INNATE SOCIAL APTITUDES OF MAN: AN APPROACH FROM EVOLUTIONARY GENETICS

W. D. HAMILTON

It has become clear that, although learning has great importance in the normal development of nearly all phases of primate behaviour, it is not a generalized ability; animals are able to learn some things with great ease and other things only with the greatest difficulty. Learning is part of the adaptive pattern of a species and can be understood only when it is seen as a process of acquiring skills and attitudes that are of evolutionary significance to a species when living in the environment to which it is adapted.

WASHBURN, JAY, and LANCASTER

SURVIVAL OF THE FITTEST

The phrase directs attention to differential survival. Darwin accepted it from Herbert Spencer as adequately expressing the idea of natural selection. While accusations of tautology seem hardly fair on this small phrase itself, it must be admitted that some descendant ideas in the theory of natural selection are open to attack. For example, the idea of measuring ability to survive and reproduce—biological ‘fitness’—has undoubtedly been useful, but a slight haziness still lingers, a lack of precise and general definition: we do not know exactly what qualities natural selection is after. I think it is doubt like this rather than doubt about the reality and effectiveness of natural selection that inspires a present spirit of caution in evolutionary biology, including caution and distrust towards Spencer’s ideogram whenever it renews aspirations to become a slogan.

A part of the difficulty, and the part I am mainly concerned with now, is that of saying exactly what are the things that natural selection is supposed to select.

The fittest what? Is it a trait, an individual, a set of individuals bearing a trait, or bearing its determinants expressed or latent? Can it be a population, a whole species, perhaps even an ecosystem? In such a confusion of possibilities (and of fervent opinions either way) the individual organism stands out as one clear and obvious choice, with the number of its offspring as the measure of its fitness. But, beyond the problem of when to count and how to weight offspring for their ages, there is the problem that in sexual species the individual is really a physical composite of contributions from two parents and it may be composite in slightly different ways for different parts. Moreover, Mendel's principles concerning the fair distribution of genes to gametes and fair competition of these in fertilization do not always hold, so that the set of offspring of a given individual may carry a biased sample from the composite. Does this matter? For safe conclusions, do we have to descend to the level of the individual gene, perhaps ultimately to that of changed or added parts of the replicating molecule? Or can we, on the contrary, confidently follow the consensus of biologists to a higher level, in believing that the generally significant selection is at the level of competing groups and species? I shall argue that lower levels of selection are inherently more powerful than higher levels, and that careful thought and factual checks are always needed before lower levels are neglected. In this I follow a recent critical trend in evolutionary thought (see also Chapter 8 and references therein). Incidentally, to a biologist, a rather similar critique seems to be invited by the supposition that cultural evolution is independent of evolution in its biological substratum: to come to our notice cultures, too, have to survive and will hardly do so when by their nature they undermine the viability of their bearers. Thus we would expect the genetic system to have various inbuilt safeguards and to provide not a blank sheet for individual cultural development but a sheet at least lightly scrawled with certain tentative outlines. The problem facing a humane civilization may be how to complete a sketch suggesting some massive and brutal edifice—say the outlines of an Aztec pyramid—so that it reappears as a Parthenon or a Taj Mahal. These ideas concerning cultural evolution will not be expanded in what follows, but I hope to produce evidence that some things which are often treated as purely cultural in humans—say racial discrimination—have deep roots in our animal past and thus are quite likely to rest on direct genetic foundations. To be more specific, it is suggested that the ease and accuracy with which an idea like xenophobia strikes the next replica of itself on the template of human memory may depend on the preparation made for it there by selection—selection acting, ultimately, at the level of replicating molecules.

Returning to the problem of units of selection, Darwin himself, vague about the process of heredity, based most of his arguments on considerations of the fitness of individuals. He made occasional exceptions, as for the social insects where he treated the 'family group' as the unit of selection. I believe even these limited concessions were incautious (Chapter 8), and value his judgement more where, discussing the evolution of courage and self-sacrifice in man, he left a difficulty apparent and unresolved. He saw that such traits would naturally be counterselected 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. This open problem which Darwin left is really the starting-point of my own argument, but it is historically interesting to note that after some initial wavering between the calls of Spencer, Kropotkin, and others, almost the whole field of biology stumped in the direction where Darwin had gone circumspectly or not at all.

Until the advent of Mendelism uncritical acceptance of group selection could be understood partly on grounds of vagueness about the hereditary process. For example, courage and self-sacrifice could spread by cultural contagion and, in so spreading, modify heredity as well. But in the event neither the rediscovery of Mendel's work nor the fairly brisk incorporation of Mendelism into evolutionary theory had much effect. From about 1920 to about 1960 a curious situation developed where the models of 'Neodarwinism' were all concerned with selection at levels no higher than that of competing individuals, whereas the biological literature as a whole increasingly proclaimed faith in Neodarwinism, and at the same time stated almost all its interpretations of adaptation in terms of 'benefit to the species'. The leading theorists did occasionally point out the weakness of this position but on the whole concerned themselves with it surprisingly little (references in Chapters 2, 6, and 8).

With facts mostly neutral and theory silent it seems that we must look to the events and the 'isms' of recent human history to understand how such a situation arose. Marxism, trade unionism, fears of 'social darwinism', and vicissitudes of thought during two world wars seem likely influences. Confronted with common social exhortations, natural selection is easily accused of divisive and reactionary implications unless 'fittest' means the fittest species (man) and 'struggle' means struggle against nature (anything but man). 'Benefit-of-the-species' arguments, so freely used during the period in question, are seen in this light as euphemisms for natural selection. They provide for the reader (and evidently often for the writer as well) an escape from inner conflict, exacting nothing emotionally beyond what most of us learn to accept in childhood, that most forms of life exploit and prey on one another.

**LEVELS OF SELECTION**

Often the problem is not acute. There are many traits like resistance to disease, good eyesight, dexterity which are clearly beneficial to individual, group, and species. But with most traits that can be called social in a general sense there is some question. For example, as language becomes more sophisticated there is
also more opportunity to pervert its use for selfish ends: fluency is an aid to persuasive lying as well as to conveying complex truths that are socially useful. Consider also the selective value of having a conscience. The more consciences are lacking in a group as a whole, the more energy the group will need to divert to enforcing otherwise tacit rules or else face dissolution. Thus considering one step (individual vs. group) in a hierarchical population structure, having a conscience is an ‘altruistic’ character. But for the next step—group vs. supergroup—it might be selfish, in the sense that the groups with high levels of conscience and orderly behaviour may grow too fast and threaten to over-exploit the resources on which the whole supergroup depends. As a more biological instance similar considerations apply to sex ratio, and here a considerable amount of data has accumulated for arthropods (Chapter 4).

A recent reformulation of natural selection can be adapted to show how two successive levels of the subdivision of a population contribute separately to the overall natural selection. The approach is not limited to Mendelian inheritance but its usefulness in other directions (e.g. cultural evolution) has not yet been explored.

Consider a population consisting of a mixture of particles, and suppose we are interested in the frequency of a certain kind of particle G. Suppose the particles are grouped: let the subscript s denote the sth subpopulation. For subpopulation and for the whole we define parameters relevant to natural selection as follows:

<table>
<thead>
<tr>
<th>Subpopulation</th>
<th>Whole population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particles</td>
<td>$N = \sum n_s$</td>
</tr>
<tr>
<td>Frequency of G</td>
<td>$q = \frac{\sum n_s q_s}{N}$</td>
</tr>
<tr>
<td>Mean fitness</td>
<td>$w = \frac{\sum n_s w_s}{N}$</td>
</tr>
</tbody>
</table>

Fitness measures the amount of successful replication of particles in one ‘generation’. Thus the total population of the next generation will be $N' = \sum n_s w_s$. The symbol $f$ (denoting ‘next generation’) is used again in the same sense in the following further addition to notation:

<table>
<thead>
<tr>
<th>Subpopulation</th>
<th>Whole population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in frequency of G in one generation</td>
<td>$\Delta q_s = q_s' - q_s$, $\Delta q = q' - q$</td>
</tr>
</tbody>
</table>

With such notation it is easy to derive:

$$w \Delta q = \sum n_s w_s (q_s' - q_s)/N + \sum n_s w_s \Delta q_s/N = \text{Covariance } (w_s, q_s) + \text{Expectation } (w_s \Delta q_s).$$  \hspace{1cm} (1)$$

where Covariance and Expectation are understood to involve weighting by the $n_s$ as indicated. This is Price’s form. The covariance term represents the contribution of intergroup selection, so quantifying the intuitive notion that high $q_s$ must cause high $w_s$ for selective change to occur. The expectation term represents the contribution of intragroup selection. It is possible to apply the formula within itself, to expand $w_s \Delta q_s$. For example, if the next level is that of diploid individuals and $i$ indexes the ith individual of the sth group we have $n_i = \sum n_s$, $q_i = \sum 2q_{ms}/n_i$, and $w_i = \sum 2w_{is}/n_i$, where these summations are understood to cover all i instead of all s as previously. Then $w_i \Delta q_i$ decomposes into two terms, one of which represents ordinary diploid selection with strictly Mendelian inheritance, while the second represents the effects of genetic ‘drift’ (random sampling effects), and ‘drive’ (non-Mendelian ratios). Even this latter term can be reformulated using equation (1), but then our ‘groups’ are the fundamental particles themselves, which, neglecting mutation, must give $\Delta q_{\text{particle}} = 0$, so that here finally the second term goes out.

An often useful rearrangement of (1), which shows the dependence of selection on the variability in its units, introduces the regression coefficient of $w_s$ on $q_s$. If $\beta_1$ is this coefficient:

$$w_i \Delta q = \beta_1 \text{ Variance } (q_s) + \text{Expectation } (w_i \Delta q_s).$$  \hspace{1cm} (2)$$

Conceptual simplicity, recursiveness, and formal separation of levels of selection are attractive features of these equations. But, of course, being able to point to a relevant and generally non-zero part of selective change is far from showing that group selection can override individual selection when the two are in conflict. Moreover, even the possibility of devising model circumstances in which a positive group-selection term (first term) outweighs a negative individual selection one (second term, assuming no further levels), gives no guarantee that ‘altruism’ can evolve by group selection: we have to consider whether the population can get into the specified state, and, if it can, whether its present trend will continue. For example, if we suppose persistent groups with no extinction and no intergroup migration it is easy to arrange that the group-beneficial effect ($\beta_1$), of frequent altruism in a certain group is so large that the rapid expansion of the group with the highest frequency of G ($q_m$ say) draws the population $q$ rapidly upwards. But $q$ will never reach or pass $q_m$, and must eventually fall, remaining below the ever-falling value of $q_m$. Admittedly, all this is reasonably obvious without the equation; but the equation does emphasize that natural selection depends on a certain variance which is in this model must at last die away as the best group increasingly predominates. This is the essential objection to an algebraic model of Haldane for selection of altruism, which other writers have wrongly treated as the first successful analytical model for altruism. In verbal discussion Haldane himself admitted the necessity of a device to maintain diversity. He suggested that if groups split on reaching a certain size, random assortment of altruists and egotists would raise
the frequency of altruists in some daughter moieties, and if the critical size was low enough and the group advantage of altruism high enough, a process having endless overall enrichment in altruism might be devised. Increasing the intergroup variance by random (or, better, associative) division of existing tribes leaves less variance within groups, which, as a development of the equation will shortly make plain, weakens the power of individual selection, and this further improves the case. But Price’s equation does not seem to lend itself to a detailed analysis of Haldane’s suggestion—indeed the lack of analysis by Haldane himself suggests that it is not easy. The value of the covariance approach lies not so much in analytical penetration as in clarifying the approach to a problem.

Therefore, noting hopeful auguries in Haldane’s tribe-splitting no-migration idea, let us now turn to a model at the opposite extreme in which groups break up completely and re-form in each generation. Suppose that on reaching maturity the young animals take off to form a migrant pool, from which groups of \( n \) are randomly selected to be the group of the next generation. Assume completely asexual reproduction (or perfect matrilinear or patrilinear inheritance of a cultural trait in an ordinary population), and assume that an altruist gives up \( k \) units of his own fitness in order to add \( K \) units to the joint fitness of his \( (n - 1) \) companions. These companions are a random selection from the gene pool and therefore, in a supposed infinite population, have the expected gene frequency of the gene pool. Thus compared to a non-altruist, the altruist is putting into the next gene pool fewer of his own genes plus a random handful from the pool of the last generation. Obviously his trait is not enriching the population with genes that cause the trait. The specification of grouping has been a mere gesture. Nevertheless it is instructive to see how equation (2) handles the matter.

With asexuality individuals are basic particles, so, as already explained, the recursive use of (2) to expand its second term gives simply:

\[
w \Delta q = \beta_1 Var(q_i) + E(\beta_0 Var_s(q_{ii})).
\]

All units are now of the same size, so \( Var(\text{inance}) \) and \( E(\text{expectation}) \) can have their conventional meanings. Since \( \beta_0 \) does not vary with group constitution,

\[
w \Delta q = \beta_1 Var(q_i) + \beta_0 E(Var_s(q_{ii}));
\]

and the expectation is what is commonly called the within-group variance.

With random grouping, the distribution of the different compositions of groups will be binomial with parameters \( (q, n) \). The variance of \( q_i \) is \( \frac{1}{n}pq \).

Likewise it is easily shown that \( E(Var_s(q_{ii})) = \frac{n - 1}{n}pq \), so that

\[
w \Delta q = \frac{1}{n}pq(\beta_1 + (n - 1)\beta_0).
\]

Thus by inspection and by subtraction, respectively

\[
\beta_1 = K - k \quad \text{and} \quad \beta_0 = -k - \frac{K}{n - 1}.
\]

Substituting in (4) we find:

\[
w \Delta q = -kpq.
\]

This confirms the earlier argument that altruism cannot progress in such a model. It seems at a first glance that the benefits dispensed by altruists have been entirely null in the working of the model, but they affect it through their involvement in mean fitness:

\[
w = p + q(K - k).
\]

This being the only involvement of \( K \), we see that the most that altruism can achieve in the model is a slowing of the rate at which natural selection reduces its frequency—an effect which I explained earlier as altruism diluting each new gene mixture by adding, as it were, handfuls taken randomly from the previous one. Apart from this minor effect the model, like Haldane’s algebraic one, is a failure, in spite of having shifted to the opposite extreme in respect of migration. It reveals a group-selection component which is not zero but which is bound in an unchanging subordination to the individual selection component. However, the relation between the two variances in this case suggests how we must change the model to make altruism succeed. \( Var(q_i) \) must be increased relative to \( E(Var_s(q_{ii})) \). As already mentioned, this can be done by making \( G \) assort positively with its own type in settling from the migrant cloud. Suppose it sorts to such a degree that the correlation of two separate randomly selected members of a group is \( F \). If this correlation is achieved by having a fraction \( F \) of groups made pure for each type and then the remainder again formed randomly, then it is easily shown that the between- and within-group variances are respectively:

\[
\frac{1}{n}pq(1 - F + nF) \quad \text{and} \quad \frac{1}{n}pq(n - 1)(1 - F)
\]
Putting these results and those of (5) into equation (3) we find as the generalization of (6):

\[ w \Delta q = p q (F K - k) \]

so that the criterion for positive selection of altruism is

\[ K - 1 \]

\[ k - F \]

Now the model can be made to work. Moreover, the simple form and the independence of group size suggest that the criterion may hold beyond the limits of the rather artificial model discussed here. Careful though confirms that this is indeed the case: the criterion is completely general for asexual models with non-overlapping generations, and also holds for sexual diploid models when the coefficient \( F \) is suitably redefined (Chapters 2 and 5). The easiest way to see the basis of generality is to notice that the benefits of altruism do not now fall on a random section of the population and therefore do not simply enlarge the existing gene pool; instead they fall on individuals more likely to be altruists than are random members of the population. Indeed, the existence of the positive correlation \( F \) could be interpreted as implying in this case that there is a chance \( F \) that the \( K \) units of fitness are definitely given to a fellow altruist, while with chance \( (1 - F) \) they are given (as they always were in the previous version) to a random member of the population.

The redefinition necessary for diploid organisms involves specifying a regression coefficient, \( b_{AB} \), representing the regression of the genotype of recipient B on genotype of donor A. Often this is the same as the correlation coefficient of such genotypes (it always is so in the haploid case), but where they differ it is the regression coefficient that gives the prediction of gene content that we need. To get the form like (7) which applies to diploid selection other changes are obviously necessary, notably dividing \( pq \) by two to get the variance of gene frequency between pairs instead of that between individuals and other more complex changes connected with dominance and details of the assortative process. However, it is striking that a criterion like \( (F K - k) > 0 \) can be shown to determine positive selection of each genotype, and can be generalized to cover cases where A distributes various effects, positive or negative, to numerous individuals B, C, D, ... all having different regressions on A. Including A himself in the list of recipients we arrive at the idea of A's 'inclusive fitness' his basic non-social fitness, plus all the effects caused by his action when each has been devolved by a regression coefficient.

The usefulness of the 'inclusive fitness' approach to social behaviour (i.e. an approach using criteria like \( b_{AB} (K - k) > 0 \)) is that it is more general than the 'group selection', 'kin selection', or 'reciprocal altruism' approaches and so provides an overview even where regression coefficients and fitness effects are not easy to estimate or specify. As against 'group selection' it provides a useful conceptual tool where no grouping is apparent—for example, it can deal with an ungrouped viscous population where, owing to restricted migration, an individual's normal neighbours and interactants tend to be his genetical kindred.

Because of the way it was first explained, the approach using inclusive fitness has often been identified with 'kin selection' and presented strictly as an alternative to 'group selection' as a way of establishing altruistic social behaviour by natural selection. But the foregoing discussion shows that kinship should be considered just one way of getting positive regression of genotype in the recipient, and that it is this positive regression that is vitally necessary for altruism. Thus the inclusive-fitness concept is more general than 'kin selection'.

Haldane's suggestion about tribe-splitting can be seen in one light as a way of increasing intergroup variance and in another as a way of getting positive regression in the population as a whole by having the groups which happen to have most altruists divide most frequently. In this case the altruists are helping true relatives. But in the assortative-settling model it obviously makes no difference if altruists settle with altruists because they are related (perhaps never having parted from them) or because they recognize fellow altruists as such, or settle together because of some pleiotropic effect of the gene on habitat preference. If we insist that group selection is different from kin selection the term should be restricted to situations of assimilation definitely not involving kin. But it seems on the whole preferable to retain a more flexible use of terms; to use group selection where groups are clearly in evidence and to qualify with mention of 'kin' (as in the 'kin-group' selection referred to by Brown), 'relatedness' or 'low migration' (which is often the cause of relatedness in groups), or else 'assortation', as appropriate. The term 'kin selection' appeals most where pedigrees tend to be unbounded and interwoven, as is so often the case with humans.

Although correlation between interactants is necessary if altruism is to receive positive selection, it may well be that trying to find regression coefficients is not the best analytical approach to a particular model. Indeed, the problem of formulating them exactly for sexual models proves difficult (Chapter 2). One recent model that makes more frequent group extinction the penalty for selfishness (or lack of altruism) has achieved rigorous and striking conclusions without reference to regression or relatedness. But reassuringly the conclusions of both this and another similar model (more general but less thorough and much less well explained) are of the general kind that consideration of regression leads us to expect. The regression is due to relatedness in these cases, but classified by approach these were the first working models of group selection.
TRIBAL FACIES OF SOCIAL BEHAVIOUR

One of the conclusions of the models just mentioned is that with a grouped population the migration between groups is crucially important in determining the general level which altruism can reach within a group. This is something which should now seem fairly obvious but which has been surprisingly overlooked in most discussions of group selection previous to Eshel’s. The less migration there is the more relatedness will build up within groups. This will permit selection of acts with low gain ratio (i.e. ratios like $K/k$) but the gain ratios must always exceed one, and this means that the act must actually aid group fitness in some way—reduce its chance of sudden extinction,$^{11,12}$ or increase its rate of emission of migrants.$^{13}$ With the last eventually it is better for altruism if the migrants get together in small groups to found new colonies than if they all enter existing groups, since entering undermines the assumption of low migration—in other words reduces intergroup variance. If groups of founder migrants are assortative so much the better, although if they are so by coming all from the same parent group this could be treated as fission. Likewise if migrant acceptance is the established mode, so much the better if groups selectively accept altruists. The ability of animals to exercise such discrimination may seem dubious when behaviour even in humans is rather indefinite in this respect, but it is noticeable that with many of the tight-knit groups of social carnivores and primates the would-be immigrant does go through a probationary period of hostile treatment and low status, which sometimes terminates his attempt to join.$^{14}$ Similar phenomena of possibly similar significance are certainly not lacking in people, witness the harsh requirements of achievement and service for an aspiring Amerindian brave (or neophyte British doctor for that matter) and the general suspicion, hostility, and low position accorded to wealthless immigrants. I should add here that the idea that such human behaviour is natural does not mean that it is right or even sensible under modern conditions. For example, the immigrants may bring new skills and aptitudes, a point to which I return later. And as regards ‘altruism’, recent tribal immigrants are likely to be net importers of this precious stuff—themselves the losers when they expose their natural communitistic generosity to civilized exploitation. On the other hand, when experience of ambient guile and cunning has taught them better, such immigrants may learn to confine this generosity again among themselves and to turn outwards a contemptuous and unsympathetic attitude which is also typically tribal; but such expected ambivalence in tribal feeling is another matter to which I must return.

I have carefully spoken of ‘migration’ rather than ‘migration rate’ so far, and in doing so I intend to emphasize that it is the number of acts of successful migration that is important for mean intragroup relatedness. The size of demes may matter surprisingly little. An indigenous village may know some of his many connections with other villagers and be aware of a plexus of relationship through the misty past. What might surprise him (as it surprised me) is that relatedness as measured here (and as manifested in physical similarity) builds up just as much eventually in a large unit, say a remote town, as it does in a village, if the same actual numbers enter and leave each generation. In other words, connections which the remote townsman does not so easily know of make up in multiplicity what they lack in close degree. Of course, a large unit usually does have more migration, and consequently less intrarelatedness, but the important thing is that it is the number of migrants rather than the size of colonies that determines this fact. For Wright’s simple island model where migrants go anywhere among infinitely many colonies the approximate formula for mean intragroup relatedness (after migration has occurred) is very simple, $b = 1/(2M + 1)$ where $M$ is the number of migrants (assumed small) per subpopulation per generation. So with one migrant exchanged every other generation we find $b = \frac{1}{2}$, the same as for siblings in a panmictic population, and we therefore expect the degree of amicability that is normally expressed between siblings. If three migrants go (and three come) every generation we get $b = \frac{3}{5}$. This is slightly more than the relatedness of outbred cousins ($b = \frac{1}{4}$), so such colonies should be slightly more in-the-amicable than groups of cousins would be. If, as normally happens, migrants tend to go mainly to neighbouring populations, then emigrant and immigrant genotypes will tend to correlate and so a given level of relatedness can be maintained with more migration. However, recent achievements with the analysis of the harder stepping-stone model,$^{15,16}$ which covers the island model as a special case, show that the difference is not very great as regards own-deme relatedness. Consider the case where the colonies are supposed spread on the plane in a square lattice. Suppose that an act of migration is either ‘distant’, with the migrant going to any deme among the infinitely many, as in the island model, or ‘close’, with the migrant going to one of the four neighbouring demes, and suppose that the odds on events of the two kinds are specified. With odds 100:1 for ‘distant’ to ‘close’ $M = 0.5$ leads to $b = 0.5$ within colonies to a good approximation: in other words, the change produced by such a small amount of local migration is negligible. If the odds are reversed to 1:100, implying local migration much more probable than distant migration, the Kimura–Weiss solutions show that the relatedness only rises to 0.68. With 10 times as much migration (i.e. about five exchanged per deme per generation) the corresponding relatednesses are 0.09 and 0.17, so relatedness still only doubles when migration is local rather than distant. A much greater contrast is apparent in the relatedness of individuals of neighbouring demes: when distant migration preponderates this relatedness tends to be extremely small, but when close migration preponderates members of a neighbour deme can easily have more than half the relatedness that applies to an own-deme member. And up to a point increasing migration reduces the contrast between own and neighbour deme, so that there are getenical as well
as cultural reasons why, in humans, intergroup migration and marriage should decrease intergroup hostility.

Two other points seem worth making about the stepping-stone model. One is that, in the one-dimensional version of this model which could apply to demes in a linear habitat such as a coastline or river, relatedness holds up much more strongly as local migration is increased, and relatedness to neighbour's deme member more strongly still (relative to the two-dimensional case). This means that hardly any extra hostility is expected to members of neighbouring demes. From this point of view, a seashore phase of hominid evolution, if it occurred, should have been particularly harmonious. The other point concerns the distribution of gene frequencies. The apparent variability of colonies is expected to change rather sharply at certain critical levels of migration. These are $M = 0.5$ for the island model and $M = 1$ for the two-dimensional stepping-stone model with close migration predominant. This means that at about the point where the colony members are related to each other like outbred sibs it should become relatively easy for individuals to detect a fairly clear difference in appearance when comparing fellow colony members with outsiders. Actually, in the stepping-stone model the possibilities with regard to patchiness andcline-like effects are complex, but, considering simultaneously several traits which are independently inherited and at most weakly selected, the complex overlap of patterns should make possible fairly accurate separation of 'us' and 'them' at the level of colonies. We shall shortly see why natural selection might favour motivation and ability so to discriminate.

What is happening to the ordinary families embedded in these supposedly endogamous colonies? Siblings, parents, and offspring will still be the individual's closest relatives. Owing to the inbreeding, their relatednesses will be above the value of $\frac{1}{8}$ that applies under outbreeding. Thus an individual should be more altruistic than usual to his immediate kin. But other neighbours who are not immediate kin are now also closely related, and it is this reduced contrast between neighbours and close kin that will give what is probably the most striking effect: we expect less nepotistic discrimination and more genuine communism of behaviour. At the boundary of the local group, however, there is usually a sharp drop in relatedness. If migrants (or whole groups) are very mobile, leading to an 'island' rather than a 'stepping-stone' situation, this drop may be such as to promote active hostility between neighbouring groups. Even though these groups have some relatedness, as practical limitations to distant migration naturally ensure, the contrast is still such that a minor benefit from taking the life of an outsider would make the act adaptive. Recent studies on hunting dogs and hyenas show strangers sometimes being killed, while within-group relations are usually amicable and even communal. The most serious wounding which Lawick-Goodall recorded in her study of chimpanzees occurred when two males combined to attack a male of another group. Bygott witnessed a fierce attack by a group of male chimpanzees on a stranger female. The female escaped but the males caught and ate her infant. Trespassers may sometimes be killed in wolves, and in rats. In lions, langurs, and probably in rats and mice there is also killing of strange young, but this is probably in a rather different category because it is done only by males whose aim seems to be to sire new offspring on the mothers they bereave.

These phenomena are reminiscent of the intercolony hostility so often observed in social insects, where again actual killing may be frequent along the frontiers. With regard to relatedness, the situation is the same except that intrarelatedness in groups is usually due to all colony members being descendents of a single queen. But polygynous ants (e.g. the common red ants of the genus Myrmica) may approximate the breeding structure of group hunting carnivores rather well, and this tempts one to apply the superorganism concept often used for social insects to the co-operative social mammals. Such a view would compare the killing of occasional trespassers to the occasional minor wound with death of cells which occurs in the restrained fighting between, say, two individual dogs.

The basis for thinking that group-hunting carnivores are highly related within groups is the known low rate of migration and the reluctance with which migrants are accepted. Why such reluctance? The probable reason has already been touched on: the group has a co-operative job to do, necessary for its survival. This job is the hunting and killing of prey which are too large for one individual to tackle alone. The more work is invested in a task prior to its fruition the more worthwhile a parasitic option of behaviour becomes—at least, until parasites are too numerous (Chapter 6). And the more co-operation is involved in any endeavour the more scope there is for the inconspicuous idler. What is to stop a hunting dog from watching the hunt from afar and trotting up, by all possible short cuts, just after the prey has been killed? Probably this has happened and probably groups over full of the offspring of such idlers have found themselves unable to kill prey and have died out. This would give a slow selection for features of pack behaviour (either cultural or genetic) that make infiltration progressively more difficult. Simply cutting down on immigration would have the desired effect through raising relatedness but, as mentioned earlier, a selective entrance requirement, with the applicant's behaviour watched through a probationary period, would be even better. It should be mentioned here that a development which closed a group's frontiers completely would probably also fail in the long run for reasons of general adaptation: complete inbreeding abandons the obviously important advantages of sexual reproduction, whatever these are.

Roughly, as we currently see it, a cunning ape-like creature onc pushed boldly out from near niches now held by baboons and chimpanzees. Whether or not (as one quite plausible view holds) it first left its less enterprising cousins to take a holiday on the seashore, eventually it reappears on the African
savana participating in the Pleistocene wildlife bonanza as a group-hunting carnivore. In spite of the now-not-so-prehensile foot which it kept all the while in the door of an omnivorous diet, it seems likely that this creature would have needed the same population structure as the other group-hunting carnivores, and for the same but more urgent reasons. It is difficult to see what was the first factor in the escalation in cunning of this particular primate line, but the choice seems mainly between tools and language. The great benefits that these could confer to a co-operative hunter through improved technique and organization would ensure rapid selection for their development. But they would also affect the social situation in significant ways and indirectly this might escalate their selection: (1) both would provide extra cultural clues to group identification; (2) tools (and later other valued artefacts) would give further scope for parasitical behaviour, first intragroup but later between groups as well. Tools and possessions could be appropriated instead of made. With language in rapid evolution, learning, experience, and even intelligence would become increasingly open to parasitism. Meanwhile, increasing intelligence would make possible a very plastic approach to parasitical and altruistic behaviour, which in turn would increase the complexity of the semi-serious deception and coalition games which are so characteristic of behaviour within primate groups. Real rewards in food and mating are the incentive to this activity and thus escalate the selection for skill in play. The main point is that intelligence, plus (1) and (2), plus what has been explained about the real differentiation of genetical relatedness suggest the development of an explosive situation. Close frontiers to migrants a little more, or slightly increase group mobility, and it is possible to imagine the sudden success of a policy which makes any frontier incident an occasion for an attempt at violent incursion by the more populous group with losers killed, enslaved, or driven off. Successful occupation of the captured territory would soon bring the victors into contact with still less related ‘stones’ of the ‘stepping-stone’ lattice, which they could attack adaptively with even less reason for restraint. Increasing foresight would mean that a group would not necessarily wait until large enough to need neighbouring territory, if attacking a weak group while it is weak helped to ensure that space could be occupied as needed. Increasing ability to abstract and generalize would enable groups to reanimate their intragroup coalition games in the more serious intergroup context. The usual and firmest coalitions would be between related groups, as is the case with coalitions of individuals (usually males) in wild turkeys,27 lions,26 and chimpanzees.25 Are group fights necessarily more serious for the species if, on the analogy of the superorganism, we are allowed to equate a few deaths to a minor wound? Perhaps not, and of course it is possible that making groups more aggressive would not ‘melt’ the lattice structure to the extent suggested. Moreover, groups might be units in supergroups that are themselves in a ‘stepping-stone’ lattice. In such cases warfare—for that is the behaviour we now survey—might carry over from intragroup behaviour (or itself spontaneously develop) quite orderly restrained procedures involving little loss of life.

In developing this admittedly speculative outline of certain cultural and genetic processes in tribal evolution, I confess a bias towards discovering the patterns of coalitions, warfare, language, and so on that are documented in certain remote peoples of the present day—for example, the Yanomamo28 and various New Guinea highlanders.29,30 Admittedly in these cases there is agriculture; it is possible to claim that most hunter-gatherers are more peaceable. For example, why not aim to derive the customs of the Kalahari Bushmen (San)? But most hunter-gatherers are certainly less peaceable than Bushmen. The record of human violence goes back far indeed, even if the earliest attributions (Darri’s cases in Australopithecus29) are doubtful. A trace of homology with the sporadic violence of chimpanzees seems not impossible. Probably with hominids, as with chimpanzees, actual violence towards outsiders would contrast with restrained violence, or mere threats, used within the group, while within the group too there would be much sharing and co-operation. One Neanderthal skeleton of Shanidar Cave had bone damage suggesting a stab wound.35 Another skeleton also had bone defects but of quite different implication: one forearm was lacking, perhaps from birth, certainly for a long time, and a healed injury to the skull showed that one eye was blind. Goodall’s chimpanzees were part hostile, part sympathetic, and part indifferent to comrades suddenly crippled with polio: they did nothing positive to support them. In contrast, the Neanderthals of Shanidar evidently supported a cripple, and on his death they buried him in the cave where, in other graves, they also sometimes buried their dead with flowers. These hints of violence and loyalty and (perhaps most purely human) of incipient love of things for themselves evoke a startlingly familiar and sympathetic portrait. Considering only the same affectional attributes in the present-day tribal and pastoral Kurds (as opposed to attributes connected to the ever-accelerating change in material culture), a recording angel perhaps notes today much the same events in Shanidar Cave as he noted an ice age ago.35,38

Probable instances of cannibalism in Homo erectus and Neanderthals have been plausibly compared to similar recent cannibalism in New Guinea. In New Guinea, it is interesting to note, this practice acted as a kind of population control, since by eating the brain of his victim, it was believed, the headhunter won a name for a child of his own—in effect, won a birthright. Other usually less drastic beliefs and practices affecting fertility that are widespread in human cultures may help to explain how they manage to be as peaceful as they are. In so far as the practices amount to effective birth control, they cut warfare at its demographic root. Unfortunately, it is possible that in doing so they also cut
an important link that has escalated the selection for intelligence.\textsuperscript{24,27} No hunt needs quite so much forethought or ability to communicate complex instructions as does a war, nor do such drastic demographic consequences hinge on the outcome.

The rewards of the victors in warfare obviously increase for peoples past the neolithic revolution. There are tools, livestock, stores of food, luxury goods to be seized, and even a possibility for the victors to impose themselves for a long period as a parasitical upper class. Hunter-gatherers, on the other hand, at most win only mates and land. It might seem that these things would not repay the expected cost of the fighting, but it has to be remembered that to raise mean fitness in a group either new territory or outside mates have to be obtained somehow. The occurrence of quasi-warlike group interactions in various higher primates\textsuperscript{36,38} (and references in Bigelow\textsuperscript{35}) strongly suggests that something like warfare may have become adaptive far down in the hominid stock. These primate examples suggest the prototype war party as an all-male group, brothers and kin, practised as a team in successful hunting and at last redirecting its skill towards usurping the females or territory of another group. Out of such cells can be built the somewhat less stable organism of the post-neolithic army. The Homeric Iliad gives a vivid inside view of the process of coalition, while the siege it describes emphasizes the existence of economic surpluses supporting the warriors on both sides (something hunter-gatherer warriors would never have). If the male war party has been adaptive for as long as is surmised here, it is hardly surprising that a similar grouping often reappears spontaneously even in circumstances where its present adaptive value is low or negative, as in modern teenage gangs.\textsuperscript{36}

Whether or not the neolithic revolution brought an increase in the per capita incidence of violence it does seem that from then on warfare looms larger in the affairs of men. The situation seems reflected in the fact that only one of various series of pre-neolithic cave paintings depicts warfare,\textsuperscript{40} whereas for most early civilizations the earliest known written records of warfare, booty, captives, and the like.

It has been argued that warfare must be a pathological development in humans, continually countered by natural selection, and this claim is sometimes based on a sweeping a priori view that habits of mortal intraspecific fighting must always endanger the survival of a species.\textsuperscript{24} While endorsing such a view as regards wars between the few frightfully armed superpowers of today, I see no likelihood for it as regards fighting of individuals or of groups up to the level of small nations. Of course, for the species as a whole, and in the short term, war is detrimental from the biological demographic point of view, but, as shown above and elsewhere, detriment to the species does not mean that a genetical proclivity will not spread. Anyway, what is bad at one level may be good at another and the cost to the species may be paid in the long run. The gross inefficiency of warfare may be just what is necessary, or at least an alternative to birth control and infanticide, in order to spare a population's less resilient resources from dangerous exploitation. Maybe if the mammoth-hunters had attacked each other more and the mammoths less they could be mammoth-hunters still. And the rich ice-age fauna of the Americas might have had time to adapt to the human predator as it adapted in Africa if fighting had induced man to draw his curtain of overkill across the continent less rapidly. Many examples in the living world show that a population can be very successful in spite of a surprising diversion of time and energy into aggressive displays, squabbling, and outright fights. The examples range from bumble bees to European nations. In case all this reads like a paean for fascism let me add one caution from the geological record. Arms and armour seem to weigh one down in the end: it is hard in the modern human world to see warfare as a stabilizing influence.

The relatively peaceful Bushmen may tell us something valuable about the aetiology of wars, but I am doubtful if they tell us much about the role of this factor in the main stream of human evolution. However, it is noteworthy that the Bantu who replaced the khosian races in much of the rest of Africa were warlike in ways that evoke comparisons from the dark ages of Europe. The Bantu were, of course, mainly pastoralists and agriculturalists, for whom, as stated above, booty would be an important additional incentive to warfare. Pastoralists tend to be particularly warlike and the histories of civilization are punctuated by their inroads. Pastoral tribes have to be mobile in following or driving their herds and this mixes tribes and reduces relatedness of neighbours. Viewed as booty, the mobility of stock is a great convenience. Both factors must contribute to the warlike propensity.

The incursions of barbaric pastoralists seem to do civilizations less harm in the long run than one might expect. Indeed, two dark ages and renaissances in Europe suggest a recurring pattern in which a renaissance follows an incursion by about 800 years. It may even be suggested that certain genes or traditions of the pastoralists revitalize the conquered people with an ingredient of progress which tends to die out in a large panmictic population for the reasons already discussed. I have in mind altruism itself, or the part of the altruism which is perhaps better described as self-sacrificial daring. By the time of the Renaissance it may be that the mixing of genes and cultures (of or cultures alone if these are the only vehicles, which I doubt) has continued long enough to bring the old mercantile thoughtfulness and the infused daring into conjunction in a few individuals who then find courage for all kinds of inventive innovation against the resistance of established thought and practice. Often, however, the cost in fitness of such altruism and sublimated pugnacity to the individuals concerned is by no means metaphorical, and the benefits to fitness, such as they are, go to a mass of individuals whose genetic correlation with the
innovator must be slight indeed. Thus civilization probably slowly reduces its altruism of all kinds, including the kinds needed for cultural creativity (see also Eshel11).

RECIPIROCATION AND SOCIAL ENFORCEMENT

The last suggestion is rather different from saying, as has sometimes been said, that civilization selects against all kinds of creative intelligence. It seems to me that there are some aspects of innate intelligence that civilization steadily promotes Mercantile operations, for example, are an inseparable part of Old World civilizations and need complex models in the minds of their operators, just as military ventures do. The main difference is in more emphasis on prudence and less on daring. It is probable that civilization has given steady selection for the intelligence needed for this mercantile kind of preparatory modelling. The intelligence that gives a good appreciation of the real principles involved in a new technology, as opposed to seeing it as a kind of magic, is probably also constantly favoured, since improvers of a technology avoid the arrows of contempt and penury that face pioneers and can do very well. However, my main reason for turning to the subject of trade is to introduce the idea of another kind of positive social arrangement which thrives in a mercantile and technological atmosphere, for which intelligence is more necessary and relatedness much less so. This refers, of course, to reciprocation.

Starting perhaps with something like the meat-sharing of chimpanzees (‘feed me while you have plenty and I’ll feed you when I have plenty’), proceeding through barter (where differing aptitudes may begin to be important), reciprocative activity branches out into all the various business-like arrangements of modern humans. The key words are ‘client’ and ‘partner’ as opposed to ‘kinsman’ and ‘friend’.

Establishing a basis for reciprocating has problems of natural selection closely similar to those of altruism as discussed so far. It is very frequently necessary for one party to execute his half of a bargain without any way of being certain that the other party will later stick to his. The best response if the other does not reciprocate is to cut off any further benefits to him.41 Unfortunately, this leaves the selfish non-reciprocator better off than the ‘altruistic’ initiator and unless the two are related this is against the habit of reciprocation, at least when the trait is rare (see discussion of Prisoner’s Dilemma in Chapter 6; also Boorman and Levitt42). However, this initial barrier to selection is a slight one if the rewards of the interaction are high, as they would be when an advantageous exchange can be repeated many times.43,44 Once the barrier is passed by genetic drift or the like, non-reciprocation finds itself in the category of maladaptive spite—harming the self to harm others more. Nepotistic altruism, of course, also has an initial barrier to pass when it first occurs by mutation. But once positive selection supervenes the resemblance between the two situations fades: reciprocal altruism of the kind described is less purely altruistic. Indeed, the term altruism may be a misnomer: there is an expectation of benefit of the initiating individual, not just an expectation of benefit to the genotype. To put the matter another way, reciprocal altruism can never be suicidal, whereas suicidal nepotistic altruism can and has evolved—it is apparent, for example, in worker sacrifices in the social insects.

Whether reciprocation involves altruism or not, we see that in so far as it involves repeated acts between the same two individuals this useful and immensely variegating type of interaction can spread genetically, given only an ability to remember individual faces of those who have helped and those who have cheated in the past. Unfortunately, by the very aid it gives to the growth and diversification of social systems, reciprocation tends to undermine the basis of its success. Situations demanding reciprocation just once between individuals destined never to meet again naturally become more common and it becomes easier for cheaters to specialize in these and to hide from retribution. Cheating can also become more subtle, especially along lines which make it hard for victims to be sure just who has cheated them.

In considering this problem I think there may be reason to be glad that human life is a ‘many-person game’ and not just a disjointed collection of ‘two-person games’. Admittedly, it may not seem so at first. At first reading the theory of many-person games may seem to stand to that of two-person games in the relation of sea-sickness to a headache. But given also a little real intragroup altruism endowed from the tribal past, it may turn out that the one is at least a partial cure for the other (see Fig. 9.1). The idea here is that for pairs in isolation the problem of cheating in a single exchange may be insoluble, and that therefore we have all evolved, more or less in proportion to our exposure to civilized (i.e. relatively panmictic) conditions, into potential cheaters. But at the same time we also have every reason to agree as to the parasitical nature of cheating as it affects the welfare of the community as a whole and to deplore its successful practice by others. So detection of cheating arouses indignation in everyone except the accused, and everyone sees a benefit to both group and self in trying to punish the cheater and in forcing restitution (some part of which, as an added incentive, may be diverted as a fee for those who administer the collective justice). The reason I believe that a little real intragroup altruism is also necessary for the evolution of efficient justice on these lines is that individuals must feel the difference between the usefulness of this behaviour and the futility of using collective power arbitrarily in ways profitless to the group. A healthy society should feel sea-sick when confronted with the endless internal instabilities of the ‘solutions’, ‘coalition sets’, etc., which the theory of many-person games has had to describe. One hears that game theorists, trying to persuade people to play even two-person games like ‘Prisoner’s Dilemma’, often encounter exasperated remarks like: ‘There ought
Inmate social aptitudes of man: an approach from evolutionary genetics

Figure 9.1 (continued) doing a player does best by playing selfishly; as in Prisoner's Dilemma the rational yet 'paradoxical' result of this is that the players all play selfishly and get one unit each, whereas had they all played co-operatively they could have four units each.

In case B, it is assumed that if two players play co-operatively while a third is selfish, the two use their united strength to 'punish' the third. His selfishly-gained five units are taken from him and divided between the two co-operators. This creates a situation where no player sees an advantage in departing from triple co-operation. Likewise no player has an advantage in departing unilaterally from the triply selfish corner; but now any two players can decide to depart from this corner by a coalition which is not only advantageous (as it was also in case A) but also secure in that neither member can benefit by defaulting.

to be a law against such games! Some of the main points of this paper can be summarized as an answer to this comment: that often, in real life, there is a law, and we can see why, and that sadly we also see the protean nature of this Dilemma, which, when suppressed at one level, gathers its strength at another.

References and notes

4. R. Fox, in Kinship and Marriage (Penguin, London, 1967), also emphasized this in discussing human kinship systems and why some which are easily conceivable never actually occur. His discussion of the incest taboo is also very pertinent to the idea that follows.
7. Although Price first pointed out the generality and usefulness of this relation, earlier partial recognition of it seems to be due to Alan Robertson. See A. Robertson, A mathematical model of the culling process in dairy cattle, Animal Production 8, 95–108 (1966).
9. For example, J. Maynard Smith, Group selection and kin selection, Nature 201, 1145–7 (1964) and Lewontin 1970 (ref. 3).


17. As pointed out elsewhere (see Chapters 5 and 6), even spiteful behavior, harming oneself in order to harm another more, is a theoretical possibility. The mean relatedness to the entire species population other than self is \(-1/N - 1\) where \(N\) is the population. If inbreeding or otherwise a colony has grouped \(n\) identical genotypes together, then relatedness to the average outsider is \(-1/(N - n)\). Thus with only a few large long-isolated groups spite is more possible.


33. A. M. Hamilton, Road through Kurdistan (Faber, London, 1937).


43. The game-theoretic situation known as Prisoner's Dilemma has as paradigm the dilemma of a criminal in custody who is offered the inducement of a light sentence if he will give state evidence to clear up an important crime, while knowing that a confederate, also in custody, has the same offer and that the pay-off situation which both face is as follows, arranged from lightest to heaviest sentence:

(c) I confess, he doesn't > (a) We neither confess > (d) We both confess >
(b) He confesses, I don't.

So that (c) can be better than (a) it is assumed that they can certainly be convicted of some minor offence whereas at least one confession is needed in order to settle the major crime. For further information on this 'game' see A. Rapoport and A. M. Chammah, Prisoner's Dilemma (University of Michigan Press, Michigan, 1965).