6

Multi-level Selection, Assortment, and the
Evolution of Altruism

Any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as well developed, or nearly as well developed, as in man.


6.1 Introduction

Evidence that altruistic and other social preferences provide the proximate causes of important aspects of human behavior poses the following challenge: In light of what we know about other animals and about human evolution, how could these preferences have become common among humans?

Our explanation of how social preferences became common hinges on three facts. First, group living is essential to human survival. Second, groups differ in their evolutionary success, some producing large numbers of copies or contributing large numbers of individuals to membership in other groups, while other groups are absorbed into more successful groups or pass out of existence in warfare or during environmental crisis. Differential group success therefore plays a central role in the evolution of human behaviors and institutions, less successful groups copying the more successful or being eliminated by them. Well-documented examples of this process include the peopling of much of the world by individuals of European ancestry and the associated spread of European customs and institutions in the past half millennium, and the spread of agriculture and its associated novel systems of social organization and behavior from the Middle East to Eu-
rope beginning ten millennia ago. Third, groups in which altruistic social preferences are common tend to cooperate, and cooperative groups tend to prevail in intergroup competition, and to survive environmental crises.

The tendency of more cooperative groups to survive and expand thus means that altruistic individuals among our ancestors might have enjoyed enhanced reproductive success, tending to proliferate altruism as a distinctive human trait. It also suggests that our altruistic dispositions might motivate us to help not only immediate genetic relatives but also other members of the groups making up human society and to include not only feelings of filial love and fraternal solidarity, but also a tendency to extend similar sentiments to those with whom we do not share a recent common ancestor.

The likelihood that the group-structured nature of populations and group competition would strongly affect evolution has long been recognized not only by sociologists and historians (Tilly 1981, Parsons 1964) but also by biologists (Lewontin 1965, 1970, Alexander 1979, Wilson 1977, Durham 1991, Dunbar 1993, Laland and Feldman 2004). But, until recently, most biologists have concluded that group-level effects that would favor the spread of genes contributing to altruistic behavior cannot offset the effects of individual within-group selection operating against altruists, except where special circumstances heighten and sustain genetic differences between groups relative to differences within the group (Williams 1966, Crow and Kimura 1970, Boorman and Levitt 1973, Maynard Smith 1976). The reason is that the speed of an evolutionary process is proportional to the differences on which it works, so in order for between-group selection to outrun within-group selection, between-group differences must be substantial. But, gene flow due to group exogamy and other sources of migration are thought to preclude this.

Beginning with Darwin, however, a number of evolutionary theorists have suggested that human evolution might provide an exception to this negative assessment of the force of multi-level selection. William Hamilton (1975:331) summarized Darwin’s view as follows: “He saw that such traits [as]…courage and self-sacrifice…would naturally be counter-selected within a social group, whereas in competition between groups the groups with the most of such qualities would be the ones best fitted to survive and increase.” In *The Causes of Evolution*, J. B. S. Haldane (1932) provided a plausible mechanism for how this might come about. He suggested that in population of small endogamous ‘tribes,’ an altruistic trait might evolve because the ‘tribe splitting’ that occurs when successful
groups reach a certain size would by chance create a few successor groups
with a very high frequency of altruists, reducing within-group differences
and increasing between-group differences, a process very similar to that
modeled and simulated in Chapters 8 and 9. The modest size of typical
human groups during most of our evolution thus could play a crucial role in
the chance occurrence of one or more groups with a high fraction of altru-
ists, which would then proliferate. Haldane concluded: “evolution in large
random-mating populations… is not representative of evolution in general,
and perhaps gives a false impression of the events occurring in less numer-
ous species…. Our ancestors were mostly rather rare creatures.” (p. 213-14)
William Hamilton (1975) took up Haldane’s suggestion, adding that if the
allocation of members to successor groups following ‘tribe splitting’ was
not random but was rather what he called ‘associative,’ (p. 137) between
group differences would be even greater and multi-level selection pressures
would be further enhanced.

More recent research also suggests that impediments to multi-level selec-
tion may be less general than was once thought (Uyenoyama and Feldman
1980, Harpending and Rogers 1987). A number of writers have pointed
out that multi-level selection (We use the terms multi-level selection and
group selection interchangeably) may be of considerably greater impor-
tance among humans than among other animals given the advanced level
of human cognitive and linguistic capabilities and consequent capacity to
maintain group boundaries and to formulate general rules of behavior for
large groups, and the resulting substantial influence of cultural inheritance
on human behavior (Alexander 1987, Cavalli-Sforza and Feldman 1973,

Among the consequences of these distinctive human capacities are the
suppression of within-group phenotypic differences through egalitarianism,
coinsurance, pair bonded monogamous mating systems, consensus decision
making and conformist cultural transmission. Insider biases and individual
preferences to interact with like individuals lead to large between-group
differences supporting high levels of assortative interactions both within
and between groups and frequent between-group conflict. Other animals
do some of these things, but none does all of them on a human scale. All
of these aspects of human social life enhance the force of between group
selection relative to within-group selection.
Here we explain how group selection works and its relationship to other biological models of the evolution of human behavior thus far introduced. Group selection may take both a strong form, in which group competition offsets the selective pressures operating against altruistic members within a group, and a weak form, in which group competition selects among stable equilibria, favoring those with a higher level of cooperation. In section 6.4 we show that this group selection model shares an important characteristic with the other biologically inspired models reviewed in chapter 3, kin altruism and reciprocal altruism: they all work because altruists are more likely than non-altruists to interact with individuals who help them. The models differ in the processes that account for this behavioral assortment benefiting the altruistic types. In part for this reason the models also differ in the proximate motives likely to account for helping behavior, providing further clues, which we consider in the concluding section, to the puzzle of how human cooperation evolved.

### 6.2 Multi-level Selection

The multi-level selection model works because members of predominantly altruistic groups have above average fitness and thus contribute disproportionately to the next generation. Here we apply multi-level selection to a process of genetic transmission, but the same model applies to any process of selection based on the differential replication of traits over time. To see how group selection works, consider a single altruistic behavior that is the expression of an ‘altruistic allele’ in a haploid genetic system, the quotation marks being a reminder that altruistic behaviors are unlikely to be the expression of a single allele.

Thus, we consider a trait that may \((A)\) or may not \((N)\) be present in each individual in a large population that is subdivided into a number of reproductively somewhat isolated groups, commonly known as demes. Suppose altruistic behavior costs the individual \(c\) and confers a total benefit of \(b\) on another member of the group, where \(c\) and \(b\) are both measured in units of fitness, and \(b > c > 0\). It follows that a member in a group composed entirely of \(A\)’s has a payoff \(b - c > 0\), greater than a member of a group composed of all \(N\)’s. In any mixed group, however, the expected payoff to \(A\)’s will be lower than that of the \(N\)’s by an amount \(c\). These payoffs are the same as in Figure 3.1.
Let $p_{ij} = 1$ indicate that individual $i$ in group $j$ is an A, with $p_{ij} = 0$ if $i$ is an N. Let $p$ represent the fraction of A’s in the population at the start of a given time period, and let $p'$ be the fraction of A’s at the start of the next period. Define $w_{ij}$ as the expected fitness of an individual of type $i$ in group $j$ and let

$$w_{ij} = \beta_o + p_j \beta_g + p_{ij} \beta_i$$  \hspace{1cm} (6.1)$$

where $\beta_g$ and $\beta_i$ are the partial effects on $w_{ij}$ of $p_j$, the frequency of A in the group and $p_{ij}$, the presence of the A allele in individual $i$, respectively, and $\beta_o$ is a baseline replication rate based on factors not considered here. In equation 6.1 the two effects, between- and within-group selection, are separable, so that the size of the group effect on an individual is independent of whether the individual is altruistic or not, and the individual effect of being an altruist is independent of the frequency of altruists in the group. The assumption that the two effects are separable is unrealistic, but it allows an illuminating simplification in this case.

George Price (1972) showed that the change in the frequency of altruists in the population, $\Delta p \equiv p' - p$, can be partitioned into between-group and within-group effects. We prove this in Appendix A6, where we show that the partition can be written as

$$w \Delta p = \var(p_j) \beta_G + \beta_i \Evar(p_{ij}).$$ \hspace{1cm} (6.2)$$

where $w$ is the population-wide average of the number of replicas, which we normalize to unity, as we assume the population size is constant. The terms $\var(p_j)$ and $\Evar(p_{ij})$ respectively are the between-deme genetic variance and weighted average within-deme genetic variance (the weights based on group size). The coefficient $\beta_G$ is the effect of variation in $p_j$ on the average fitness of members of deme $j$ (that is, on $w_j$), which is determined as follows. If we sum 6.1 over the individuals in a group and divide by group size, we see that

$$w_j = \beta_o + p_j (\beta_g + \beta_i).$$ \hspace{1cm} (6.3)$$

So $\beta_G = \partial w_j / \partial p_j = \beta_i + \beta_g$.

Recall that a behavior is altruistic if adopting it, hypothetically switching from an N to an A, lowers one’s expected fitness while increasing the average fitness of one’s group. Our definition of altruism thus maps neatly onto the terms of the Price equation, and we are interested in the case where $\beta_i \equiv \partial w_{ij} / \partial p_{ij} < 0$ and $\beta_G > 0$. 
The separability assumption now allows us to represent equation 6.2 in terms of the payoffs: \( \beta_i = -c \) and \( \beta_g = b \). Thus equation 6.2 can be written

\[
\Delta p = \text{var}(p_j)(b - c) - c \text{var}(p_{ij}).
\] (6.4)

The first term captures the group effect, which is positive, by the definition of altruism, while the second represents the effect of within-group selection, which is negative, also by the definition of altruism. Setting aside degenerate cases such as zero variances, it follows that the frequency of the trait will be stationary (i.e., \( \Delta p = 0 \)) where these two terms are of equal absolute magnitude (assuming that the \( \beta \)'s and variances making up these terms are themselves stationary). Because the second term is negative, the frequency of the A-trait within all groups will fall over time. But as \( b - c \) is positive, this tendency will be offset by the decline in the size of groups with low frequencies of the trait and the expansion of groups with many altruists.

Let us define the variance ratio \( F \) as the ratio of the between-group variance in the fraction of altruists to the total population variance, which is the within-group plus the between-group variance of the fraction of altruists, or

\[
F = \frac{\text{var}(p_j)}{\text{Evar}(p_{ij}) + \text{var}(p_j)}
\] (6.5)

The variance ratio is thus a population-wide measure of the degree of non-randomness in who interacts with whom, resulting from the tendency of altruists to find themselves disproportionately in groups with many other altruists. The variance ratio \( F \) is Wright’s inbreeding coefficient \( F \) (Wright 1943), measuring the degree of genetic differentiation among groups.

This ratio measures the difference between the probability of being paired with an altruist conditional on being an altruist, \( P(A|A) \), and conditional on being a non-altruist, \( P(A|N) \), that arises because the population is group-structured (Crow and Kimura 1970). Thus,

\[
F = P(A|A) - P(A|N).
\] (6.6)

An example illustrating this result is given in §A6. Being “paired with an altruist” means being the recipient of the randomly assigned benefit, \( b \), that altruists confer on fellow group members. This is more likely to occur if there are many A’s in one’s group. If A’s tend to be in groups with many other A’s, they enjoy an advantage over N’s. \( F \) measures the extent of this
advantage. Rearranging terms in (6.2), we see that the condition for \( \Delta p \) to be zero so that the fraction of altruists in the population is stationary, is

\[
F = \frac{\beta_i}{\beta_g + \beta_i},
\]

or

\[
F = \frac{c}{b}
\]

If \( F > c/b \), the fraction of altruists may be expected to increase and if \( F < c/b \), it decreases.

Equation 6.8 also indicates the most costly form of altruism that may proliferate by this method. When the variance among group means is zero (so that \( F = 0 \)), A’s no longer have the advantage of being in groups with disproportionately many A’s. In this case multi-level selection is inoperative, so only a costless form of group beneficial behavior could proliferate. By contrast when \( \text{var}(p_{ij}) = 0 \) for all \( j \), groups are either all A or all N, and one meets only one’s own type, independently of the composition of the total population. In this case, within-group selection is absent and between-group selection is the only selective force at work. Thus, the force of multi-level selection will depend on the magnitude of the group benefit relative to the individual cost (\( b \) and \( c \)) and the degree to which groups differ in their frequency of the trait, relative to the within-group variance of the trait.

Figure 6.1 shows how the group structure of the population may overcome the disadvantage of bearing the costs of altruistic behaviors and indicates the variance ratio \( F \), namely the difference \( P(A|A) - P(A|N) \), that is just sufficient to equate the expected fitness of the two types and thus to maintain a stationary value of \( p \). As is evident from the figure, this is the \( F \) that satisfies equation 6.8, namely \( Fb = c \). The slope of both expected fitness functions is \( b \) and the distance between them is \( c \). How large \( P(A|A) - P(A|N) \) must be depends, as we have seen, and as the figure makes clear, on the costs and benefits of the altruistic behavior, \( b \) and \( c \). It is clear from the figure that for a given level of \( b \) (slope of the lines), the greater is the cost of altruism, \( c \) (the vertical distance between the lines) the greater must be the degree of positive assortment \( P(A|A) - P(A|N) \) in order for altruism to proliferate in the population. In Appendix A6 we provide an numerical example that will clarify how the model works.

The Price equation does not represent a complete dynamical system giving the movements of \( p \). This would require a set of equations giving the
movements over time of the between- and within-group variances. It is easy to check, for example, that the variances given in the numerical example in the Appendix that render \( p \) stationary are not themselves stationary. Because they will be different the next period, the values of \( b \) and \( c \) that made \( p \) stationary in the period under study will not ensure stationarity in subsequent periods. Except in degenerate cases of little interest, equations giving the movement of the relevant variances over time are not attainable. This is the reason that exploring evolutionary processes under the influence of group selection typically requires that we resort (in the next four chapters) to agent-based simulations.

![Figure 6.1. The Evolution of an Altruistic Trait in a Group-structured Environment](image)

The vertical axis measures expected fitness minus the baseline fitness \( \beta_o \). If the population structure’s variance ratio is such that the difference in the conditional probabilities of being paired with an A, \( P(A|A) - P(A|N) \), is as shown, \( p \) is stationary, because the expected fitness of the two types, \( w_j^A = bP(A|A) - c - \beta_o \) and \( w_j^N = bP(A|N) - \beta_o \) are equal.

### 6.3 Equilibrium Selection

The above model illustrates what we call strong multi-level selection, which occurs when the second term in the Price equation is negative, indicating that those with group-beneficial traits are altruistic and hence would attain higher fitness by abandoning their helping behaviors. But the second term may be zero, so that those who confer benefits on others suffer no disadvantage within groups while profiting from the group beneficial effects of their behaviors. When this is the case, we say weak multi-level selection is operative. The cooperative traits that may be supported in the equilibria
selected by this process are not altruistic, because cooperators suffer no disad-
advantage relative to others. There are two main cases of weak multi-level se-
lection.

The first type of weak multi-level selection occurs when both the A-trait
and the N-trait are evolutionarily stable strategies, so that two stable within-
group equilibria exist, one with all N’s and the other with all A’s, each of
which cannot be invaded by a small fraction exhibiting the other type of
behavior. This was the case, as we have seen, when the conditions for
the proliferation of cooperation by means of reciprocal altruism, indirect
reciprocity or costly signaling obtain. Note that in this case the second term
of the Price equation is zero because if \( p_j = 1 \) or \( p_j = 0 \), then \( \text{var}(p_{ij}) = p_j(1 - p_j) = 0 \). This would be the case, for example in the repeated game
setting in which both unconditional defection and conditional cooperation
are mutual best responses. But here, we model the case of weak group
selection using a above model of a one-shot interaction between A’s and
N’s. Two stable equilibria could exist in this case, for example if, following
an A-N interaction, the A’s in the group collectively attempted to punish the
N. The A’s would then be akin to the strong reciprocators whose preferences
were described in the previous chapter. Suppose their attempt to punish the
N succeeds with a likelihood equal to the fraction of A’s in the population,
and if successful the cost imposed on the N is \( c_p \). Assume each A incurs
a cost of \( k \) in attempting to punish the N. Then, the expected fitness of the
A’s in group \( j \) is

\[
w_j^A = p_j b - c - k(1 - p_j) + \beta_o,
\]

while the fitness of the N’s is

\[
w_j^N = p_j (b - p_j c_p) + \beta_o,
\]
giving the expected fitness functions within a single group illustrated in
upper panel of Figure 6.2, for the case where \( c_p > c \), so the target of the
punishment bears a greater cost than does an individual punisher. Note
that if \( p > p^* \), A’s will have higher fitness and thus eventually eliminate
the N’s, while if the reverse inequality holds, N’s will have higher fitness
and thus eventually eliminate the A’s. A population composed of many
groups with this interaction structure could contain some with all A’s and
some with all N’s. Because groups are homogeneous, they would remain
so, but the all A groups would have higher average fitness and if the total
population is constant, the \( N \)'s would eventually be eliminated. Boyd and Richerson (1985) showed that conformist cultural transmission, a tendency to copy the more common behaviors, can have the same effect, giving rise to homogeneous groups among which weak group selection can support the proliferation of an altruistic trait.

In the second type of weak multi-level selection, cooperative individuals coexist with others in a stable within-group equilibrium. Multi-level selection works because groups in which the equilibrium number of cooperators is greater prevail in competition with other groups. In this case, at the equilibrium for a particular group, the second term of the right hand side of the price equation is zero because the fitness of the two types must be equal (otherwise it is not an equilibrium) and as a result, the A trait suffers no fitness deficit. An example in a model of cultural rather than genetic evolution is one in which non-altruists may turn into altruists, by being socialized through group rituals to behave altruistically, and altruists can revert back to nonaltruists, attracted by the possibility of not paying the cooperation cost \( c \). Equilibrium now occurs when the two movements from altruist to nonaltruist and back lead to equal fitness of both phenotypes. This is shown in the lower panel of Figure 6.2, which depicts the model just described. We present a more complete model along these lines in Chapter 11.

We assume all agents are genetically identical and there is clonal reproduction. We assume also that altruist phenotypes (A's) pass their cultural preference on to their offspring, as do the non-altruists (N's). As this is a cultural model, the term fitness means the number of replicas that an individual will make in the next period which may be one, if the individual retains the status quo trait or greater, if the individual is copied by another, or zero if the individual abandons the trait. The fitness of an N due to material payoffs is now \( m^N_j = b p_j \), where \( p_j \) is the fraction of A's in group \( j \), and the fitness of an A due to material payoffs (we term this "material fitness") is \( m^A_j = b p_j - c \). Suppose \( w^A_j \) is the average fitness of A's, including the contribution \( m^A_j \), as well as the net probability that N's switch to A's through socialization and from A's by imitation. Similarly, \( w^N_j \) is the average fitness of N's, including the contribution \( m^N_j \), as well as the net probability that A's switch to N's through imitation. Let \( \gamma_a \) be the rate per non-altruist at which N's are socialized into A's, and let \( \gamma_n \) be the rate per unit of payoff difference per altruist at which A's revert to N's through imitation. Since the payoff difference between A's and N's is \( m^N_j - m^A_j = c \),
we have

\[ w_j^A = b p_j + \gamma_a (1 - p_j) - \gamma_n c p_j - c + \beta_0 \]  
(6.9)

\[ w_j^N = b p_j - \gamma_a (1 - p_j) + \gamma_n c p_j + \beta_0 \]  
(6.10)

where \( \beta_0 \) is baseline fitness. There is then a long-run phenotypic equilibrium when \( w_j^A = w_j^N \) with a fraction \( p_j^* \) of altruists within the group given by

\[ p_j^* = \frac{2\gamma_a - c}{2(\gamma_a + c\gamma_n)} \]

The average material fitness of a member of group \( j \), which depends only on the material payoffs and not the phenotypic movements, is then

\[ p_j^* m_j^A + (1 - p_j^*) m_j^N = \frac{(b - c)(2\gamma_a - c)}{2(\gamma_a + c\gamma_n)} + \beta_0. \]

Because in equilibrium this is greater than \( \beta_0 \) provided \( \gamma_a > c/2 \), a group with an effective socialization process will have higher mean material fitness than a group without such a process. Therefore any inter-group dynamic that favors wealthier groups will lead to the spread of the socialization process by group selection.

In the upper panel, depicting the first model in which both cooperation and non-cooperation are evolutionarily stable strategies, \( F = 1 \), reflecting the fact that all of the variance is between groups and none is within, so the second term on the right hand side of Price equation is zero. This, or equivalently condition 6.8, tells us that altruism can proliferate no matter how small is the excess of benefits to others over cost to self. In the lower panel in equilibrium (that is, at \( p_j^* \)) \( \beta_i = 0 \), as there is no within-group selection operating against the A’s, the second right hand side term in the Price equation is zero in this case too and equation 6.8 tells us that any positive \( F \) will be sufficient to promote the spread of altruism.

### 6.4 Positive Assortment

In the multi-selection model, as in the kin altruism models, the evolution of helping behaviors requires that those with a predisposition to help others receive help from those with whom they interact more frequently than would occur by chance. How this positive assortment comes about differs
Figure 6.2. Weak Multi-level Selection. In the upper panel, \( p_j^* \) is an unstable equilibrium and constitutes the boundary between the basin of attraction of the all-N and the all-A equilibria. In the lower panel, \( p_j^* \) is a stable equilibrium. In both cases within-group selection will result in a stationary distribution of A’s and N’s with no within-group selection operating against the A’s. By contrast, Figure 6.1 illustrates strong multi-level selection.

from case to case. Eshel and Cavalli-Sforza (1982) provide a glimpse of the variety of causes of assortment: “Kin, deme, niche, and social group structure, neighborhood effect, idiosyncratic behavior, and discrimination in the choice of companions are some of the possible sources of deviation from randomness...critical for understanding the evolutionary stability of social structure.”

For the group selection model, we have equation 6.8, which can be written:

\[
P(A|A) - P(A|N) > \frac{c}{b}.
\]

This is a generalization of Hamilton’s rule for the degree of positive assortment permitting an altruistic trait to proliferate when rare. This can be seen by returning to Figure 6.1 and noting that the same model applies to kin selection using the fact that the degree of genetic relatedness between the two members of an interacting pair \( P(A|A) - P(A|N) \), which is equal to \( r \). \( P(A|A) \) is the probability that the bearer of an altruistic allele will inter-
act with another bearer of that allele, while $P(A|N)$ is the probability that the bearer of the non-altruistic allele is paired with a bearer of the altruistic allele. The former will exceed the latter, reflecting positive assortment, if interactions are among individuals of common recent genealogical descent. Because the processes contributing to positive assortment among close descendant kin on the one hand and in group-structured populations on the other differ in important ways, we refer to evolutionary process based on the former as kin selection, and use the relatedness term $r$, while we refer to the latter as multi-level selection, and use the variance ratio $F$.

Surprisingly, the same expression gives the condition for reciprocal altruism to evolve even when the process of dyadic pairing to interact is random. This has been shown in a number of ways, beginning with Queller (1985) and including Nee (1989) and Fletcher and Zwick (2006). Here, the altruists are the Nice Conditional Cooperators, while the non-altruists unconditionally defect. In this interaction, $A$’s and $N$’s will be paired with an $A$ with the same probability, $p$. But positive assortment of the $A$ allele and the cooperative phenotype occurs nonetheless because an $N$ paired with an $A$ will benefit from the partner’s cooperation for just one period because the $A$, a Conditional Cooperator, will switch to defect after the first round. By contrast, the $A$ paired with an $A$ will benefit from the partner’s cooperation for as many periods as the interaction endures. The key to the evolutionary success of the Conditional Cooperator in this situation is (Fletcher and Zwick 2006): 253

“there must be sufficient positive assortment between individuals with the altruistic genotype...and the helping phenotypes of others they interact with”

Thus, suppose we observe a large randomly-paired population in which the fraction of Conditional Cooperators is $p$, and we count the periods in which an $A$ and an $N$ respectively benefit from the cooperation of a partner as a fraction of the total periods of interaction for each type. The former will exceed $p$ while the latter will be less than $p$. The longer the interaction endures, the greater will be the degree of positive assortment between altruistic genotypes and helping phenotypes measured by this difference.

It remains to show that the condition for Nice Conditional Cooperator to be an evolutionary stable strategy is also just another version of $P(A|A) - P(A|N) > c/b$. Recall that this condition (equation 3.2) is $\delta > c/b$ where $\delta$ is the probability of continuation of the interaction at the end of
each interaction, and $b$ and $c$ are respectively the benefits and costs of the cooperative act undertaken by the reciprocal altruist ($A$). Suppose that the expected total duration of an interaction from initiation to termination is normalized to one and that opportunities to alter one’s action occur at the end of every period, each of which is a fraction, $1 - \delta$, of the expected duration of the interaction. Thus there are $1/(1 - \delta)$ periods in the expected duration, and the probability of continuation at the end of each is $\delta$. We measure the advantageous pairing of the $A$ types by the fraction of the expected total duration of the interaction in which they experience helping from the individual with whom they are paired. To determine if $A$ (all Nice Conditionally Cooperators) is an ESS, we study the difference in the number of periods in which an $A$ and an $N$ respectively will experience cooperation from its partner. The duration of cooperation enjoyed by an $A$ is $p$, because with probability $p$ the $A$ is paired with an $A$ in which case they both cooperate for the expected duration of the interaction, which is 1. The corresponding duration of cooperation enjoyed by an $N$ is $p(1 - \delta)$ because with probability $p$ the $N$ will be paired with an $A$, who cooperates during the first period, which is a fraction $(1 - \delta)$ of the expected duration, and then defects for all subsequent periods. Setting $p = 1$ to study the evolutionary stability of the all Conditionally Cooperator equilibrium and letting $P(C|A)$ and $P(C|N)$ represent the expected duration in which an $A$ and an $N$ respectively experience cooperation from its partner, we have: $P(C|A) - P(C|N) = \delta$. Thus $\delta$ is a measure of positive assortment given by the difference in the frequency of cooperative actions by one’s partner, conditional on one’s type. So we can rewrite the condition for Conditionally Cooperate to be an ESS as $P(C|A) - P(C|N) > c/b$.

This demonstrates that the condition for helping behaviors to evolve by repeated interactions is the same as the condition for altruistic cooperation to evolve by means of kin selection or group selection. An ingenious set of papers (Ohtsuki et al. 2006) suggests yet another example of this logic, and one that captures essential aspects of human society. Until now we have studied just two types of social structure: families, meaning kin of recent common descent, and demes, that is sub-populations within a larger population. But, within any group, some individuals are closer to some than to others, and this social proximity should make a difference in evolutionary dynamics. Ohtsuki and coauthors show that it does.

Suppose Cooperators and Defectors are arrayed on a network and each plays a public goods game with all of the $k$ other “neighbors” with whom
they are directly linked. As above, at a cost of $c$, each cooperative individual contributes an amount $b$ to be shared among its $k$ neighbors, whether they are Cooperators (C) or Defectors (D). Defectors bear no costs and contribute no benefits. Individuals periodically may alter their strategy by adopting a strategy of a neighbor, with a probability proportional to the relative fitness of that neighbor. Thus, the individual will adopt C with probability equal to the total payoffs of neighboring C’s divided by the total payoffs of all neighbors.

Through extensive simulations using a wide range of network structures, Ohtsuki and his coauthors find that the Cooperators proliferate if $1/k > c/b$ where $k$ is the average number of neighbors (the degree) of the players. This surprising result occurs because, for reasons that are not fully understood as yet, by chance there will be one more C in the neighborhood of C’s than in the neighborhood of D’s independently of the size of the neighborhood. Thus, the smaller the neighborhood, the more different in their compositions are they. It seems likely that the fact that smaller neighborhoods are (relatively) more different one from another than are larger ones arises by a process similar to sampling error where differences in the means of small samples exceed differences in the means of larger samples. Whatever the explanation of this regularity, the probability that any given one of my $k$ neighbors will be a C differs by $1/k$ depending on whether I am a C or a D. Thus, $1/k$ is exactly equal to $P(A|A) - P(A|N)$, affirming the positive assortment logic already demonstrated in the other models.

Whether this model explains observed forms of human cooperation is open to some question, however. The networks of foraging and other small scale societies are typically quite large often including virtually all adult members of a group. If average degree ($k$) of an individual were 20, altruism could not proliferate by this mechanism unless costs did not exceed five percent of the benefits. The average number of exchange partners in the Southern African !Kung coinsurance (so called hxaro) networks described in §7.2 is 17, while the families among whom food is shared on a regular basis among both the Paraguayan Aché exceeds ten even for small packages, and 20 for large packages (Wiessner 2002, Kaplan and Hill 1985a). Average degree in the coinsurance networks among the Pokot and Himba herders are much larger than this (Bollig 2006). Thus, in order for a form of altruistic cooperation to be spread by this mechanism, it would have to have truly extraordinary ratios of benefit to cost.
Not surprisingly, positive assortment is also at work in the indirect reciprocitvity and costly signaling models. In all cases the degree of positive assortment determines the critical cost benefit ratio for the proliferation of altruism. The differences among the models arise in the way that positive assortment comes about. These results are summarized in Table 6.1.

<table>
<thead>
<tr>
<th>Model (citation)</th>
<th>Rule</th>
<th>Interpretation</th>
<th>Reason for Assortment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kin Altruism (Hamilton, 1964)</td>
<td>$r = \text{average degree of genetic relatedness among interacting individuals}$</td>
<td>$r = \text{difference in the probability that the bearer of an } A \text{ allele or the bearer of the } N \text{ allele (resp.) will interact with another bearer of the } A \text{ allele}$</td>
<td>Limited dispersal of close relatives</td>
</tr>
<tr>
<td>Group Selection (Aoki, 1992)</td>
<td>$F &gt; c/b; F = \frac{\text{fraction of variance of } A \text{ allele that is between demes}}{\text{var}(p_j)}$</td>
<td>$F = \text{difference in the probability that } A \text{ and an } N \text{ will be paired with an } A \text{ arising from the deme structure}$</td>
<td>Deme structure of population; associative deme fission</td>
</tr>
<tr>
<td>Reciprocal Altruism (Trivers 1971)</td>
<td>$\delta &gt; c/b; \delta = \frac{\text{probability that the interaction is not terminated at end of a given period}}{\text{var}(p_{ij}) + \text{var}(p_{ji})}$</td>
<td>$\delta = P(C</td>
<td>A) - P(C</td>
</tr>
<tr>
<td>Network Reciprocity (Ohtsuki, Hauert, Liberman and Nowak 2006)</td>
<td>$\frac{1}{k} &gt; c/b; k = \text{average degree of the network}$</td>
<td>$k = \text{difference in number of interactions in which one’s partner cooperates, conditioned on strategy } A \text{ or } N \text{ when } p = 1$</td>
<td>Similar to deme structured population; chance plus small size generates substantial $P(A</td>
</tr>
</tbody>
</table>

Table 6.1. Positive Assortment and Altruism
6.5 Conclusion

Each model explains why humans might have come to behave cooperatively while differing in the sociological, strategic and demographic mechanisms accounting for positive assortment and remaining silent as to the proximate motives leading them to engage in helping, and sometimes altruistic, behaviors. We have seen that positive assortment is a common feature of all of these models, so the distinctive character of each must lie in the processes explaining how positive assortment comes about, or in the kinds of social preferences likely to result.

Thus, for example, if kin altruism were the main reason for the evolutionary success of helping behaviors, the motives involved in helping would include a love for one’s children and other close genetic relatives and a concern for their well-being sufficient to motivate self-sacrifice. Similarly, we would expect that if the sole mechanisms contributing to the evolution of helping behavior were reciprocal altruism, indirect reciprocity or costly signaling, the proximate motives for helping and the cognitive processes activating them would involve individual advantage, accompanied by detailed behavioral bookkeeping about one’s fellows. Finally, we will see in subsequent chapters that the most plausible model of group selection involves warfare among groups, and hence strong group selection. If this is indeed the main evolutionary mechanism accounting for cooperative behaviors, it would most likely favor helping to be motivated by feelings of solidarity and generosity towards the members of one’s group extending beyond relatives of recent common descent, accompanied by the lack of such feelings towards members of other groups. A common feature of all of these preferences is that helping is conditional: on close genetic relatedness, on the likelihood of future interactions, on the possibility of reputation building or signaling, or on group membership.

The fact that helping behaviors are indeed motivated by this wide range of proximate motives, from maternal love, to enlightened self interest, to solidarity with one’s co-ethnics or co-nationals, is consistent with our view that in all likelihood each of the mechanisms we have identified here has played a significant role in human evolution.

Not surprisingly, given this account, recent (and therefore provisional) experiments in behavioral neuroscience have found that in some settings helping behaviors are correlated with activation in brain regions associated with cognitive operations such as the prefrontal cortex, as the reciprocal
altruism, indirect reciprocity, and signaling models would predict, while in other settings brain activities associated with social reward processing are involved, consistent with models of kin altruism and multi-level selection.

All of the models in table 6.1, and no doubt other mechanisms as well, were involved in the evolution of human cooperation, the importance of each depending on the forms of cooperation under consideration and the ecological and social conditions under which ancestral humans interacted. We will see in the next chapter that what can be known or conjectured about these ancestral human conditions from genetic, archaeological and other data suggests that the two workhorse models of evolutionary biology, helping close family members and reciprocal altruism, do not provide an adequate account of the emergence of this cooperative species. In subsequent chapters we will show that group selection models based on gene culture coevolution contribute substantially to a convincing explanation.