Soc Just Res
DOI 10.1007/s11211-008-0064-1

A Review of Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies, edited by J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, & H. Gintis & Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life, edited by H. Gintis, S. Bowles, R. Boyd & E. Fehr

Michael E. Price

Abstract Two books edited by members of the MacArthur Norms and Preferences Network (an interdisciplinary group, mainly anthropologists and economists) are reviewed here. These books in large part reflect a renewed interest in group selection that has occurred among these researchers: they promote the theory that human cooperative behavior evolved via selective processes which favored biological and/or cultural group-level adaptations as opposed to individual-level adaptations. In support of this theory, an impressive collection of cross-cultural data are presented which suggest that participants in experimental economic games often do not behave as self-interested income maximizers; this lack of self-interest is regarded as evidence of group selection. In this review, problems with these data and with the theory are discussed. On the data side, it is argued that even if a behavior seems individually-maladaptive in a game context, there is no reason to believe that it would have been that way in ancestral contexts, since the environments of experimental games do not at all resemble those in which ancestral humans would have interacted cooperatively. And on the theory side, it is argued that it is premature to invoke group selection in order to explain human cooperation, because more parsimonious individual-level theories have not yet been exhausted. In summary, these books represent ambitious interdisciplinary contributions on an important topic, and they include unique and useful data; however, they do not make a convincing case that the evolution of human cooperation required group selection.

Keywords Evolution of cooperation · Altruism · Strong reciprocity · Experimental economics

M. E. Price (✉)
Department of Psychology, School of Social Sciences, Brunel University, West London, Uxbridge, Middlesex UB8 1BL, UK
e-mail: Michael.Price@brunel.ac.uk
URL: http://www.people.brunel.ac.uk/~hsstmep/

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

Both these books were edited by members of the Norms and Preferences Network, a MacArthur Foundation-funded group of behavioral scientists who share a strong interest in understanding the evolutionary origins of human cooperation. Over the past several years, the editors have produced a large number of publications promoting their perspective on how human cooperation evolved. The gist of this perspective is that people are more cooperative than the standard economic and evolutionary theory would predict. According to the editors, while standard theory predicts that all people will cooperate only when it is in their self-interest to do so, the truth is that many people will cooperate even when they have no chance of recovering the costs of cooperation via reciprocity or any other means. The editors somewhat confusingly term this unselfish cooperativeness "strong reciprocity," in order to contrast it with regular reciprocity, which is what occurs when people cooperate in order to acquire return benefits from those whom they have assisted.

For years prior to the publication of these two volumes, the editors had been developing their empirical and theoretical case for strong reciprocity. Fehr had been especially active in making the empirical case; studies by himself and colleagues (e.g., Fehr & Gächter, 2000; Fehr & Fischbache, 2004) suggested that a large fraction of undergraduate subjects in industrialized societies would spend their own money both to cooperatively benefit others, and to punish non-cooperators, even when there was no chance of recouping these expenditures via reciprocity. Meanwhile, on the theoretical side, Gintis, Bowles, Boyd, Henrich and other formal modelers were explaining how such strong reciprocity could evolve (e.g., Boyd, Gintis, Bowles & Richerson, 2003; Gintis, 2000). Since strong reciprocity is fitness-damaging to the individual strong reciprocator, they need to invoke some kind of selective process operating above the level of the individual. They argued, therefore, that human cooperative behavior must have evolved by biological and/or cultural group selection, and they backed up their arguments with mathematical models and computer simulations.

These two volumes represent recent efforts by the editors to test, extend, and apply the theory of strong reciprocity. *Foundations* is more ambitious and unusual of the two books; in order to determine whether people behave as strong reciprocators not just in college environments in industrialized societies, but also in small-scale cultures around the world, the editors enlisted a small army of anthropologists to conduct experimental economic games in 15 societies on several different continents. In addition, though *Moral Sentiments* is a more typical edited volume, consisting of invited chapters on topics related to strong reciprocity and to the origins of cooperation more generally, it is unusual in its interdisciplinary scope: contributors include economists, anthropologists, primatologists and political scientists. The range of topics covered by these diverse contributors is wide, as one might expect, and includes outstanding and data-rich reviews of cooperation in primate groups (by Silk) and of food sharing in small-scale societies (by Kaplan & Gurven).

Strength of both volumes is the importance of their topic. Humans are capable of cooperating with unrelated individuals to an extent that is unprecedented among animals, and human society would be unrecognizable without this ability:

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

we would have no trade, no moral or legal systems, and no universities, religions, unions, armies, political parties, or organizations of any kind. However despite the centrality of cooperation to human social life, the issue of how human cooperativeness evolved is a contentious one that has been regarded as a major unsolved scientific puzzle for several decades. The basic puzzle is this: in a cooperative interaction that produces shared resources (e.g., a collective action), lower-contributing interactants ('free riders') will, all else equal, reap higher net benefits than higher-contributors (Olson, 1965). Unless the higher-contributors can solve this free-rider problem, they will be exploited to extinction, and cooperation will fail to evolve. The free-rider dilemma has been one of the more longstanding and intractable problems in behavioral science, and especially in behavioral biology, and the editors seem determined to solve it once and for all.

Another strength of these books is an emphasis on cross-cultural hypothesis testing. In order to thoroughly test any theory about the evolution of an aspect of human nature, it is necessary to study humans in as wide a range of cultures as possible, rather than studying only undergraduate samples. The contributors to *Moral Sentiments* review studies from a diverse sample of non-Western societies, and the 15-culture study chronicled in *Foundations* is one of the most expansive and ambitious cross-cultural research projects in history. Also admirable is the extent to which both books involve contributions by researchers from diverse fields. These contributors were able to see over the arbitrary boundaries that divide their disciplines, boundaries which continue to exist for historical, administrative, and professional reasons, but which have a little or nothing to do with scientific goals and the production of knowledge. When the topic is as basic to behavioral science as human cooperation, this kind of interdisciplinary approach is especially essential.

Despite these considerable strengths, however, these books suffer from a fundamental flaw: the theory of strong reciprocity, on which both books are based, is probably wrong. Understanding why this is so requires a thorough consideration not just of strong reciprocity itself, but also of the standard evolutionary theories of cooperation that strong reciprocity was designed to replace.

Strong Reciprocity Defined

In Foundations, the editors say "[s]trong reciprocity is a predisposition to cooperate with others, and to punish (at personal cost, if necessary) those who violate the norms of cooperation, even when it is implausible to expect that these costs will be recovered at a later date" (p. 8; italics in original). The above definition makes clear that strong reciprocity is actually decidedly non-reciprocal, in the respect that strong reciprocators do not expect to be compensated by the beneficiaries of their altruism. Hence, why was the term strong reciprocity chosen at all, instead of a seemingly more accurate term such as 'non-reciprocal altruism'? Apparently because, as the above definition also makes clear, strong reciprocity is meant to refer to the negative reciprocity that strong reciprocators exhibit when they punish non-cooperators. The

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

punishment of non-cooperators is considered a kind of second-order cooperation, because the threat of such punishment often convinces would-be shirkers to contribute their fair share.

The Evidence for Strong Reciprocity Comes Mainly from Experimental Economics

For evidence of strong reciprocity, the editors draw primarily on behavioral data from economics experiments such as dictator, ultimatum, and public good games. In all of these games, participants are endowed with a sum of money (or tokens that can be exchanged for money at the end of the game), and they are given an opportunity to either keep this money for themselves, or to give some or all of it to co-participants. For example, in the dictator game, a participant is endowed with (for example) £10 and then dictates how much of this amount will be gifted directly to one recipient co-player; what is not gifted is kept by the dictator. The ultimatum game is identical to the dictator game, except that the recipient of the gift has the opportunity either to accept or reject the offering; if the recipient accepts then the deal goes through, but if the recipient rejects then neither the recipient nor the offerer receives any money at all. In the public good game, participants play not with just one other co-player but with a small group of other co-players, and whatever portion of their endowment they choose to donate is multiplied by some factor >1.0 and then divided equally among all group members, including themselves. These games are typically conducted on computers in a university laboratory setting. All players usually play the game simultaneously in the same lab room, but the identity of each individual player is masked, so that a player cannot know the specific identity of his or her partner(s). Players can usually identify their co-players only by arbitrarily assigned anonymous titles such as Player 1, Player 2, etc. Therefore players never know which other participant in the lab is actually their partner in any one round.

All three of the above games are structured such that if a player is solely interested in achieving the maximum monetary payoff, then he or she should give the minimum possible amount to co-players. For example, in the dictator game, the dictator achieves maximal personal gain by offering the recipient as little as possible, and in the ultimatum game, the offerer gains maximally by offering the recipient as little as possible, and the recipient gains maximally by accepting whatever offer is made (because recipients who reject offers make nothing). In the public good game, all donations to the group are multiplied by some factor that is greater than 1.0 but less than the number of group members. As a result, while it is best for the group as a whole if each member donates as much as possible, each individual member has a personal incentive to donate as little as possible and to free ride on co-member donations. As the editors point out, any theory that regards people as being interested solely in maximizing their personal monetary gain would predict that participants in these games will behave as selfishly as the game allows. However, a large majority of players in

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

fact do not behave as income-maximizers. Players typically share from 40% to 50% of their endowments with their co-players in the ultimatum game, for example, and from 40% to 60% in the initial rounds of the public good game.

Is Generosity in Economics Games the Result of Hidden Reciprocal Motives and/or Western Cultural Bias?

One plausible explanation for this generosity is that it is actually the product of reciprocal motives that are selfish in the long run: perhaps players realize that their co-players will behave generously only as long as they themselves do the same, so they contribute in order to maximize the amount that they can acquire from co-players; or, similarly, perhaps they contribute to signal their generosity to co-players, in order to build their own reputation for magnanimity so that their future partners will return this magnanimity. However, such claims are apparently only partially correct. These games can be structured so as to deny players the opportunity to acquire reciprocal returns, and when they have this structure, players do behave less generously. For example, when the game is one-shot (i.e., after a player has interacted with a particular co-player, there is no possibility that he or she will interact with that co-player again), and/or when players are given a new arbitrary identity after every round of the game (which makes it impossible to build a reputation for magnanimity), generosity is reduced. However, even under these reciprocity-precluding conditions, players continue to behave significantly more generously than would be predicted by any theory that casts humans as income-maximizers.

The tendency of economic game players from several different societies to engage in strong reciprocity (i.e., to continue to contribute above the required minimum, even when reciprocity and reputation are impossible) raised the question of whether strong reciprocity might be some kind of evolutionary adaptation. However, up until a few years ago, most economic games had involved only undergraduates from industrialized societies as subjects, which raised doubts about the extent to which strong reciprocity was culturally universal. In order to address this question, the editors recruited a research team of anthropologists and charged them with the task of conducting experimental games in a wide variety of 15 small-scale, non-western, non-industrialized societies. This effort became the research project that is detailed in *Foundations*. Members of the project team conducted ultimatum games in all 15 of the cultures, as well as public good games in six cultures and dictator games in three. The results of this research showed that players in all of these cultures, like undergraduates in industrialized societies, behave more generously than an income-maximization theory would predict: in no culture did players consistently contribute only the required minimum. However, there was some cross-cultural diversity in how the game was played, and some deviations from the Western undergraduate pattern of contributing roughly half of one's endowment initially. For example, mean offers in the ultimatum game ranged from 25% among the Quichua of Ecuador to 57% among the Lamalera of Indonesia. In order to explain the variation of mean offers in the ultimatum game, the cultures were ranked in terms of two factors: market integration (i.e., how much exposure the culture has to western-style market economies), and

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

payoffs to cooperation (i.e., the extent to which economic life in that culture depends on cooperation with non-kin). These two factors together explained 47% of the variance in mean ultimatum game offers.

Strong Reciprocity is Theorized to have Evolved by Group Selection

The cross-cultural data presented in *Foundations* suggests to the editors that people all over the planet (and not just undergraduates in industrialized societies) act as "strong reciprocators" as opposed to income-maximizers, and that strong reciprocity could therefore be a universal aspect of human nature. The editors' theoretical work focuses primarily on how strong reciprocity could have evolved as a behavioral adaptation in human ancestral environments. Strong reciprocity would seem problematic for standard theories of behavioral biology, because most evolutionary biologists assume that genes promote their own replication, by encoding adaptations that promote the inclusive fitness (hereafter 'fitness') of the individuals who bear them. Since strong reciprocators are believed to cooperate even when they cannot recoup the costs of cooperating, they should be at a fitness disadvantage relative to the lesser-sacrificing beneficiaries of their altruism, and standard evolutionary theory would seem unable to explain their existence. Therefore, the editors see a need to develop an alternative to the theory that selfish genes promote individual fitness.

Their alternative theory, presented most completely in Chapter 7 of *Moral Sentiments*, suggests that although strong reciprocity harms the fitness of the individual strong reciprocator, its benefits to the strong reciprocator's group are sufficient to permit its evolution. In other words, strong reciprocity may have evolved via a process of "group selection" whereby groups with more strong reciprocators are "less prone to extinction… because more cooperative groups are more effective in warfare, more successful in co-insuring, more adept at managing commons resources, or other similar reasons" (p. 217). The editors include a formal model suggesting that a strongly reciprocal strategy—which both cooperates and punishes non-cooperators—can more successfully evolve than a strategy that cooperates without punishing. They emphasize that "…the process modeled here is likely to be much less important for genetic evolution than for cultural evolution" (p. 239), because their model assumes a level of migration between groups that would seriously weaken the effects of biological group selection. However, in Chapter 1 of *Moral Sentiments*, the editors suggest "strong reciprocity, like kin altruism and reciprocal altruism, has a significant genetic component" (p. 22). While the theory is always presented as group selectionist, it not always clear whether the editors wish to argue for cultural or biological group selection.

What Theories is Strong Reciprocity Designed to Replace?

According to the editors, the fact that subjects in economic games tend to give more than the required minimum runs contrary to the "selfishness axiom" of traditional

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

economics and evolutionary biology, and suggests that these disciplines' "canonical model" (*Foundations*, p. 49) should be replaced by the theory of strong reciprocity. However, is it fair to say that standard economic and evolutionary theory is characterized by this selfishness axiom, or are the editors attempting to demolish a pair of straw men?

The editors' attack on standard economic theory would be more impressive if they cited examples of the selfish axiom being invoked by contemporary economists. I did not notice a single such reference in either book. In their *Behavioral and Brain Sciences* article (essentially a précis of *Foundations*), the editors do cite F. Y. Edgeworth as saying that the "first principle of economics is that every agent is actuated only by self-interest," (Henrich et al., 2005, p. 839) but Edgeworth does not qualify as a contemporary foil since this quote dates from 1881. The current standard perspective from economics is that people consistently "behave *as though* maximizing something… [called] utility… [but it is not] axiomatic that utility is the same as income" (Binmore, 2005, p. 817, italics in original). In other words, most contemporary economists would not expect for subjects in economic games to behave as income-maximizers. However, if there are any who would make this prediction, they will find that it has been falsified cross-culturally in *Foundations*.

Nor does the selfishness axiom appear to be a current characteristic of orthodox evolutionary theory. However, behavioral biology has been strongly associated with this axiom for several decades, particularly since Dawkins decided to give his bestselling 1976 popularization of Hamilton's theories the title *The Selfish Gene*. This title refers to the fact that genes which promote their own replication, by benefiting the fitness of the individuals in whom they are contained, will tend to proliferate in a population. Since genes often promote individual fitness by enabling individuals to cooperate with others, selfish genes often produce cooperative, rather than selfish, individuals: "Animals are sometimes nice and sometimes nasty, since either can suit the self-interest of genes at different times. That is precisely the reason for speaking of 'the selfish gene' rather than, say, 'the selfish chimpanzee'" (Dawkins, 1998, p. 212). However, Dawkins himself has at times seemed confused on this point; for example when he stated that if you want people to "cooperate generously and selflessly towards a common good, you can expect little help from biological nature. Let us try and teach generosity and altruism because we are born selfish" (Dawkins, 1989 [1976], p. 3). Dawkins' seemingly contradictory statements suggest how easy it can be to confuse selfish genes with selfish people. Nevertheless, while most behavioral biologists would certainly agree that genes replicate primarily by promoting individual fitness (and would therefore disagree with the editors' group selectionism), they would also agree that a crucial way in which genes promote fitness in humans is by building genuinely cooperative individuals.

Problems with the Data: Minds are Executors of Adaptations, Rather than Maximizers of Income, Fitness, or Anything Else

As noted above, the editors' evidential case for strong reciprocity rests almost entirely on data from economic games. While these data are potentially useful, one

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

must be cautious in interpreting their meaning, especially with regard to what they imply about human psychological and behavioral adaptations for cooperative behavior. The editors seem to believe that if mainstream evolutionary theory subscribes to the selfish axiom, then in games designed to make reciprocity and reputation impossible, this theory should expect players to act as income-maximizers. In other words, they imply that evolutionary biologists would expect behavior in these games to be enabled by some kind of cognitive mechanism which simply calculates the range of monetary payoffs given the game's parameters, and then acts to maximize this payoff.

However, this view of the mind as an income-maximizing (or utility- or fitness-maximizing) device is certainly not the mainstream perspective among those who study the mind from an evolutionary perspective. Evolutionary psychologists regard the mind as an *executor of adaptations*, as opposed to a maximizer of fitness or of any other currency (Tooby & Cosmides, 1992). A cognitive adaptation can be thought of as a semi-autonomous 'if, then' mechanism: if a specific kind of informational input is present in the environment, then the mechanism generates a certain kind of psychological or behavioral output (for example, 'if you see a snake-like object, then become wary of this object'). Although the output would have been, on average, an adaptive response to the input in environments in which the input–output mechanism evolved, the mechanism itself can only execute its protocol and does not know how to generate adaptive behavior per se. A cognitive adaptation will produce its output in any environment that provides the kind of informational input that would have stimulated this execution in ancestral environments, even if some novel aspect of that environment causes the adaptation to fail to actually produce a fitness-enhancing outcome (Burnham & Johnson, 2005). This view of the mind has been standard among ethologists for decades (Burnham & Kurzban, 2005), and has always characterized modern mainstream evolutionary psychology. For example, in what many consider to be the first major contribution to evolutionary psychology to emerge following biology's 'new synthesis' in the 1970s, Symons (1979) discusses the toad that continues to eat insect-sized metal pellets that an experimenter rolls past it, until it becomes "a living beanbag" (Symons, 1979, p. 21), and the herring gull that cares for eggs that are not its own (and that are not even the correct color), and that have been placed in its nest by an experimenter. These species did not evolve in environments characterized by pellet-rolling or egg-introducing scientists, and they both possess cognitive adaptations that fail to function adaptively in novel, experimental environments.

When subjects play economic games, their psychological adaptations for social behavior are being deployed in environments that are radically different from those in which they evolved. As a result, these adaptations may produce behaviors which seem maladaptive for individual fitness (if income-maximization is seen, as it is by the editors, as a proxy for fitness-maximization). However, this result 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.

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

In what important ways are the environments of economic games different from those of ancestral human social interaction? To answer this question, consider that, as discussed above, these games are often designed so as to minimize the extent to which subjects can gain via reciprocity and reputation. For example, the ultimatum game that was played in 15 different cultures was anonymous and one-shot, so players could not know the identities of their partners, and would expect to interact with each partner only once. The editors intended this design to eliminate the possibility that players would cooperate only in order to elicit more generous contributions from co-players, as the result of either direct reciprocity or indirect (i.e., reputation-dependent) reciprocity. After thus ruling out all cooperation motivated by long-term selfishness, the editors believed that any remaining cooperativeness would constitute evidence for truly unselfish strong reciprocity. In addition after observing significant levels of this remaining cooperativeness, the editors concluded that just as modern humans engage in strong reciprocity in experimental anonymous one-shot interactions, ancestral humans also engaged in strong reciprocity in real-life anonymous one-shot interactions; and although this behavior was individually fitness-damaging for ancestral strong reciprocators, it evolved nonetheless because it was favored by group selection. For the editors' reasoning to be accurate, it must be true that the anonymous one-shot interactions they created were similar to the anonymous one-shot interactions that occurred naturally in ancestral environments; otherwise, there is no reason to believe that their subjects' behavior in these interactions reveals anything about how ancestral humans behaved in these interactions.

It is unlikely, however, that ancestral humans would have routinely encountered anonymous one-shot interactions of the kind modeled in experimental games. In the first place, in the small, face-to-face social environments of ancestral humans, encounters with complete strangers would have been rare. Rarer still would have been encounters with strangers that would have been truly anonymous or one-shot. For example, say an ancestral individual encounters an unfamiliar person one day while they were both out hunting for gazelle in a remote location, and he interacts in some way with the person. For the interaction to be and to remain anonymous, he would have to be certain that the stranger did not already know who he was; and even if he could be certain of this, he would have to also be certain that the stranger would never see him again, discover his identity, or be able to describe him (and his actions) to anyone who might know his identity. A truly one-shot interaction would also be unlikely, because no experimenter would be present to ensure the termination of the interaction. If the interaction left one hunter feeling grateful (or angry), for example, he may decide to repay (or attack) the other, either on the spot or at some later date. As Trivers notes, "[s]ocial interactions are intrinsically repeat interactions - at least over short periods of time and usually over much longer periods as well. So the natural assumption is that in life we will respond to social situations as if there were later repercussions to our actions" (Trivers, 2004, p. 965). In light of the social reality that exists outside of experimental labs, it seems unlikely that ancestral individuals would have had much opportunity to engage in social interactions that they could safely expect to be truly anonymous and one-shot. In natural social environments, the only way to thoroughly ensure that one's

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

interactant partner does not tell others about one's actions is to completely destroy that partner's ability to communicate with others (e.g., via homicide). However, the editors do not theorize that anonymity must be maintained through anything like homicide; they just assume that ancestral individuals could often have correctly expected interactions to be anonymous and one-shot, and that strong reciprocity evolved (by group selection) under such conditions.

For the sake of argument, however, imagine that relatively anonymous one-shot interactions did sometimes occur in ancestral environments, and that modern humans do have psychological adaptations for negotiating such social landscapes. If so, then experimenters would be able to observe the deployment of these adaptations, if they designed an experiment that convincingly simulated the conditions under which these ancestral interactions occurred. Would such an ecologically valid experiment be at all similar to the one designed by the editors? Probably not. Ancestral interactions would have been relatively anonymous and one-shot only if they occurred in socially remote locations in the absence of witnesses, and especially of witnesses who were particularly interested in interactant behavior. Economic game environments do not resemble such social scenarios at all: games are conducted in lab or lab-like spaces that are crowded with subjects and experimenters, and subjects know that experimenters are intensely interested in their decisions and behaviors (after all, the whole point of the experiment is to evaluate behavior). Even if players are told that the games are anonymous, and even if they consciously believe that their identities will be shielded from co-players (and also from experimenters, if the design is double-blind), the semi-autonomous psychological mechanisms which regulate players' cooperative behavior might not recognize the situation as one that would ensure anonymity. Unless these mechanisms are given information from the game environment that they are operating in the kind of social situation that would have afforded relative anonymity in the ancestral past, there is no reason to predict that players should act as if there actions will be relatively invisible to others. There is thus little reason to expect that players in the editors' games should behave as if they are participating in an ecologically valid, anonymous one-shot encounter.

The editors are aware of the general criticism that economic games do not resemble real-life ancestral social interactions, and that human minds should therefore not be expected to be adapted for such games. They respond by saying that players must be able to understand the cues to social context that are provided by the game environment, because when these cues are more indicative of an anonymous one-shot interaction, players behave more selfishly. For example, in a game where players have repeated interactions with the same player, they are more generous than in a one-shot game (presumably because the repeated game allows them to selfishly sustain a reciprocal relationship); or in a game where players are identifiable to co-players only via ID numbers which are reassigned arbitrarily after every round, thus making it impossible to build reputations over the course of the game, players are more selfish than in games where reputation-building is possible. While such evidence supports the view that people are indeed motivated by 'ultimately selfish' concerns about reputation and reciprocity, the editors attempt to turn this evidence on its head by emphasizing that players still contribute significantly more than nothing, even in games where reciprocity and reputation are

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

impossible, behaviors which they interpret as group-selected strong reciprocity. In other words, since players can apparently tell the difference between anonymous and non-anonymous games and between iterated and one-shot games, they should be able to detect anonymous one-shot conditions; and if their cooperative behavior were the product only of individual-level psychological adaptations, they would be completely unmotivated to contribute under such conditions.

The editors are incorrect, however, in asserting that individual-level cooperative adaptations should necessarily contribute nothing in games that are designed to prohibit reciprocity and reputation. Again, just because the experimenters are convinced that they have designed a truly anonymous one-shot game, their subjects' cognitive mechanisms for cooperation may not be similarly 'convinced.' There is no reason to expect that the baseline rate of cooperativeness in such games should be zero, even if the mechanisms producing cooperative behavior are individual-level adaptations.

Problems with the Theory: Group Selection has Never Worked as an Evolutionary Explanation for Human Cooperation

As noted above, the editors are impressed by the fact that subjects in economics experiments continue to behave somewhat generously, even after experimentalists have attempted to rule out the possibility that subjects could be motivated by concerns related to reciprocity and reputation. In both *Foundations* and *Moral Sentiments*, the editors advocate a group selectionist theory of strong reciprocity in order to account for this residual cooperativeness. Their theory is by no means the first attempt to explain apparently altruistic behavior in terms of group selection; even Darwin (1871) resorted to group selectionism in order to explain human self-sacrifice, and 'naïve' group selectionist accounts of behavior were relatively common in behavioral biology prior to Williams' (1966) deeply influential critique of these approaches. The general problem with group selectionist approaches, as Williams and many since have pointed out, is that they leapfrog lower levels of selection in order to explain adaptation at higher levels, even though selection's power to sculpt adaptations is much stronger at lower levels. Parsimony demands that we attempt to exhaust all plausible lower-level adaptationist explanations for behavior, before we resort to higher-level, group selectionist explanations: "When recognized, adaptation should be attributed to no higher a level of organization than is demanded by the evidence" (Williams, 1966, p. v).

That is not to say that all forms of group selection have necessarily been irrelevant in shaping aspects of human psychology and culture. On the contrary, some form of cultural group selection seems central to explaining why certain cultural characteristics come to proliferate in the world. For example, cultural traits associated with a relatively powerful society may spread throughout the world, as that society imposes its culture on other societies via conquest, and as it is imitated by other societies who wish to emulate its power. But while cultural group selection is probably important for explaining the origins of some human behaviors, it cannot provide an adequate theory for the ultimate origins of human cooperation (Burnham

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

& Johnson, 2005). The ability to engage in cooperative interactions, including complex ones such as collective actions, is a fundamental attribute of human sociality that probably has biological roots. By being well adapted for cooperation in groups, ancestral individuals would have been able to engage in a variety of fitness-promoting activities, for example, cooperative hunting and foraging, and group defense against predators (Price, Cosmides & Tooby, 2002), while at the same avoiding being exploited to extinction by free riders. To discover the deepest origins of the human ability to cooperate, the editors should be focusing not on cultural evolution but on human biological nature.

Fortunately, since biology's new synthesis in the second half of the twentieth century, behavioral scientists have been gifted a new set of tools with which to solve the puzzle of human cooperation. The most crucial of these tools are the theories of genic self-favoritism (Hamilton, 1964) and reciprocal altruism (Trivers, 1971). Since we already have these tools at hand, there is no need to reinvent the wheel in order to solve the puzzle, let alone a square wheel like group selectionism, which has always been a non-starter as an explanation for human cooperation. There is a need, however, for more research that extends and synthesizes the theories of Hamilton and Trivers, in order to explain the more complex forms of human *n*-person cooperation, such as collective action. While Hamilton and Trivers' theories have successfully explained several general types of animal cooperation (e.g., altruism among close genetic kin, reciprocal exchange between dyadic partners), their potential to explain complex human cooperation is only now beginning to be fully exploited (Brown & Moore, 2002; Brown, Palameta & Moore, 2003; Frank, 2005; Tooby, Cosmides & Price, 2006). The basic prediction of Hamilton and Trivers is that collective action participants ought to *cooperate with other cooperators*, that is, preferentially form groups with individuals whom they expect will cooperate, and behave cooperatively only to the extent that their co-members reciprocate their cooperativeness. In this way, cooperators can avoid being exploited to extinction by lower-contributing free riders, and cooperation can evolve via individual-level adaptation. A growing body of experimental and field data suggests that cooperation with other cooperators is indeed the most common behavioral pattern of collective action participants cross-culturally (review in Price, 2006), and formal models suggest that such behavior could evolve via individual-level adaptation (Johnson, Price, & Takezawa, in press; Takezawa & Price, in preparation).

Conclusion

Much of the content of these volumes rests on a wobbly theoretical foundation that is unlikely to succeed as an explanation for the ultimate origins of human cooperation. Nevertheless, these books' contributions should not be underestimated. *Foundations* is one of the most ambitious cross-cultural research projects ever in behavioral science, and *Moral Sentiments* includes several outstanding chapters from an interdisciplinary group of experts on cooperation. Finally, both books draw attention to one of the most longstanding and fundamental puzzles in social science and behavioral biology: the evolution of complex cooperation in humans. All of

Journal : Small-ext 11211 | Dispatch : 19-1-2008 | Pages : 13
Article No. : 64 | LE | TYPESET
MS Code : 64 | CP | DISK

behavioral science will benefit when this puzzle is finally solved, and it benefits today from the discussion and controversy generated by such serious attempts to solve it.

References

Binmore, K. (2005). Economic man—or straw man? *Behavioral and Brain Sciences, 28*, 817–818.

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

Brown, W. M., & Moore, C. (2002). Smile asymmetries and reputation as reliable indicators of likelihood to cooperate: An evolutionary analysis. In S. P. Shohov (Ed.), *Advances in psychology research* (Vol. 11, pp. 59–78). New York: Nova Science Publishers.

Brown, W. M., Palameta, B., & Moore, C. (2003). Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior, 1*, 42–69.

Burnham, T., & Johnson, D. D. P. (2005). The evolutionary and biological logic of human cooperation. *Analyse & Kritik, 27*, 113–135.

Burnham, T. C., & Kurzban, R. (2005). On the limitations of quasi-experiments. *Behavioral and Brain Sciences, 28*, 818–819.

Darwin, C. R. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Dawkins, R. (1989 [1976]). *The selfish gene* (2nd ed.). Oxford: Oxford University Press.

Dawkins, R. (1998). *Dawkins, Richard. Unweaving the rainbow: Science, delusion and the appetite for wonder*. London: Penguin.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*, 980–994.

Fehr, E., & Fischbacher, U. (2004). Third party punishment and social norms. *Evolution and Human Behvavior, 25*, 63–87.

Frank, R. H. (2005). Altruists with green beards: Still kicking? *Analyse & Kritik, 27*, 85–96.

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

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

Henrich, J., Boyd, R., Bowles, S., Gintis, H., Fehr, E., Camerer, C., McElreath, R., Gurven, M., Hill, K., Barr, A., Ensminger, J., Tracer, D., Marlow, F., Patton, J., Alvard, M., Gil-White, F., & Henrich, N. (2005). 'Economic Man' in cross-cultural perspective: Ethnography and experiments from 15 small-scale societies. *Behavioral and Brain Sciences, 28*, 795–816, 838–855.

Johnson, D. D. P., Price, M. E., & Takezawa, M. (In press). Renaissance of the individual: Reciprocity, positive assortment, and the puzzle of human cooperation. In Crawford & Krebs (Eds.), *Foundations of evolutionary psychology*.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge: Harvard University Press.

Price, M. E. (2006). Monitoring, reputation and "greenbeard" reciprocity in a Shuar work team. *Journal of Organizational Behavior, 27*, 201–219.

Price, M. E., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23*, 203–231.

Symons, D. (1979). *The evolution of human sexuality*. Oxford: Oxford University Press.

Takezawa, M., & Price, M. E. (In preparation). Revisiting 'the evolution of reciprocity in sizeable groups'.

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., & Price, M. E. (2006). Cognitive adaptations for *n*-person exchange: The evolutionary roots of organizational behavior. *Managerial and Decision Economics, 27*, 103–129.

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

Trivers, R. (2004). Mutual benefits at all levels of life. *Science, 304*, 964–965. .

Williams, G. C. (1966). *Adaptation and natural selection*. Princeton: Princeton University Press.

Journal : **Small-ext 11211**	Dispatch : **19-1-2008**	Pages : **13**
Article No. : **64**	□ LE	□ TYPESET
MS Code : **64**	☑ CP	☑ DISK