culture coevolution. To see this, suppose we have a vector g of genetic variables, a vector c of cultural variables, and a vector e of environmental variables, including the prevalence of predators and prey, weather, and the like. In an evolutionary model, the rate of change of variables is a function of the variables, so we have

$$\dot{\mathbf{g}} = F(\mathbf{g}, \mathbf{c}, \mathbf{e}) \tag{1}$$

$$\dot{\mathbf{c}} = G(\mathbf{g}, \mathbf{c}, \mathbf{e}) \tag{2}$$

$$\dot{\mathbf{e}} = H(\mathbf{e}). \tag{3}$$

The contention that culture is an effect of genetic fitness maximization in this framework is the assertion that c can be eliminated from these equations. Under what conditions can this occur? Taking the derivative of the first equation, and substituting the second and third into the first, we get

$$\ddot{\mathbf{g}} = F_{\mathbf{g}}(\mathbf{g}, \mathbf{c}, \mathbf{e})F(\mathbf{g}, \mathbf{c}, \mathbf{e}) + F_{\mathbf{c}}(\mathbf{g}, \mathbf{c}, \mathbf{e})G(\mathbf{g}, \mathbf{c}, \mathbf{e}) + F_{\mathbf{e}}(\mathbf{g}, \mathbf{c}, \mathbf{e})H(\mathbf{e}). \tag{4}$$

If c is to be absent from this second-order differential equation, the derivative of the right-hand side of Eq. 4 with respect to c must be identically zero. Thus, we have

$$0 \equiv F_{\mathbf{gc}}F + F_{\mathbf{g}}F_{\mathbf{c}} + F_{\mathbf{cc}}G + F_{\mathbf{c}}G_{\mathbf{c}} + F_{\mathbf{ec}}H. \tag{5}$$

All five of the above terms must then be identically zero, so $F_{bf\,c} \equiv 0$, implying that c does not enter on the right-hand side of the defining equations (1); i.e., genes are not a function of culture. This is obviously not appropriate for humans, since both genes and culture are functions of culture. For instance, the repositioning of the larynx to permit linguistic communication involves an interaction of genes and culture (Deacon, 1998).

Altruism and Ancestral Human Environments

Price argues that individual selection is incompatible with altruistic behavior because sacrificing on behalf of others who do not share one's genes must lower one's fitness as compared with non-altruist competitors. However, there is no conflict between individual fitness calculation and group or population-level calculations, since however the accounting is done, we must take into account the fact that fitness losses that accrue to individual altruistic acts may be compensated at a higher social level, so that the fitness of the altruism-generating genes are not compromised at the level of the population (Kerr % Godfrey-Smith, 2002). Evolutionary psychologists in the Cosmides–Tooby tradition appear to reject this analysis, although without warrant or explanation.

In its place, evolutionary psychologists suggest that the altruism that people exhibit in the laboratory, and presumably in everyday life under conditions of anonymity and non-repetition, are evolutionary mistakes. In Price's (2008) words, "[strong reciprocity] does not necessarily falsify the theory that these adaptations evolved because they promoted individual fitness in ancestral environments; instead, it may merely suggest that these adaptations will continue to execute their protocols, even in environments where experimenters have introduced novel conditions that make it impossible for these protocols to lead to adaptive outcomes." We agree with Price that strong reciprocity must have promoted individual fitness, or it could not have evolved. Our contention is that strong reciprocity enhanced relative fitness because groups with a high frequency of altruism survived and prospered at a higher rate than groups with a low frequency of altruism.

Is Human Morality an Adaptation?

We have argued that human morality is the product of gene-culture coevolution and hence has a long evolutionary history. Price (2008) has advanced the view that moral behavior is the legacy of an evolutionary past in which individuals behaving prosocially simply had higher fitness than other group members, and hence their prosocial behavior is selfish, not altruistic. Thus, Trivers (2007) reasons that "unfair arrangements...may exact a very strong cost in inclusive fitness. In that sense, an attachment to fairness or justice is self-interested" (p. 77). If Price and Trivers are correct, fair-mindedness could have become common among humans by inclusive fitness maximization in repeated interactions, allowing fair-minded individuals to gain reputations that advanced their genetic interests. The same reasoning may be applied to generosity, bravery on behalf of ones' associates, and punishing those who transgress social norms.

Thus, it could be that humans became cooperative because in our ancestral environments we rarely engaged non-kin in short-term interactions and instead interacted frequently with the same group of individuals, among whom tit-for-tat and other self-regarding strategies consistent with Trivers' reciprocal altruism were sufficient to support cooperative outcomes. Cooperation thus emerged as a form of mutualism. Others, including Price (2008) and Tooby and Cosmides (1992) share this view of the evolutionary origins of social preferences, but in contrast to Trivers hold them to be fitness-reducing (i.e., maladaptive) in modern settings. Dawkins (2007, p. 222) explains it this way: "the lust to be generous and compassionate...is the misfired consequence of ancestral village life."

We have no disagreement concerning the importance of reputation-building in moral behavior (Gintis, Smith, & Bowles, 2001). But, we think it unlikely that either alone or together with inclusive fitness maximization is sufficient to explain the evolutionary origins of human cooperation. The problem with the maladaptation view, in brief, is that it fails to explain many common forms of human cooperation for which there exist more plausible models based on gene-culture coevolution and multilevel selection.

First, explanations based on reputation-building are contradicted by the evidence from controlled laboratory settings and in real life that people routinely engage in acts of generosity toward unknown and unrelated others, and punish those who transgress social norms, even when no material benefit can be gained thereby. Second, the maladaptation view suggests that our evolutionary history has not left us mentally equipped to distinguish clearly between one-shot and repeated interactions, or between situations of anonymity versus non-anonymity. This "maladaptation" would, of course, explain the experimental and other observed instances of human altruism. But, in fact, in the modern world people distinguish rather acutely between long-term and one-shot interactions, and they behave quite differently when they think they are being observed as compared to being shrouded in anonymity. In experiments, behavior responds to whether the interaction is with an on-going partner or one with whom future interactions have been precluded by experimental design. Moreover, experimental subjects not only punish those who have treated them ungenerously in a one-shot interaction, but also those who have

treated *others* ungenerously. The first might be explained, by a stretch of imagination, by Price's notion that it is, on average, fitness enhancing to punish those who have treated one poorly, and subjects may believe that even in an anonymous experiment that this rule of thumb should be followed. But, anonymously punishing those who have treated *others* poorly is an order of magnitude harder to reconcile with the self-interested cooperation paradigm.

The third problem is that there is good reason to believe that early humans did indeed engage in fitness-relevant interactions with non-kin of sufficiently short duration that only extraordinarily beneficial cooperation could have been supported by tit-for-tat and related strategies in these interactions.

Neither the likely size of groups, nor the degree of genetic relatedness within groups, nor the typical demography of foraging bands is favorable to the view that kin altruism and mutualism provide an adequate account of late Pleistocene human cooperation. The median of the 235 hunter-gather groups recorded in Binford (2001) is 19, and even if we count this as, say, seven adult decision makers, reciprocal altruism will evolve for groups this large only for implausibly low rates of error and extraordinarily high benefit cost ratios of the altruistic behavior. Moreover, though parent–child and sibling altruism will be supported for plausible benefit cost ratios, relatedness within typical foraging bands will not support cooperation among band members generally unless the benefit cost ratio is extraordinarily high. In contrast, there is considerable support for explanations based on cultural and/or genetic group selection (Soltis, Boyd, & Richerson, 1995; Bowles, 2007).

The high levels of mortality and frequent population crashes and dispersals that appear to have characterized the late Pleistocene are also unfavorable to the evolution of reciprocal altruism, even in dyads. Paleo-demographic data suggest that the life expectancy at age 15 during the late Paleolithic may have been about 17 years and a bit less for seven contemporary hunter-gatherer populations (Weiss, 1973). Assuming that cooperation requires the participation of adults (not those younger than 15) the relevant annual hazard rate is thus about 0.06, giving a lower bound on the rate of time preference of 6%, even in the absence of myopia, weakness of will, and the like. If older members of the group are critical to the success of cooperation then the maximal rate of time preference will be even higher. Due to frequent group conflicts and extraordinary climactic instability, group longevity was also probably quite limited, so that even if members of a group survived, they were unlikely to remain together over very long periods. The very low rate of growth of the human population over the whole prehistoric period, plus the high rate of human population growth in even poor contemporary foraging societies in good times, suggests periodic crises occurred in the past.

Summing up a pattern of demographic instability likely to have been faced by the foraging bands of the late Pleistocene, Gajdusek (1964, p. 121) writes:

> Migrations, murders and suicides, warfare, and massacres, splitting and fragmentation of communities and bands, or amalgamations and sudden mergers of groups, sudden social changes in mating practices and prohibitions and such natural accidents and catastrophes as earthquakes, floods, typhoons,

    volcanic eruptions, droughts, famines, and plagues all have a major influence
    in determining the genetic composition of small groups.

These are hardly conditions under which reciprocal altruism could be expected to flourish (Gintis, 2000).

    Nor is it likely that fitness-relevant interactions were confined to intimates. Far more consistent with what is known is the view that individuals had frequent encounters, both friendly and hostile, with members of other groups especially during migrations. Using evidence from a large number of existing and recent hunter-gatherer peoples, Fehr and Henrich (2004) find that individuals typically traveled hundreds of kilometers to sustain insurance networks, acquire livelihoods, and to find mates, encountering thousands of strangers as a result. Moreover, group conflict was likely to have been both frequent and lethal.

    As early as 75,000 to 90,000 years before the present, early-modern humans in eastern and southern Africa traded over hundreds of kilometers. Using evidence from a large number of existing and recent hunter-gatherer peoples, Fehr and Henrich (2004) point out that individuals typically traveled hundreds of kilometers to sustain insurance networks, acquire livelihoods, and to find mates, encountering thousands of strangers as a result. In these environments those who failed to distinguish between long-term and short-term or one-shot interactions would be at a significant fitness disadvantage as a result.

    Additional evidence that our ancestral hunter-gatherers frequently interacted with those outside their immediate group comes from studies of genetic differentiation among groups. Among aboriginal Australians, for example, distance is a very weak predictor of genetic relatedness, suggesting "that local groups traveled large distances and often came into contact with one another for the purposes of trade, which sometimes included exchange of people from each group" (Keats, 1977, p. 327). One can also infer that effective group size was substantial and that there was frequent between-group migration that one may infer from the degree of genetic differentiation among sub-populations. Among small partially reproductively isolated sub-populations, drift tends to increase between group genetic differentiations for selectively neutral loci, while migration attenuates differences. The equilibrium level of differentiation, that which balances these two effects, depends on the migration rate and a measure of group size. Equilibrium differentiation among groups is enhanced by population crashes, group fission along lineage lines, inequalities in reproductive success within groups, and non-random migration (Bowles, 2007). Taking account of these reasons for elevated genetic differentiation, the observed measures of genetic differentiation are inconsistent with the view that ancestral demes were small and reproductively isolated. These data suggest that our ancestors did not live in closed worlds in which interactions were confined to a small number of relatives or lifelong close associates.

    Thus, while reciprocal altruism, kin altruism, and reputation building no doubt played a role in early human cooperation, the view that Pleistocene conditions were such that these mechanisms would have been sufficient for the emergence and proliferation of a uniquely cooperative species is not supported by what is known about that period. In contrast, explanations of the emergence and proliferation of

cooperative behaviors based on gene-culture coevolution and multilevel selection are quite plausible.

Conclusive evidence about the origins of human cooperation may remain elusive given the paucity of the empirical record and the complexity of the dynamical processes involved. As in many problems of historical explanation, perhaps the best that one can hope for is a plausible explanation consistent with the known facts. This is what we have attempted to provide.

References

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390–1396.

Binford L. (2001). *Constructing frames of reference: An analytical method for archeological theory using hunter-gatherer and environmental data sets*. Berkeley: University of California Press.

Bowles, S. (2007). Genetic differentiation among hunter gatherer groups. Santa Fe Institute.

Bowles, S., Choi, J., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology, 23*, 135–147.

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

Dawkins, R. (1989). *The selfish gene* (2nd edn.). Oxford, England: Oxford University Press.

Dawkins, R. (2007). *The god illusion*. Boston and New York: Houghton-Mifflin Press.

Deacon, T. W. (1998). *Symbolic species: The co-evolution of language and the brain*. New York: Norton.

Farkas, S., & Robinson, J. (1996). *The values we live by: What Americans want from Welfare Reform*. New York: Public Agenda.

Fehr, E., & Henrich, J. (2004). Is strong reciprocity a maladaption? On the evolutionary foundations of human altruism. In: P. Hammerstein (Ed.), *Genetic and cultural origins of cooperation*. Cambridge, MA: The MIT Press.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature, 13*, 1–25.

Fong, C. M., Bowles, S., & Gintis, H. (2005). Reciprocity and the welfare state. In: H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiments and material interests: On the foundations of cooperation in economic life*. Cambridge, MA: The MIT Press.

Frank, S. A. (1995). Mutual policing and repression of competition in the evolution of cooperative groups. *Nature, 377*, 520–522.

Frank, S. A. (1998). *Foundations of social evolution*. Princeton, NJ: Princeton University Press.

Frolich, N., & Oppenheimer, J. A. (1992). *Choosing justice: An experimental approach to ethical theory*. Berkeley: University of California Press.

Gadagkar, R. (1991). On testing the role of genetic asymmetries created by haplodiploidy in the evolution of eusociality in the hympenoptera. *Journal of Genetics, 70*, 1–31.

Gajdusek, D. C. (1964). Factors governing the genetics of primative human populations. *Cold Spring Harbor Symposium on Quantitative Biology, 29*, 121–135.

Ghiselin, M. T. (1974). *The economy of nature and the evolution of sex*. Berkeley: University of California Press.

Gilens, M. (1999). *Why Americans hate Welfare*. University of Chicago Press.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206*, 169–179.

Gintis, H. (2003). The hitchhiker's guide to altruism: Genes, culture, and the internalization of norms. *Journal of Theoretical Biology, 220*, 407–418.

Gintis, H. (2007). A framework for the unification of the behavioral sciences. *Behavioral and Brain Sciences, 30*, 1–61.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology, 213*, 103–119.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution Human Behavior, 24*, 153–172.

Hamilton, W. D. (1964). The genetic evolution of social behavior I II. *Journal of Theoretical Biology, 7*(1–16), 17–52.

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

Henrich, J., McElreath, R., Barr, A., Ensimger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science, 312*, 1767–1770.

Keats, B. (1977). Genetic structure of the indigenous populations in Australia and New Guinea. *Journal of Human Evolution, 6*, 319–339.

Kerr, B., & Godfrey-Smith, P. (2002). Individualist and multi-level perspectives on selection in structured populations. *Biology and Philosophy, 17*, 477–517.

Kreps, D. M. (1988). *Notes on the theory of choice*. London: Westview.

Luttmer, E. F. P. (2001). Group loyalty and the taste for redistribution. *Journal of Political Economy, 109*, 500–528.

Maynard Smith, J., & Szathmary, E. (1997). *The major transitions in evolution*. Oxford, England: Oxford University Press.

Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (2003) *Niche construction: The neglected process in evolution*. Princeton, NJ: Princeton University Press.

Price, M. E. (2008). The resurrection of group selection as a theory of human cooperation. *Soc. Justice Res., 21*, this issue.

Richerson, P. J., & Boyd, R. (2004). *Not by genes alone*. Chicago: University of Chicago Press.

Smith, V. (1982). Microeconomic systems as an experimental science. *American Economic Review, 72*, 923–955.

Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural group selection: An empirical test. *Current Anthropology, 36*, 473–483.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In: J. H. Barkow, L. Cosmides, & J. Tooby (Eds.) *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). New York: Oxford University Press.

Tooby, J., & Cosmides, L. (1997). Evolutionary psychology: A primer. http://www.psych.ucsb.edu/research/cep/primer.html.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57.

Trivers, R. L. (2007). Reciprocal altruism: 30 years later. In: P. K. Kappeler, & C. P. van Schaik (Eds.), *Cooperation in primates and humans: Mechanisms and evolution* (pp. 67–85). Berlin: Springer.

Weaver, R. K., Shapiro, R. Y., & Jacobs, L. R. (1995). Poll trends: Welfare. *Public Opinion Quarterly, 39*, 606–627.

Weiss, K. M. (1973). *Demographic models for anthropology*. Washington, DC: Society for American Archeology.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton, NJ: Princeton University Press.

Wilson, E. O., & Holldobler, B. (2005). Eusociality: Origin and consequences. *PNAS, 102*, 13367–13371.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology, 51*, 110–116.