Evolution of Social Behavior: Individual and Group Selection

Theodore C. Bergstrom

"A selector of sufficient knowledge and power might perhaps obtain from the genes at present available in the human species a race combining an average intellect equal to that of Shakespeare with the stature of Carnera. But he could not produce a race of angels. For the moral character or for the wings he would have to await or produce suitable mutations."

—J.B.S. Haldane (1932, p. 110)

What can our evolutionary history tell us about human motivations and social behavior? The genes that influence our own behavior are copies of copies of copies of genes that have guided our ancestors to survive and to reproduce. If selfish behavior leads to reproductive success, then we should expect our genes to impel us toward selfishness. Thus, Richard Dawkins (1976, pp. 2–4) observed:

> If we were told that a man lived a long and prosperous life in the world of Chicago gangsters, we would be entitled to make some guesses as to the sort of man he was . . . . Like successful Chicago gangsters, our genes have survived, in some cases for millions of years, in a highly competitive world . . . . If you look at the way natural selection works, it seems to follow that anything that has evolved by natural selection should be selfish.

Sir Alexander Carr-Saunders (1922), a sociologist who pioneered studies of demography and social evolution, took an opposing view. Carr-Saunders maintained that evidence from primitive cultures suggests that prehistoric humans clustered into groups that inhabited well-defined areas between which migration

■ *Theodore C. Bergstrom is the Aaron and Cherie Raznick Professor of Economics, University of California at Santa Barbara, Santa Barbara, California.*

was infrequent. These groups usually managed to avoid overpopulation and the attendant scourges of war, famine and disease. Population densities remained roughly constant and close to the levels that maximized per capita food consumption. To support these claims, Carr-Saunders offered evidence that in existing primitive societies, fertility is deliberately restrained by means of abortion, infanticide and long-term sexual abstinence. He argued that such reproductive restraint is inconsistent with selfish maximization of individual fertility and must somehow be explained by "group selection."

In an encyclopedic study of the spatial dispersion of animal populations, V.C. Wynne-Edwards (1962) proposed that this principle has far more ancient roots. He believed that in much of the animal kingdom, species maintain population densities at an "optimal level for each habitat that they occupy." Wynne-Edwards argued that group selection is possible because "animal (and plant) species tend to group into more or less isolated populations" who depend on food resources of a "localized, immobile" character and that "the local stock conserves its resources and thereby safeguards the future survival of its descendants."

Some biologists viewed Wynne-Edwards' book as academic heresy. George C. Williams (1966) argued that the behavior that Wynne-Edwards cites as evidence for group selection is consistent with individuals maximizing their own long-run reproductive interests or those of close relatives. For example, individuals may find it in their own reproductive interests to vary their birth rates inversely with population density. It also may be in an individual's selfish interest to defend territories that exceed minimum requirements in normal years, because the extra territory will be critically important in bad years. Other biologists have attempted to provide formal underpinnings for Wynne-Edwards's rather loosely argued proposition. A detailed and interesting account of this controversy can be found in Sober and Wilson (1999).

The polar theories of individual and group selection make dramatically different predictions about social interactions. Individual selection theory suggests a world populated by resolutely selfish *homo economicus* and his zoological (and botanical) counterparts. By contrast, a world shaped by group selection would likely be inhabited by natural socialists, with an instinctive "Kantian" morality toward other members of their group. Of course the localism that leads to group selection would also tend to produce unsavory impulses toward xenophobia and intertribal warfare.

Games and Social Interactions

Prisoners' Dilemma Games

Evolutionary biologists, game theorists and anthropologists have frequently used the prisoners' dilemma game as a research vehicle to explore the polar regions of individual and group selection theory and the interesting terrain that lies between. A multiplayer prisoners' dilemma is a game in which individuals may take

actions that are, in the words of J.B.S. Haldane (1932), "socially valuable but individually disadvantageous." Specifically, we consider a game that has two possible strategies for each player, cooperate and defect, where the payoff to each player depends on her own strategy and the number of other players who play cooperate. An *N-player prisoners' dilemma* is defined to be a game in which 1) all *N* players are better off if all play cooperate than if all play defect, and 2) given the actions of others, every individual necessarily gets a higher payoff from playing defect than from playing cooperate.

The prisoners' dilemma is especially useful for exploring the alternative theories of individual and group selection, because for this game, the two theories predict starkly different outcomes. If individuals are selected to act in their own self-interest, all will defect. If they are selected to act in the group interest, all will cooperate.

Much of the literature of evolutionary biology focuses on the special class of *N*-player prisoners' dilemma games in which each player's payoff depends linearly on the number of players who play cooperate. This game was introduced to biologists in 1932 by J.B.S. Haldane, one of the founders of modern population biology.[1] Economists will recognize Haldane's game as the linear "voluntary contribution to public goods" game, much studied in experimental economics. (See Ledyard, 1995, for a good survey of this work.) In this paper, we will call this game the *linear public goods game.*

In an *N*-player linear public goods game, a cooperator bears a cost of c and by so doing confers a benefit of b/N on each of the *N* players, including herself.[2] A defector bears no costs and confers no benefits, but receives the benefits conferred by cooperators in her group. Thus, if the fraction of cooperators in the group is x, each cooperator gets a payoff of $bx - c$, and each defector gets a payoff of bx. For a linear public goods game to be a prisoners' dilemma, it must be that all players are better off if all cooperate than if all defect and that given the actions of others, each player gets a higher payoff from defecting than from cooperating. The first condition is satisfied if and only if $b > c$. By cooperating, an individual adds b/N to her own payoff at a cost of c. Thus, the second condition is satisfied if and only if $b/N < c$. It follows that a linear public goods game is an *N*-player prisoners' dilemma game if and only if $b/N < c < b$.

Stag Hunt Games

In one-shot prisoners' dilemma games, the socially optimal action is never a best response for selfish individuals. But in many social interactions, the action that

---

[1] The notation used here is that of Cohen and Eshel (1976), who credit Haldane (1932) in pp. 207–210 of their appendix.

[2] David S. Wilson (1975) studies a variant of this game in which at a cost of c, a cooperator confers expected benefits of b/N on every player *except for herself.* Results for one variant translate easily into corresponding results for the other, since Wilson's formulation of the game with costs c is equivalent to the Haldane formulation with costs $c + b \setminus N$.

Table 1
A Stag Hunt Game
(entries are payoffs to row player)

		Column Player's Strategy	
		Cooperate	Defect
Row Player's Strategy	Cooperate	4	0
	Defect	3	3

best serves one's self-interest depends on the actions taken by others. This suggests the usefulness of a second exploratory vehicle, a simple two-person game, known as the *stag hunt.* This game formalizes a story told by Jean Jacques Rousseau (1755 [1950], p. 428) of two hunters who could cooperate by jointly hunting a stag or defect by individually hunting hare.[3] Table 1 is a game matrix for a stag hunt game, where entries represent payoffs to the row player. In contrast to a prisoners' dilemma, where defect is the best response regardless of the other's strategy, in stag hunt games, defect is the best response to defect, but cooperate is the best response to cooperate. Thus, the stag hunt has two equilibria, one where both players cooperate and one where both defect.

Evolutionary Dynamics and Group Formation

We will explore the evolutionary dynamics of populations in which individuals are "programmed," perhaps genetically or perhaps by cultural experience, to play either cooperate or defect in a game. We assume that the dynamics are *payoff-monotonic* (Weibull, 1995), which means that for any two player types, the type receiving the higher payoff in the game will reproduce more rapidly.

If the entire population participates in a single multiperson prisoners' dilemma game, the prediction of payoff-monotonic dynamics is simple. Since defectors always receive higher payoffs than cooperators, they will reproduce more rapidly than cooperators, and eventually, the population will consist almost entirely of defectors.[4]

Population biologists, like J.B.S. Haldane (1932) and Sewall Wright (1945), proposed that cooperation is most likely to evolve in populations where social interaction takes place within relatively small subpopulations, called *demes,* where there is occasional, but relatively infrequent migration between subpopulations. To investigate this conjecture, biologists have studied a class of dynamic models known

---

[3] An engaging paper by Brian Skyrms (forthcoming) makes a strong case that social thinkers should pay more attention to the stag hunt game.

[4] The result that the proportion of defectors converges to one is less obvious than the result that this proportion increases monotonically. A proof is found in Weibull (1995), who credits this result to John Nachbar (1990).

as *haystack models.*[5] In haystack models, random group formation produces some groups with more cooperators than others. Although cooperators have fewer offspring than defectors in their own group, all members of groups with more cooperators reproduce more rapidly. Haystack models are designed to explore circumstances under which the between-group advantage of cooperation can overwhelm the within-group advantage of defection.

In a haystack model, time is divided into discrete periods. Each period has a reproductive phase and a dispersal phase. The reproductive phase begins with a population that is partitioned into a large number of groups. During the reproductive stage, which may continue for several generations, asexual reproduction takes place within each group. Descendants of any individual are of the same type as their ancestor. The number of descendants of each founding group member present at the end of the reproductive phase depends on that individual's type and on the distribution of types in the founding group. At the end of the reproductive phase comes the dispersal phase, where the entire population is pooled and new groups are randomly formed from the pooled population. The process is repeated as the newly formed groups reproduce and disperse once again.

In a haystack model, we define group formation to be *nonassortative* if the probability distribution of types of the other members of one's group is independent of one's own type. Where individuals in the overall population can be of either type C or type D, let us define $p_i(M, N)$ to be the probability that a type i individual is assigned to a group of size N, in which M of the *other* group members are type Cs. The assignment is defined to be nonassortative if $p_C(M, N) = p_D(M, N)$ for all relevant M and N. If the overall population is large and groups are formed by random sampling without replacement from this population, then matching will be almost nonassortative. There will be a slight difference between $p_C(M, N)$ and $p_D(M, N)$ because the other members of a cooperator's group are drawn from the overall population less one cooperator, while the other members of a defector's group are drawn from the overall population less one defector. For a large population size, this difference can be shown to be negligible.

Where Not to Look for Cooperation

In a survey of the literature on group selection, Alan Grafen (1984) states that "with random selection there is no selection for altruism." Grafen does not offer a proof of this assertion.[6] However, if we define "altruism" to mean playing cooperate in a prisoners' dilemma and "random" to mean nonassortative, then Grafen's claim is equivalent to the following theorem.

   Proposition 1: Iron Rule of Selfishness. In a haystack model, if matching is nonassortative and if reproductive rates are determined by a prisoners'

---

[5] The first haystack model was introduced by John Maynard Smith (1964) and will be discussed below.
[6] An earlier paper by David S. Wilson (1975) showed for a quite specific model that "random" formation of groups must result in the elimination of altruism.

dilemma game among group founders, then the only asymptotically stable outcome is a population consisting entirely of defectors.

To prove Proposition 1, we show that in each time period, at the dispersal stage, the expected number of descendants of a defector exceeds that of a cooperator.[7] Thus, each time the haystacks are dispersed and reformed, it will be with a higher proportion of defectors. It follows that eventually the proportion of defectors in the population becomes arbitrarily close to one.

It is important to understand that Proposition 1 does not rule out the possibility that group selection can sustain cooperative behavior. Instead, this proposition allows us to classify environments where selection can favor cooperation as follows: In a haystack model, cooperation will survive only if either a) the game being played is not a prisoners' dilemma, or b) matching is assortative.

Haystack Models of Group Selection

Maynard Smith's Mice

John Maynard Smith (1964) introduced the first formal model of "group selection" in which seemingly altruistic behavior is sustained in equilibrium, even though the matching process is nonassortative. The key to Maynard Smith's result is that groups may remain intact for several generations before the population is dispersed and randomly regrouped. Maynard Smith motivates his model with a charming story of "a species of mouse which lives entirely in haystacks."

A meadow contains many haystacks, each of which is colonized every year by exactly two mice. These mice and their descendants reproduce asexually throughout the summer until harvest time when the haystacks are cleared.[8] The dislodged resident mice scramble out into the meadow, and when new haystacks are built in the next year, each haystack is colonized by exactly two mice randomly selected from the survivors of last year's breeding season. If there are more than enough mice to provide two founders for each new haystack, the extra mice are consumed by predators.

---

[7] The expected number of descendants of a cooperator is $\sum_M \sum_N p_C(M, N) \Pi_C(M, N)$, which is the sum over all possible group configurations of the probability that a cooperator belongs to that type of group times her number of descendants in this case. Similarly, the expected number of descendants of a defector is $\sum_M \sum_N p_D(M, N) \Pi_D(M, N)$. Since matching is nonassortative, it must be that $p_C(M, N) = p_D(M, N)$ for all M and N. Therefore, the difference between the expected number of descendants of a cooperator and that of a defector is equal to $\sum_M \sum_N p_C(M, N)(\Pi_C(M, N) - \Pi_D(M, N))$. If the game is prisoners' dilemma, it must be that $\Pi_C(M, N) - \Pi_D(M, N) < 0$ for all M and N. It follows that $\sum_M \sum_N p_C(M, N)(\Pi_C(M, N) - \Pi_D(M, N)) < 0$, and therefore, in every period, the expected number of offspring of a cooperator is less than that of a defector.

[8] In Maynard Smith's (1964) telling, mice reproduce as sexual diploids. However, his model is mathematically equivalent to one with asexual reproduction. To ease exposition, and to make this model directly comparable to extensions introduced by Cohen and Eshel (1976), I present the equivalent asexual model here.

Table 2

The Haystack Game

(entries are payoffs to row player)

		Column Player's Strategy	
		Cooperate	Defect
Row Player's Strategy	Cooperate	$r(1+K)/2$	0
	Defect	r	$r/2$

In a haystack settled by two timid mice, all descendants are timid, and in a haystack settled by two aggressive mice, all descendants are aggressive. In a haystack settled by one mouse of each type, the descendants of the aggressive mouse eliminate the descendants of the timid mouse, and the number of its descendants at harvest time is the same as the number in a haystack colonized by two aggressive mice. Although timid mice do poorly when matched with aggressive mice, haystacks inhabited entirely by timid mice produce more surviving offspring at harvest time than do haystacks inhabited by aggressive mice. Thus, a haystack colonized by two timid mice produces $1 + K$ times as many descendants as does a haystack with aggressive mice.

Since the reproduction rate enjoyed by a founding mouse depends on its own type and that of its cofounder, these rates can be represented as the payoffs in a game between the two mice who colonize each haystack. If two aggressive mice colonize a haystack, they will have a total of r descendants, half of whom are descended from each founder. Thus, each mouse has $r/2$ descendants. If an aggressive mouse and a timid mouse colonize a haystack, the timid mouse will have no descendants, and the aggressive mouse will have r descendants. If two timid mice colonize a haystack, they will have a total of $r(1 + K)$ descendants, and each will have $r(1 + K)/2$ descendants. In the game played by cofounders, payoffs to the row player are shown in Table 2.

If $0 < K < 1$, the haystack game is a prisoners' dilemma, since regardless of its cofounder's type, an aggressive mouse will have more offspring than a timid mouse. If $K > 1$, the haystack game is not a prisoners' dilemma, but a stag hunt. If matched with a timid mouse, a mouse will have more offspring if it is timid than if it is aggressive. But if matched with an aggressive mouse, a mouse will have more offspring if it is aggressive than if it is timid.

For the prisoners' dilemma case with $K < 1$, the only equilibrium is a population made up entirely of defectors. For the stag hunt case, with $K > 1$, there are two distinct stable equilibria, one in which all mice are timid and one in which all are aggressive. We demonstrate this as follows: Let the proportion of timid mice in the population at time t be x_t. Since matching is random, any mouse is matched with a timid cofounder with probability x_t and with an aggressive cofounder with probability $1 - x_t$. Given the payoffs in Table 2, the expected reproduction rate of an aggressive mouse is $x_t r + (1 - x_t)r/2$, and the expected reproduction rate of

a timid mouse is $x_t r(1 + K)/2$. Subtracting the latter expression from the former, we find that the difference between the expected reproduction rates of timid mice and of aggressive mice is proportional to $x_t K - 1$. Therefore, timid mice reproduce more rapidly than aggressive mice if $x_t K > 1$, and aggressive mice reproduce more rapidly if $x_t K < 1$. These dynamics are illustrated by Figure 1. The graph on the left shows the case where $K > 1$, so that the game played between founders is a stag hunt. In this case, when the proportion of timid mice is large, timid mice reproduce more rapidly than aggressive mice, and when the proportion of timid mice is small, aggressive mice will reproduce more rapidly. Thus, there are two stable equilibria, one in which all mice are timid and one in which all are aggressive. (There is also an unstable equilibrium where the fraction $1/K$ of mice are timid.) The graph on the right shows the case where $K < 1$, so that the game played between founders is a prisoners' dilemma. In this case, aggressive mice always reproduce more rapidly than timid mice, and there is a unique equilibrium in which all mice are aggressive.

According to Proposition 1, a population of cooperators will be stable only if either a) the game played is not a prisoners' dilemma, or b) matching is assortative. We see that the survival of cooperation in Maynard Smith's haystack model falls into the first case. If $K > 1$, the game played between founders is a stag hunt rather than a prisoners' dilemma and there are two stable equilibria, one with cooperators only and one with defectors only. Where $K < 1$, the game played between founders is a prisoners' dilemma, and as Proposition 1 predicts, the only stable equilibrium is a population of defectors.

The Cohen-Eshel Linear Public Goods Model

Dan Cohen and Ilan Eshel (1976) generalize Maynard Smith's (1964) haystack model in an instructive way. They study haystack models with cooperators and defectors, where founding groups may be of arbitrary size and where reproduction rates vary continuously with the proportion of cooperators in the group. Groups remain intact for a fixed length of time T and then are dispersed, and new groups are formed with nonassortative matching. We will pay special attention to one of the cases that they study. This is a linear public goods game in which $x(t)$ is the proportion of cooperators in a group at time t and where cooperators in the group reproduce at the instantaneous rate $a + bx(t)$ and defectors reproduce at the rate $a + bx(t) - c$. For this model, Cohen and Eshel are able to solve the resulting differential equations and thereby determine when the more rapid growth rates of groups with more cooperators is sufficient to offset the fact that cooperators reproduce less rapidly within groups than defectors. Their answer is expressed in the following proposition.

Proposition 2: Cohen-Eshel Theorem. In the Cohen-Eshel linear public goods model, where groups are of size N and where T is the length of time for which groups remain intact: i) If T is small, there is a unique asymptotically stable equilibrium. The equilibrium population is made up entirely of defectors if $b/N < c$ and entirely of cooperators if $b/N > c$. ii) If T is sufficiently large,

Figure 1
Dynamics of the Haystack Model
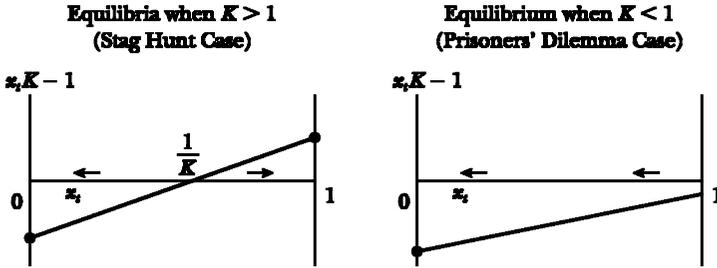
and $b > c > 0$, there exist two distinct stable equilibria: one with selfish players only and one with altruists only.

A heuristic explanation of part (i) is that if T is small, the reproduction rate over the life of the group is approximately bx for defectors and $bx - c$ for cooperators, where x is the proportion of cooperators among the founders. If $b/N < c$, this game is a prisoners' dilemma. If $b/N > c$, then the game is not a prisoners' dilemma, since the private benefits that an individual gets from cooperating exceed the cost of cooperation, and so cooperate rather than defect is a dominant strategy. Thus, we see that part (i) of Proposition 2 is consistent with Proposition 1.

Part (ii) is the more surprising result. Where groups are sufficiently long-lived, a population of cooperators can be sustained as a stable equilibrium, even if the game played between contemporaries is a prisoners' dilemma. To see how this can happen, let us suppose that T is large and ask whether a population of cooperators can be invaded by defectors. A defector who joins $N - 1$ cooperators in founding a group will have fewer expected offspring than would a cooperator in a group made up entirely of cooperators. The reason is that for large T, a group that has at least one defector among its founders will eventually consist almost entirely of defectors, and thus, the reproduction rate of every member of that group will eventually be lower than the reproduction rate of the normal population of cooperators who live in communities founded by cooperators only. This implies that if groups are sufficiently long-lived, a mutant defector will have fewer descendants than the normal cooperators who live among cooperators. Thus, mutant defectors cannot invade the population.

How does part (ii) of the Cohen-Eshel theorem square with Proposition 1? If we consider the game played between founders and if the survival duration T of groups is large, then as we have seen, the game that determines reproductive rates over the lifetime of a group is not a prisoners' dilemma. Thus, there is no conflict with Proposition 1. It is also instructive to look at the situation another way. We note that the linear public goods game that determines instantaneous reproduction

rates is a prisoners' dilemma if $b/N < c$. But after a group has been intact for some length of time, the individuals who participate in this game will not be nonassortatively matched, but will tend to share common ancestry.

Absolute and Relative Payoff Within Groups

There is an interesting class of games in which every player gets a higher payoff from cooperating than from defecting, but where, paradoxically, it is also true that within each group, defectors receive higher payoffs than cooperators. The root of this paradox is the difference between absolute and relative payoffs. Consider an N-player linear public goods game, where in a group with the fraction x of cooperators, cooperators' payoffs are $bx - c$ and defectors' payoffs are bx. By cooperating rather than defecting, a player increases x by $1/N$ and hence confers a benefit of b/N on all group members, including herself. If $b/N > c > 0$, the private benefits she gets from cooperating exceed the cost of cooperation, so cooperating increases her absolute payoff. But defectors enjoy all of the benefits from cooperators' efforts and bear none of the costs. Therefore, in any group, the defectors have higher payoffs than the cooperators.

D.S. Wilson (1979) noticed this possibility and suggested that someone who cooperates when $b/N > c$ but not when $b/N < c$ be called a "weak altruist," while someone who cooperates whenever $b > c$ be called a "strong altruist." Wilson argued that "many, perhaps most, group-advantageous traits such as population regulation, predation defense, and role differentiation" may be explained by weak altruism. Grafen (1984) responded that Wilson's "weak altruism" is not really altruism, since weak altruists cooperate only when it is in their self-interest as measured by absolute (but not necessarily relative) payoffs. Grafen suggests that Wilson's weak altruism would be better called "a self-interested refusal to be spiteful."

Whether we speak of "weak altruism" or "refusal to be spiteful," it is interesting to determine whether natural selection favors maximization of absolute payoff or of relative payoffs when these two objectives conflict. Cohen and Eshel's Proposition 2 offers an answer. Where the lifetime T of groups is short, a stable equilibrium population of cooperators can be supported if and only if $b/N > c$. Thus, natural selection favors maximization of absolute payoffs, even at the expense of relative payoffs. However, if groups remain intact for a long time, maximization of absolute payoffs is not unambiguously favored. The Cohen-Eshel result is that for large T, there exist two distinct equilibria—one populated by cooperators only and one by defectors only.

Variations on the Haystack Model

In the haystack model, groups survive in perfect isolation until they are simultaneously disbanded. More realistic models would allow migration between groups and would have asynchronous extinctions and resettlements. Such models have been studied by Eshel (1972), Levins (1970), Levin and Kilmer (1974) and Boorman and Levitt (1980). In these models, defectors reproduce more rapidly

than cooperators within their own group, but groups face a higher probability of group extinction the greater the proportion of defectors. For some parameter values, a population of cooperators will be sustained in equilibrium. This is more likely if the size of founding populations is relatively small, if the migration rate is relatively small and if the difference between the extinction rates of selfish and altruistic groups is relatively large.

Assortative Group Formation

In the prisoners' dilemma, defect always yields a higher payoff than cooperate, but everyone gets a higher payoff if matched with a cooperator rather than a defector. Proposition 1 tells us that defection will prevail if types are matched at random to play prisoners' dilemma. But if matching is assortative rather than random, the cost of cooperation may be repaid by a higher probability of playing with another cooperator.

The Index of Assortativity

Sewall Wright (1921) defined the assortativeness of mating with respect to a trait as "the coefficient of correlation m between two mates with respect to their possession of the trait." Cavalli-Sforza and Feldman (1981) pointed out that Wright's correlation coefficient is equivalent to a matching process where the fraction m are assigned to their own type with certainty and the remaining fraction $(1 - m)$ mate at random.[9] For a population of cooperators and defectors, Bergstrom (2002) defined the *index of assortativity* $a(x)$ to be the difference between the expected fraction of cooperators that a cooperator will encounter in her group and the expected fraction of cooperators that a defector will encounter. He shows that this definition reduces to Wright's correlation coefficient m when matching is pairwise and $a(x)$ is constant. As we will see, in some important applications $a(x)$ is constant, but there are interesting cases where it is not.

Recall that in an N-player linear public goods game, a player who cooperates increases the fraction of cooperators in her group by $1/N$ and thus adds b/N to the payoff of each player, including herself. Therefore, the net cost of cooperating is $c* = c - b/N$, and the game is a prisoners' dilemma whenever $c* > 0$. Proposition 3 tells us that if the index of assortativity is positive, then in haystack models, individuals in equilibrium will act as if they "discount" benefits to other group members at a rate equal to the index of assortativity.

---

[9] Let I_i be an indicator variable that takes on value 1 if mate i has the trait and 0 otherwise. Wright's assortativeness m is the correlation between I_1 and I_2. Thus, $m = (E(I_1 I_2) - E(I_1) E(I_2))/(\sigma_1 \sigma_2)$, where σ_i is the standard deviation of I_i. Now $E(I_1 I_2) = xp(x)$, and for $i = 1, 2$, $E(I_i) = x$ and $\sigma_i = \sqrt{x(1 - x)}$. Therefore, $m = (xp(x) - x^2)/x(1 - x)$. Rearranging terms, one finds this expression equivalent to $p(x) = m + x(1 - m)$.

Proposition 3: In a haystack model in which the game played among founders is an *N*-player linear public goods game, where $a(x)$ is the index of assortativity, a population of cooperators will be stable if and only if $a(1)b - c^* > 0$, and a population of defectors will be stable if and only if $a(0)b - c^* < 0$, where $c^* = c - b/N$.

The proof of Proposition 3 is sketched as follows: Recall that where x is the fraction of cooperators in the population, the index of assortativity $a(x)$ is the difference between the expected fraction of cooperators encountered by cooperators and that encountered by defectors. Therefore, the expected difference between the total benefits experienced by cooperators and by defectors is just $a(x)b$, and since c^* is the net cost of cooperating, the reproduction rate of cooperators will exceed that of defectors if $a(x)b - c^* > 0$. A large population of cooperators is stable if a mutant defector in this population will reproduce less rapidly than cooperators. This happens if $a(1)b - c^* > 0$. A large population of defectors is stable if a mutant cooperator in a population of defectors will reproduce less rapidly than defectors. This will happen if $a(0)b - c^* < 0$.

Bergstrom (1995, 2002) generalizes Proposition 3 from the special class of linear public goods games to games in which payoffs are not linear in the number of cooperators. For the case of pairwise matching with a constant index of assortativity, a, the equilibrium behavior can be described as follows: Take the action that you would choose if you believed that with probability a, your partner will mimic your action and with probability $1 - a$, your partner will mimic a random draw from the population at large. Bergstrom calls this the *semi-Kantian rule*, since it can be viewed as a probabilistic version of Kant's categorical imperative.

Family Groups and Assortative Matching

Families are conspicuous examples of nonrandomly formed groups. William Hamilton (1964) developed *kin selection theory,* which predicts the willingness of individuals to make sacrifices to benefit their relatives according to a benefit-cost principle known as *Hamilton's rule.* Hamilton's rule states that individuals are willing to reduce their own expected number of offspring by c in order to add b to that of a relative if and only if $br > c$, where r is *the coefficient of relatedness* between the two relatives. Biologists define the coefficient of relatedness between two individuals as the probability that a rare gene found in one of them also appears in the other. In a population without inbreeding, the coefficient of relatedness is one-half for full siblings, one-fourth for half siblings and one-eighth for first cousins.

Hamilton (1964) obtained his result for a simplified genetic model, known as *sexual haploidy.* In sexual haploids, an inherited trait is determined by a single gene that is inherited from a random draw of the heir's two parents.[10] The sexual haploid model seems appropriate for studying cultural transmission, where behav-

---

[10] See Boorman and Levitt (1980) and Bergstrom (1995) for similar results applying to sexual diploids.

ior is passed from parents to children by imitation. Bergstrom (2002) and Bergstrom and Stark (1993) show that if mating between parents is not assortative and if inheritance follows the sexual haploid model, then the index of assortativity between two relatives is equal to their coefficient of relatedness. For example, two full siblings inherit their behavior from the same parent with probability $1/2$ and from different parents with probability $1/2$. The difference between the probability that a cooperator has a cooperative sibling and the probability that a defector has a cooperative sibling is $a(x) = 1/2$, which is also their degree of relatedness. Since Hamilton's model of interactions is equivalent to a two-player linear public goods game,[11] Hamilton's rule is a special case of Proposition 3.

Bergstrom (2002) shows that the index of assortativity can be calculated under a variety of alternative assumptions about cultural or genetic inheritance. These include cases where parents mate assortatively, where children may copy a randomly chosen stranger rather than one of their parents, where children preferentially copy mother or father and where marriage is polygamous.

Assortative Matching with Partner Choice

In a multiplayer prisoners' dilemma game, everyone prefers being matched with cooperators rather than defectors. Thus, if groups could restrict entry and players' types were observable, cooperators would not admit defectors to their groups, two types would be strictly segregated, and cooperators would reproduce more rapidly than would defectors. In more realistic environments, type detection is imperfect, and the index of assortativity falls between zero and one.

Bergstrom (2002) studies players who are labeled with an imperfect indicator of their type. Everyone sees the same labels, and so when players choose partners, there are only two distinguishable types: players who are labeled cooperator and players who are labeled defector. Although everyone realizes that the indicators are imperfect, everyone prefers to match with an apparent cooperator rather than with an apparent defector. Therefore, with voluntary matching, all individuals are matched with others of the same label. The expected proportion of true cooperators encountered by true cooperators will exceed that encountered by true defectors. The index of assortativity $a(x)$ is shown to vary with the proportion of cooperators in such a way that $a(0) = a(1) = 0$, while $a(x)$ is positive for intermediate values. Where groups play a linear public goods game, there turns out to be a unique stable polymorphic equilibrium, that includes positive fractions of both types. Frank (1987) presents an alternative model with two types who emit partially informative type indicators and also finds a unique polymorphic equilibrium.

Skyrms and Pemantle (2000) study dynamic formation of groups by reinforcement learning. Individuals initially meet at random and play a game. The payoffs determine which interactions are reinforced, and a social network emerges. The

---

[11] In Hamilton's formulation, the benefit b accrues only to one's relative and not to oneself. This model is equivalent to a linear public goods game with $c^* = c$.

networks that tend to form in their model consist of small interaction groups within which there is partial coordination of strategies and which can support cooperative outcomes.

Assortative Matching Induced by Spatial Structure

Evolutionary biologists have stressed the importance of spatial structure on the spread of mutations, genetic variation and the formation of species. Wright (1943) studied the degree of inbreeding in a population distributed over a large area where individuals are more likely to mate with those who live nearby. Kimura and Weiss (1964) studied genetic correlations in a one-dimensional "stepping stone model" with colonies arrayed along a line and where "in each generation an individual can migrate at most 'one step' in either direction." Nowak and May (1993) ran computer simulations with agents located on a two-dimensional grid playing the prisoners' dilemma with their neighbors. After each round of play, each site is occupied by its original owner or by one of its neighbors, depending on who had the highest score in the previous round. This process generates chaotically changing spatial patterns in which the proportions of cooperators and defectors fluctuate about long-term averages.

Bergstrom and Stark (1993) considered a population of farmers located on a road that loops around a lake. Each farmer plays the prisoners' dilemma with his two neighbors. The farmers' sons observe the strategies and payoffs of their fathers and their immediate neighbors and imitate the most successful. In this case, it turns out that an arrangement is stable if cooperators appear in clusters of three or more and defectors in clusters of two or more. With slightly different rules, they show that spatial patterns of behavior "move in a circle" around the lake. Thus, a long-lived chronicler, who observed behavior at a single farm would see cyclic behavior in which spells of cooperation are interrupted by defection according to a regular temporal pattern.

Eshel, Samuelson and Shaked (1998) studied a similar circular setup, but allowed the possibility of random mutations. They discovered that in the limit as the mutation rate becomes small, the only stationary states that have positive probability are those in which at least 60 percent of the population are cooperators. As the authors explain: "Altruists can thus invade a world of Egoists with only a local burst of mutation that creates a small string of Altruists, which will then subsequently grow to a large number of Altruists. Mutations can create small pockets of egoism, but these pockets destroy one another if they are placed too close together, placing an upper bound on the number of Egoists that can appear."

Eshel, Sansone and Shaked (1999) constructed another circular model in which each individual plays the prisoners' dilemma with her k nearest neighbors and observes the payoffs realized by her n nearest neighbors. They showed that the population dynamics can be determined from an explicitly calculated index of assortativity $r(k, n)$ for critical players who are located on the frontier between a string of cooperators and a string of defectors. Where the game played between

neighbors is a linear public goods game, cooperation will prevail if $r(k, n) \, b > c$ and defection will prevail if $r(k, n) \, b < c$.

Cooperation without the Prisoners' Dilemma

Ken Binmore (1994b) has observed: "If our Game of Life were the one-shot Prisoners' Dilemma, we should never have evolved as social animals." Binmore argues that the "Game of Life" is best modeled as an indefinitely repeated game in which reciprocal rewards and punishments can be practiced. As Binmore points out, this idea is not new. In the seventh century before Christ, Hesiod (1929) stated the maxim, "Give to him who gives, and do not give to him who does not." David Hume (1739 [1978], p. 521) used language that is suggestive of modern game theory: "I learn to do service to another, without bearing him any real kindness, because I foresee, that he will return my service in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me and others. And accordingly, after I have serv'd him . . . he is induced to do his part, as foreseeing the consequences of his refusal."

The Equilibrium Selection Problem

Several game theorists in the 1950s nearly simultaneously discovered a result known as the *folk theorem,* which tells us that in indefinitely repeated games, almost any pattern of individual behavior can be sustained as a Nash equilibria by a stable, self-policing norm. Such a norm prescribes a course of action to each player and includes instructions to punish anyone who violates his prescribed course of action. For game theorists, this is discouraging news, since it means that the usual tools of game theory have little predictive power. Those who want further guidance from theory must seek some way to distinguish "plausible" Nash equilibria from implausible ones. Game theorists call this the "equilibrium selection problem."

The repeated prisoners' dilemma, like other repeated games, has many Nash equilibria, some of which are Pareto superior to others. Group selection can play a powerful role here, as suggested by Boyd and Richerson (1990, 1992, 2002) and Binmore (1992, 1994a, b). A mechanism that allows groups with higher total payoffs to "reproduce" more rapidly will not be directly opposed by individual selection within groups. As Boyd and Richerson (1990) explain: "Viewed from the within-group perspective, behavior will seem egoistic, but the egoistically enforced equilibria with the greatest group benefit will prevail."

Since norms and the amount of cooperation differ greatly across societies, it seems that the attainment of efficient cooperation by group selection is neither swift nor inevitable. Soltis, Boyd and Richerson (1995) propose a model of group selection that requires that there is variation in norms among groups, that extinction of groups is fairly common and that new groups are formed by fission of existing groups. Using data on group extinction rates of New Guinea tribes, the authors estimate that the amount of time needed for a rare advantageous cultural

attribute to replace a common cultural attribute is of the order of 500 to 1000 years. However, Boyd and Richerson (2002) show that the spread of socially beneficial norms will be much faster if individuals occasionally imitate strategies of inhabitants of successful neighboring groups.

The Punishment Problem

Although we know from the folk theorem that punishment norms can maintain cooperation as Nash equilibria, it is not obvious that evolutionary processes will sustain the willingness to punish. As Henrich and Boyd (2001) put it: "Many students of human behavior believe that large-scale human cooperation is maintained by threat of punishment . . . . However, explaining cooperation in this way leads to a new problem: why do people punish noncooperators? . . . Individuals who punish defectors provide a public good, and thus can be exploited by non-punishing cooperators if punishment is costly."

The standard game theoretic answer is that equilibrium strategies include instructions to punish others who are "supposed to punish" and fail to do so. These instructions for punishing include instructions to punish those who won't punish others when obliged to, and so on ad infinitum. From an evolutionary point of view, this resolution seems unsatisfactory. It does not seem reasonable to expect individuals to be able to keep track of higher order failures to punish deviations from punishment norms. Moreover, if a society is in an equilibrium where all attempt to obey the norm, direct violations would be sufficiently infrequent that selection for punishing violators would be weak.

"The Viability of Vengeance," by Friedman and Singh (1999), has a nice discussion of the evolutionary stability of costly punishment. They suggest that within groups, one's actions are observed and remembered. A reputation for being willing to avenge harmful actions may be sufficient compensation for the costs of retribution. In dealing with outsiders, however, one is remembered not as an individual but as a representative of her group. Accordingly, a willingness to avenge harm done by outsiders is a *public good* for one's group, since it deters outsiders from uncooperative behavior to group members. They propose that failure to avenge wrongs from outsiders is punished (costlessly) by one's own group, through loss of status.

Henrich and Boyd (2001) propose an interesting explanation for the viability of vengeance. They argue that "the evolution of cooperation and punishment are a side effect of a tendency to adopt common behaviors during enculturation." Since it is not possible for humans to analyze and to "solve" the complex social games that they play, imitation becomes important. It is often easier to observe a player's actions than to observe both actions and consequences. Thus, Henrich and Boyd suggest that imitation may take the form of "copy-the-majority" rather than "copy-the-most-successful."

Henrich and Boyd (2001) show that it takes only a slight tendency toward conformity to maintain an equilibrium that supports punishment strategies. In a simplified version of their model, community members engage in a two-stage game.

The first stage is a linear public goods game where players decide to cooperate or defect, but with a small probability, a player who intends to cooperate inadvertently defects. In the second stage, individuals decide whether to punish first-stage defectors. Punishing is costly both to the punisher and the punished. Consider a population in which initially everyone intends to cooperate at the first stage and where in the second stage, everyone intends to punish those who defected in the first stage. Since everyone intends to cooperate, the only defections observed are mistakes. Since everyone intends to punish those who defect in the first stage, the payoff to those who defect in the first stage is lower than the payoff to those who cooperate. Individuals who fail in the second stage to punish first-stage defectors get higher payoffs than those who cooperate by punishing first-stage defectors, but only slightly higher since there are very few first-stage defections. Since most individuals conform to the norm, the expected cost of punishing observed defections is low, and thus, even a very small weight on the conformity measure is sufficient to overcome the payoff loss from punishing. Henrich and Boyd show that when higher levels of punishment are accounted for, an even smaller weight on conformity suffices to maintain cooperation at all stages.

Fehr and Gächter (2000) find interesting experimental evidence of the viability of vengeance. In their experiment, subjects play a linear public goods game repeatedly with anonymous partners who are reshuffled after each play. The game also has a punishment stage in which at a cost to themselves, players can impose punishments on those who have defected. Fehr and Gächter find that although subjects are unlikely to reencounter those whom they punish, they frequently punish defectors. In consequence, in later rounds of play, most subjects cooperate fully.

Cultural versus Genetic Transmission

There is room to question whether the visceral, seemingly irrational anger that people feel when they are cheated or otherwise violated can be explained by cultural transmission rather than as genetic hard-wiring. Cosmides and Tooby (1989) offer experimental evidence indicating that people are much better at solving logical problems that are framed as "cheater detection" problems than at solving equivalent problems in other frameworks. In their view, this is evidence that humans have evolved special mental modules for solving such problems.

There is, however, evidence that culture influences readiness to anger. Experiments by Nisbett and Cohen (1996) subjected male college students to rude and insulting behavior in the laboratory. Using questionnaires, behavioral responses and checks of testosterone levels, they found that students raised in the American South become much angrier and more ready to fight than those raised in the North. The authors attribute this difference to the existence of a "culture of honor" in the South that is not present in the North.

Economists and anthropologists have conducted a remarkable cross-cultural series of experiments in which subjects play the *ultimatum game.* In the ultimatum game, two players are matched to divide a fixed sum of money. The first player, "the

proposer" offers a portion of the total to the second player, "the responder." If the responder accepts the offer, the money is divided as proposed. If the responder rejects, both players receive nothing. If this game is played by rational players who care only about their own monetary payoffs, then in equilibrium the proposer offers a very small share, and the responder accepts. In laboratory experiments, proposers typically offer a share of about one half, and this is accepted. When proposers attempt to capture a significantly larger share, responders usually reject the proposal, in effect foregoing a small gain in order to "punish" a greedy proposer. A study conducted in the United States, Israel, Japan and Slovenia found similar results across these countries (Roth, Prasnikar, Okuno-Fujiwara and Zamir, 1991). But very different results were found when the ultimatum game was conducted with the Machiguenga, a hunter-gatherer group who live in small extended family hamlets scattered through the tropical forests of the Amazon. Henrich (2000) found that the modal share offered by the Machiguenga was only 15 percent. Despite the fact that responders were offered a small share, they accepted these offers about 95 percent of the time. Henrich reports that in ordinary Machiguenga life, "cooperation above the family level is almost unknown." A recent study reports on ultimatum game experiments conducted in a total of 15 "small-scale societies," including hunter-gathers, pastoralists, farmers and villagers (Henrich et al., 2001). The studies found wide divergence among these societies.

The great diversity of norms across societies indicates that the details of prescribed behavior must be transmitted culturally rather than genetically. On the other hand, it appears that the emotional and intellectual prerequisites for the support of cultural norms are part of the common genetic endowment of humans. By analogy, humans raised in different environments grow up to speak different languages, but all normal humans appear to be genetically endowed with an innate ability to learn language and manage grammar and syntax. It would be interesting to know more about just which abilities and instincts are genetically transferred and which are culturally transmitted. While the qualitative character of evolutionary dynamics may be roughly the same for either transmission mechanism, the rates of mutation and of selection for traits transmitted culturally are likely to be much faster than for genetically transmitted traits.

Final Remarks

Further Reading

The literature on social evolution is large, diverse and multidisciplinary, and I have confined my attention to a small portion of this work. For those who want to read further in this area, here are a few works that I have found especially stimulating.

Cavalli-Sforza and Feldman's *Cultural Transmission and Evolution* (1981) pioneered formal modeling of this subject. Their introductory chapter presents a clear-headed formulation of the implications of mutation, transmission and natural

selection for culturally transmitted characteristics. There is also an empirical study of the transmission from parents to children of such cultural behavior as religious beliefs, political affiliation, listening to classical music, reading horoscopes and high salt usage.

Rajiv Sethi and R. Somanathan's "Understanding Reciprocity" (2002) lucidly presents much interesting work not discussed here. Sober and Wilson's (1999) book *Unto Others* contains an extensive history of theoretical controversies between group and individual selectionists. They also offer a fascinating survey of field observations of group norms in a sample of 25 cultures selected from anthropological studies.

H. Peyton Young's (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Social Institutions* contains a remarkably accessible introduction to the mathematical theory of stochastic dynamics and to its applications in the study of the evolution of social institutions. Young maintains that a proper treatment of the very long run must directly incorporate the stochastic process into the laws of motion, and he shows that in models with multiple equilibria, "long run average behavior can be predicted much more sharply than that of the corresponding determinate dynamics."

Skyrms's (1996) *Evolution of the Social Contract* is a beautifully written application of evolutionary dynamics to the study of bargaining games and the evolution of notions of fairness and social contracts.

Finally, my own thinking about the evolutionary foundations of social behavior has been much influenced by Ken Binmore's (1994a, b) two-volume work, *Game Theory and the Social Contract*. These books combines social philosophy, political theory, evolutionary theory, anthropology and modern game theory with great depth and subtlety.

Conclusion

I have attempted to map the territory between the opposing poles of individual and group selection theory. The former predicts outcomes that are Nash equilibria in games played by selfish individuals, while the latter predicts outcomes in which individuals act so as to maximize the total reproductive success of their own groups.

Haystack models occupy interesting intermediate terrain. In this setting, variations in reproductive success depend both on individuals' own actions and on the composition of the groups to which they are assigned. In haystack models, if groups are assembled nonassortatively and if the reproductive success of founding group members is determined by a multiperson prisoners' dilemma, then cooperation cannot be sustained in equilibrium. But the assumptions that ensure the defeat of cooperation are not always satisfied, and there are many important social situations in which at least some cooperation is sustained.

Stable groups of moderate size are likely to foster social interactions where, unlike in prisoners' dilemmas, individual self-interest is consistent with behavior that maximizes group success. Interaction in such groups is best modeled as a repeated game. In repeated games, where one's actions can be observed and

remembered by others, almost any pattern of individual behavior, including behavior that maximizes group payoff, can be sustained by social norms that include obligations to punish norm violations by others. Where many equilibria are possible, group selection is likely to play a major role in determining which equilibrium will obtain. In a population where different groups maintain different internally stable equilibria, each supported by a different norm, those groups following norms that lead to higher group success may be expected to reproduce more rapidly, in which case the behavior predicted by group selection models may predominate.

Evolutionary theory, laboratory experiments and field observations indicate that humans are "social animals" who take a strong interest in the effects of their actions on others and whose behavior is not always explained by simple models of selfish behavior. Does this mean that our familiar analytic tool, selfish old *homo economicus,* is an endangered species? I don't think his admirers have reason to worry. Among modern humans, and probably among our distant ancestors, matching is far from perfectly assortative. While reciprocity and the presence of norms can support a great deal of cooperation, much human activity and most human motivation is impossible for others to observe and, hence, lies beyond the reach of punishment or reward. If you seek empirical evidence that *homo economicus* survives, you need only venture onto a congested freeway, where you will observe his close relatives piloting their gargantuan sports utility vehicles.

References

Bergstrom, Theodore C. 1995. "On the Evolution of Altruistic Ethical Rules for Siblings." *American Economic Review.* 85:1, pp. 58–81.

Bergstrom, Theodore C. 2002. "The Algebra of Assortative Encounters and the Evolution of Cooperation." *International Game Theory Review.* Forthcoming.

Bergstrom, Theodore and Oded Stark. 1993. "How Altruism Can Prevail in an Evolutionary Environment." *American Economic Review.* 83:2, pp. 149–55.

Binmore, Ken. 1992. *Fun and Games.* Lexington, Mass.: D.C. Heath.

Binmore, Ken. 1994a. *Game Theory and the Social Contract. I. Playing Fair.* Cambridge, Mass.: MIT Press.

Binmore, Ken. 1994b. *Game Theory and the So-*cial Contract. II. Just Playing.* Cambridge, Mass.: MIT Press.

Boorman, S.A. and P.R. Levitt. 1980. *The Genetics of Altruism.* New York: Academic Press.

Boyd, Robert and Peter Richerson. 1990. "Group Selection among Alternative Evolutionarily Stable Strategies." *Journal of Theoretical Biology.* August, 145, pp. 331–42.

Boyd, Robert and Peter Richerson. 1992. "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups." *Ethology and Sociobiology.* May, 13, pp. 171–95.

Boyd, Robert and Peter Richerson. 2002. "Group Beneficial Norms can Spread Rapidly in Structured Populations." *Journal of Theoretical Biology.* Forthcoming.

Carr-Saunders, A.M. 1922. *The Population Prob-*

lem: A Study in Human Evolution. Oxford: Claren-don Press.

Cavalli-Sforza, L.L. and M.W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, N.J.: Princeton University Press.

Cohen, D. and I. Eshel. 1976. "On the Founder Effect and the Evolution of Altruistic Traits." *Theoretical Population Biology.* December, 10, pp. 276–302.

Cosmides, Leda. 1989. "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason?" *Cognition.* April, 31:3, pp. 187–276.

Cosmides, Leda and John Tooby. 1989. "Evolutionary Psychology and the Generation of Culture, Part II: A Computational Theory of Exchange." *Ethology and Sociobiology.* January, 10, pp. 51–97.

Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

Edwards, V.C. Wynne. 1962. *Animal Dispersion in Relation to Social Behaviour*. Edinburgh and London: Oliver and Boyd.

Eshel, Ilan. 1972. "On the Neighbor Effect and the Evolution of Altruistic Traits." *Theoretical Population Biology.* September, 3, pp. 258–77.

Eshel, Ilan, Larry Samuelson and Avner Shaked. 1998. "Altruists, Egoists, and Hooligans in a Local Interaction Model." *American Economic Review.* March, 88:1, pp. 157–79.

Eshel, Ilan, Emilia Sansone and Avner Shaked. 1999. "The Emergence of Kinship Behavior in Structured Populations of Unrelated Individuals." *International Journal of Game Theory.* November, 28:4, pp. 447–63.

Fehr, Ernst and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review.* September, 90:4, pp. 980–94.

Frank, Robert H. 1987. "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?" *American Economic Review.* September, 77:4, pp. 593–604.

Friedman, Daniel and Nirvikar Singh. 1999. "The Viability of Vengeance." Technical Report, University of California at Santa Cruz.

Grafen, Alan. 1984. "Natural Selection, Group Selection, and Kin Selection," in *Behavioural Ecology, Second Edition*. J.R. Kreb and N.B. Davies, eds. London: Blackwell, chapter 3, pp. 62–80.

Haldane, J.B.S. 1932. *The Causes of Evolution*. New York and London: Harper & Brothers.

Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior, Parts I and II." *Journal of Theoretical Biology.* July, 7, pp. 1–52.

Henrich, Joseph. 2000. "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon." *American Economic Review.* September, 90:4, pp. 973–79.

Henrich, Joseph and Robert Boyd. 2001. "Why People Punish Defectors." *Journal of Theoretical Biology.* 208:1, pp. 79–89.

Henrich, Joseph et al. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review.* May, 91:2, pp. 73–78.

Hesiod. 1929. H. Evelyn Waugh, ed. *Hesiod: The Homeric Hymns and Homerica*. London: Heineman.

Hume, David. 1739. *A Treatise of Human Nature, Second Edition*. Oxford: Clarendon Press. L.A. Selby-Bigge, ed. Revised by P. Nidditch. 1978.

Kimura, Motoo and George H. Weiss. 1964. "The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance." *Genetics.* April, 49, pp. 561–76.

Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*. John Kagel and Alvin Roth, eds. Princeton, N.J.: Princeton University Press, chapter 2, pp. 111–81.

Levin, B.R. and W.L. Kilmer. 1974. "Interdemic Selection and the Evolution of Altruism." *Evolution.* December, 28:4, pp. 527–45.

Levins, R. 1970. "Extinction," in *Some Mathematical Problems in Biology*. M. Gerstenhaber, ed. Providence: American Mathematical Society, pp. 77–107.

Maynard Smith, John. 1964. "Group Selection and Kin Selection." *Nature.* March 14, 201, pp. 1145–147.

Nachbar, John. 1990. "'Evolutionary' Selection Dynamics in Games: Convergence and Limit Properties." *International Journal of Game Theory.* 19:1, pp. 59–89.

Nisbett, Richard E. and Dov Cohen. 1996. *Culture of Honor: The Psychology of Violence in the South*. Boulder, Colo.: Westview Press.

Nowak, Martin A. and Robert M. May. 1993. "Evolutionary Games and Spatial Chaos." *Nature.* October 29, 359, pp. 826–29.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara and Shmuel Zamir. 1991. "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." *American Economic Review.* December, 81:5, pp. 1068–095.

Rousseau, Jean Jacques. 1755 [1950]. *Discourse on the Origin and Foundation of Inequality Among Men (Second Discourse)*. Everyman's Library, N.Y.: Dutton.

Sethi, Rajiv and E. Somanathan. 2002. "Understanding Reciprocity." *Journal of Economic Behavior and Organization.* Forthcoming.

Skyrms, Brian. 1996. *Evolution of the Social Contract.* Cambridge: Cambridge University Press.

Skyrms, Brian. Forthcoming. "The Stag Hunt." *Proceedings and Addresses of the American Philosophical Association.*

Skyrms, Brian and Robin Pemantle. 2000. "A Dynamic Model of Social Network Formation." *Proceedings of the National Academy of Science.* August, 97:16, pp. 9340–346.

Sober, Eliot and David Sloan Wilson. 1999. *Unto Others.* Cambridge, Mass.: Harvard University Press.

Soltis, Joseph, Robert Boyd and Peter Richerson. 1995. "Can Group-Functional Behaviors Evolve by Cultural Group Selection? An Empirical Test." *Current Anthropology.* June, 36:3, pp. 473–83.

Weibull, Jörgen. 1995. *Evolutionary Game Theory.* Cambridge, Mass.: MIT Press.

Williams, George C. 1966. *Adaptation and Natural Selection A Critique of Some Current Evolutionary Thought.* Princeton, N.J.: Princeton University Press.

Wilson, David Sloan. 1975. "A Theory of Group Selection." *Proceedings of the National Academy of Sciences.* January, 72:1, pp. 143–46.

Wilson, David Sloan. 1979. "Structured Demes and Trait-Group Variation." *American Naturalist.* January, 113, pp. 157–85.

Wright, Sewall. 1921. "Systems of Mating." *Genetics.* March, 6:2, pp. 111–78.

Wright, Sewall. 1943. "Isolation by Distance." *Genetics.* January, 28, pp. 114–38.

Wright, Sewall. 1945. "Tempo and Mode in Evolution: A Critical Review." *Ecology.* October, 26, pp. 415–19.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton, N.J.: Princeton University Press.