Hamilton's Rule Applied to Reciprocal Altruism

Jeffrey A. Fletcher
Systems Science Ph.D. Program
Portland State University
Portland, OR 97207
Email: jeff@pdx.edu

Martin Zwick
Systems Science Ph.D. Program
Portland State University
Portland, OR 97207
Email: zwick@pdx.edu

Abstract—Reciprocal altruism and inclusive fitness are generally considered alternative mechanisms by which cooperative, altruistic traits may evolve. Here we demonstrate that very general versions of Hamilton's inclusive fitness rule (developed by Queller) can be applied to traditional reciprocal altruism models such as the iterated Prisoner's Dilemma. In this way we show that both mechanisms rely fundamentally on the same principle—the positive assortment of helping behaviors. We discuss barriers to this unified view, including phenotype/genotype differences and non-additive fitness (or utility) functions that are typical of reciprocal altruism models. We then demonstrate how Queller's versions of Hamilton's rule remove these obstacles.

I. INTRODUCTION

Over two decades ago Axelrod and Hamilton [1, 2] suggested two alternative mechanisms for the evolution of cooperative (altruistic)[1] traits: 1) when the benefits of altruism fall to relatives, inclusive fitness [3-5] is the likely mechanism; and 2) when benefits fall to non-relatives, reciprocal altruism [6] is the likely mechanism. Around this same time researchers began to show that multilevel (group) selection theory [7, 8], a third alternative not emphasized by Axelrod and Hamilton, is mathematically equivalent [9-14] to Hamilton's inclusive fitness theory. This theory shows that Hamilton's coefficient of relatedness can be treated as a measure of positive assortment among altruistic traits (or behaviors) rather than strictly as a measure of kinship [5] and that the differential productivity of groups constitutes a group-level selection process [15]. Other researchers have gone further in claiming that reciprocal altruism also can be unified with multilevel selection and inclusive fitness theories. Queller [10] originally suggested that a generalization he developed of Hamilton's rule could be applied to reciprocal altruism and more recently Sober and Wilson [14] have claimed such a unification is possible, yet there has not been a clear demonstration of *how* to apply Queller's equations to traditional reciprocal altruism models. In fact in a subsequent expansion of his original result Queller [11, 12] does not mention its application to reciprocal altruism.

Here we take up this thread of inquiry and demonstrate that Queller's equations do indeed provide a foundation for the unification of reciprocal altruism, inclusive fitness, and multilevel selection theories. We do this by using

---

1 We will use *cooperation* and *altruism* synonymously because in the examples used in this paper, the cooperate behavior (C) always involves an altruistic sacrifice in fitness (or utility).

Hamilton's inclusive fitness rule to predict whether a reciprocal altruism strategy will increase or not within a classic iterated Prisoner's Dilemma (PD) model. This model is similar to those used by Axelrod and Hamilton in their original work. In this way we support the idea that, rather than relying on alternative mechanisms, these theories offer different ways of understanding a common principle by which self-sacrificing traits can increase. This principle is that altruism can evolve if there exists: 1) sufficient positive assortment among altruistic behaviors such that on average altruistic genotypes benefit more from these behaviors than average population members, 2) sufficient non-zero-sum benefits of cooperation such that on average the benefit provided to others by an altruist exceed the cost to the altruist, and 3) heritability of these behaviors. The "sufficient" positive assortment and benefit to cost values necessary for altruism to increase are given by Queller's generalizations of Hamilton's rule. This rule applies whether the source of positive assortment among altruistic behaviors is due to conditional strategies, higher interaction levels among relatives, or other causes of population grouping (even among non-relatives).

II. OBSTACLES TO APPLYING HAMILTON'S RULE TO RECIPROCAL ALTRUISM

There are two major difficulties in unifying reciprocal altruism with inclusive fitness and multilevel selection theories—one stemming from the difference between genotype and phenotype and the other from non-additive fitness functions. The first obstacle is that by definition there is not a one-to-one relationship between behavior and genotype in reciprocal altruism models. In contrast, inclusive fitness and multilevel selection models most often assume that there is a cooperation or altruism "gene" (or some additive combination of genetic determinants) and that the frequency of the genetic trait is proportional to the frequency of the cooperative or altruistic behavior. With reciprocal altruism models, however, behaviors are conditioned on environmental circumstances (e.g. the behaviors of others) and are not generally proportional to genetic frequencies. For example, if there is a trait such that cooperation is conditioned on what others did in previous encounters, then the frequency of this trait does not directly predict the frequency with which an individual displays cooperative behaviors. For this one must know the previous behaviors of those with whom the individual interacts.

Because the round-robin tournament structure used by Axelrod and Hamilton ensures no positive assortment among player types (ignoring the play against ones "shadow" self [2]), it seems that altruism cannot evolve via either inclusive fitness or multilevel selection. Both Hamilton [16] and Wilson [7] (working independently) have shown that altruism does not evolve when groups are formed randomly because there is not enough positive assortment in the altruistic trait. Recently, however, we showed that altruism *can* evolve in randomly formed groups if these groups are multigenerational [17]. Here we will demonstrate that altruism can also evolve in single generation groups (pairings) when *behaviors* are positively assorted—even in the absence of assortment between the interacting *genetic types*.

The second obstacle to unification is that the fitness (or utility) values used in reciprocal altruism models are not necessarily additive and therefore are not decomposed into the familiar benefits and costs commonly employed in inclusive fitness and multilevel selection models. This non-additivity is true of the most common payoff matrices used in iterated PD experiments (Fig. 1), including those of Axelrod and Hamilton. We explain this further in the section on non-additivity.

Both of these obstacles, as we demonstrate below, are effectively handled by Queller's generalizations of Hamilton's rule. We begin by briefly reviewing the iterated PD experiments of Axelrod and Hamilton as well as Hamilton's rule and Queller's contributions. We then demonstrate how Hamilton's inclusive fitness rule can be applied to an additive PD experiment involving a conditional strategy, and then finally do the same when fitness payoffs are non-additive. Table I lists the symbols that will be used in our discussion.

TABLE I. SYMBOLS USED

Symbol	Meaning
PD	Prisoner's dilemma
TFT	Tit-For-Tat genotype
ALLD	Always-Defect genotype
C	cooperate behavior
D	defect behavior
G_P	fraction of TFT genotype in population
G_A	fraction of TFT genotype in an actor (1 or 0)
P_P	fraction of C behaviors in population
P_A	fraction of C behaviors by an actor
P_O	fraction of C behaviors by actor's opponents
i	number of iterated games in a pairing
b	benefit provided by a C behavior
c	cost paid for a C behavior
d	deviation from additivity for CC interaction
f_{TT}	fraction of TFT-TFT pairings
f_{TD}	fraction of TFT-ALLD pairings
f_{DD}	fraction of ALLD-ALLD pairings

III. THE ITERATED PRISONER'S DILEMMA

The prisoner's dilemma (PD) game captures a fundamental problem of social life—individually rational behavior may lead to a collectively irrational and deficient outcome. Fig. 1 illustrates a typical 2-player PD payoff matrix (and the one used by Axelrod and Hamilton [1, 2]). The game is symmetric and each player has a dominant strategy to defect (D). This mutual defection results in each player receiving a payoff of 1, but if they had both cooperated (C), they could have each received a payoff of 3. In this game it is presumed that players choose their strategy simultaneously and there is no knowledge or guarantees about what the other player will do. The dilemma is that cooperation makes a player vulnerable to exploitation—in this case the exploiting defector gets 5 while the lone cooperator gets 0.

opponent's behavior

		C	D
actor's behavior	C	3	0
	D	5	1

Figure 1. Shows typical PD utility (or fitness) values for the actor given its own and opponents behaviors. Behaviors are either cooperate (C) or defect (D).

An n-player version of the PD in which cooperators contribute to the common good (also called *whole-group* altruism [18]) models familiar social dilemmas such as the "tragedy of the commons" [19] and the "free-rider problem" [20, 21]. Elsewhere we have demonstrated that an n-player PD with fixed strategies played across multiple groups captures fundamental features of the evolution of altruism via multilevel selection (or equivalently inclusive fitness) [22-24]. Here we will show that a 2-player game with conditional strategies also fits into this broader framework.

Although in a PD situation it is individually rational to defect in single plays of the game, Axelrod and Hamilton [1, 2] showed that conditional cooperative strategies can do well overall when games are iterated. This combination of iterated games and conditional play can create positive assortment among *behaviors* even when there is none among *player types*. In their experiments many different types of players (submitted by researchers from many fields) competed in a round-robin tournament in which each submission played all others an average of 200 times. The most successful strategy in Axelrod and Hamilton's experiments was also one of the simplest. Submitted by social scientist Anatol Rapoport, it was called Tit-For-Tat (TFT). TFT always cooperates with an opponent on the first move and then in all subsequent moves simply plays whatever the opponent did in the last game. Even though TFT never came out ahead in any

single match of 200 iterated games [14, 25], its overall score was superior as it was able to minimize exploitation by defecting strategies such as Always-Defect (ALLD), while taking advantage of mutual cooperation when it met other "nice" strategies. When viewed as an evolutionary process where utility payoffs of the PD are used to determine the number of each player type in the next generation, Axelrod and Hamilton showed that both ALLD and TFT are Evolutionarily Stable Strategies (ESS). That is, they cannot be invaded by individual mutations or migrants exhibiting a different strategy. Since these original experiments more than 20 years ago, much research has been done on the iterated PD (e.g. see [26] for review]).

IV. HAMILTON'S RULE

Hamilton's rule gives the condition necessary for an altruistic trait to increase in a subsequent generation and is deceptively simple [3]:

$$rb > c \qquad (1)$$

where b is the average benefit provided by an altruist and c is the average cost to the altruist. The complications lie in the meaning of the r term which has been progressively generalized over the years. Originally thought of as a simple measure of relatedness via descent [3], Hamilton (after interacting with Price [15]) broadened the meaning of r to be a measure of the interaction of genetic types regardless of relatedness:

$$r = \frac{\text{cov}(G_A, G_O)}{\text{var}(G_A)} \qquad (2)$$

where G_A is the genotype or breeding value with respect to the altruistic trait for each potential actor and G_O it the average genotype or breeding value of those that interact with the actor. After several examples of Hamilton's rule failing were reported in the literature [e.g. 27], Queller showed that these were due to phenotype/genotype differences and that it was the altruistic *behaviors* of others and itself that determined whether the actor's genotype increased, not the *genotype* of others. Queller generalized Hamilton's r term as follows [10-12]:

$$r = \frac{\text{cov}(G_A, P_O)}{\text{cov}(G_A, P_A)} \qquad (3)$$

where P_O is the average phenotypic value of others interacting with the actor and P_A is the average phenotypic value of the actor.

Note that Eqn. (3) is more general than Eqn. (2). In the case of no genotype/phenotype difference ($P_O = G_O$ and $G_A = P_A$) Eqn. (3) reduces to Eqn. (2). The absence of any G_O component in this most general interpretation of Hamilton's r term by Eqn. (3) is especially noteworthy

because it challenges the selfish gene [28, 29] view of inclusive fitness in which genes encode altruistic behaviors only because these behaviors help other copies of the genes residing in those receiving the altruism.

Based on the definition of covariance, Eqn. (3) can also be written as:

$$r = \frac{\sum (G_A - G_P)(P_O - P_P)}{\sum (G_A - G_P)(P_A - P_P)} \qquad (4)$$

where the summation is over each individual in the population (each potential actor) and P_P is the average population phenotypic value. We use this equation to calculate r below. Note that here there is no mutation and the genetics are haploid, but Frank [13] has expanded Queller's equations to accommodate different degrees of fidelity in transmission from one generation to the next.

V. APPLYING HAMILTON'S RULE TO RECIPROCAL ALTRUISM

Here we provide a simple example of applying Hamilton's rule to reciprocal altruism using a population consisting of two classic evolutionarily stable (ESS) types, TFT and ALLD. Because one of the types (TFT) uses conditional behaviors we must measure genotypes and phenotypes separately. Fig. 2 illustrates a general view of the PD in which an altruist provides a benefit value of b to its opponent at a cost c to itself. w_0 is the base fitness value unrelated to the altruistic trait. The d term is a deviation from additivity and is discussed in the next section; in this section $d = 0$. Behaviors are either cooperate (C) which has a phenotypic (P) value of 1 or defect (D) which has a phenotypic value of 0. Note that the fitness values in Fig. 2 depend only on the phenotype of the actor (P_A) and the opponent (P_O), not on their genotypes. We focus on the TFT type and thus give it a genotypic (breeding) value of 1 and ALLD a value of 0.

		opponent's behavior	
actor's behavior		C ($P_O = 1$) contributes b	D ($P_O = 0$) contributes 0
	C ($P_A = 1$) sacrifices c	$w_0 + b - c$ $(+ d)$	$w_0 - c$
	D ($P_A = 0$) sacrifices 0	$w_0 + b$	w_0

Figure 2. Shows PD utility (or fitness) values *for the actor* given its own and opponents behavior. The fitness (or utility) values for the actor are represented as the sum of additive contributions from the opponent and its own sacrifice. The w_0 term is an additive base fitness value uncorrelated with C and D behaviors. The d term is the deviation from additivity (addressed in next section).

In this population of two types there will be three possible parings each with set values for G_A, P_O, and P_A. Table II gives the values for each of the player types when

it is the potential actor (A) and its opponent (O) is either the same type or different.

TABLE II. G_A AND PREDICTED P_A, AND P_O VALUES

Actor (A)	Opponent (O)	G_A	P_A	P_O
TFT	TFT	1	1	1
TFT	ALLD	1	$1/i$	0
ALLD	TFT	0	0	$1/i$
ALLD	ALLD	0	0	0

The value $1/i$ in Table II represents the fact that when playing an ALLD player, TFT will cooperate once on the first play out of i total plays in this pairing. To use Eqn. (4), we need additionally the values of i, G_P, and P_P. We take i and G_P to be parameters of the model, but the overall fraction of cooperate behaviors, P_P, can be calculated from the frequencies (f) with which different pairings occur:

$$P_p = \frac{f_{TT}2i + f_{TD}}{2i}, \qquad (5)$$

where the numerator of Eqn. (5) represents $2i$ cooperate behaviors in all TFT-TFT pairings plus one cooperate behavior in all TFT-ALLD pairings. The expected frequency of pairing types can be calculated from G_P and population size for various ways of forming pairs (e.g. a binomial or hypergeometric distribution). The round-robin tournaments used by Axelrod and Hamilton are equivalent to the expectation for a hypergeometric distribution (sampling without replacement) if one ignores the ad hoc modifications of having each player play a "shadow" self and an opponent playing randomly. For a round-robin tournament (hypergeometric distribution) of large population size, the frequencies of pair types approaches those of a binomial distribution. Here for simplicity we will assume random pairings from an infinite population and therefore use a binomial distribution. In this case $f_{TT} = (G_P)^2$, $f_{TD} = 2G_P(1 - G_P)$, and $f_{DD} = (1 - G_P)^2$. We thus have $P_P(i)$ and also from Table II, G_A, $P_A(i)$, and $P_O(i)$. Now using only G_P and i as parameters to this simple model, we can calculate r from Eqn. (4) and use it in Hamilton's rule (Eqn. (1)) to predict whether TFT will increase or not.

In order to check the predictions of Hamilton's rule, we also calculate the fraction of TFT in a subsequent generation, G_P', using the above frequencies (f) of pairings and b and c. This assumes that the number of each player type in the new generation is proportional to the fitness (or utility) values gained in the previous generation. A generation is defined as one round of pairings each with i games between paired players. G_P' is then given by:

$$G_P' = \frac{f_{TT}2i(w_0 + b - c + d) + f_{TD}(iw_0 - c)}{f_{TT}2i(w_0 + b - c + d) + f_{TD}(2iw_0 + b - c) + f_{DD}2w_0} \quad (6)$$

where for now $d = 0$.

Fig. 3 shows the change in the fraction of TFT in the population after one generation (or tournament), $\Delta G_P = G_P' - G_P$, as a function of benefit level where cost and base fitness are held constant at $c = 1$ and $w_0 = 1$. Calculation results are shown for a few different starting G_P and i values. Note the general trends—ΔG_P is positive for higher initial G_P and more game iterations.

For our purposes, the main point is to notice that the precise value of b needed to get an increase in TFT, $\Delta G_P > 0$, is given by Hamilton's rule. Rearranging Eqn. (1) gives the condition $b > c/r$. Table III gives the c/r values for the five example set of parameters used in Fig. 3 (where cost is held at $c = 1$). Each c/r value precisely predicts the benefit level needed for TFT to increase, i.e. $\Delta G_P > 0$.
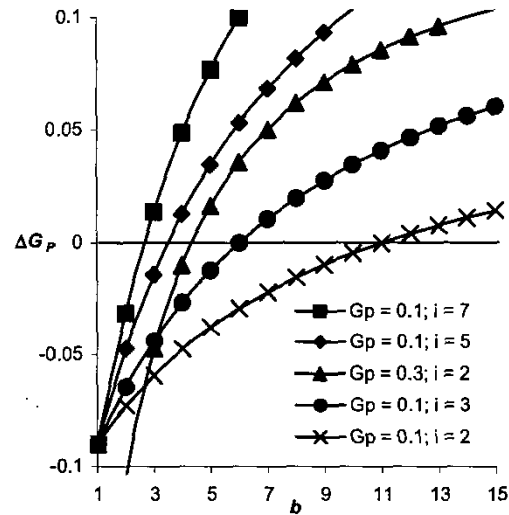

Figure 3. The change in the fraction of TFT players in population after one generation, ΔG_P, as a function of benefit level, b. Data is shown for indicated initial G_P and i. Cost c and base fitness w_0 are both held constant at 1.

TABLE III. HAMILTON'S RULE MIN. b FOR $\Delta G_P = 0$ IN FIG. 3

G_P	i	Minimum $b = c/r$
0.1	7	2.6667
0.1	5	3.5000
0.3	2	4.3333
0.1	3	6.0000
0.1	2	11.0000

VI. HAMILTON'S RULE AND NON-ADDITIVITY

The version of Hamilton's rule given in Eqn. (1) relies on the assumption that fitness (utility) benefits and costs are additive and due to the independent effects of individual behaviors. But note that it is not possible to pick positive values for w_0, b, and c in Fig. 2 that will sum to the values found in Axelrod and Hamilton's classic

version of the PD in Fig. 1. This latter PD (Fig. 1) is non-additive. One way to view non-additivity is as the presence of some positive or negative synergy for mutual action (where defection is considered non-action). As an example of positive synergy, imagine a cooperative trait that involves hunting for one's paired partner—when both hunt simultaneously the catch may exceed two times the individual result. Negative synergy can be seen in terms of diminishing returns. In a different ecological context two hunters might interfere with each other or be exploiting the same limited resource—their total might then be less than the sum of individual efforts. The d term in Fig. 2 gives the deviation from additivity. Now we can use the values $b = 4$, $c = 1$, $w_0 = 1$, and $d = -1$ to arrive at the familiar PD in Fig. 1.

Again we can rely on Queller [10-12] who also developed a version of Hamilton's rule that accommodates a non-linear term. In this case the condition for an altruistic trait to increase is [10-12]:

$$rb + r_{dev}d > c \qquad (7)$$

where r is still given by Eqn. (4) and r_{dev} is given by:

$$r_{dev} = \frac{\text{cov}(G_A, P_A P_O)}{\text{cov}(G_A, P_A)} = \frac{\sum (G_A - G_P)(P_A P_O - \overline{P_A P_O})}{\sum (G_A - G_P)(P_A - P_P)} \qquad (8)$$

Notice that the difference between r_{dev} and r is that P_O is replaced by the interaction term $P_A P_O$ in the numerator. This term (which incorporates the effect of d on G_A in Eqn. (7)) is only non-zero when there is mutual cooperation.

Again, to verify Hamilton's rule in this reciprocal altruism model we will need to calculate G_P' using the deviation in the payoff for mutual cooperation, which happens only in TFT-TFT pairings (see Eqn. (6)). Fig. 4 shows the effect of non-additive deviations on ΔG_P where parameters are set to resemble those in Axelrod and Hamilton's iterated PD experiments, $b = 4$, $c = 1$, $w_0 = 1$, and $i = 200$. Again, when $d = -1$ this corresponds to the PD used in their experiments (Fig. 1). There are again general trends such that ΔG_P is easier to make positive for higher initial G_P and higher values of d. Note also that TFT can gain in the population even when relatively rare, given the high number of iterated interactions, $i = 200$. The more linear appearance of the curves in Fig. 4 (compared to Fig. 3) is due to scale differences—here we start with much smaller G_P values and plot much smaller changes.

To check the application of Hamilton's rule under these non-additive conditions we rearrange Eqn. (7) to get the condition $d > (c - rb) / r_{dev}$. Table IV shows the value of $(c - rb) / r_{dev}$ for the parameters used in Fig. 4. For each set of parameters, these values precisely predict the threshold value of d where ΔG_P becomes positive, thus, demonstrating that this version of Hamilton's rule can be

successfully applied to reciprocal altruism models, even when fitness values are non-additive.
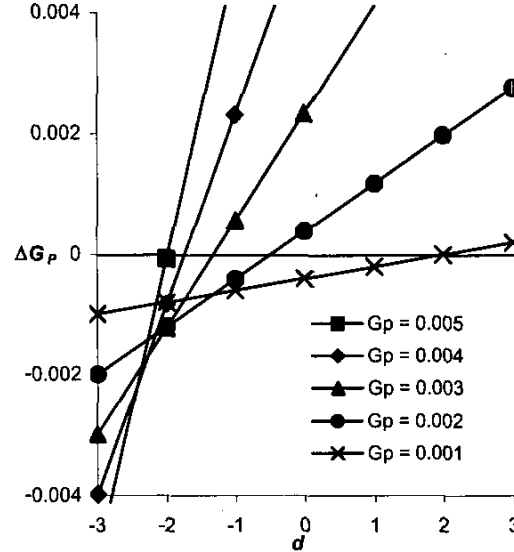

Figure 4. The change in the fraction of TFT players in a population after one generation, ΔG_P, as a function of the deviation from additivity for mutual cooperation, d. Data is shown for indicated initial G_P. Cost c and base fitness w_0 are both held constant at 1, $b = 4$, and the number of iterated games $i = 200$.

TABLE IV. HAMILTON'S RULE MIN. d FOR $\Delta G_P = 0$ IN FIG. 4

G_P	Minimum $d = (c - rb) / r_{dev}$
0.005	-1.985
0.004	-1.735
0.003	-1.318
0.002	-0.485
0.001	2.015

VII. CONCLUSION

We have demonstrated how reciprocal altruism can be analyzed using Hamilton's inclusive fitness rule. The key is to use general versions of Hamilton's rule appropriate to the situation under study. In reciprocal altruism with its conditional strategies and repeated interactions there can be a positive assortment *among behaviors*, even when there is no positive assortment *among genetic types*. Therefore these differences must be accommodated in the analysis. Similarly, when the fitness consequences of individual behaviors are non-additive, this deviation must be accounted for. Queller [10-12] has provided versions of Hamilton's rule that handle both of these situations.

Our example has been minimal with only two different types, a fixed number of interactions, and predictable fractions of interacting pair types from an infinite binomial distribution. But these techniques can be used

998

for more complicated situations if one uses the actual frequency of phenotypes and genotypes observed in all interactions.

Our demonstration of an underlying unity in mechanisms for the evolution of altruism broadens the meaning of both reciprocal altruism and inclusive fitness. Fundamentally, the evolution of altruism depends on the positive assortment of social behaviors such that "nice" types help each other more than they help average population members. In this sense, altruism requires reciprocity in order to evolve whether based on conditional strategies or some other source of positive assortment. The notion of inclusive fitness must also be broadened beyond the selfish gene viewpoint [28, 29] when phenotype differs from genotype. For instance, if one imagines a single Always-Cooperate (ALLC) individual interacting in a population where all remaining individuals are TFT, the fitness of the ALLC player would be equal to the fitness of a TFT player—not because it selfishly helps copies of its own genetic type (ALLC), but because it successfully exchanges cooperate behaviors with the TFT type. Axelrod and Hamilton emphasized that symbiotic interactions were an example of cooperation between different species that could not be based on relatedness. In the above example one could view the ALLC and TFT types as different species. Cooperation across species is a clear demonstration that the "selfish-gene" of kin-selection theory is only a special case of a much more general phenomenon. Although it sounds strange to say, inclusive fitness as captured by Hamilton's rule can involve different species as well as non-relatives of the same species.

Reciprocal altruism is not fundamentally different from inclusive fitness or multilevel selection. As we have shown using Queller's equations [10-12], inclusive fitness can be generalized to encompass reciprocal altruism as a special case. As already noted, inclusive fitness and multilevel selection have been shown to be mathematically equivalent. It thus follows that all three theories offer different accounts of the same phenomenon. In fact there are many specific mechanisms by which altruism can evolve, including passive ones such as foraging in non-uniform resource distributions [30], continuous population viscosity with periodic environmental disturbances [31], the coevolution of group joining and cooperative behaviors [21], the presence of non-participants [32], and multigenerational randomly-formed groups [17], as well as active methods such as kin recognition [33], conditional behavior based on the reputations [34, 35] or past behaviors [1, 2, 6, 26] of others, policing [36, 37], punishment of non-altruists [38-40], the coevolution of cultural institutions that constrain individual behaviors [41], and even recognition of arbitrary tags [42]. What all these mechanisms have in common is: 1) sufficient positive assortment among altruistic behaviors, 2) sufficient non-zero-sum benefits of cooperation, and 3) heritability of these behaviors.

REFERENCES

[1]   R. Axelrod and W. D. Hamilton, "The Evolution of Cooperation," *Science*, vol. 211, pp. 1390-1396, 1981.
[2]   R. Axelrod, *The Evolution of Cooperation*. New York: Basic Books, Inc, 1984.
[3]   W. D. Hamilton, "The Genetical Evolution of Social Behavior I and II," *Journal of Theoretical Biology*, vol. 7, pp. 1-52, 1964.
[4]   W. D. Hamilton, "Selfish and Spiteful Behavior in an Evolutionary Model," *Nature*, vol. 228, pp. 1218-1220, 1970.
[5]   W. D. Hamilton, "Altruism and Related Phenomena, Mainly in Social Insects," *Annual Review of Ecology and Systematics*, vol. 3, pp. 193-232, 1972.
[6]   R. L. Trivers, "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology*, vol. 46, pp. 35-57, 1971.
[7]   D. S. Wilson, "A Theory of Group Selection," *Proceedings of the National Academy of Science, USA*, vol. 72, pp. 143-146, 1975.
[8]   M. J. Wade, "A Critical Review of the Models of Group Selection," *The Quarterly Review of Biology*, vol. 53, pp. 101-114, 1978.
[9]   M. J. Wade, "Kin Selection: Its Components," *Science*, vol. 210, pp. 665-667, 1980.
[10]   D. C. Queller, "Kinship, Reciprocity and Synergism in the Evolution of Social Behavior," *Nature*, vol. 318, pp. 366-367, 1985.
[11]   D. C. Queller, "Quantitative Genetics, Inclusive Fitness, and Group Selection," *The American Naturalist*, vol. 139, pp. 540-558, 1992.
[12]   D. C. Queller, "A General Model for Kin Selection," *Evolution*, vol. 46, pp. 376-380, 1992.
[13]   S. A. Frank, *Foundations of Social Evolution*. Princeton: Princeton University Press, 1998.
[14]   E. Sober and D. S. Wilson, *Unto Others, The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press, 1998.
[15]   G. R. Price, "Selection and Covariance," *Nature*, vol. 277, pp. 520-521, 1970.
[16]   W. D. Hamilton, "Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics," in *Biosocial Anthropology*, R. Fox, Ed. New York: John Wiley and Sons, 1975, pp. 133-155.
[17]   J. A. Fletcher and M. Zwick, "Strong Altruism Can Evolve in Randomly Formed Groups," *Journal of Theoretical Biology*, (in press) 2004.
[18]   J. W. Pepper, "Relatedness in Trait Group Models of Social Evolution," *Journal of Theoretical Biology*, pp. 355-368, 2000.
[19]   G. Hardin, "The Tragedy of the Commons," *Science*, vol. 162, pp. 1243-48, 1968.
[20]   J. McMillan, "The Free-Rider Problem: A Survey," *Economic Record*, vol. 55, pp. 95-107, 1979.
[21]   L. Avilés, "Solving the Freeloaders paradox: Genetic associations and frequency-dependent selection in the evolution of cooperation among nonrelatives," *Proceedings of the National Academy of Science, USA*, vol. 99, pp. 14268-14273, 2002.

[22] J. A. Fletcher and M. Zwick, "Altruism, the Prisoner's Dilemma, and the Components of Selection," presented at IEEE Systems, Man, and Cybernetics Conference, Tucson, AZ, 2001.

[23] J. A. Fletcher and M. Zwick, "N-Player Prisoner's Dilemma in Multiple Groups: A Model of Multilevel Selection," presented at Artificial Life VII Workshop Proceedings, Portland, Oregon, 2000.

[24] J. A. Fletcher and M. Zwick, "Simpson's Paradox Can Emerge from the N-Player Prisoner's Dilemma: Implications for the Evolution of Altruistic Behavior," presented at Proceedings of The World Congress of the Systems Sciences and ISSS 2000, Toronto, Canada, 2000.

[25] A. Rapoport, "Ideological Commitments and Evolutionary Theory," *Journal of Social Issues*, vol. 47, pp. 83-99, 1991.

[26] L. A. Dugatkin, *Cooperation Among Animals, An Evolutionary Perspective*. New York: Oxford University Press, 1997.

[27] S. Karlin and C. Matessi, "The Eleventh R. A. Fisher Memorial Leture: Kin Selection and Altruism," *Proceedings of the Royal Society London B*, vol. 219, pp. 327-353, 1983.

[28] R. Dawkins, *The Selfish Gene*. New York: Oxford University Press, 1976.

[29] R. Dawkins, *The Extended Phenotype: The Gene as the Unit of Selection*. Oxford: W.H. Freeman and Co, 1982.

[30] J. W. Pepper and B. B. Smuts, "Assortment through Environmental Feedback," *The American Naturalist*, vol. 160, pp. 205-213, 2002.

[31] J. Mitteldorf and D. S. Wilson, "Population Viscosity and the Evolution of Altruism," *Journal of Theoretical Biology*, 2000.

[32] C. Hauert, S. De Monte, J. Hofbauer, and K. Sigmund, "Volunteering as Red Queen Mechanism for Cooperators in Public Goods Games," *Science*, vol. 296, pp. 1129-1132, 2002.

[33] G. J. Gamboa, H. K. Reeve, and W. G. Holmes, "Conceptual issues and methodology in kin-recognition research: A critical discussion," *Ethology*, vol. 88, pp. 109-127, 1991.

[34] M. A. Nowak and K. Sigmund, "Evolution of Indirect Reciprocity by Image Scoring," *Nature*, vol. 393, pp. 573-577, 1998.

[35] K. Panchanathan and R. Boyd, "A tale of two defectors: the importance of standing for evolution of indirect reciprocity," *Journal of Theoretical Biology*, vol. 224, pp. 115-126, 2003.

[36] S. A. Frank, "Mutual Policing and Repression of Competition in the Evolution of Cooperative Groups," *Nature*, vol. 377, pp. 520-522, 1995.

[37] S. A. Frank, "Perspective: Repression of Competition and the Evolution of Cooperation," *Evolution*, vol. 57, pp. 693-705, 2003.

[38] R. Boyd and P. J. Richerson, "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups," *Ethology and Sociobiology*, vol. 13, pp. 171-195, 1992.

[39] E. Fehr and S. Gächter, "Altruistic Punishment in Humans," *Nature*, vol. 415, pp. 137-140, 2002.

[40] R. Boyd, H. Gintis, S. Bowles, and P. J. Richerson, "The evolution of altruistic punishment," *Proceedings of the National Academy of Science, USA*, vol. 100, pp. 3531-3535, 2003.

[41] S. Bowles, J.-K. Choi, and A. Hopfensitz, "The co-evolution of individual behaviors and social institutions," *Journal of Theoretical Biology*, vol. 223, pp. 135-147, 2003.

[42] R. L. Riolo, M. D. Cohen, and R. Axelrod, "Evolution of cooperation without reciprocity," *Nature*, vol. 414, pp. 441-443, 2001.