The social evolution of human cooperation via group competition and conflict

Brendan O'Connor
Honors Thesis, Symbolic Systems Program
March 28, 2006

I certify that I have read this thesis and that it is fully adequate to receive departmental honors in Symbolic Systems.

______________________________ Noah Mark (advisor)

______________________________ James Fearon (second reader)

Table of Contents

List of Figures

Abstract

Why do humans cooperate within large non-kin societies? A simple answer is that the ancestors of humans genetically evolved cooperative tendencies through millions of years of living in small groups where small-scale cooperation would be advantageous for survival. However, of the 50,000 year history of modern humans, only the last 10,000 have seen explosive growth in the size of human societies, from societies of several dozen members to societies of millions. I seek to explain the evolution of large-scale human cooperation as this increase in social order. I examine two previously underemphasized mechanisms to support cooperation: (1) institutional variance, and (2) war between societies. These two phenomena have been prevalent through the last 10,000 years of human history, and are unique to humans. In contrast to earlier work predominantly concerned with reciprocation and kin selection to support the evolution of cooperation, I develop evolutionary and rational choice models in which between-group competition and conflict help support within-group cooperation.

I. Introduction

Many aspects of our daily lives are products of massive social cooperation. Walking into any of thousands of stores in a city, I am often confident I will not be cheated by a shopkeeper I have never met and may never see again. Walking out of the store, I enjoy public streets paid for, constructed by, and maintained by the joint work of thousands of people. Hundreds of cars zoom past me, their drivers often obeying traffic laws, successfully coordinating to not smash into each other or pedestrians. And if someone fails to cooperate, hitting and injuring me, all is not lost: my life might be saved by a blood bank system, composed of donations by volunteers who receive little benefit in return. And after I survive, I can find recourse through a legal system whose directives millions of people jointly obey.

We live in societies with cooperative and coordinating behavior among millions of members. In the several million years of human history, this is an anomaly. Just

10,000 years ago, no human lived in a society of more than several hundred people; no large-scale cooperation was possible. Why does this astounding level of cooperation exist today?

Social order – i.e., large-scale cooperation – makes modern large-scale societies possible. If people stopped obeying traffic laws, stopped basic trust in everyday commerce, stopped paying taxes, or stopped settling disputes through peaceful means, society would stop working or lose efficiency. But how did this evolution of large-scale cooperation come about in the first place? First, I’ll provide some context by summarizing the literature on the more abstract question of how cooperation can evolve in a world without central authority.

The evolution of cooperation in generic game theoretical agents

In the past few decades, researchers in the social and biological sciences have created a tremendous literature on the evolution of cooperation. The “evolution of cooperation” usually means the emergence of cooperation from a world of selfish non-cooperators, and the stability of cooperation over time. The problem is very sharply defined in evolutionary biology. Imagine a population of selfish, individualistic organisms – the sort one would expect as the result of survival of the fittest. It seems counterintuitive that over time they would ever evolve cooperative or altruistic tendencies. Helping others at a cost to yourself only hurts your own relative reproductive fitness; therefore, cooperative behavior should disappear from an evolutionary system.

The most problematic cases of cooperation can be abstractly characterized as a prisoners’ dilemma (PD). Two players must simultaneously decide to play nice, or look out for themselves – cooperate or defect. Figure 1 illustrates the payoffs for the four different outcomes of a typical 2-person PD. Its first important property is that mutual cooperation has the highest collective payoff – the sum of the players’ payoffs at (C,C) is higher than under any other outcome. If both sides cooperate, they do better as a group than if either or both defect. However, there is a strong temptation for each individual to defect. If the other side is cooperating, you can do better if you defect. If the other side is defecting, you also do better defecting. Therefore, no matter what the other player does, it is better for you to choose defection. Individual incentives run counter to the social good.

Figure 1: A Prisoners’ Dilemma

Each cell has the pair: (*1’s payoff, 2’s payoff*)

		2	
		C	D
1	C	2, 2	0, 3
	D	3, 0	1, 1

If this still seems counterintuitive, imagine you are player one. If your opponent is going to cooperate, then you can choose between the upper left and lower left outcomes. Your choice of C vs. D is a choice between a payoff of 2 vs. 3. If your opponent is going to defect, then you’re in the second column and your choice of C vs. D is a choice between 0 vs. 1. Either way, choosing defection is a better strategy – thus with no knowledge of what your opponent will do, you should choose defection.

Consider if two players meet once, must play a PD, then never meet again. In such a one-shot PD, cooperation should be very difficult. We would expect both players to defect. The outcome of mutual defection is a self-enforcing (Nash) equilibrium: if both are thinking to play D, it is harmful for either player to unilaterally switch their strategy to C. The only Nash equilibrium for the PD is defection by both players.

A common theme in the cooperation literature is to examine how cooperative outcomes can emerge from a world of all defectors. Two major mechanisms have been noted that allow cooperation to outperform defection: kin selection (Hamilton 1964) and reciprocal altruism (Trivers 1971).

The most well-known work along the lines of reciprocity is Robert Axelrod’s study of the evolution of strategies for the iterated prisoners’ dilemma (Axelrod and Hamilton 1981, Axelrod 1984). In an IPD, the two players play a series of PD’s against each other. In contrast to the PD’s two strategies – C vs. D – the IPD has many more, since the decision to cooperate or defect can be conditioned on the history of previous moves. For example, the strategy "tit-for-tat" (TFT) cooperates on the first move, then does whatever the opponent did in the last round. It rewards cooperation and punishes defection. If you are playing against TFT, it is true that you can play D to exploit it for one round, but you then incur a cost: it punishes you with a subsequent D move, which guarantees you a low score in the next round. If you sufficiently concerned care about payoffs in future rounds, you would do best to keep cooperating and therefore maintain the C,C outcome. Thus reciprocity can support cooperation. (TFT is only one example of this, but it has received tremendous attention.)

The evolution of strategies can be modeled via evolutionary game theory (Maynard Smith 1982), where payoffs represent reproductive success. Consider an ecology of IPD players, each with its own strategy. Every round they all randomly pair off to play IPD's. Strategies that tend to win higher payoffs become more common in the population. In an ecology split between the saintly ALL C and the aggressive ALL D, ALL D always has a better fitness payoff, and thus outreproduces ALL C and eventually takes over the population. An ecology mixed between ALL D and TFT, however, may sometimes be taken over by TFT. When paired against an ALL D, it gets exploited only once on the first round, and ties it for the rest. When paired against another TFT, however, both players will profit from mutual cooperation. If the proportion of TFT is high enough, the cooperation benefits from TFT-TFT matchups outweigh the minimal costs of first-round exploitation in TFT-ALL D matchups, thus TFT eventually takes over the population. This is not the end of the story, since TFT itself could be invaded by another strategy (Boyd and Lorderbaum 1987); however, the point is that reciprocal strategies can perform quite robustly and manage to establish high levels of cooperation.

Besides the famous example of IPD reciprocity, many other mechanisms have been found to support the evolution of cooperation from a world of all defectors: altruistic punishment (e.g. Boyd, Gintis, Bowles and Richerson 2002), shared reputation information (e.g. Milgrom, North and Weingast 1990), geographical clustering (e.g. Nakamaru, Matsuda and Iwasa 1997), self-similar biases (e.g. Riolo, Cohen and Axelrod 2001), inclusive fitness such as kin and group selection, and many more. We will examine inclusive fitness in Section II.

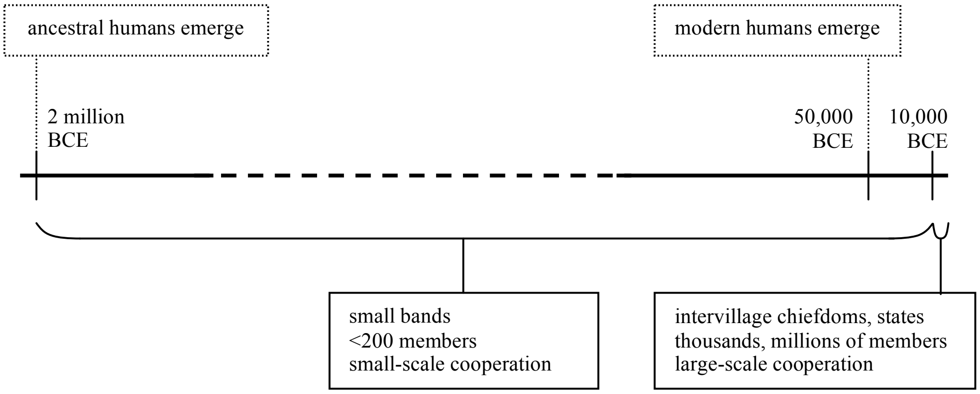
A very recent 10,000 years of large-scale human cooperation

What do these findings mean when applied to the question of the evolution of human cooperation? Compared to all other animal species, humans possess uniquely complex cognition, linguistic communication, and culture. The behavior of all animals, including humans, can adapt through the process of genetic evolution. But the behavior of humans can also adapt to a through cultural learning and transmission – and do it magnitudes faster than biological evolutionary time. If we want to understand the evolution of human cooperation, we need to consider both the biological evolution of prosocial tendencies, and the cultural evolution of cooperative behaviors.

The mechanisms leading to the emergence of cooperative strategies are useful ways to understand many specific human social situations. They also can be useful to understand genetic evolution in biological organisms, and perhaps even human biology.

However, in this paper I will focus on one very important example of human cooperation: human societies. Human societies predate the emergence of homo sapiens, and homo sapiens have lived in societies for as long as they have lived.

Figure 2: Timeline: human biology and social organization imperfectly correlate



50,000 years ago, the most modern human species started becoming widespread, and only 10,000 years ago, the first village and intervillage societies formed (Lenski 1970, Diamond 1999). Since then, human societies have been growing larger and larger. 10,000 years ago, no human lived in a society of more than a few hundred people. Today, almost all humans live in societies of thousands, millions, or even billions of people.

Societies can exist only with social order – public goods and cooperation. Large societies are a consequence of large-scale cooperation. This dramatic growth of societies is only possible with larger and larger scales of cooperation. Why did this ever happen? Why did not humans just keep living as hunter-gatherers for another mere 10,000 years? How did this example of the evolution of cooperation come about? The previously described literature on the evolution of cooperation is somewhat unhelpful here. Can TFT invade an all defector population? That is irrelevant because there never was an ALL D phase of human history – not among individuals. Individual humans always lived together in hunter-gatherer groups. ALL D is implausible as the setting before the evolution of large-scale societies, since basic innate propensities to cooperate had already genetically evolved by 10,000 years ago. In fact, basic prosocial cultural norms probably existed as well. Modern-day hunter-gatherer societies display high levels of interdependence and exchange, which suggests that ancestral humans did as well (e.g., egalitarian meat-sharing practices, reviewed in Wilson 1998). Starting from internally cooperative small bands and villages, how did we get intervillage and larger societies?

Carneiro (1970, 1998) provides a compelling account based on conflict between societies. He notes that early human settlement was often characterized by the existence of many village societies together in one geographical region. When villages bump up against each other, raids and conflict can happen between them. If resources and space are limited, then increasing population guarantees overcrowding and more conflict. Many original states did in fact arise in geographically circumscribed areas, where surrounding mountains or ocean confined humans in narrow habitable areas.

Increasing levels of conflict, in turn, led to larger political units in at least two ways. (1) An aggressive village can conquer another (Carneiro 1970) and thus form an intervillage society – a chiefdom. An alternative explanation is (2) military alliances between villages evolve into political units, headed by the military leader of the joint forces (Carneiro 1998). Carneiro cites suggestive contemporary and historical examples of warfare at the time of chiefdom formation.

In a world of scarcity-fueled conflict, large societies are selected for. Without conflict, human groups would remain at a fairly optimal size for hunting and gathering. Since very few hunter-gatherer societies were larger than a few hundred people, it stands to reason that larger sizes groups would have been ineffective at survival; they would go extinct or break up. But conflict introduces an incentive to grow larger, to beat out other societies and defend better against larger ones.

If cooperation is key to societal growth and/or strength, then cooperative societies are selected for. Societies with social disorder go extinct, since they cannot compete against more internally cooperative societies. And when some societies become more cooperative and larger, other societies have to get even more cooperative and even larger to compete.

In Section III, I present Wilson's ([1989] 1994) model to show that variance of cooperativeness across groups can encourage the evolution of cooperation. Since cultural institutions have tremendous variance across societies, this can to a certain extent explain increases in cooperation among competing societies. In Section IV, I introduce between-group conflict into the model and analyze how it could encourage within-group cooperation. Conflict functions as a way of intensifying group competition. I also develop a complementary decision-making model of how conflict can solve within-group prisoners' dilemmas. I offer both the rational and evolutionary models of cooperative behavior as possible formal explanations of large-scale cooperation and human societies along the lines of Carneiro’s theory.

In this paper I hope to make three contributions. The first is to recast the evolution of human society in formal terms of the evolution of cooperation. This seems natural since society is an important example of large-scale cooperation. Second, I want to reinterpret and modify formal cooperation evolution theories into explanations of the evolution of large-scale human cooperation – that is, make a model that explains stylized facts we know. My third goal is to explore two particular mechanisms to encourage within-group cooperation: institutional variance and between-group conflict. Of course, the models and analysis here do not fulfill this entire research agenda, but hopefully are advances on which future work can build.

II. Evolutionary Dynamics and Human Cooperation

In this section I justify why it is reasonable to model changes to human societies as an evolutionary process, and explain the specific formulation used in this paper, the proportional fitness rule. I also review the inclusive fitness explanations of cooperation – kin selection and group selection – and argue that group selection is a useful explanation for the cooperation in human societies, especially during their dramatic increases in size, complexity, and cooperation in the last 10,000 years.

Human societies can change over time via evolutionary processes

As has been stated before, human societies have demonstrated tremendous changes in the last 10,000 years. One type of change is growth. Societies have become larger in terms of both population and occupied land.

Another type of change concerns a society’s internal institutions, structure, and culture. Human behavior is heavily influenced by culture, norms, beliefs, skills, and social structure. Culture and institutions can change over time, which changes people’s behavior over time.

Finding a systematic explanation of these changes is a goal of this thesis. The explanation explored here is social evolution, meaning a process in which certain societies and cultural practices survive and become more prevalent.

A few clarifying notes. First, social evolution has nothing to do with biological evolution of genes. Social evolution consists of changes to cultural and social institutions. It acts fast enough to have substantial effects over 10,000 years, whereas genetic evolution does not.

Secondly, we need a definition of the term “evolution.” It can sometimes mean “gradual change over time,” as in “pre-revolutionary society started to evolve unstable practices.” This is a general notion of “evolution.” When I use the term in this paper, I intend to give it a much sharper meaning: evolution describes a process in which *the most fit types of reproductive units become more common or prevalent later in time* (Bendor and Swistak 1997). This definition includes Darwinian natural selection among organisms, where fitness is measured as reproductive success: the most fit organisms have the most offspring, and thus that species becomes more common in the next generation.[1]

With some care, this definition of evolution can also be applied to societies and cultural changes. Societies with higher fitness tend to survive or grow at the expense of others. Societies with low fitness have a harder time surviving; societies that stagnate, shrink, or become destroyed in conflict all have low fitness.

Fitness is reproductive success, but we are not concerned with the odd concept of societal reproduction. Rather, societal *traits* reproduce themselves when their societies survive or grow. Such traits become more prevalent. Thus, traits that cause survival or growth are societal traits with high fitness. If we say a society has high fitness, that merely means its traits tend to become more prevalent in the world.[2]

Size and culture can affect a society's fitness. Size can increase fitness[3] by providing more land and human resources with which a society can compete against others – getting scarce resources, growing faster, or dominating through conflict. Since size causes survival, larger societies tend to survive, and thus become relatively more common as smaller societies become less common. Similarly, certain cultures and institutions can increase a society's fitness. All other things equal, a strong government that coerces members to contribute to efficient public goods would increase the fitness of a society. Or, certain norms of altruism might increase efficiency in production. For example, a norm of helping people in need might solve inefficiencies caused by the fact that people with disease need temporary help to survive. Caring for your society's sick members has a cost to caregivers, but a high benefit of production from members after

---

[1] Besides natural selection, a number of other forces can cause genetic change over time; for example, genetic drift (random changes of genetic makeup that have no bearing on reproductive fitness) and genetic migration (caused by movement of populations). Evolution is sometimes considered as the totality of all these forces; here we only examine selection.

[2] To be precise: a trait's prevalence in the world is defined as its frequency among societies, weighted by society size. If you select from the world a random individual, the probability that individual has trait X is trait X's prevalence in the world. Thus, doubling one society's size has the same impact on societal trait prevalences as the appearance of an new identical society of the same size.

By this definition, a society's survival positively influences the prevalence of its traits (relative to its not surviving). But furthermore, if a society grows, that also positively influences the prevalence of its traits. This definition allows us to simultaneously examine societal survival and growth (or extinction/shrinking).

[3] This is not tautological. By our definition, one way a society is "fit" is if it grows/survives. Therefore growth causes size, but the reverse is not necessary: large societies could conceivably be selected against.

they get well. Thus, this particular norm would increase the society’s fitness. It is easy to think of (and dispute) many cultural and institutional elements that could increase or decrease a society’s fitness.

The proportional fitness rule describes a class of evolutionary processes

We have defined an evolutionary process to be one where greater fitness leads to greater prevalence in the population. But this is vague: how much more prevalent does a more fit trait become?

Before considering social evolution, it is instructive to consider the biological evolution case. Many biological evolutionary models use the specific formulation of linear replicator dynamics, specified as follows. A simple form consists of a population of asexually reproducing organisms, which transmit their genetic traits perfectly to children. Imagine a population of 100 such organisms. 50 have a genetically transmitted strategy A that gives them a high fitness value of 2. The other 50 do not have the trait (‘B’), and reproduce only at a fitness value of 1. After one generation, the A’s double in population to 100, while the B’s are still at 50. The proportion of the population with the successful trait will changed from 1/2 to 2/3.

Let p be the percentage of the population with the trait A. The fitness of trait A is $W_A = 2$, and the fitness of not having the trait is $W_B = 1$. The average fitness of the group is the fitness of each trait, weighted by their proportions in the population: $W = pW_A + (1 - p)W_B$. The initial average is $W = .5*2 + .5*1 = 1.5$.

The update rule for the prevalence of strategy ‘A’ is called the Proportional Fitness Rule (PFR):

$$p' = p\frac{W_A}{W}$$

(so here, our $p' = .5 * 2/1.5 = 2/3$)

The PFR follows from a description of the replicator dynamics of population reproduction, but can be generalized. It does not mention the population size, but only posits that whatever is more successful will appear more in the future at a rate proportional to its fitness, normalized by the relative fitness of all strategies in the population. A trivial example: say that everyone in a constant sized population independently decides at each round whether to be A or B. If they choose A at probability pW_A / W then the expected population dynamics follow the PFR, though individuals do not reproduce.

Noting that the biological population origins of the PFR do not seem applicable to human social phenomena, Bendor and Swistak (1997) suggest a justification for the PFR as modeling behavior transmission. If actors wish to change to more optimal strategies, but determine optimality by evaluating current and past predominance in the population, then strategies in the population can be selected for via the PFR.[4]

This paper sticks to the PFR to analyze the evolution of size and culture within human societies. This might seem problematic, but at least it is a model well-understood in the evolutionary literature. There are a number of mathematical models of cultural transmission that have more realistic features (e.g. Cavalli-Sforza and Feldman 1981, Rogers 1988, Boyd and Richerson 1985), but the PFR is extremely simple and amply developed.[5]

Group and kin selection can explain cooperative behavior

The previous description of a population under evolution had selection working at the individual level: individual organisms reproductively competed against one another. However, selection can occur among reproductive units at different levels. Richard Dawkins argues that biological evolution is best viewed in terms of the "selfish gene" (Dawkins 1976), where evolution occurs not as selection among organisms, but rather as selection among genes. It is true that ultimately all biological evolutionary processes are grounded in competition among genes. However, it makes sense to talk of individual

---

[4] Of course, this explanation of replicator dynamics only works if people evaluate current and past performance in specific ways detailed by Bendor and Swistak. On the topic of what sorts of adaptive rules mimic replicator dynamics, see (Weibull 1995).

[5] Besides, it is not entirely clear that there exists an effective and ready-to-use set of dynamics to better model the evolution of societies. Greif (2005) examines a number of different frameworks to use for his theory of institutional change, including rational choice (classic game theory), evolutionary game theory, and individual learning game theory. Noting weaknesses in all of them, he builds his own version of classic game theory plus changing game parameters ("quasi-parameters").

selection among organisms because each individual is a collection of genes whose survival is bound together with the fate of the organism. A particular gene's survival is heavily correlated with the individuals it is present in, and thus the other genes it lives with. They are all in the same boat.[6] Thus talking of the fitness of an organism is a useful summary of gene fitnesses.

Since altruism is so hard to explain via individual selection, group selection has been proposed as a level of selection among groups of individual organisms. This can explain altruism: if you live in a group with other individuals, it can be beneficial to help the group even at a cost to yourself. Consider a proportional fitness rule among groups: groups with more cooperation outperform groups with less cooperation. For example, an internally cooperative colony of bacteria would outcompete other bacteria colonies.

This form of group selection is regarded by most biologists as too weak for genetic evolution in many cases. Small amounts of migration between groups or substantial levels of selection within a group will easily derail group selection. Critiques of group selection (e.g. Williams 1966; summary in Henrich 2004) have been influential in genetic evolution theory.[7]

Kin selection is an alternative altruism explanation. It pays to behave altruistically towards your kin because they share genes with you. Kin selection is an excellent explanation for extreme altruism observed in the social insects (bees, ants), but not so great for human cooperation, where cooperation often occurs among non-kin. This is especially true among the larger and more complex societies that have developed in the last 10,000 years.

---

[6] In this view, advocated in the beginning chapters of Sober and Wilson (1998), the gene level is lower than the individual organism level. Yet it is also higher than the individual level, since a particular gene type has copies among many different individuals. (Many genes are in the same boat, though each gene is in many boats.) Kin selection can be viewed as a particular gene instance helping its copies in other organisms. This is an important consideration when dealing with many different genes influencing a single individual; however, the models developed here will have only one reproductive trait, predilection for cooperation.

[7] I will not attempt an in-depth treatment of the debate, but it should be noted that many of these critiques are intended specifically for certain issues in genetic evolution and do not readily apply to human cultural evolution. For example, cultural change is much faster than genetic change, and between-group cultural differences are quite large. See (Henrich 2004) for a more thorough discussion of group selection for cultural evolution, and (Sober and Wilson 1998) for a group selectionists' overview of the biological debate.

Besides kin selection, the other major explanation of the evolution of cooperation put forth has been reciprocity: cooperate with others as long as they keep cooperating, and punish those who do not cooperate. Reciprocity can explain cooperation among non-kin, but it does not scale well to big group games like the n-person prisoners’ dilemma, since it becomes harder to punish individuals (Boyd and Richerson 1988). In a big society where you do not have the chance to retaliate against individuals, altruistic punishment – and more generally reciprocity – cannot work without formal structures to enforce punishment against defectors.

Group selection among societies can explain human cooperation

Group selection among human societies can explain the evolution of cooperation because (1) humans live in societies whose members share cultural and social traits that help determine cooperativeness, (2) there is variance in these traits across different societies, (3) cultural and social traits can persist in a society over time, and (4) societies are groups in competition with one another, while within-group cooperation increases group fitness.

Societies are a reasonable evolutionary unit for cultural institutions. As (1) and (2) state, culture (and other intergenerationally transmitted determinants of behavior, such as political institutions) is shared much more within members of one society than between societies. Given two random members of the same society, the expected difference in cultural norms should be less than for two people selected from different societies. Furthermore, culture is transmitted within a society across generations, and persists beyond individual persons, since new members are socialized into the group’s culture. Cultural institutions can persist for hundreds or even thousands of years, organizational cultures in businesses persist despite employee turnover, and so forth.

Since societies are a reasonable unit of selection, the next necessary element for group selection is competition and variance among groups. This is true of the evolution of societies in the last 10,000 years, where competition and cultural variance have been the norm.

Human societies exhibit a bewildering variance in cultural institutions and practices, including variance in cultural norms for cooperation. A series of experimental field studies of ultimatum and public goods games across 15 small-scale societies found tremendous cultural variance in rates of altruistic gift giving and norms of fairness (Henrich et al 2005). Also keep in mind that the cultural variance seen in today's well-known societies should be less than cultural variance through history, since today's observations already select for the dependent variable, survival. Societies with non-cooperative cultural practices have already gone extinct.

Competition and conflict seem to be endemic to the relations between human societies. The number of human societies has been drastically decreasing for the last 10,000 years. Extinction or absorption has been the fate of most societies. In the societies that mangae to survive, we should see the result of enormous competitive pressures.

Keely (1996) finds that warfare was prevalent among prehistoric humans – death rates from conflict were much higher in prehistoric societies than in the modern ones. Soltis et al. (1995) survey empirical data of many New Guinea societies, and find numerous examples of conflict causing group social extinction. An often cited example of cultural societal selection is the 19th century Nuer conquest of the Dinka in eastern Africa. The Nuer and Dinka had similar technology, habitat and resources, but different cultural practices. The Nuer's bridewealth customs helped maintain larger tribal systems and therefore larger military power; this and other differences in cultural practices led to expansion at the Dinka's expense (Kelly 1985).

We have seen that reciprocity and kin selection mechanisms cannot explain the dramatic increases in large-scale cooperation and society growth of the last 10,000 years. Group selection for cooperative culture and society size, however, seems one reasonable mechanism.[8] In the following sections, we will interpret a standard group selection model as a model of societal competition, and build a mechanism of conflict-driven competition to explain within-society cooperation.

---

[8] From the evolutionary psychology perspective, other authors have used the hunter-gatherer group conflict/selection argument to explain human genetic evolution of cooperative tendencies (Henrich 2004, Boyd and Richerson 2005), or even social intelligence and genetic tendencies toward certain political preferences (Rubin 2001). I make no such claims; the argument here is solely about social and cultural evolution, not human genetic evolution.

III. Cooperator Variance

Group-level selection is an intensely debated explanation for the evolution of cooperation. Since cooperators within a group help one another, could group selection explain the development of large-scale cooperation in human societies? The most well-known group selection mechanism is periodic dispersion and regrouping of individuals – “haystack” models (Maynard Smith 1964, reviewed in Bergstrom 2002). This model is inspired by the reproductive and social habits of various organisms, but does not seem to fit the case of human societies. On the other hand, perhaps other mechanisms could sustain human societal cooperation through group selection. As a baseline, I briefly explore random institutional change: if groups randomly change their cultural institutions influencing cooperation, that could sustain between-group cooperator variance to promote cooperation. Therefore, a variety of variance-supporting mechanisms should be plausible.

This section has another purpose: the group selection model developed here will be generalized in Section IV. Group selection is just one example of group competition effects.

Defectors always win in the basic evolutionary one-shot PD

The following is an adaptation of D.S. Wilson's model of cooperation via gift-giving altruists ([1989] 1994; Sober and Wilson 1998). A group of size N is comprised of altruists (cooperators) and defectors. p is the proportion of cooperators. All individuals have a base reproductive fitness of X. An individual altruist gives a gift to everyone in the group. The gift provides a fitness benefit of b/N to each individual, and incurs a personal cost c to the altruist. Defectors receive benefits just as much as cooperators, but do not contribute anything themselves. Assume $b>c>0$ so that cooperation is Pareto-efficient: the group does better if everyone is a cooperator, even though an individual does better as a defector. (This is an n-person PD.) Thus the reproductive fitnesses for cooperators and defectors within the same group are

$$W_A = X + bp - c$$
$$W_D = X + bp$$

Defectors always do better than cooperators, and under the PFR will take over any mixed population given enough time. There is only one evolutionarily stable equilibrium, at all defectors (p=0).[9]

Segregation can sustain cooperation in the short term

A counterintuitive result illustrated by Wilson is that in a system of multiple groups, it is possible to have an increase in the altruist proportion of the global population, though cooperator ratios are declining in each group. This can be achieved through concentrating cooperators in one group: that group will grow faster than a relatively defector-heavy group.

Consider two equally sized groups:

$$N_1 = 100,\ \ p_1 = 0.8$$
$$N_2 = 100,\ \ p_2 = 0.2$$

At first, the global cooperator count, defector count, and cooperator proportions are

$$\#A_{all} = 100,\ \ \#D_{all} = 100,\ \ p_{all} = 0.5$$

With X=1, c=0.1, b=0.5, at the next timestep, internal proportions of altruists decline: p'_1=0.78 and p'_2=0.19. However, the global proportion of altruists increases:

$$\#A'_{all} = 124,\ \ \#D'_{all} = 116,\ \ p'_{all} = 0.517$$

---

[9] Another stipulation: $b/N < c$ otherwise in small populations it may be the case that an individual does better as an altruist than as a defector. The parameters used here fulfill that. Sober and Wilson (1998) and Cooper and Wallace (2004) avoid this problem by having altruists give to everyone but themselves, so per-agent benefits are $b/(N-1)$. This complicates the fitness equations, so I have chosen the other route (used by Wilson (1975)) here.

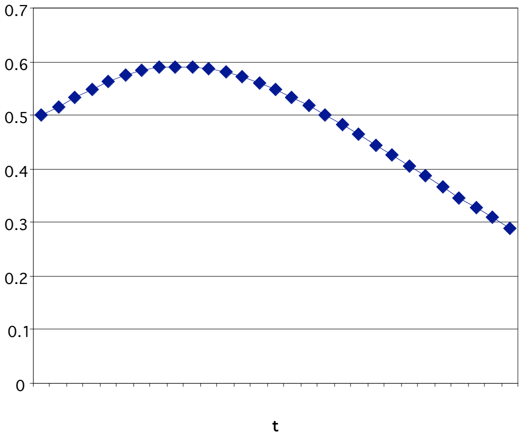
This is possible because the cooperator-heavy first group grows substantially more than the second group:

$$N'_1 = 132,\ N'_2=108$$

Since most of the cooperators in the system live in group 1, most of the cooperators receive the benefits. If we characterize each *b/N* gift delivery as a social interaction tie, then segregation maximizes the number of A-A and minimizes A-D interaction ties. This ensures that most of the gift benefits end up being received by fellow altruists instead of freeriding defectors. This is the same reason cooperation is encouraged in studies that embed agents in a lattice, where they only interact with neighbors. Clusters of cooperators can result.

These increases, of course, cannot continue indefinitely, since both groups must approach all-defector. In this example, the overall altruist proportion peaks after 10 rounds.

Figure 3: Cooperation increases for a time



How much segregation is needed? The distribution equation

We can more precisely specify the notions of altruist concentration and segregation that can help sustain cooperation, at least in the short run.

The previously introduced variables A_{all} and D_{all} for global cooperator and defector populations allowed us to think about the population of a certain type of individual across all groups. We can also think of the distribution of a certain type into each group. Let a_i signify the proportion of the system's cooperators that live in group i, and similarly d_i is the proportion of all defectors there. These variables let us view the population of cooperators across the group boundaries that constrain their interactions.

We now can define the average altruist reproductive fitness as the weighted average of altruist fitness across both groups:

$$W_{A,all} = \sum_{i=1}^{m} a_i W_{A_i} = \sum (a_i X + b a_i p_i - c a_i)$$

$$W_{D,all} = \sum_{i=1}^{m} d_i W_{D_i} = \sum (d_i X + b d_i p_i)$$

for an arbitrary m number of groups. Cooperation is increasing globally when

$$\begin{aligned}
W_{A,all} &> W_{D,all} \\
\sum [a_i X + b a_i p_i - c a_i] &> \sum [d_i X + b d_i p_i] \\
\sum [(a_i - d_i)X + (a_i - d_i) b p_i] &> \sum c a_i \\
X\left(\sum a_i - \sum d_i\right) + b \sum (a_i - d_i) p_i &> c \\
\sum (a_i - d_i) p_i &> \frac{c}{b} \qquad (1)
\end{aligned}$$

The left hand side of (1), $\sum (a_i - d_i) p_i$, is the aggregate of per-group altruist-defector distribution differentials, weighted by altruist proportion. It measures the “unevenness"” of the altruist and defector distributions, and weights them by their significance per group.

For the two group case, this becomes

$$(a_1 - d_1)p_1 + ((1 - a_1 - (1 - d_1))p_2 > c/b$$
$$(a_1 - d_1)(p_1 - p_2) > c/b$$

Say $a_1 > d_1$ and $N_1 = N_2$ so both terms are positive. Cooperation is better sustained the more altruists are concentrated in group 1, and the more defectors are concentrated in group 2.

Alternate formulation via statistical variance (the Price equation)

A standard method of performing this distributional analysis of reproductive traits across groups is via the Price equation (1970, 1972), which decomposes the selection of any trait into individual and group selection components.

E is statistical expectation, *Cov* is covariance:[10]

$$\Delta p_{all} = E[\Delta p_i] + Cov[W, p] / E[W]$$

The first term is the average change in the trait proportion in each group, weighted by size of the group, and the second term measures the extent to which a group's fitness co-occurs with a predominance of the trait. Consider the group selection of genes at the organism level, where it is beneficial for individual genes in an organism to cooperate. Since an organism is a group of associated genes, if the organism does well, its component genes does well. This type of group selection is strong because a particular gene's survival is heavily associated with the survival of organisms it is present in. Price's second term exactly specifies "association": a particular trait's survival is covariant with the fitness of the groups it tends to belong to, normalized by the average fitness of an individual of the population. If those groups where it often appears do well, the gene does well.

For our case, group fitness is a direct linear function of the occurrence of the altruist trait. *Np* altruists each add *b* and take away *c* from the group reproductive fitness;

---

[10] In Section IV, $E[W]$ is also W_{all}.

thus a particular group's fitness $W = X + (b\text{-}c)p$. Therefore, the $Cov[W,p]$ term measures the variance in the p distribution. Specifically,

$$\begin{aligned} Var[W + p] &= Var[W] + Var[p] + 2\,Cov[W,p] \quad \text{(from definition of covariance)} \\ Cov[W,p] &= 1/2\ \ (Var[W+p] - Var[W] - Var[p]) \\ &= 1/2\ \ (Var[X + (b-c+1)p] - Var[X + (b-c)p] - Var[p]) \end{aligned}$$

Since for any constants a,b and random variable R, $Var[aR+b] = a^2\ Var[R]$,

$$\begin{aligned} Cov[W,p] &= 1/2\ \ ((b-c+1)^2 Var[p] - (b-c)^2 Var[p] - Var[p]) \\ &= 1/2\ \ ((b-c+1)^2 - (b-c)^2 - 1)\,Var[p] \\ &= 1/2\ \ ((b-c)^2 + 2(b-c) + 1 - (b-c)^2 - 1)\,Var[p] \\ &= (b-c)\ Var[p] \end{aligned}$$

Therefore we have change in total cooperator proportion as

$$\Delta p_{all} = E[\Delta p_i] + (b-c)\,Var[p]\,/\,E[W]$$

The second term – the change to p_{all} due to group selection – increases with greater statistical variance (expected squared distance from the mean) of per-group altruist proportions. The first term, describing individual selection, will always be negative, since this is a PD and defectors always beat out cooperators within a group.

In order for cooperation to increase, the p_i distribution must have substantial variance.

Periodic regrouping and dispersion can sustain cooperation via group selection

An entire literature of “haystack models” has analyzed populations where individuals form small groups to reproduce, then disperse and regroup again. (See Bergstrom 2002 for an overview.) The term was coined by Maynard Smith (1964) who imagined a population of mice with either cooperator or defector strategies for the PD. The mice would be separated into a number of haystacks at varying ratios of cooperators, and reproduce inside them for several generations. After several generations, the groups would disperse, then form back up into groups again. During the haystack phases, the group selection effect leads to a global rise in cooperators. This is what happens on the upward slope of the p_{all} curves shown in the previous section. The dispersal and regrouping phase prevents defectors from taking over the heavily cooperator groups – which is what causes the decline in p_{all} here. To be precise, say every 5 rounds you redistribute cooperators and defectors into the m groups again, but at the old a_i and d_i proportions. In that case, since p_{all} has increased, every p_i will be higher than before. For each group i, $(a_i - d_i)p_i$ must increase, so the left side of the $\sum(a_i - d_i)p_i > c/b$ condition only increases. This dispersal and regrouping mechanism therefore can sustain cooperator growth forever.[11]

While this model is very elegant, it is not clear that it can explain human cooperation. While there is some evidence for society dispersals in some very small societies (Soltis, Boyd, and Richerson 1995), it is probably too rare an event to sustain the evolution of cooperation in this way. If the typical person lives in the same society for his/her entire life, haystack models should not apply.

Furthermore, unbiased random migrations or unbiased regroupings do not cause the haystack effect: any between-society movements or regroupings have to have the effect of segregating cooperators and defectors. It is not clear that any human social mechanism does this at the societal level.

---

[11] Bergstrom (2002; summarized in Skyrms 2004: 6-9) analyzes the game played by the founding members of a group, where payoffs are the total offspring accumulated over the within-group iterations. This game is a Stag Hunt, where cooperation can be a stable equilibrium – unlike the PD. Section IV presents another example of transforming a PD into a Stag Hunt.

Random cultural/institutional changes could sustain cooperation via group selection

The proportional fitness dynamics of gift-giving, and the eventual invasion of defectors, could be interpreted in two ways. The first is reproductive selection. Recall that we presented replicator dynamics as asexual genetic transmission from parents to offspring. A better interpretation for human cooperation is the following. Each agent in the population represents an entity that is capable of intergenerationally transmitting a bias towards cooperation – say, a family. The group then represents many of these entities playing a public goods game with each other – say, a village of many families. Families that do not cooperate tend to do better than ones that do, and their numerous offspring (later generations) are similarly selfish.

Another interpretation of the PFR within groups is cultural learning, as explained in section II summarizing Bendor and Swistak (1997). Over time, people start to realize being selfish pays off, so they start switching to defection. With this interpretation, we would still have a replicator-like explanation for the PFR with intergroup selection: groups with higher cooperation rates grow faster (more production) than other groups.

Considering this cultural learning interpretation, it is evident that individuals do not always slowly learn over time. Occasionally large shifts can happen in cultural institutions. New leaders might alter policies or attitudes, or they may change for exogenous reasons. These exogenous shifts could change the propensities of cooperation in the population.

Such random variations would have the effect of increasing cooperator variance across groups. Higher cooperator variance means $\sum (a_i - d_i) p_i$ is high, and thus cooperators grow more. Unfortunately, I have found it difficult to formally model a process to generate cooperation in this manner. When random variations cause an increase in the global cooperator proportion, it is often partially due to a gross increase in cooperators from selecting a large, low-cooperation group for a random increase. More work is needed to disentangle the effects of cooperator variance increases versus changes to cooperator population directly caused by the random changes.

Besides the random institutional change interpretation, random variance should be only a baseline existence proof: if random variance, which does not favor cooperator-heavy groups, can cause cooperation, then potentially many variance mechanisms that favor cooperator-heavy groups should be able to do it too.

IV. Between-Group Conflict

In trying to explain the evolution of large-scale cooperation in the last 10,000 years, the previous section noted two facts – that humans are grouped into societies within which they cooperate, and that societies have cultural institutions that vary between societies and across time – and analyzed how these facts can encourage the growth of cooperation. We now note a third fact: human societies are often in conflict with one another. This can elevate levels of cooperation since cooperation is essential to prevailing in conflict.

First, I examine how conflict can lead to rational decisions for cooperation. Individuals may decide to cooperate when their well-being is tied to their group's performance in conflict. Second, I explain how group conflict is a particular example of a more general phenomenon of group competition. The group selection model developed in section III is analyzed as a group competition model similar to the rational cooperation/conflict model. Finally, I add explicit conflict to the group selection model, and find that higher levels of conflict lead to more cooperation. There is selection for groups with high internal cooperation that do better in conflict, whereas groups with low cooperation do poorly and may be destroyed.

Evolutionary and rational choice frameworks illustrate different aspects of human behavior. Rational frameworks emphasize individuals' powerful cognitive capabilities to find and exploit the best decisions. Evolutionary[12] frameworks also emphasize individual efficiency, but movement towards equilibria takes time, mediated by rates of learning and cultural transmission. Since conflict can encourage cooperation under both, that suggests it is a possible mechanism for the evolution of human cooperation.

Conflict can lead to decisions for within-group cooperation in the PD

To illustrate how conflict can promote decisions to cooperate, consider the rational choice characterization of a game. Players make decisions to maximize payoffs. In a two person PD, both players choose to defect. For either player, no matter what the other

---

[12] An evolutionary model can be interpreted as social/cultural evolution, learning, or genetic evolution. Like section III, this section uses the social/cultural interpretation, not the genetic one.

player does, he/she does better defecting. The best reply arrows in the payoff matrix illustrate this fact. When player 1 cooperates, player 2 would like best to defect, thus a right arrow on the top row. When player 1 defects, player 2 would like best to defect, thus a right arrow on the bottom row. The same reasoning yields the downwards best reply arrows for player 1.

Figure 4: PD illustrated with best-reply moves

$$\begin{tikzcd}
 & & 2 & \\
 & & C & D \\
1 & C & 2,\ 2 \arrow[r] \arrow[d] & 0,\ 3 \arrow[d] \\
 & D & 3,\ 0 \arrow[r] & 1,\ 1
\end{tikzcd}$$

Figure 5: More general version of section III’s gift-giving characterization, with b>c>0.
Using X=1, b=4, c=3 for the W_A and W_D fitness equations
yields the above payoff matrix we use here.

$$\begin{tikzcd}
 & & 2 & \\
 & & C & D \\
1 & C & \begin{array}{c} X + b - c, \\ X + b - c \end{array} \arrow[r] \arrow[d] & \begin{array}{c} X + b/2 - c, \\ X + b/2 \end{array} \arrow[d] \\
 & D & \begin{array}{c} X + b/2, \\ X + b/2 - c \end{array} \arrow[r] & X,\ X
\end{tikzcd}$$

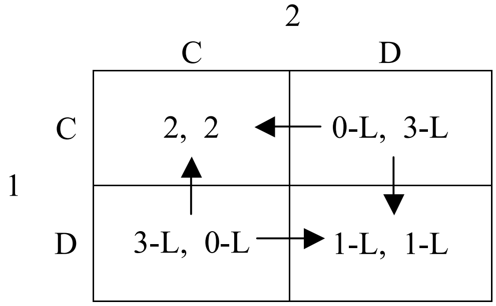
Multiple groups. Following Hausken (1995), consider two pairs of individuals. The two members of each pair play a within-group, 2-person PD with one another. For each group, the Nash equilibrium is the standard (D,D). We can also look at this scenarios as a 4-person game. This game also has only one equilibrium: (D,D, D,D).

Between-group conflict can increase the possibilities of cooperation under two conditions: (1) conflict hurts the losing side, and/or (2) conflict helps the winning side. To start, say an all-cooperator group attacks if the other group has at least one defector, and the conflict’s only effect is to harm the players in the defending group. This is Conflict Rule #1:

> CR1: If one group is all-cooperator and the other is not, the CC group attacks, and each player in the second group incurs a loss of L>1, or more generally L > c - b/N. (c - b/N is the freeriding gain: the benefit to switch to defection in a group size N.)

Under this rule, if the other side is cooperating, and your side is also cooperating, it is unwise to switch to defection. Since L>1, the personal cost of triggering warfare outweighs the gains of freeriding. You will do better than your partner, but not as well as if you had stuck to cooperation. If your partner is cooperating and you are defecting, you should switch to cooperation to stop the attack. The within-group game is now a Stag Hunt, where the cooperative outcome is a second self-enforcing equilibrium.

Figure 6: Within-group stag hunt under CR1 when other group is (C,C)



This is contingent on the second group being (C,C). Instead of thinking of two subgames for each group, it useful to view the entire four person game. Though we cannot show the 2x2x2x2 payoff matrix (16 total states), we can draw a best reply diagram among the 6 unique states of the game: (CC CC), (CD CC), (DD CC), (CD CD), (CD DD), (DD DD). Figure 8 illustrates the best reply (unilateral move) for each player at each state. The six states are ordered from the highest number of cooperators on the left, to lowest number of cooperators on the right. Therefore a best reply arrow to the left represents a shift to cooperation, and a best reply arrow to the right represents a shift to defection. Figure 7 illustrates this notation for a standard PD. CR1 makes (CC CC) self-enforcing, and adds a best reply from (DC CC) to it as well.

Figure 7: Best replies in the three unique states of a standard 2-player PD

One NE: (D,D)

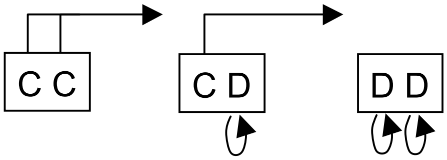


Figure 8: Best replies in the 4-player game with no conflict. Cooperators always switch to defection.

One NE: (DD DD).

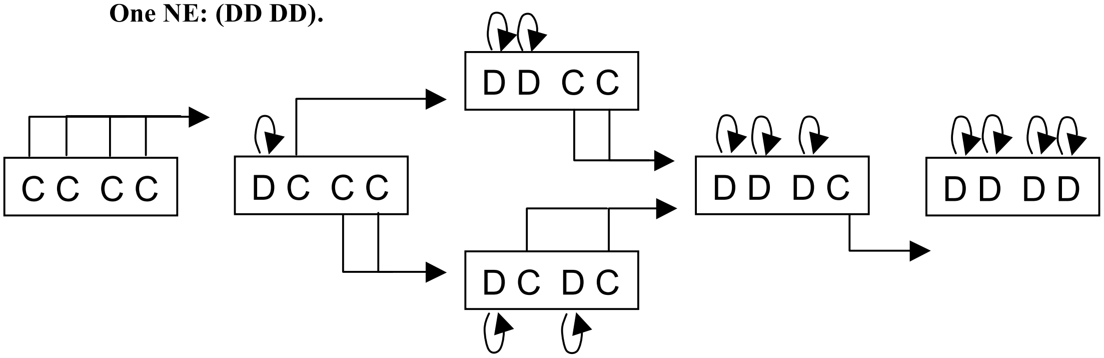


Figure 9: Best replies in the 4-player game with CR1.

Two NE's: (CC CC) and (DD DD)

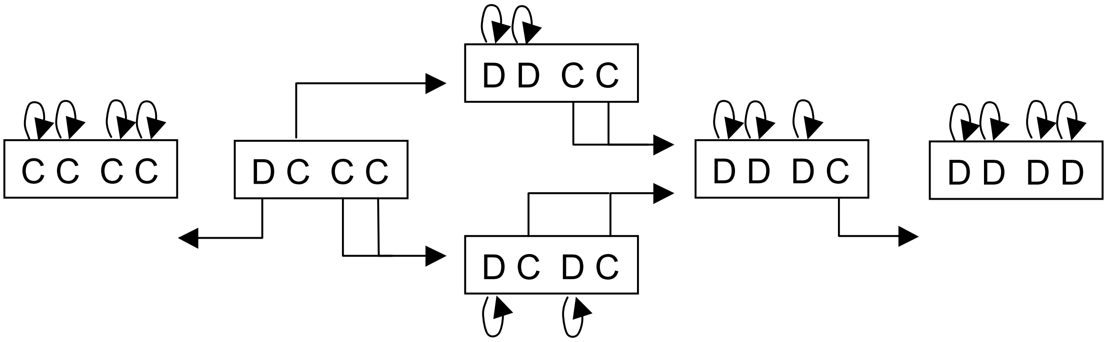
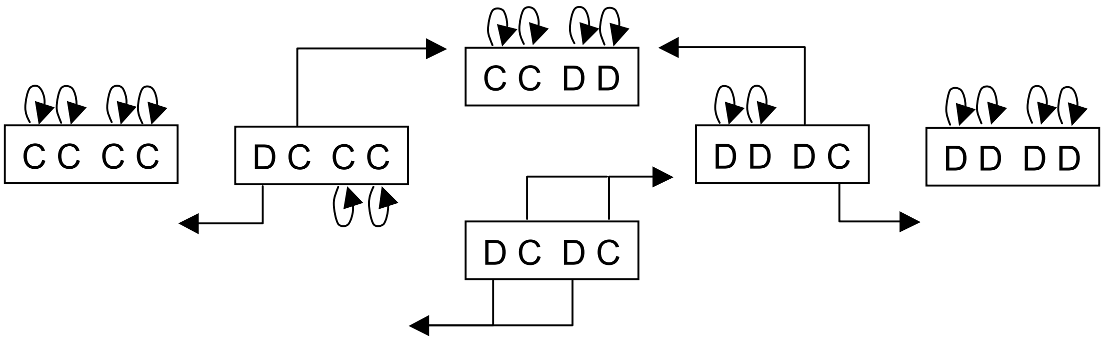


Figure 10: Best replies in the 4-player game with CR1 & CR2.

Three NE's: (CC CC) (CC DD) (DD DD)



Under CR1, switching to cooperation is self-beneficial only to stop an attack. However, if attacking also gives a reward V to the attacker that is higher than the benefit of freeriding, then there is an incentive to create a cooperative group if the other side is vulnerable. CR2 provides this condition:

> CR2: each member of the attacking group gains V>1, or more generally
> V > c - b/N.

Figure 10 illustrates that with both CR1 and CR2, (C,C) within a group is always self-enforcing. In the 4-person game, a new equilibrium (CC DD) results.

These two rules encourage cooperation, but do not handle a number of cases with incremental differences in cooperation between the groups. If the conflict rule is gradated so that incremental increases in a group’s cooperation incrementally decrease harm from losing a conflict, and the marginal harms from conflict outweigh marginal gains from freeriding, then every outcome with cooperator amounts tied in both groups is an equilibrium. (Another interpretation is that higher levels of cooperation increase the probability of victory. In that case, decision making is contingent on marginal expected harms and gains.) As in CR1, there is no incentive to increase cooperation to initiate an attack; but unlike CR1, partial levels of within-group cooperation can be sustained. A rule that does this is CR3, which is illustrated in the 4-person game in Figure 11. Under CR3, if the other group has more cooperators than yours, every defector should want to switch to cooperation until cooperators are tied with the other group.

> CR3: Let G be the absolute value of the difference between the number of cooperators in each group. If G>0, every player in the group with fewer cooperators suffers a loss of G*h, where h>1, or more generally h > c - b/N.

And if the conflict rule is similarly incremental for gains, it always pays off to switch to cooperation. As CR3 provides for marginal losses in conflict, CR4 describes marginal gains in conflict:

> CR4: Let G be the absolute value of the difference between the number of cooperators in each group. If G>0, every player in the group with more cooperators gains G*h, where h>1, or more generally h > c - b/N.

With the rules combined, in every situation each player wants to move towards cooperation, since their move alleviates losses or strengthens gains in the war.

Figure 11: 4-person game with just CR3. Losing-side defectors switch upward to defend incrementally better against an attacker. Winning-side cooperators slack off downward.

Three NE's: (CC CC) (CD CD) (DD DD) [NE iff cooperator counts tied]

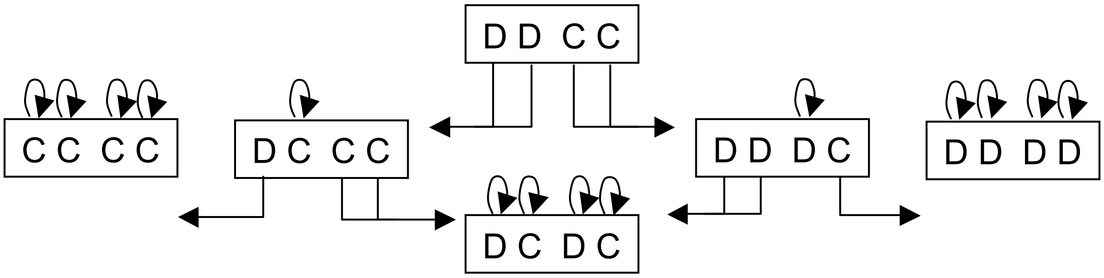
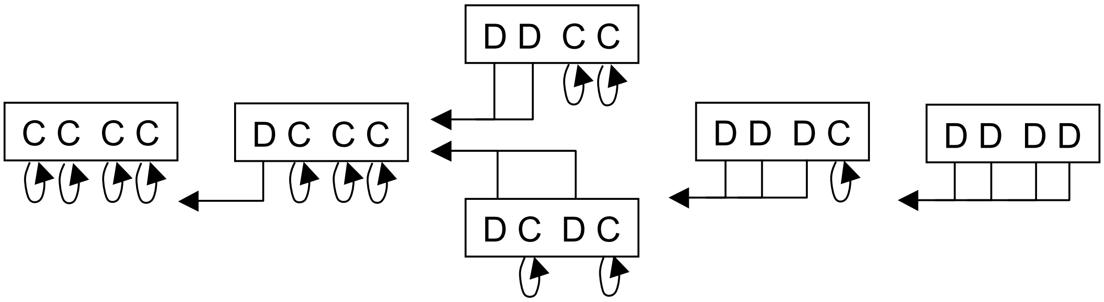


Figure 12: 4-person game with CR3 & CR4. Defectors always switch to cooperate.

One NE: (CC CC)



This conflict model illustrates a decisional explanation for the evolution of cooperation. Warfare provides incentives to individuals to cooperate within their own societies. As conflict intensifies and becomes more crucial to the well-being of groups – L and V increase past freeriding benefits – the incentives for cooperation grow. Unlike the evolutionary model of behavior given in section III, cooperation is not a fleeting phenomenon: it is robust and sustained given high levels of conflict. If early humans made rational decisions whether to cooperate, and conflict rewarded cooperating groups, then cooperation could have developed when intensity of conflict increased due to environmental circumscription and overpopulation, as in Carneiro's account. The next logical steps down this line of analysis would be to take conflict as an endogenous variable caused by geographic constraints and population changes.

But there is another significance to this model: it relates to our earlier story of group selection for cooperation. The conflict rules in this game are a way of introducing between-group competition based on relative levels of cooperation. Individuals in a group do better if their group's level of total cooperation is high relative to the other group's.

The between-group PFR is another form of group competition

Group competition is already present in the multiple n-person PD evolutionary model developed in Section III. If a group's fitness is higher, the group's share of the total population increases; the group is selected for. Furthermore, since its share of the population increases and the collective share of all other groups decreases, each other group's share is influenced downward. This group selection story is a way of describing competition between groups. We derived it earlier in non-competition terms, imagining some groups' raw populations growing faster than others', but it is zero-sum competition in terms of total share of system population.

It is instructive to view the share of the total population that each group has. Groups with a higher altruist ratio will grow faster than others; thus their share of the total population will rise, and other groups' shares will be hurt. The relatively more cooperative a group is, its members have higher relative payoffs compared to the rest of the individuals in the system. Similarly, in the evolutionary game individuals tend to have a higher fitness compared to the average fitness of an individual in the total population across groups when their group has higher cooperation than other groups.

The following analysis should make these facts clearer. Recall that the proportional fitness rule governs the evolutionary dynamics of the system. Within a single group, the proportion of altruists p_i is updated via the PFR:

$$p'_i = p_i \frac{W_{A,i}}{W_i}$$

where

$$W_{A,i} = \text{the fitness of group } i\text{'s altruists} = X + bp_i - c, \text{ and}$$

$$\begin{aligned} W_i &= \text{the average fitness of group } i \\ &= \text{mean fitness across its subgroups, weighted by size of each subgroup} \\ &= \text{the expected fitness of an individual from group } i \\ &= p_i W_{A,i} + (1 - p_i) W_{D,i} \\ &= X + (b - c) p_i \end{aligned}$$

In section III, we assumed that the altruist and defector populations in individual groups reproduced via linear replicator dynamics, causing the group they belonged to grow in size. N_i represents the size of a particular group, so the update rule for the absolute growth of a group via replicator dynamics is $N'_i = W_i N_i$. If we want to examine the competition among groups reproducing via linear replication, we can describe that via the proportional fitness rule. This dynamics of competition between groups is the same as competition of subgroups within a group. Let n_i be group *i*'s share of the total population: $n_i = N_i / N_{all}$ * n_i changes proportionally with group *i*'s relative fitness compared to all groups:

$$n'_i = n_i \frac{W_i}{W_{all}}$$

where

W_{all} = the average fitness of a group
= $E[W]$ by the statistical notation of Section III,
= mean fitness across all groups, weighted by size of each group
= the expected fitness of an individual from the total population
= $\sum_i n_i W_i$

The higher one particular group's reproductive fitness is, the more it grows the next round – since W_i increases. But also, the higher its reproductive fitness is, the less other groups grow. As its W_i increases, that increases the denominator W_{all} for everyone else.

So under our linear replication assumption, the PFR for between-group selection demonstrates that groups with proportionally higher collective in-group payoffs reproduce faster compared to the other groups.

However, the between-group PFR makes no commitments to the actual population of the system, just the populations of groups relative to one another. Say that population is at environmental carrying capacity, and is therefore constrained to be constant.[13] At each timestep, reproduction might happen, but if the environment kills off extra population in a random process back down to the old population size, then the new *n'* and *p'* values won't change. Thus *n'* and *p'* have changed from *n* and *p* via the PFR, but total population of the system has not changed. In such a world, population size is population share, so groups are in a zero-sum competition for population – groups can grow only at the expense of others.

Conflict provides more competition beyond the PFR

In Section III, we saw that group selection can encourage cooperation by rewarding the fitness of groups high in cooperators, and penalizing fitness for groups low in cooperators. It should be evident this is a particular instance of a more general class of between-group processes that help groups with more cooperators and hurt groups with fewer. Call the rules that create such processes competition rules. Conflict rules CR1-4 are additional examples of competition rules for the group decision game. In the group

---

[13] As in Carneiro's scenario of environmental circumscription leading to overpopulation and conflict to form the first states.

selection model, competition occurred due to the between-group PFR; in the group decision game, competition occurred through CR1-4.

Any competition rule that rewards groups with more cooperators will encourage cooperation. CR1-4 explicitly do this. Between-group PFR does this implicitly: more cooperators mean greater growth and a greater share of the population. The details of conflict can provide more competition beyond PFR-based group selection, as illustrated in the next subsection.

A conflict rule in addition to PFR group selection for the n-person PD encourages cooperation

While the evolutionary multi-group n-person PD model ultimately ends in defection, we can analyze the early portion of a run in which cooperators are increasing globally, and introduce a conflict rule that can encourage cooperation.

In Section III's model, the size of a group can be interpreted as its population. It could also be a group's land if land fills up immediately with new population, and more population enables expansion. Or alternatively, individuals could represent political sub-entities like families or villages, and being a cooperator means working together in a politically unified fashion. Whatever the case, the following formula defines military strength as the number of cooperators in the society (*Np*). Large societies with very low cooperation can be beaten by small, highly cooperative societies.[14] In a conflict, the winning side steals land/people/resources from the other. For the warring societies *i* and *j*, $\Delta_{F,i,j}$ denotes the change to N_i due to fighting with *j*. The rule is

$$\Delta_{F,i,j} = \alpha(N_i p_i - N_j p_j)$$

α represents the stakes or intensity of conflict; a higher α means there is greater transfer of land/people/resources through fighting. So for $\alpha = 1\%$, a matchup of 300 vs. 100 leads to a transfer of 2 units to the larger. A matchup of 10,000 vs. 9000 leads to a

---

[14] An simpler alternative is to have military strength be only the society size *N*. When run this way, the simulation yields similar results to the ones presented here.

transfer of 10 land. A matchup of 10,100 vs. 100 pushes the smaller exactly to extinction.[15]

Add this conflict rule to the n-person PD model of *m* groups by adding a conflict phase before every reproduction phase, in which every group enters a conflict with every other group simultaneously. Thus the update rule for the group-level is now[16]

$$N'_i = \text{(update due to replication)} + \text{(change due to conflict)}$$

$$N'_i = W_i N_i + \sum_{j=1}^{m} \alpha(N_i p_i - N_j p_j)$$

Note that this model is somewhat different than the decision-making conflict model. There, only internal levels of cooperation were relevant, since group sizes were uniform. Here, a larger group size increases military power.

The introduction of this conflict rule can dramatically increase the success of cooperation. In a world of 10 villages, with cooperator proportions evenly spaced out from {0.0, 0.1, .. 0.9}, the first plot of Figure 13 shows the progress of p_{all} without conflict. With conflict at $\alpha = 0.05$, the total population's proportion of cooperators increases faster and peaks higher, as shown on the right.

---

[15] This function actually has to be piecewise in the case that population is hurt so much it goes negative: instead, the population transfer is just enough to bring the smaller group to population 0.

[16] Since in Section II, I advocated looking at the n'_i PFR update rule instead of the N'_i replicator rule, it would be sensible to solve for the proportional fitness version

$$n'_i = n_i \frac{W_i}{W_{all}} \beta_{F,i}$$

But I am not sure what is the multiplier due to conflict $\beta_{F,i}$ for this conflict rule, or even if it can be represented in this manner. It may be possible to formulate a different conflict rule that can be solved in this form.

Figure 13: Individual runs for $\alpha = 0$ **vs.** $\alpha = 0.05$

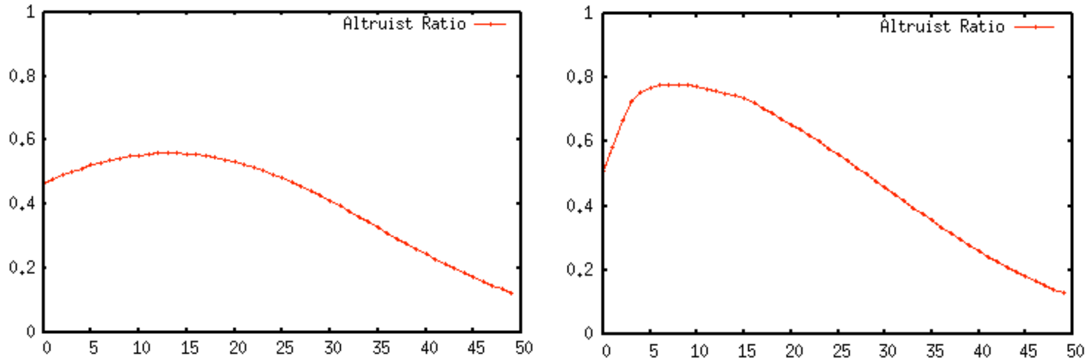
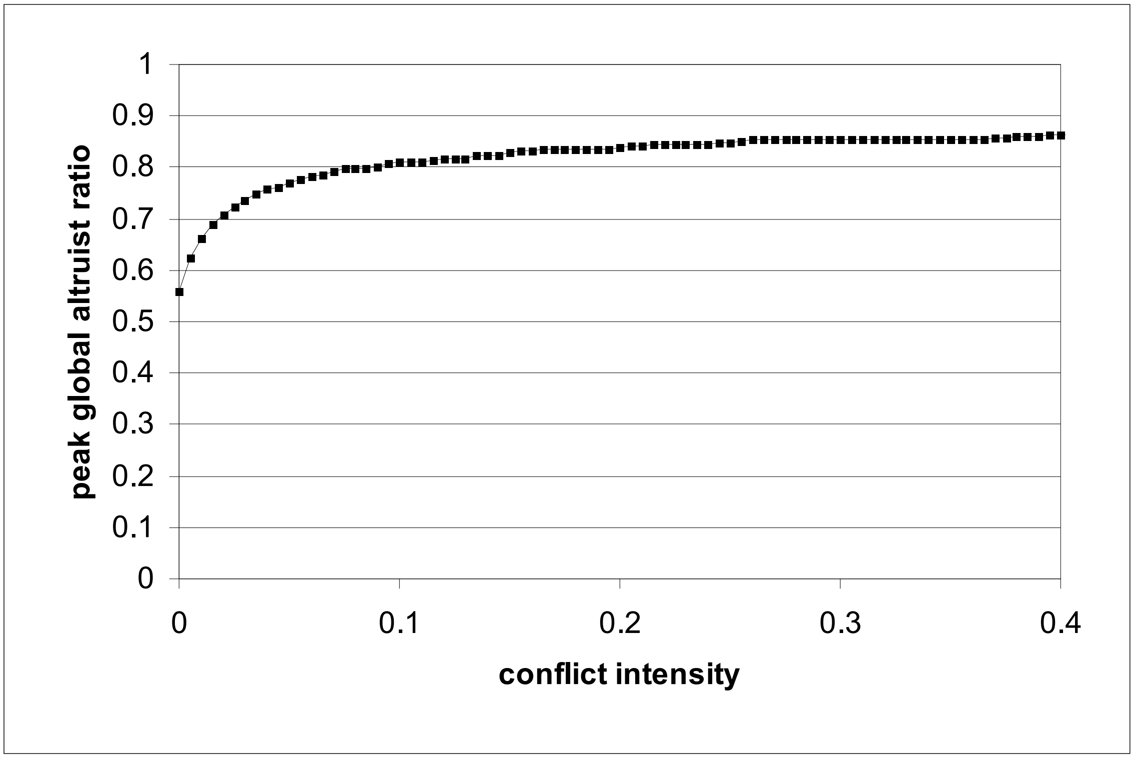


Figure 14 examines a range of conflict intensity values, plotting the peak cooperator proportion achieved during the run against each α level. With the same 10 village setup, higher values of α create a higher maximum cooperator proportion. Thus increasing intensity of conflict encourages more and more cooperation.

Figure 14: Peak system cooperator proportion per conflict intensity



V. Discussion and Conclusion

Our original motivation for examining these evolutionary models was to develop formal explanations of the evolution of cooperation that agreed with known facts about a particular example of the evolution of large-scale cooperation: the last 10,000 years' dramatic growth of human societies. I intend this analysis to bring together ideas from the evolution of cooperation literature and anthropological theories of social evolution. By extending formal cooperation modeling to the specific case of social order, we can help answer fundamental questions about the existence, efficacy, and evolution of human society.

The discursive theory that has provided most of the motivation for these models has been Carneiro (1970, 1998). His story is fairly intricate and subtle. It speaks most directly of the assembly of larger political units out of smaller ones – something that the evolutionary conflict model developed here does not do, unless one interprets stealing land/population/resources as stealing entire polities. This is somewhat reasonable: groups can go extinct in the model, when other groups have won so many military victories that they have stolen all their land. This interpretation is analogous to Carneiro (1970) in which winning villages swallow up their neighbors.

It is much more difficult to model a process of alliance building that solidifies into unitary political entities. The relationship of cooperation to political entity-building is complex. Cooperation can include cooperating to build a state or nation – for example, social contract theory's view of the origin of the state was that people agreed (i.e., cooperated) to give power to a government, which would maintain the social order, the quintessential public good. Therefore groups with higher levels of cooperation should be viewed as more unified entities. Besides the social contract story, there are numerous political and biological examples where cooperation makes larger entities. A system with components that are heavily cooperating – and therefore highly interdependent – looks like a unitary system of only a few components. If the 50 states of the United States are very highly interdependent and have few conflicts, it makes sense to think of them as one unit. If a fungus and algae are engaged in intense energy and food exchange, it is useful to think of them a single organism, lichen. If a mitochondrion – an organelle within certain cells that is theorized to have originally been a separate organism (Margulis 1993) – is engaged in heavy energy exchange and dependence with the rest of the cell, it is useful to think of it as part of the that organism. The same goes for colonies of many single-celled organisms that formed into single multi-celled organisms. Axelrod (1995) models the building of international alliances into aggregate political actors; it would be useful to rework this analysis into the rise of social cooperation.

The extremely simple C vs. D model of group populations used here fails to capture the phenomenon of creating larger entities out of smaller ones, or how social breakdown can be related to fracturing of political entities, i.e. in civil war. An example of accommodating the latter would be to add a rule that a very high defector ratio would cause the group to split into new groups.

Still, this analysis says something about conflict and group competition. Boyd and Richerson (2005) aim to examine social norms and cooperation as a result of the coevolution of genes and cultures among humans during the hunter-gatherer phase. The explanation here shares the use of group selection for culture, but in the much different context of the last 10,000 years of social evolution. Genetic evolution is probably impossible over such short time. Given saturated human populations in fertile areas and concomitantly increasing conflict, selection for more cooperative, larger, and complex societies can emerge. Selection can also occur for social factors like total size of population. Since the last 10,000 years are a unique phase in human history, unique factors should be taken into account when explaining the evolution of cooperation and societies during that time.

Currently the evolutionary models only demonstrate possible effects that merely encourage increases in cooperation, but do not sustain them. A logical next step to take with the models would be to close the loop and find the evolution of cooperation to stable equilibria. It may be necessary to move away from the rather harsh prisoners' dilemma model of cooperation to secure that. Soltis, Boyd, and Richerson (1995) note that cultural transmission can be conformant – pulling individuals to have cultural traits that are popular, not necessarily the most fit. Therefore, it is possible to have multiple equilibria for within-group cooperation. If this is true, then group selection should eliminate groups stuck at less productive equilibria. And beyond cultural conformity, richer models of cultural transmission may incorporate other important aspects of cultural change beyond what the simplistic and unrealistic proportional fitness rule has to offer.

Understanding the evolution of human cooperation, altruism, and society is a challenging task. Hopefully the consideration of group competition and conflict will be helpful in understanding the particular case of cooperation in large, complex societies.

References

Axelrod, Robert, and William D. Hamilton. 1981. “The evolution of cooperation.” *Science,* 211, 1390-1396.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, Robert. 1995. “A Model of the Emergence of New Political Actors.” In Nigel Gilbert and Rosaria Conte (eds.), *Artificial Societies: The Computer Simulation of Social Life.* London: University College Press, 19-39.

Bendor, Jonathan, and Piotr Swistak. 1997. “The Evolutionary Stability of Cooperation.” *American Political Science Review,* 91, 290-307.

Bergstrom, Theodore C. “Evolution of Social Behavior: Individual and Group Selection.” *Journal of Economic Perspectives,* 16, 67-88.

Boyd, Robert, Herbert Gintis, Samuel Bowles, and Peter J. Richerson. 2003. “The evolution of altruistic punishment.” *Proceedings of the National Academy of Sciences* (USA), 100:3531-3535. Reprinted in Boyd, R., and P.J. Richerson, 2005, *The Origin and Evolution of Cultures*.

Boyd, Robert, and Jeffrey Lorberbaum. 1987. “No Pure Strategy is Evolutionarily Stable in the Repeated Prisoner’s Dilemma Game.” *Nature,* 327:58-59.

Boyd, R. and Richerson, P.J. 1988. “The evolution of reciprocity in sizable groups.” *Journal of Theoretical Biology*, 132:337-356.

Boyd, Robert, and Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Oxford University Press.

Carneiro, Robert L. 1970. “A Theory of the Origin of the State.” *Science*, 169, 733-8.

Carneiro, Robert L. 1998. “What happened at the flashpoint? Conjectures on Chiefdom Formation at the Very Moment of Conception.” *Chiefdoms and Chieftaincy in the Americas*, ed. E.M. Redmond. 18-42. Gainesville: University Press of Florida.

Cooper, Ben and Chris Wallace. 2004. “Group selection and the evolution of altruism.” *Oxford Economic Papers,* 56, 307-330.

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

Diamond, Jared. 1999. *Guns, Germs, and Steel.* New York: W. W. Norton & Company, Inc.

Greif, Avner. 2005. *Institutions: Theory and History*. March 2005 manuscript; forthcoming Cambridge University Press. Draft online: http://www-econ.stanford.edu/faculty/Greif_Instutions/GreifBook.html. Retrieved October 31, 2005.

Hamilton, William D. 1964. “Genetical evolution of social behavior I and II.” *Journal of Theoretical Biology*, 7, 1-52.

Hausken, Kjell. 1995. “Intra-level and Inter-level Interaction.” *Rationality and Society,* 7, 465-88.

Henrich, Joseph. 2004. “Cultural Group Selection, Coevolutionary Processes and Large-scale Cooperation.” *Journal of Economic Behavior and Organization*, 53, 3-35 and 127-143. Online: http://www.anthropology.emory.edu/FACULTY/ANTJH/Website/Papers/HenrichplusCommentsJebo2003.pdf

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton, Natalie Smith, and David Tracer. 2005. “'Economic Man' in Cross-cultural Perspective: Behavioral Experiments in 15 Small-scale Societies.” *Behavioral and Brain Sciences*, in press. Online versions: http://www.bbsonline.org/Preprints/Henrich/Henrich.html, http://www.sscnet.ucla.edu/anthro/faculty/boyd/MacGamesBBSFinal.pdf

Keeley, Lawrence H. 1996. *War Before Civilization.* New York: Oxford University Press.

Kelly, Raymond C. 1985. *The Nuer Conquest: The Structure and Development of an Expansionist System.* Ann Arbor: University of Michigan Press.

Lenski, Gerhard. 1970. *Human Societies: An Introduction to Macrosociology*. (First Edition.) New York: McGraw-Hill, Inc.

Margulis, Lynn. 1993. *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons*. New York: W.H. Freeman & Co.

Maynard Smith, John. 1982. *Evolution and the Theory of Games*. Cambridge University Press.

Milgrom, Paul R., Douglass C. North, and Barry R. Weingast. 1990. “The role of institutions in the revival of trade: the law merchant, private judges, and the champagne fairs.” *Economics and Politics*, 2, 1, 1-23.

Morrow, James D. 1994. *Game Theory for Political Scientists*. Princeton: Princeton University Press.

Nakamaru, M., H. Matsuda, and Y. Iwasa. 1997. “The evolution of cooperation in a lattice-structured population.” *Journal of Theoretical Biology*, 184, 65-81.

Price, George. 1970. “Selection and covariance.” *Nature*, 227, 520-1.

Price, George. 1972. “Extension of covariance selection mathematics.” *Annals of Human Genetics*, 35**,** 485–90.

Riolo, Rick L., Michael D. Cohen, and Robert Axelrod. 2001. “Evolution of cooperation without reciprocity.” *Nature,* 414, 441-443.

Rogers, Alan R. 1988. “Does biology constrain culture?” *American Anthropologist,* 90, 819-831.

Rubin, Paul H. 2001. “The state of nature and the evolution of political preferences.” *American Law and Economics Review*, 3, 1, 50-81. Earlier version available online: http://www.economics.emory.edu/Working_Papers/wp/rubin_98_08_paper.pdf

Sober, Elliott, and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press.

Soltis, Joseph R., Robert Boyd, and Peter J. Richardson. 1995. “Can group functional behaviors evolve by cultural group selection? An empirical test.” *Current Anthropology*, 36, 473-494.

Trivers, Robert L. 1971. “The Evolution of Reciprocal Altruism.” *The Quarterly Review of Biology*, 46, 1, 35-57.

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

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

Wilson, David Sloan. 1989. “Levels of Selection: An Alternative to Individualism in Biology and the Human Sciences.” *Social Networks*, 11, 257-252. Reprinted in Elliott Sober (ed.), *Conceptual issues in Evolutionary Biology*, 143-154. Cambridge, Massachusetts: MIT Press, 1994.

Wilson, David Sloan. 1998. “Hunting, Sharing, and Multi-Level Selection: The Tolerated Theft Model Revisited.” *Current Anthropology*, 39, 73-97.