

## The Evolution of Cooperation in Heterogeneous Populations

Samuel Bowles  
Herbert Gintis

May 13, 2003

### Abstract

How do human groups maintain a high level of cooperation despite a low level of genetic relatedness among group members? We suggest that many humans have a predisposition to punish those who violate group-beneficial norms, even when this reduces their fitness relative to other group members. Such altruistic punishment is widely observed to sustain high levels of cooperation in behavioral experiments and in natural settings. It is known that if group extinctions are sufficiently common, altruistic punishment may evolve through the contribution of norm adherence to group survival. Additionally, those engaging in punishment of norm violators may reap fitness benefits if their punishment is treated as a costly signal of some underlying but unobservable quality as a mate, coalition partner, or opponent. Here we explore a different mechanism in which neither signaling nor group extinctions plays a role. Rather, punishment takes the form of ostracism or shunning, and those punished in this manner suffer fitness costs.

We offer a model of this behavior, which we call *strong reciprocity*: where members of a group benefit from mutual adherence to a social norm, strong reciprocators obey the norm and punish its violators, even though they receive lower payoffs than other group members, such as selfish agents who violate the norm and do not punish, and pure cooperators who adhere to the norm but free-ride by never punishing. Our agent-based simulations show that, under assumptions approximating some likely human environments over the 100,000 years prior to the domestication of animals and plants, the proliferation of strong reciprocators when initially rare is highly likely, and that substantial frequencies of all three behavioral types can be sustained in a population.

## 1 Introduction

How do human groups maintain a high level of cooperation despite a low level of genetic relatedness among group members?<sup>1</sup> The hypothesis we explore is that cooperation is maintained because many humans have a predisposition to punish those who violate group-beneficial norms, even when this reduces their fitness relative to other group members. Compelling evidence for the existence and importance of such altruistic punishment comes from controlled laboratory experiments, particularly the study of public goods, common pool resource, trust, ultimatum, and other games (Yamagishi 1986, Ostrom, Walker and Gardner 1992, Fehr and Gächter 2002), from ethnographic studies of simple societies (Knauff, 1991; Boehm, 1984,1993), from historical accounts of collective action (Moore 1978, Scott 1976, Wood 2003), as well as from everyday observation. Several plausible resolutions to the evolutionarily puzzle posed by altruistic punishment have been offered. If group extinctions are sufficiently common, altruistic punishment may evolve through the contribution of norm adherence to group survival (Boyd, Gintis, Bowles and Richerson 2003). Also, those engaging in punishment of norm violators may reap fitness benefits if their punishment is treated as a costly signal of some underlying but unobservable quality as a mate, coalition partner, or opponent (Gintis, Smith and Bowles 2001). Here we explore a different mechanism in which neither signaling nor group extinctions plays a role. Rather, punishment takes the form of ostracism or shunning, and those punished in this manner suffer fitness costs.

We hypothesize that where members of a group benefit from mutual adherence to a norm, individuals may obey the norm and punish its violators, even when this behavior incurs fitness costs by comparison to other group members who either do not obey the norm or do not punish norm violators, or both. We call this *strong reciprocity*. Strong reciprocity is altruistic, conferring group benefits by promoting cooperation, while imposing upon the reciprocator the cost of punishing shirkers.

Our model assumes strong reciprocity is passed intergenerationally from parent to child (for simplicity we assume a single-parent transmission process, the genetic component of which is therefore haploid), through some combination of genetic and cultural mechanisms, the genetic component of which is an adaptation that has emerged over the course of human evolution. As the late Pleistocene is the only

---

<sup>1</sup>We would like to thank Christopher Boehm, Jung-Kyoo Choi, Marcus Feldman, Robert Boyd, Leda Cosmides, Steven Frank, Alan Grafen, Kristin Hawkes, Hillard Kaplan, Olof Leimar, Peter Richerson, Rajiv Sethi, Eric Alden Smith, E. Somanathan, Leigh Tesfatsion, Polly Wiessner, and Peyton Young for their help with this paper, and the John D. and Catherine T. MacArthur Foundation for financial support and the Santa Fe Institute for a supportive research environment. Institutional affiliations: Santa Fe Institute. The authors can be contacted at bowles@santafe.edu, <http://www.santafe.edu/~bowles> and hgintis@attbi.com, <http://www-unix.oit.umass.edu/~gintis>.

period in the history of anatomically modern humans long enough to account for a significant development in modern human gene distributions, we base our model on the structure of interaction among members of the mobile hunter-gatherer bands in this period. Modern accounts of these societies record considerable variety in social organization and livelihood (Kelly 1995, Binford 2001). But widespread participation in joint projects such as hunting and common defense as well as the sharing of food, valuable information, and other sources of survival among many of these societies in the modern world is well established. A good case can be made that these cooperative projects were at least as important among our Late Pleistocene ancestors as they are among modern mobile foraging bands (Boehm 2002).<sup>2</sup>

Our model therefore reflects the following empirical considerations. First, groups are sufficiently small that members directly observe and interact with one another, yet sufficiently large that the problem of shirking in contributing to public goods is present. Second, there is no centralized structure of governance (state, judicial system, Big Man, or other) so the enforcement of norms depends on the participation of peers. Third, there are many unrelated individuals, so altruism cannot be explained by inclusive fitness. Fourth, status differences are quite limited, especially by comparison to agricultural and later industrial societies, which justifies our treatment of individuals as homogeneous other than by behavioral type and by the group to which they belong. Fifth, the sharing on which our model is based—either of food individually acquired or of the common work of acquiring food, for example—is characteristic of these societies. Sixth, the individuals in our model do not store food or accumulate resources. This, too, is a characteristic of at least those hunter-gather bands based on what Woodburn (1982) calls an “immediate return” system of production.

Seventh, we take the major form of punishment to be *ostracism* and we treat the cost of being ostracized as endogenously determined by the amount of punishment and the demographic parameters of the model. This manner of treating punishment reflects a central aspect of hunter-gatherer life: since individuals can often leave the group to avoid punishment, the cost of being ostracized is among the more serious penalties that can be levied upon an individual group member.<sup>3</sup> Finally, behavioral

---

<sup>2</sup>We have relied on Woodburn (1982), Kaplan, Gurven, Hill and Hurtado (2003), Binford (2001) and Boehm (2000) for overviews and interpretations, and on Balıkcı (1970), Chagnon (1977), Lee (1979), Cashdan (1980), Boehm (1982), Kaplan, Hill, Hawkes and Hurtado (1984), Kaplan and Hill (1985b), Kaplan and Hill (1985a), Blurton Jones (1987), Woodburn and Barnard (1988), Endicott (1988), Kent (1989), Knauff (1989), Knauff (1991), Hawkes (1992), Boehm (1993), Hawkes (1993), Damas (1972), Kelly (1995), and Wiessner (2002) for ethnographic field studies. For a particularly detailed study of ostracism, see (Mahdi 1986).

<sup>3</sup>There being relatively little individually held property, individuals cannot be severely punished

heterogeneity is an emergent property of populations in our model, one that corresponds to what we know from the ethnographic record of foraging bands, as well as from the experimental evidence on both hunter gatherers (Henrich, Boyd, Bowles, Camerer, Fehr and Gintis 2003) and modern market based societies (Loewenstein, Thompson and Bazerman 1989, Andreoni and Miller 2002).

It might be thought that *Cooperators*, who unconditionally cooperate but never punish would outcompete *Reciprocators*, who bear the cost of punishing norm violators. But as *Cooperators* replace *Reciprocators* in a group, the fraction of *Selfish* agents increases, and with *Reciprocators* less frequent in the group, the *Selfish* agents increasingly shirk, thereby attaining higher fitness than the *Cooperators*, and eventually replacing them. We model this dynamic below (see Figure 1). For this reason, *Cooperators* do not displace *Reciprocators* in the long run. Moreover, once a low level of shirking is established, the expected costs of punishing others become quite low, so the within-group selection pressures operating against *Reciprocators* is weak.

## 2 Equilibrium Working, Shirking and Punishing Within a Group

Consider a population in which agents can live and work alone, in which case each has fitness  $\phi_0 < 0$ .<sup>4</sup> Agents can also work cooperatively in a group, each producing an amount  $b$  at cost  $c$  (all benefits and costs are in fitness units). We assume that output of the group is shared equally by the agents, so if all group members work, each has a net group fitness benefit  $b - c > 0$ .

The group consists of three type of actors. The first type, whom we call *Reciprocators*, work unconditionally and punish shirkers. The second type, whom we call *Selfish*, maximize fitness. They never punish shirkers, and work only to the extent that the expected fitness cost of working exceeds the expected fitness cost of being punished. The third type, whom we call *Cooperators* work unconditionally but never punish shirkers.

Parents pass on their type to their offspring with probability  $1 - \epsilon$ , and with probability  $\epsilon/2$ , an offspring takes on each of the other two types. We call  $\epsilon$  the *rate of mutation*. Also, with probability  $1 - \epsilon$ , *Selfish* agents inherit the estimate of

---

by having their wealth confiscated; there being no fixed residence, individuals cannot be jailed or otherwise confined; extreme physical harm can be meted out against norm violators, but such measures are generally reserved for such serious crimes as adultery and murder.

<sup>4</sup>By the *fitness* of an agent we mean the expected rate of growth of the agent plus his offspring in one period or, equivalently, the expected number offspring produced by the agent in one period minus the probability the agent dies in that period. Fisherian fitness is thus one plus fitness in our sense. We assume that a constant is added to each agent's fitness so that the population remains constant in total size across periods.

$s > 0$  (the cost of being ostracized) from their parents, but with probability  $\epsilon$  receive a distinct  $s$ , drawn from a uniform distribution. Note that  $s$  is thus an endogenous variable that changes over time. Selfish agents with an  $s$  that is different from the objective fitness cost of being ostracized may shirk too little or too much, leading to suboptimal fitness, a selective pressure for a modification of  $s$ .<sup>5</sup>

Suppose a Selfish agent shirks a fraction  $\sigma_s$  of the time, so the average rate of shirking is given by  $\sigma = (1 - f_r - f_c)\sigma_s$ , where  $f_r$  is the fraction of Reciprocators in the group, and  $f_c$  is the fraction of Cooperators. The fitness value of group output is  $n(1 - \sigma)b$ , where  $n$  is the size of the group. Since output is shared equally, each member receives  $(1 - \sigma)b$ . The loss to the group from a Selfish agent shirking is  $b\sigma_s$ . The fitness cost of effort function, which can be written as  $\lambda(1 - \sigma_s)$ , is increasing and convex in its argument, with  $\lambda(1) = c$  and  $\lambda(0) = 0$ . Expending effort always benefits the group more than it costs the workers, so  $(1 - \sigma_s)b > \lambda(1 - \sigma_s)$  for  $\sigma_s \in (0, 1]$ , so at every level of effort,  $1 - \sigma_s$ , working helps the group more than it hurts the worker.

We assume that group size  $n$  is sufficiently large that  $(1 - \sigma_s)b/n < \lambda(1 - \sigma_s)$  for  $\sigma_s \in (0, 1]$ , so if there is no punishing of shirkers, Selfish agents would completely shirk ( $\sigma_s = 1$ ). Thus in the absence of punishment of norm violators the members face a public goods problem (i.e., an  $n$ -player prisoner's dilemma), in which the dominant strategy is to not contribute.

We model punishment as follows. The fitness cost to a Reciprocator of punishing a shirker is  $c_p > 0$ . A member shirking at rate  $\sigma_s$  will be punished with probability  $f_r\sigma_s$ , where  $f_r$  is the fraction of agents in the group who are Reciprocators. Punishment consists of being ostracized from the group. We assume an ostracized agent works alone for a period of time before being readmitted to a group.

Selfish agents, given their assessment  $s$  of the cost of being ostracized, and with the knowledge that there is a fraction  $f_r$  of Reciprocators in their group, choose a level of shirking  $\sigma_s$  to maximize expected fitness.<sup>6</sup>

Writing the fitness cost of working,  $g(\sigma_s)$ , as the cost of effort plus the expected cost of being ostracized, plus the agent's share in the loss of output associated with

<sup>5</sup>In fact, in our model, Selfish agents with very high  $s$  behave exactly like Cooperators, except that if there are zero Reciprocators in the group, they supply zero effort.

<sup>6</sup>Either of two informational assumptions justify our model. The first is that the type of an agent is unknown, but the fraction of Reciprocators is known to Selfish agents when they join a group. The second is that the identity of Reciprocators is known to Selfish agents after they join a group, but this cannot affect the probability of being caught shirking. It does not matter for our model whether or not Reciprocators can distinguish between Cooperators and Selfish agents, or whether they know the expected cost of ostracism used by a Selfish agent to determine the agent's shirking level. All of these informational assumptions can be weakened, at the cost of increased model complexity, but none is crucial to its operation.

one's own shirking, we have

$$g(\sigma_s) = \lambda(1 - \sigma_s) + sf_r\sigma_s + \sigma_s b/n, \quad (1)$$

where  $n$  is the size of the group.<sup>7</sup> Then Selfish agents select a fitness-maximizing  $\sigma_s^*$ , namely, one that minimizes the cost of working (1). Assuming an interior solution, this is given by

$$g'(\sigma_s^*) = \lambda'(1 - \sigma_s^*) + f_r s + b/n = 0, \quad (2)$$

requiring the shirker to equate the marginal fitness benefits of expending less effort on work (the first and third terms) with the marginal costs of greater shirking, namely the increased likelihood of bearing the fitness cost of ostracism (the second term). Since  $\lambda''(1 - s_s) > 0$ , there is at most one solution to this equation, and it is a minimum. This first order condition shows that Selfish agents who inherit a large  $s$  (i.e., who believe the cost of ostracism is very high) will shirk less.

The expected contribution of each group member to the group's population in the next period is equal to the member's fitness minus the likelihood of ostracism. This gives

$$\pi_s = (1 - \sigma)b - \lambda(1 - \sigma_s) - sf_r\sigma_s, \quad (3)$$

$$\pi_c = (1 - \sigma)b - c, \quad (4)$$

$$\pi_r = (1 - \sigma)b - c - c_p(1 - f_r - f_c)\sigma_s, \quad (5)$$

where the subscripts  $s$ ,  $c$ , and  $r$  refer to Selfish agents, Cooperators, and Reciprocators. The final term in the expression for  $\pi_r$  follows because each Reciprocator chooses a random agent to monitor; this agent is Selfish with probability  $(1 - f_r - f_c)$  and this agent shirks with probability  $\sigma_s$ ; thus the expected cost to the Reciprocator is  $c_p(1 - f_r - f_c)\sigma_s$ .

We assume that at the end of each period, ostracized members join a common pool of solitary individuals, and groups admit a number of new members equal to a fraction  $\mu$  of their existing numbers. Candidates for immigration into groups are the pool of solitary individuals, plus a fraction  $\gamma$  of current group members who want

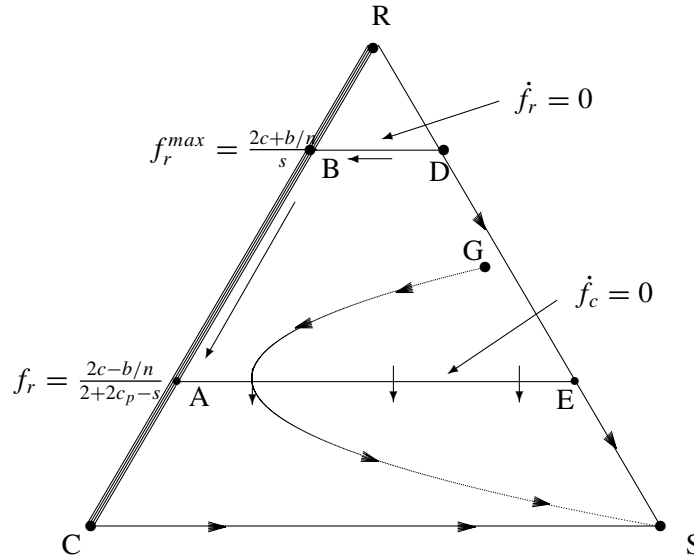
<sup>7</sup>We obtain the second term by assuming each Reciprocator randomly chooses one agent to monitor in each period. Then the probability of begin ostracized is

$$\sigma_s \left[ 1 - (1 - 1/(n - 1))^{f_r n} \right].$$

We approximate this by  $\sigma_s f_r$ , for simplicity. The approximation is very good for  $f_r < 0.4$ , but for higher  $f_r$ , ours is an overestimate. For instance, when  $f_r = 1$ , the actual cost is  $0.66s\sigma_s$ , but our estimate is  $s\sigma_s$ . This is a harmless but useful approximation, since when  $f_r > 0.4$ , there is always universal cooperation in our simulations, so Selfish agents rarely pay the cost of being ostracized.

to emigrate for exogenous reasons (e.g., to find a mate, or to end a personal dispute with another group member). If the number of candidates exceeds the number of places, a random sample of these candidates emigrate to groups. We assume that group members that desired to emigrate but did not find a receptive group remain in their current group.

The number of groups and the total population is fixed throughout. A simulation starts with all groups of the same size, which we call the *initial group size*. Individual group size will, of course, change from period to period. We assume that if a group falls below some  $n_{\min}$  in size, it disbands, the remaining members migrating to the pool. We further assume that the vacated site is repopulated by migrants randomly chosen from the most populous remaining groups, so that the initial group size is restored.



**Figure 1:** Within Group Dynamics with Ostracism but No Migration. This diagram is based on  $s > 2c$  (the cost of being ostracized is greater than twice the cost of working for one period). The value of  $f_r$  for which  $\dot{f}_r = 0$  and the value of  $f_r$  for which  $\dot{f}_c = 0$  assume the particular cost of effort function (6).

A phase diagram of within-group dynamics, abstracting from ostracism and group dissolution, appears in Figure 1. In this figure, each point in the simplex is a distribution of behavioral types in the population. The vertex S refers to the all-shirking group composition ( $f_r = f_c = 0$ ), vertex R refers to the all-Reciprocator composition ( $f_r = 1$ ), and C refers to the all-Cooperator case ( $f_c = 1$ ). For the parameter values illustrated in the figure, there is some fraction of Reciprocators

(the line BD) at and above which Selfish agents do not shirk at all, but below which they shirk at a strictly positive rate. Thus in the triangle RBD all three types have equal payoffs, equal to  $b - c$ . Along the whole CR segment, Reciprocators and Cooperators do equally well, since there are no Selfish agents to punish (the payoffs are again  $b - c$ ). Along the CS segment, Reciprocators are absent, so Selfish agents do better than Cooperators. On the segment DS, Selfish agents do better than Reciprocators. Selfish agents optimize, so when  $\sigma_s \in (0, 1)$ , we know from (2) that an increase in  $f_s$  holding the frequency of Selfish agents constant, must entail a decline in  $\sigma_s$ . Thus lowering the fraction of Selfish agents increases the payoff to Reciprocators relative to Selfish agents in the area CBDS. Moreover, Cooperators always have higher payoffs than Reciprocators in the interior of this area. We conclude that the only asymptotically stable equilibrium of the system is the all-Selfish point S, and its basin of attraction consists of all interior points below the line BD in Figure 1.

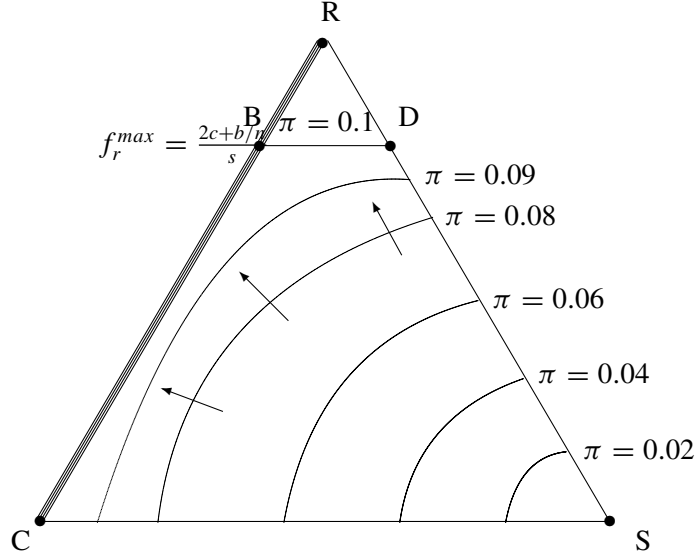
Clearly, then, if cooperation is to be sustained, it must be because Cooperators, who undermine the cooperative equilibrium by driving out Reciprocators, must themselves be harmed by shirking Selfish agents when Reciprocators are rare. Figure 1 shows that for suitable parameter values exactly this process will obtain, with Selfish agents proliferating at the expense of both Cooperators and Reciprocators once the frequency of Reciprocators falls below the line AE in the figure. We shall show by simulation that this is indeed the case for a wide range of plausible parameter values.

We can gain more insight into the dynamics of this system by choosing a specific function  $\lambda(1 - \sigma)$  satisfying the conditions  $\lambda(1) = c$ ,  $\lambda(0) = 0$ ,  $\lambda'(1 - \sigma) < 0$ , and  $\lambda''(1 - \sigma) > 0$ . Extensive simulations suggest that the exact form of this function is unimportant, and the simplest function satisfying these conditions is

$$\lambda(1 - \sigma) = c(1 - \sigma)^2. \quad (6)$$

Given this function, and using appropriate simulation parameters, Figure 2 shows the group average payoffs as a function of the distribution of types within a group. The curved lines are iso-group-average-fitness loci, showing clearly that average group fitness increases as we move away from the unique stable equilibrium S.

In the next section we consider the evolution of a number of groups, each occupying a site allowing cooperative production as described above. To this we add a pool of ostracized agents who work alone and in each period seek entrance to an established group. We allow for migration across groups, and the dissolution of groups which fall below a minimum group size. The site freed up by the dissolution of a group is repopulated by members of successful groups with large numbers of members. We also allow for mutation among the various types of agents. The



**Figure 2:** Average Payoffs and Group Composition. This diagram is based on  $b = 0.2$ ,  $c = c_p = 0.1$ , and  $s = 0.3$ . The arrows point towards increasing payoffs. For all frequencies in BDR,  $\pi = 0.1$

resulting model is too complex to admit an analytical, closed form, solution, so in we provide a series of simulations that illustrate its characteristics over an appropriate range of parameter values.

### 3 Simulating Strong Reciprocity

For our baseline simulation, we set up twenty groups, each starting out with twenty members, and an empty pool. We set the initial frequency of Selfish agents at 100%, and we assigned each Selfish agent a cost of being ostracized ( $s$ ), using a random number drawn from a uniform distribution on the interval  $[0, 1]$ . Using (6), Selfish agents choose  $s_s$  to minimize the total cost

$$\lambda(1 - \sigma_s) + f_r s \sigma_s + \sigma_s b/n.$$

It is easy to check that Selfish agents then shirk according to the function

$$\sigma_s(f_r) = \begin{cases} 1 - \frac{f_r sn - b}{2cn} & \text{for } f_r \leq f_r^{\max} = \frac{2cn + b}{sn} \\ 0 & \text{for } f_r > f_r^{\max}. \end{cases} \quad (7)$$

where  $c$  is the cost of working. We assume that immigration into a group occurs at a rate of 12% per generation, which we take to be four simulation periods. Therefore the immigration rate is  $\mu = 0.03$ .

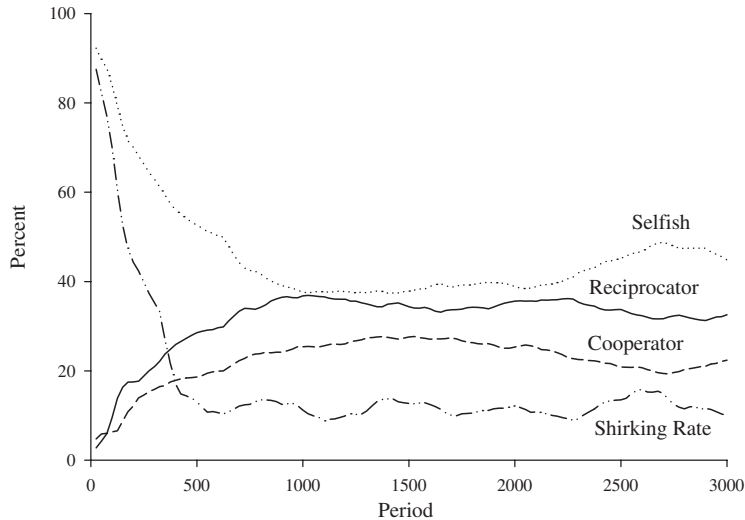
We assume a desired emigration rate of  $\gamma = 0.05$  per period, so more agents would like to emigrate, either from the pool or from other groups, than are admitted into groups. We set the fitness cost of being in the pool to be equal to the cost of working (as Figure 5 shows, this figure is not at all critical). We also set individual productivity, the gain to the group when an individual does not shirk, equal to the cost of punishing a shirker. Simulations show that it matters little how large this parameter is, and we set it to 0.1 in our baseline simulations. Since we assume the cost of working is  $c = 0.1$ , the cost of punishing becomes  $c_p = 0.1$  and individual productivity is  $b = 0.2$ . The reason absolute productivity does not matter is that we cull or augment the population randomly at the end of each period to ensure that the total population size is constant. All fitness figures are therefore relative. The baseline parameters are listed in Table 1.

Value	Description
0.2	Output per Agent, no Shirking ( $b$ )
0.1	Cost of Working, no Shirking ( $c$ )
0.1	Cost of Punishing ( $c_p$ )
0.05	Emigration Rate ( $\gamma$ )
0.03	Immigration Rate ( $\mu$ )
20	Initial Group Size ( $n$ )
20	Number of Groups
-0.1	Fitness in Pool ( $-\phi_0$ )
6	Minimum Group Size
[0, 1]	Initially Seeded Expected Cost of Ostracism ( $s$ ) Uniformly Distributed on this Interval
0.01	Mutation Rate ( $\epsilon$ )

**Table 1:** Baseline Parameters. These parameters are used in all simulations, unless otherwise noted. All simulations start with a homogeneous population of Selfish agents with a uniform distribution of expected cost of ostracism.

Using these parameters, Figure 3 shows the evolution of the distribution of agent types and the average shirking rate for the whole population on a typical run. The results shown are backward-moving averages over 100 periods. To determine the typical behavior of the model, we ran the simulation 25 times with the baseline parameters for 50,000 periods, and calculated the average fraction of each type and the average shirking rate, averaged over the last 1000 periods. These averages are reported in Table 2. There was remarkably little variation across the runs—in all cases the standard error of the frequencies reported is less than 1.14 percentage points

Why does the long-run behavior of the system involve roughly equal numbers of the three types of agents? Table 2 show that over this period, Cooperators were



**Figure 3:** Single Simulation Run. As is clear from the figure, and as in all our simulations, the initial population consisted of Selfish agents alone. The baseline parameters are as in Table 1.

slightly more likely (0.48%) than Reciprocators (0.38%) to find themselves in the pool of solitary agents, while Selfish agents were more than an order of magnitude more likely to be in this position (10%). This is in part because Cooperators are more likely than Reciprocators to be in disbanded groups (by a ratio of 1.21 to 1), and the Selfish agents are more likely than Reciprocators to be in disbanded groups (by a ratio of 3.4 to 1). But ostracism, not disbanding, is overwhelmingly important in populating the pool with Selfish agents. Indeed, we have found that even when groups are disbanded when only 1 agent is left in the group (at which time it is no better than solitary production), similar long-run values of the major variables obtain. Thus, the dispersion of members of very small groups, while empirically realistic, is not crucial to the model's workings.

To illustrate that the dynamics depicted in Figures 1 and 2 are operative, in Figures 4 we trace a bit of the history of a single group, from “birth” in period 320 (repopulation of a site after a group having disbanded) to “death” in period 390 (disbanding because group size fell below the minimum sustainable group size). This group began with roughly equal numbers of the three agent types, but Selfish agents were driven out by the Reciprocators and were replaced by Cooperators. Starting in period 330, Cooperators began displacing Reciprocators until, about

Long-run Values	
Value	Description
37.2%	Fraction of Reciprocators
24.6%	Fraction of Cooperators
38.2%	Fraction of Selfish Agents
11.1%	Average Shirking Rate
4%	Fraction of Population in Pool
0.38%	Fraction of Reciprocators in Pool
0.48%	Fraction of Cooperators in Pool
10%	Fraction of Selfish agents in Pool
4%	Fraction of Pool who are Reciprocators
3%	Fraction of Pool who are Cooperators
93%	Fraction of Pool who are Selfish
1.21	Ratio of Cooperators to Reciprocators in Disbanded Groups
3.4	Ratio of Selfish Agents to Reciprocators in Disbanded Groups

**Table 2:** Long Run Simulation Statistics. The baseline parameters are as in Table 1.

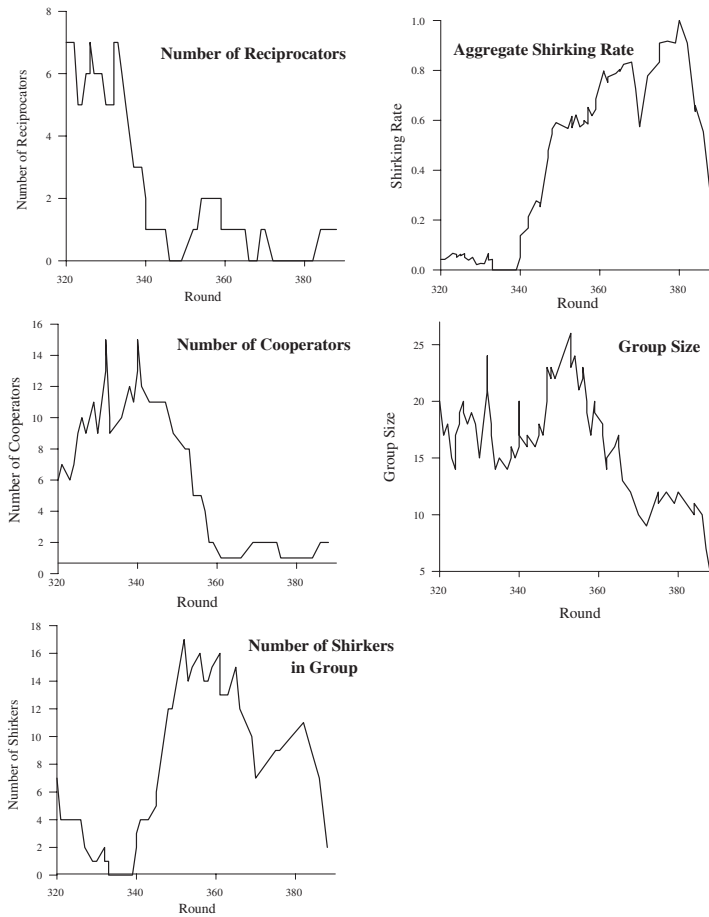
period 340, the group consisted of many Cooperators and a few Selfish agents. With few Reciprocators present, the expected cost of shirking fell dramatically, and the Selfish therefore outperformed the Cooperators. The numbers of Selfish therefore grew rapidly from period 340 to 350, displacing Cooperators. This process reproduces the dynamic illustrated by the curved arrow QS in Figure 1. From this point, average group fitness was low (the group composition placed it close to the S vertex in Figure 2) and hence, despite a small chance infusion of immigrant Reciprocators, the group loses members until it disbands in period 390. Figure 4 also show the effect of these demographic movements on the shirking rate and group size. We see that the shirking rate remains very low until period 340, and then climbs steadily until just prior to the end of the group's life. Near the end there is an increase in group fitness because a couple of Reciprocators migrate into the group and the Selfish respond by shirking less, but the cost of increasing group fitness, ostracism, kills the group before it can recover.

The first three panels of Figure 5 show that the model responds in the expected way to changes in parameter values, and that high levels of cooperation are sustained for a quite large parameter space. These panels show the average shirking value, plotted against the fitness of agents per period in the pool of solitary agents, the cost to Reciprocators of punishing shirkers, and initial group size. These figures are the average of the last 1000 periods of ten runs of 30,000 periods. In each case, we see that the shirking rate, which is the best aggregate measure of group cooperation,

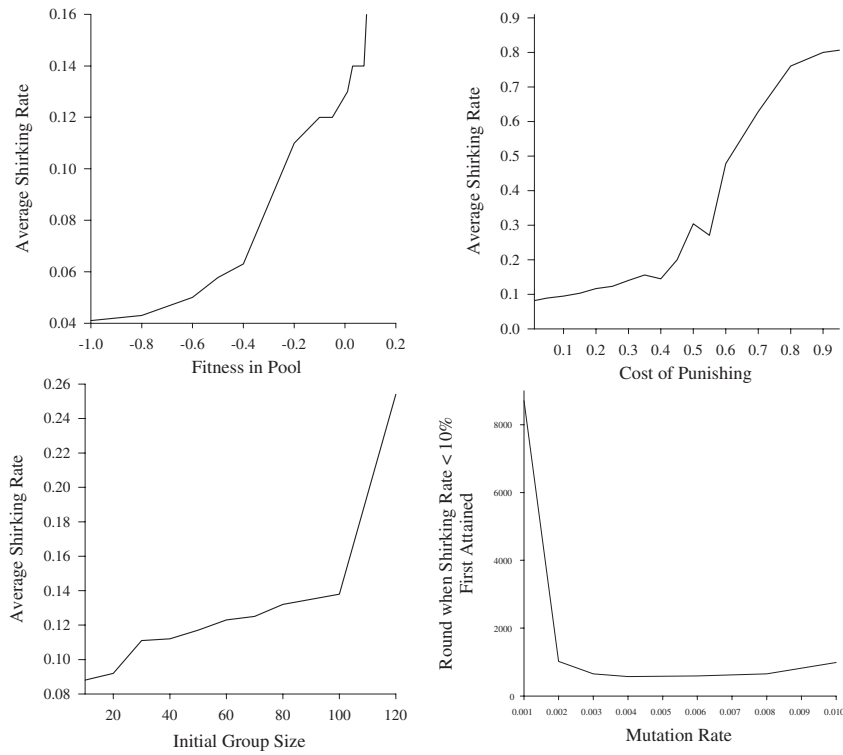
is small or moderate in size over a large range of parameters, and moves in the expected direction with parameter changes. Our results are quite insensitive to variations in other parameters, including the immigration rate, the emigration rate, and minimum initial group size. In each case, we find cooperation to be robust over a wide range of parameters.

Our simulations make it clear that the ostracism mechanism promotes high levels of cooperation in groups and that a substantial fraction of the population are Reciprocators in the long term steady state of this model under a wide range of parameter values. But how could Reciprocators come into existence *de novo*? We speculate that it could have emerged through a rather trivial modification of fitness-enhancing behaviors. For example, a good case could be made that strong reciprocity among kin could emerge and proliferate through as a form of kin-based altruism, and then be generalized to unrelated individuals. Another possibility is that strong reciprocity arose through a modification of the individual fitness enhancing strategy of reciprocal altruism. In this case the modification is trivial: simply ignore the future payoffs to current behavior. Like the extension of kin-based strong reciprocity to non-kin, the mutation or mutations to convert reciprocal altruist strategy to a strong reciprocal one involves a reduction in complex discrimination rather than an increase in complexity. These strategy conversions thus might occur with high probability.

As we have seen, the process of mutation we have assumed is sufficient for the emergence of strong reciprocity after as few as 500 periods. This occurs because at the rate of mutation assumed, it does not take many periods before at least one group will have enough Reciprocators to implement a high level of cooperation. When this occurs, the cooperative group grows in size, and as a result it seeds other groups by migration and repopulates the sites of disbanded groups. The rate at which this process takes place obviously depends on the number of groups in the population. In the fourth panel of Figure 5 we present data in a population with 50 groups. A high level of cooperation is achieved (on the average) after about 1500 periods, for mutation rates as low as  $\epsilon = 0.002$ . Simulations with just 20 groups, by contrast, show that for mutation rates lower than 0.002, it takes on the average 30,000 periods before the shirking rate first falls below ten percent. We think that our simulations with 50 groups probably overstate the obstacles to the emergence of strong reciprocity for the simple reason that in order to proliferate the behavior need only emerge once, and there were far more than 50 bands of early foraging humans.



**Figure 4:** Snapshot of the Life History of a Single Group, from Repopulation of a Site to Disbanding: Membership Composition, Size and Efficiency. The baseline parameters are as in Table 1.



**Figure 5:** The Effect of the assumed values of Fitness in Pool, Cost of Punishing, Initial Group Size, and the Mutation Rate on Shirking. The baseline parameters are otherwise as in Table 1. For initial group size larger than about 120, Reciprocators rarely invade an initially Selfish population. The Mutation Rate panel shows the average over 25 runs of the number of periods until the shirking rate, averaged over previous 1000 periods, first fell below ten percent.

## 4 The Evolution of Strong Reciprocity

Can this model illuminate a process by which strong reciprocity might have become common in human populations? Do the interactions modeled here capture the relevant aspects of the social and physical environments of *Homo sapiens sapiens* during the 50,000 or so years prior to the advent of agriculture?<sup>8</sup> To answer this question we turn to recent and contemporary accounts of societies generally thought to resemble the foraging bands that were common during this period, among them the !Kung of Botswana and Namibia, the Ache of Paraguay, Batek of Malaysia, Hadza of Tanzania, Pandaram and Paliyan of South India, the Inuit of the Northwest territories, and the Mbuti Pygmies of Zaire. On the basis of this reading, we believe that our model may be illuminating.<sup>9</sup>

There is evidence that in some contemporary simple societies the lazy and the stingy are punished. Balikci (1970):177 reports the following concerning the Netsilik, an isolated tribe of Arctic hunters living on the Arctic coast:

...there is a general rule...according to which all able bodied men should contribute to hunting, and the returns of the hunt should be shared according to established custom. Any activity in exception to this rule was bound to provoke criticism, various forms of conflict, and frequently social ostracism. (176)...lazy hunters were barely tolerated by the community. They were the objects of back biting and ostracism...until the opportunity came for an open quarrel. Stingy men who shared in a niggardly manner were treated similarly. (177)

And Lee (1979):458 reports that

The most serious accusations one !Kung can level against another are the charge of stinginess and the charge of arrogance. To be stingy, or far-hearted, is to hoard one's goods jealously and secretively, guarding them "like a hyena." The corrective for this is to make the hoarder give "till it hurts"; that is to make him give generously and without stint until everyone can see that he is truly cleaned out. In order to ensure compliance with this cardinal rule the !Kung browbeat each other constantly to be more generous and not to hoard.

<sup>8</sup>This is the time span of anatomically modern humans reported by Klein (1989):344. Foley's (1987):22 estimate is 100,000 years. The horticultural societies that eventually replaced foraging bands almost everywhere appeared 12-10,000 years ago. Even Klein's lower limit for the appearance of modern humans leaves ample time for significant change in gene distributions to have taken place under the kinds of selection pressures at work.

<sup>9</sup>Our main sources are listed in footnote 2. The difficulty in making inferences about simple societies during the late Pleistocene on the basis of contemporary simple societies is stressed by Foley (1987):75-78 and Kelly (1995). See Boehm (2002) for an opposing view.

Lethal violence among the !Kung is quite high, so the costs of these conflicts must sometimes be borne by those seeking to uphold norms of sharing (Lee 1979).<sup>10</sup> More extensive evidence of punishment of norm violators is provided by Christopher Boehm's (1993) survey of the many studies in this area.

...intentional leveling linked to an egalitarian ethos is an immediate and probably an extremely widespread cause of human societies' failing to develop authoritative or coercive leadership. (226)

Bruce Knauft (1991):393,395 adds:

In all ethnographically known simple societies, cooperative sharing of provisions is extended to mates, offspring, and many others within the band. ...This sharing takes place well outside the range of immediate kin, viz. among the diverse array of kin and non-kin who constitute the typical residence group of 25+ persons. Archeological evidence suggests that widespread networks facilitating diffuse access to and transfer of resources and information have been pronounced at least since the Upper Paleolithic...The strong internalization of a sharing ethic is in many respects the *sine qua non* of culture in these societies.

Using data from forty-eight surviving simple societies, Boehm (1993):228 concluded that

the primary and most immediate cause of egalitarian behavior is a moralistic determination on the part of a local group's main political actors that no one of its members should be allowed to dominate the others.

Boehm further sought to determine whether intentional behavior (notably, social sanctioning) that had a leveling effect was widespread in such societies and more specifically whether it had any significant effects in suppressing the growth of authoritarian leadership. He found evidence that arrogant members of the group are constrained by public opinion, criticism and ridicule, disobedience, and extreme sanction:

...assassination is reported in 11 out of the 48...behaviors that terminated relations with an overly assertive individual or removed him from a leadership role involved 38 of the 48 societies, while in an additional

---

<sup>10</sup>Most of the lethal violence documented by Lee concerns such problems as adultery, not stinginess. By contrast to the reports of Lee and Balikci, Endicott (1988):118 reports horror expressed by a Batek informant at the thought of exiling a member whose laziness had caused some resentment.

28 instances the person was manipulated by social pressure...the great majority of these misbehaviors involve dominance or self-assertion. (231)

among simple foragers, ...group execution of overassertive persons seems to be rather frequent. (239)

We have modelled punishment simply as ostracism from the group. But in the ethnographic record it takes several forms, including group fissioning to minimize interacting with Selfish agents and the withdrawal of cooperation from Selfish agents who remain co-resident. Those who have violated a norm may also leave a group in anticipation of more serious punishment. Extensions of the model to include these forms of punishment are straightforward.

Our model is also readily extended to other aspects of the enforcement of norms that have a public goods character. To see this, suppose there is norm that restricts copulations to monogamous pairs, which when violated leads to strife within a group or lessens its effectiveness in acquiring food, insuring against stochastic events, sharing information, or defending itself, all of which reduce fitness levels of group members. Those who violate the norm, however, enhance their fitness. Reciprocators never violate the norm and punish those who do. Cooperators adhere to the monogamy norm but never punish. Selfish agents violate the norm when the expected cost of punishment is sufficiently low. This model could easily be cast in the form that we have just simulated. We therefore believe the model as we have developed it is applicable to a wide range of concrete problems of norm adherence likely to arise in small stateless groups.

Our reading of the ethnographic and paleoanthropological evidence is that our model may capture the salient social and ecological conditions of the late Pleistocene.

## 5 Discussion

Explanations of the evolution of cooperation among unrelated humans sometimes fail to explain why similar behaviors are seldom observed in other animals. Our model, however, relies on cognitive, linguistic, and other capacities unique to our species (Gintis 2003). The *moralistic aggression* (Robert Trivers apt term) that provides the motivational underpinnings of altruistic punishment requires uniquely human cognitive and linguistic abilities in the formulation of behavioral norms, the achievement of group consensus that the norms ought to be followed, the communication of their violations, and the coordination of the often collective nature of the punishment of miscreants. Additionally, uniquely human capacities to inflict

punishment at a distance, through projectile weapons, reduced the cost of ostracizing a norm violator. If our speculation that strong reciprocity emerged through a modification of reciprocal altruist behaviors is correct, this provides another reason why strong reciprocity might be uniquely human, given that reciprocal altruism appears to be very rare, if not completely absent, in other species (Clements and Stephens 1995, Connor 1995, Dugatkin and Mesterton-Gibbons 1996, Dugatkin 1996, J.Godin and Davis 1995, Milinski 1996, Clutton-Brock, O’Riain, Brotherton, Gaynor, Kansky, Griffin and Manser 1999).

Our proposed explanation of human cooperation contrasts with the more standard interpretation stressing *reciprocal altruism* (Trivers 1971, Axelrod and Hamilton 1981). The canonical status of this view notwithstanding, there is little evidence that cooperation in prisoners’ dilemma type situations among non-human animals is explained by the opportunities for inflicting costs on non-cooperators offered by repeated interactions (Stephens, McLinn and Stevens 2002). Among humans however, we do not doubt the importance of repeated interactions and other structures that reward cooperators with higher fitness or other payoffs, rendering seemingly selfish acts a form of mutualism. While an important part of the explanation of human cooperation, there are several reasons for doubting the adequacy of this explanation. First, reciprocal altruism fails when a social group is threatened with dissolution, since members who sacrifice now on behalf of group members do not have a high probability of being repaid in the future (Gintis 2000). Second, many human interactions in the relevant evolutionary context took the form of  $n$ -person public goods games—food sharing and other co-insurance, as well as common defense—rather than dyadic interactions. Even if repeated with high probability,  $n$ -person public goods (or common pool resource) interactions make cooperation difficult to sustain by means of the standard tit-for-tat and other reciprocal behaviors.<sup>11</sup> Third, the contemporary study of human behavior has documented a large class of prosocial behaviors inexplicable in terms of reciprocal altruism. For instance, there is extensive support for income redistribution in advanced industrial economies, even among those who cannot expect to be net beneficiaries (Fong, Bowles and Gintis 2002). Under some circumstances group incentives for large work teams are effective motivators even when the opportunity for reciprocation is absent and the benefits of cooperation are so widely shared that a self-interested group member would gain from free-riding on the effort of others (Ghemawat 1995, Hansen 1997, Knez and Simester in press). Finally, laboratory and field experiments show that non-selfish motives are frequently robust predictors of behavior, even in one-shot, anonymous

<sup>11</sup>See Boyd and Richerson (1988) and Joshi (1987). Alexander (1987) has proposed a more general “indirect reciprocity” mechanism more amenable to large group interactions, and this has been formalized by Nowak and Sigmund (1998). However, for reasons given by Leimar and Hammerstein (2001), we doubt that indirect reciprocity is a general explanation of human sociality.

setting. This research has been summarized in Ostrom (1998) and Fehr and Gächter (2000) for industrial societies, and Henrich, Boyd, Bowles, Camerer, Fehr, Gintis, and McElreath (2001) for less-developed societies, including hunter-gatherers. This evidence bids us consider a broader range of human motivations.

Our model differs from other explanations of cooperation among unrelated individuals in several ways. Most models of reciprocity treat interactions among pairs of agents (Boorman and Levitt 1980, Axelrod and Hamilton 1981, Kreps, Milgrom, Roberts and Wilson 1982, Axelrod 1984). Since strong reciprocity is exhibited in such collective situations as group food-sharing and defense, these models not suited to explaining this phenomenon. By contrast, we model  $n$ -agent groups (where  $n$  is on the order of ten to 100) in a series of production periods that are effectively one-shot, since the only inter-period influences are those involving the biological and cultural reproduction of new agents. Moreover, in contrast to other models of cooperation in groups, (Robson 1990, Nowak, Page and Sigmund 2000, Wedekind and Milinski 2000), we assume Reciprocators cannot gain from being phenotypically identified as such, or by establishing a reputation for reciprocation across production periods.<sup>12</sup>

Nor in our model can Reciprocators use their altruistic behavior as a costly signal of superior fitness (Zahavi 1975, Bliege Bird, Smith and Bird 2001, Gintis et al. 2001). Finally, while most models of strong reciprocity depend on group extinctions (Gintis 2000, Boyd et al. 2003, Bowles, Choi and Hopfensitz in press), ours does not.<sup>13</sup>

Our approach is related to the model of Aviles, Abbot and Cutter (2002), as applied to tree-killing bark beetles (Raffa and Berryman 1987) and other species in which there are strong fitness benefits associated with social living. They posit a minimum group size and positive fitness effects of group size for smaller groups. While these conventional Allee effects play no role in our model, they are approximated by our assumption living in cooperative groups confers fitness benefits and that solitary individuals bear fitness costs. Like Trivers (1971), Hirshleifer and Rasmusen (1989), Boyd and Richerson (1992) Sethi and Somanathan (1996), and Friedman and Singh (2001), we stress the importance of altruistic punishment. Hirshleifer and Rasmusen (1989) and Friedman and Singh (2001) develop models of

<sup>12</sup>Even though our stage game is a one-shot, it may in fact involve behaviors that take place over several, or even many years (e.g., a hunting season, of which one period in our model may comprise several). Our treatment of being “detected shirking” is compatible with individuals’ building reputations during the course of the game that at some point trigger punishment. Our “probability of being detected shirking” is a summary description of this process, which may occur over time even in a one-shot game.

<sup>13</sup>In our simulations, when a group falls below a certain minimum size, it disbands and is repopulated by agents randomly chosen from successful groups, but our results hold when this minimum size is as small as one.

team production in which the threat of ostracism deters shirking. But, because they assume that ostracizing is not costly to the individual ostracizer, their models, unlike ours, do not explain the persistence of altruistic behaviors. Our paper is most closely related to Sethi and Somanathan (1996). Moreover, in Sethi and Somanathan's model the equilibrium frequency of Reciprocators is zero and cooperation is complete. By contrast, our model supports a positive (indeed, quite high) fraction of Reciprocators and a significant level of non-cooperation in the long run.

We think that our model, suitably extended, can capture the environments that may have supported high levels of cooperation among our ancestors living in mobile foraging bands during the late Pleistocene. We do not know that a human predisposition to strong reciprocity evolved as we have described. But our simulations suggest that it could have. Our results do not suggest that these genetic influences are sufficient to account for the forms and extent of altruistic punishment and other evidence of strong reciprocity that appear to be common in modern human populations. The predispositions to punish norm violators also evolved under the influence of cultural transmission as did the content of the relevant norms.

#### REFERENCES

- Alexander, R. D., *The Biology of Moral Systems* (New York: Aldine, 1987).
- Andreoni, James and John H. Miller, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica* 70,2 (2002):737–753.
- Aviles, Leticia, Patrick Abbot, and Asher D. Cutter, "Population Ecology, Nonlinear Dynamics, and Social Evolution. I. Associations among Nonrelatives," *The American Naturalist* 159,2 (February 2002):115–128.
- Axelrod, Robert, *The Evolution of Cooperation* (New York: Basic Books, 1984).
- and William D. Hamilton, "The Evolution of Cooperation," *Science* 211 (1981):1390–1396.
- Balikci, Asen, *The Netsilik Eskimo* (New York: Natural History Press, 1970).
- Binford, Lewis, *Constructing Frames of Reference: An Analytical Method for Archeological Theory Using Hunter-Gatherer and Environmental Data Sets* (Berkeley, CA: University of California Press, 2001).
- Bliege Bird, Rebecca L., Eric A. Smith, and Douglas W. Bird, "The Hunting Handicap: Costly Signaling in Human Foraging Strategies," *Behavioral Ecology and Sociobiology* 50 (2001):9–19.
- Blurton Jones, Nicholas G., "Tolerated Theft: Suggestions about the Ecology and Evolution of Sharing, Hoarding, and Scrounging," *Social Science Information* 26,1 (1987):31–54.

- Boehm, Christopher, "The Evolutionary Development of Morality as an Effect of Dominance Behavior and Conflict Interference," *Journal of Social and Biological Structures* 5 (1982):413–421.
- , *Blood Revenge: The Enactment and Management of Conflict in Montenegro and Other Tribal Societies* (Philadelphia, PA: University of Pennsylvania Press, 1984).
- , "Egalitarian Behavior and Reverse Dominance Hierarchy," *Current Anthropology* 34,3 (June 1993):227–254.
- , *Hierarchy in the Forest: The Evolution of Egalitarian Behavior* (Cambridge, MA: Harvard University Press, 2000).
- , "Variance Reduction and the Evolution of Social Control," 2002. Department of Anthropology, University of Southern California.
- Boorman, Scott A. and Paul Levitt, *The Genetics of Altruism* (New York: Academic Press, 1980).
- Bowles, Samuel, Jung-kyoo Choi, and Astrid Hopfensitz, "The Co-evolution of Individual Behaviors and Social Institutions," *Journal of Theoretical Biology* (in press).
- Boyd, Robert and Peter J. Richerson, "The Evolution of Reciprocity in Sizable Groups," *Journal of Theoretical Biology* 132 (1988):337–356.
- and —, "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizeable Groups," *Ethology and Sociobiology* 113 (1992):171–195.
- , Herbert Gintis, Samuel Bowles, and Peter J. Richerson, "Evolution of Altruistic Punishment," *Proceedings of the National Academy of Sciences* 100,6 (March 2003):3531–3535.
- Cashdan, Elizabeth A., "Egalitarianism among Hunters and Gatherers," *American Anthropologist* 82 (1980):116–120.
- Chagnon, Napoleon A., *Yanomamö: The Fierce People* (New York: Holt, Rinehart and Winston, 1977).
- Clements, Kevin C. and David W. Stephens, "Testing Models of Non-kin Cooperation: Mutualism and the Prisoner's Dilemma," *Animal Behaviour* 50 (1995):527–535.
- Clutton-Brock, T. H., J. J. O'Riain, P. N. M. Brotherton, D. Gaynor, R. Kansky, A. S. Griffin, and M. Manser, "Selfish Sentinels in Cooperative Mammals," *Science* 284 (1999):1640–1644.
- Connor, Richard C., "The Benefits of Mutualism: a Conceptual Framework," *Biological Reviews* 70 (1995):427–457.
- Damas, David, "Central Eskimo Systems of Food Sharing," *Ethnology* 11,3 (1972):220–240.

- Dugatkin, Lee Alan, "Tit for Tat, Byproduct Mutualism and Predator Inspection: A reply to Connor," *Animal Behavior* 51 (1996):455–457.
- and Michael Mesterton-Gibbons, "Cooperation Among Unrelated Individuals: Reciprocal Altruism, Byproduct Mutualism, and Group Selection in Fishes," *Biosystems* 37 (1996):19–30.
- Endicott, Kirk, "Property, Power and Conflict among the Batek of Malaysia," in T. Ingold, D. Riches, and J. Woodburn (eds.) *Hunters and Gatherers* (New York: St. Martin's Press, 1988) pp. 110–127.
- Fehr, Ernst and Simon Gächter, "Cooperation and Punishment," *American Economic Review* 90,4 (September 2000):980–994.
- and —, "Altruistic Punishment in Humans," *Nature* 415 (10 January 2002):137–140.
- Foley, Robert, *Another Unique Species: Patterns in Human Evolutionary Ecology* (New York: John Wiley and Sons, 1987).
- Fong, Christina M., Samuel Bowles, and Herbert Gintis, "Reciprocity and the Welfare State," in Jean Mercier-Ythier, Serge Kolm, and Louis-André Gérard-Varet (eds.) *Handbook on the Economics of Giving, Reciprocity and Altruism* (Amsterdam: Elsevier, 2002).
- Friedman, Daniel and Nirvikar Singh, "Evolution and Negative Reciprocity," in Y. Aruka (ed.) *Evolutionary Controversies in Economics* (New York: Springer, 2001).
- Ghemawat, Pankaj, "Competitive Advantage and Internal Organization: Nucor Revisited," *Journal of Economic and Management Strategy* 3,4 (winter 1995):685–717.
- Gintis, Herbert, "Strong Reciprocity and Human Sociality," *Journal of Theoretical Biology* 206 (2000):169–179.
- , "The Hitchhiker's Guide to Altruism: Genes, Culture, and the Internalization of Norms," *Journal of Theoretical Biology* 220,4 (2003):407–418.
- , Eric Alden Smith, and Samuel Bowles, "Costly Signaling and Cooperation," *Journal of Theoretical Biology* 213 (2001):103–119.
- Hansen, Daniel G., "Individual Responses to a Group Incentive," *Industrial and Labor Relations Review* 51,1 (October 1997):37–49.
- Hawkes, Kristen, "Sharing and Collective Action," in E. A. Smith and B. Winterhalder (eds.) *Evolutionary Ecology and Human Behavior* (New York: Aldine de Gruyter, 1992) pp. 269–300.
- , "Why Hunter-Gatherers Work: An Ancient Version of the Problem of Public Goods," *Current Anthropology* 34,4 (1993):341–361.

- Henrich, Joe, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis, *Foundations of Human Sociality: Ethnography and Experiments in Fifteen Small-scale Societies* (Oxford: Oxford University Press, 2003).
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath, "Cooperation, Reciprocity and Punishment in Fifteen Small-scale Societies," *American Economic Review* 91 (May 2001):73–78.
- Hirshleifer, David and Eric Rasmusen, "Cooperation in a Repeated Prisoners' Dilemma with Ostracism," *Journal of Economic Behavior and Organization* 12 (1989):87–106.
- J.Godin, Jean-Guy and Scott A. Davis, "Who Dares, Benefits: Predator Approach Behaviour in the Guppy *Poecilia reticulata* Deters Predator Pursuit," *Proceedings of the Royal Society of London B* 259 (1995):193–200.
- Joshi, N. V., "Evolution of Cooperation by Reciprocation Within Structured Demes," *Journal of Genetics* 66,1 (1987):69–84.
- Kaplan, Hillard and Kim Hill, "Food Sharing among Ache Foragers: Tests of Explanatory Hypotheses," *Current Anthropology* 26,2 (1985):223–246.
- and —, "Hunting Ability and Reproductive Success among Male Ache Foragers: Preliminary Results," *Current Anthropology* 26,1 (1985):131–133.
- , —, Kristen Hawkes, and Ana Hurtado, "Food Sharing among Ache Hunter-Gatherers of Eastern Paraguay," *Current Anthropology* 25,1 (1984):113–115.
- , Michael Gurven, Kim Hill, and A. Magdalena Hurtado, "The Natural History of Human Food Sharing and Cooperation: A Review and a New Multi-Individual Approach to the Negotiation of Norms," in Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr (eds.) *The Moral Sentiments: Origins, Evidence, and Policy* (???: ??, 2003) pp. —.
- Kelly, Robert L., *The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways* (Washington, DC: The Smithsonian Institution, 1995).
- Kent, Susan, "And Justice for All: The Development of Political Centralization among Newly Sedentary Foragers," *American Anthropologist* 93,1 (1989):703–712.
- Klein, Richard G., *Human Career: Human Biological and Cultural* (Chicago: University of Chicago Press, 1989).
- Knauff, Bruce, "Sociality versus Self-interest in Human Evolution," *Behavioral and Brain Sciences* 12,4 (1989):12–13.
- , "Violence and Sociality in Human Evolution," *Current Anthropology* 32,4 (August–October 1991):391–428.

- Knez, Marc and Duncan Simester, "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines," *Journal of Labor Economics* (in press).
- Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory* 27 (1982):245–252.
- Lee, Richard Borshay, *The !Kung San: Men, Women and Work in a Foraging Society* (Cambridge, UK: Cambridge University Press, 1979).
- Leimar, O. and P. Hammerstein, "Evolution of Cooperation through Indirect Reciprocity," *Proc. Royal. Soc. Lond. B* 268 (2001):745–753.
- Loewenstein, George F., Leigh Thompson, and Max H. Bazerman, "Social Utility and Decision Making in Interpersonal Contexts," *Journal of Personality and Social Psychology* 57,3 (1989):426–441.
- Mahdi, Niloufer Qasim, "Pukhtunwali: Ostracism and Honor Among the Pathan Hill Tribes," *Ethology and Sociobiology* 7 (1986):295–304.
- Milinski, Manfred, "Byproduct Mutualism, Tit-for-tat Reciprocity and Cooperative Predator Inspection," *Animal Behavior* 51 (1996):458–461.
- Moore, Jr., Barrington, *Injustice: The Social Bases of Obedience and Revolt* (White Plains: M. E. Sharpe, 1978).
- Nowak, Martin A. and Karl Sigmund, "Evolution of Indirect Reciprocity by Image Scoring," *Nature* 393 (1998):573–577.
- , Karen M. Page, and Karl Sigmund, "Fairness Versus Reason in the Ultimatum Game," *Science* 289 (8 September 2000):1773–1775.
- Ostrom, Elinor, "A Behavioral Approach to the Rational Choice Theory of Collective Action," *American Political Science Review* 92,1 (March 1998):1–21.
- , James Walker, and Roy Gardner, "Covenants with and without a Sword: Self-Governance Is Possible," *American Political Science Review* 86,2 (June 1992):404–417.
- Raffa, K. F. and A. A. Berryman, "Interacting Selective Pressures in Conifer-bark Beetle Systems: a Basis for Reciprocal Adaptations?," *The American Naturalist* 129 (1987):234–262.
- Robson, Arthur J., "Efficiency in Evolutionary Games: Darwin, Nash, and the Secret Handshake," *Journal of Theoretical Biology* 144 (1990):379–396.
- Scott, James C., *The Moral Economy of the Peasant: Rebellion and Subsistence in Southeast Asia* (New Haven, CT: Yale University Press, 1976).
- Sethi, Rajiv and E. Somanathan, "The Evolution of Social Norms in Common Property Resource Use," *American Economic Review* 86,4 (September 1996):766–788.

- Stephens, W., C. M. McLinn, and J. R. Stevens, "Discounting and Reciprocity in an Iterated Prisoner's Dilemma," *Science* 298 (13 December 2002):2216–2218.
- Trivers, R. L., "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971):35–57.
- Wedekind, Claus and Manfred Milinski, "Cooperation Through Image Scoring in Humans," *Science* 289 (5 May 2000):850–852.
- Wiessner, Polly, "Hunting, Healing and Hxaro Exchange: A Long Term Perspective on !Kung Large-game Hunting," *Evolution and Human Behavior* 23 (2002):407–436.
- Wood, Elisabeth Jean, *Insurgent Collective Action and Civil War in El Salvador* (Cambridge,: Cambridge University Press, 2003).
- Woodburn, James, "Egalitarian Societies," *Man* 17,3 (1982):431–451.
- and Alan Barnard, "Property, Power and Ideology in Hunter-Gathering Societies: An Introduction," in T. Ingold, D. Riches, and J. Woodburn (eds.) *Hunters and Gatherers* (New York: St. Martin's Press, 1988) pp. 4–31.
- Yamagishi, Toshio, "The Provision of a Sanctioning System as a Public Good," *Journal of Personality and Social Psychology* 51 (1986):110–116.
- Zahavi, Amotz, "Mate Selection—A Selection for Handicap," *Journal of Theoretical Biology* 53 (1975):205–214.