

Bateson, P. (1988) The biological evolution of cooperation and trust . In *Trust: Making a breaking cooperative relations*, ed. by D. Gambetta, pp. 14-30. Blackwell, Oxford.

THE BIOLOGICAL EVOLUTION OF COOPERATION AND TRUST

Patrick Bateson

We do not know whether animals trust each other, but we may ask whether the conditions of their social life favoured the evolution of what we understand to be trust in humans. More directly, we can discover whether they behave in ways that benefit others. Since they often do so, we may also ask what favoured the evolution of such behaviour and under what conditions cooperation is likely to occur in those animals that show it. The answers to such questions are important because so many non-biologists believe that the social behaviour of animals consists of nothing but strife and manipulation. If animal nature is the result of the competitive process of Darwinian evolution, then surely it must also be ruthlessly competitive. Such a conclusion does not follow from an acceptance of Darwinism. It is also a travesty of what biologists have observed and what most now believe.

In this chapter I shall discuss how the observed cooperation seen between animals may be interpreted in terms of the most widely accepted explanations for biological evolution. I shall begin with a description of the character of neo-Darwinian explanations and then deal with three ways in which the observed cooperation might have

arisen. The first explanation is that, at least in the past, the aided individuals were relations; cooperation is like parental care and has evolved for similar reasons. The second is that the cooperative behaviour generated characteristics in a collection of individuals that, under special conditions, favoured such groups over those that did not cooperate so effectively. Finally, cooperating individuals jointly benefited even though they were not related; the cooperative behaviour has evolved because those who did it were more likely to survive as individuals and reproduce than those that didn't. The three evolutionary explanations are not mutually exclusive.

These questions about the evolution of cooperation do not bear directly on the issue of trust, though they may give pause to anyone who supposes that trust is required for effective cooperation (defined simply in terms of working with and assisting others). To maintain such a position it would be necessary to argue that, for example, honey bees trust each other. Even though the study of cooperation in animals seems irrelevant to an understanding of trust in humans, careful analysis of the conditions in which cooperative behaviour is expressed suggests that many animals are exquisitely sensitive to the behaviour of others. This observation suggests an explanation for the evolution of the mental state that we recognise as trust in ourselves. Discussions about evolution also draw attention to the explicit distinctions which biologists are forced to make between what

animals do and the immediate and long-term consequences of their behaviour. In such discussions the meaning of "interest" is not in much doubt¹. I shall amplify on this point since it is so easy to conflate the consequences of a dynamic process with the way in which those consequences are generated. I am going to start, though, with a simple description of Darwin's explanation, because opportunities for confusion abound and even those who are supposedly expert in the area are sometimes responsible for a great deal of muddle.

1. Modern Evolutionary Theory

Virtually every biologist who cares to think about the subject believes that all living matter has evolved. Existing species were not created in their present form at the beginning of life on this planet. The modern scientific debates are about how the changes came about, not about whether or not they happened. Increasingly, the theorists agree that the evolution of life requires several distinctly different explanations. Stable forms may often arise by chance, for instance; and extinctions and opportunities for further change may result from environmental

¹ The presumption is usually that the single currency of evolution is the genetic endowment that is required to build the bodies and behaviour patterns and which was effective in doing so in the past. It certainly does not follow that the expression of the genes in any one individual is inevitable or independent of other conditions (see Bateson, 1986).

catastrophe. However, chance and catastrophe are unsatisfying and inadequate as explanations when we try to understand the numerous and exquisite examples of correlations between the characters of organisms and their physical and social environments. For instance, carnivorous and herbivorous mammals have strikingly different types of teeth and the differences are readily related to methods of feeding. To take another dental example, the large canine teeth found in carnivorous animals are also seen in non-carnivorous animals that use them as weapons in fights with members of their own species. Male polygynous primates, that fight with other males for females, have much larger canines than male primates that are characteristically monogamous (Harvey, Kavanagh & Clutton Brock, 1978). Character-environment correlations of this type are known as adaptations. They grab our attention because they seem so well designed for the job they perform. Much the most coherent explanation for the evolution of such phenomena is still Darwin's. Indeed, Darwin's proposal is much better seen as a theory about the origin of adaptations than as a theory about the origin of species.

Darwin's proposed mechanism depends crucially on two conditions. First, variation in a character must exist at the outset of the evolutionary process. Second, offspring must resemble their parents

with respect to such a character¹. The short-term steps in the process involve some individuals surviving or breeding more readily than others. If the ones that survive or breed most easily carry a particular version of the character, the character will be more strongly represented in future generations. If the character enabled them to survive or breed more readily, then the long-term consequence is that the character will generally be correlated with the conditions in which it worked. If differences between individuals depends on differences in their genes, Darwinian evolution results in changes in the frequencies of genes.

Darwin used the metaphor of selection to describe the evolutionary process of adaptation because he had in mind the activities of human plant and animal breeders. People, who want to produce a strain of pigeons with longer tail feathers than usual, pick from their flock those birds that have the longest feathers and exclusively use them for breeding purposes. This is artificial selection of the long-tailed pigeon by animal breeders. By analogy Darwin referred to the differential survival of the characters that adapt an organism to its environment as *natural* selection. The Darwinian formulation

¹ In modern versions of Darwin's theory, the correlation in the character may exist between collateral genetic relations such as siblings or uncles and nephews. Usually, the correlates will exist because genetic relations are more likely to share copies of the same rare genes, but the similarities may arise for other reasons. Genetic inheritance is not a requirement for Darwinian evolution to work.

emphasizes that the precise way in which a character enables an individual to survive or breed better is part of the process¹. The way in which the character works is not the same as the long-term changes in the frequencies of genes that are required for the expression of the character. Moreover, the evolutionary process does not require the postulation of an unconscious motive for propagating genes (let alone a conscious one)². Gene propagation is merely the result of a character, such as a particular form of behaviour, that worked better than another version.

I should locate a potential source of confusion before going any further. To state that the gene that makes the difference between a character that works (and is thereby transmitted to the next generation) and one that doesn't says nothing about the other conditions that are required for the expression of that character. It certainly does not mean that the expression of the character is

¹ It is because an explanation for the evolution of a particular character has to refer to the specific way in which it has worked (and by implication has done better than other characters) that the formulation is non-circular. "Survival of the survivors" is, of course, vacuous.

² It is unfortunate that in many of the writings about evolutionary biology, intention is allowed to intrude as in the phrase: "Animals attempt to maximise their inclusive fitness". Biologists who do this will claim that they use a shorthand for: "Animals behave as if they were attempting to maximise their reproductive success." Nonetheless, the phraseology does cause confusion - even in the minds of its authors.

inevitable, nor does it imply that its form is predictable if the conditions in which it evolved change.

If motives are unimportant, what about the "selfishness" that is sometimes claimed for genes? Richard Dawkins (1976, 1982) has argued vividly that individual organisms do not survive from one generation to the next, while on the whole their genes do. He proposed that, therefore, Darwinian evolution has acted on the genes. Dawkins' approach to evolution was presented in characteristically entertaining form when he suggested that the organism is "... a robot vehicle blindly programmed to preserve its selfish genes".

In order to understand Dawkins' particular brand of teleology, it may be helpful to forget biology for a moment and think about the spread of a new brand of biscuit in supermarkets. Consider it from the perspective of the recipe. While shoppers select biscuits and eat them, it is the recipe for making desirable biscuits that survives and spreads in the long run. A word in the recipe might specify the amount of sugar to be added and makes the difference between a good and a bad biscuit. Because it serves to perpetuate itself, that word is selfish in Dawkins' sense¹. This novel way of looking at things is unlikely to

¹ In my analogy, a word in the recipe corresponds to a gene in biology, the biscuit to the individual organism, the supermarket to the environment, and the shopper to the short-term process which leads to greater survival and reproductive success by some organisms relative to others.

mislead anyone into believing that what shoppers really do in supermarkets, when they pick a particular brand of biscuit off the shelves, is select a word in the recipe used for making the biscuits. It is odd, then, that the selfish gene approach has encouraged people to run together the crucial differences between individuals that survive and those that don't with the genetic consequences of differential survival in later generations. It has been a muddling conflation.

We should return to the strict meaning of Darwin's metaphor and make an explicit distinction between the short-term causes of differential survival and the long-term effects of differential survival on the frequency of the genetic replicators, as indeed Dawkins (1982) himself has done. Once made the distinction saves much muddle in modern discussions of evolution. It also serves a valuable role in drawing attention back from a preoccupation with single genes to the ways in which genes work together. Each gene depends on the outcome characteristics of the whole gene "team". Furthermore, special combinations of genes work particularly well together, and the gene that fits into one combination may not fit into another. The concept of the coadaptation of genes is helpful in re-establishing that organisms do, indeed, exist as entities in their own right.

2. The Evolution of Cooperation

If, as seems likely, a great deal of biological evolution involved differential survival, the outcome of the competitive evolutionary process is often social cooperation. Emperor penguins huddle to conserve warmth. Cattle press tightly together to reduce the surface exposed to biting insects. In many species, individuals clean each other. Male lions cooperate to defend females from other males. Mutual assistance may be offered in hunting; for instance, cooperating members of a wolf pack will split into those that drive reindeer and those that lie in ambush. As a result all of them are believed to get more to eat¹. In highly complex animals aid may be reciprocated on a subsequent occasion (Trivers, 1971, 1985). So, if one male baboon helps another to fend off competition for a female today, the favour will be returned at a later date (Packer, 1977). What is usually obvious about such cases is that all the participating individuals benefit by working together².

¹ The belief that cooperating predators obtaining more food than those that hunt on their own is attractive. However, Packer (1986) argues that the evidence does not stand up to careful scrutiny. There may be benefits in terms of obtaining more food, but they remain to be convincingly demonstrated. Meanwhile, biologists need to keep their minds open to alternative explanations for hunting in groups.

² Sometimes behaviour may be successfully manipulated as when a nestling cuckoo is able to obtain care from a hapless pair of reed warblers. Alexander (1974) has suggested that, even within a species, offspring may not operate in their own best interests and have been manipulated by their parents in such a way as to further the long-term interests of their parents.

Three evolutionary explanations have been proposed for non-manipulative social cooperation: (a) The individuals are closely related; (b) the surviving character is property of many individuals; and (c) the individuals mutually benefit. These explanations do not exclude each other, but it is helpful to deal with them separately. The first explanation has been the domain of the subject, called sociobiology by E. O. Wilson (1975), and has been extensively and critically discussed (e.g. Kitcher, 1985). The idea of "kin selection" is an extension of the intuitively obvious point that animals will often put themselves at risk and do things that are bad for their health in the production and care of offspring¹.

The use of the term "altruism" in sociobiological discussions was unfortunate because of its moral connotations. The evolutionary principle can be perceived more clearly, perhaps, when a non-behavioural example is used. Consider those insects, like wasps, that

¹ After a calculation made in a London pub, J.B.S. Haldane declared that he would lay down his life for two brothers or eight first cousins. However, it would have only been sensible for him to lay down his life if he especially wanted to perpetuate the habit of self-sacrifice. And even if he had wanted to do that, he would have needed to be sure that the difference between the presence and the absence of the self-sacrificial tendency was associated with a difference in a single gene. If it was two genes, presumably he would have needed to save at least four brothers or 64 first cousins (and also assume that they would all breed as much as he would have done himself). In general, Hamilton's (1964) formalisation of inclusive fitness applies to whole organisms when the difference between two types is that one type helps other individuals and the other does not and the types differ in only one gene.

are conspicuously marked and are unpalatable to their predators. Birds that eat wasps are unlikely to repeat the experience, since birds learn quickly. This does not help the wasps that died. However, in the ancestral condition, the few wasps that were conspicuously marked were likely to be closely related. Those that died provided protection for those that survived by making them less prone to predation. As a consequence, conspicuous yellow and black abdomens may have spread until all wasps were conspicuously marked in the same way. It is not difficult to see how a precisely similar argument can be mounted for care directed towards close relations. The point is that the giving of aid to a relative may evolve simply because the expression of that character increases the probability that it will recur in later generations.

3. Behaviour of Groups

The second evolutionary explanation for cooperation is the most controversial, largely because a good argument has been confused with a bad one. The bad argument is that animals ought to be nice to each other for the good of the species¹. The idea is inadequate because any individual that breaks the rule and behaves in a way that

¹ "Good for the species" arguments have a long history and can even be found in some of Darwin's writing (e.g. Darwin, 1871). They were clearly present in Kropotkin's (1902) famous book on mutual aid and recurred in another celebrated book by Allee (1949). More recently, Lorenz (1966) explained the restraint on aggression within a species in the same terms.

benefits itself at the expense of other members of its species will eventually populate the world with individuals that behave in the same self-serving way. The good argument is that some assemblages of individuals may, through their concerted efforts, generate an outcome that puts their group at an advantage over other groups. This argument becomes more obvious once observed characters are separated conceptually from their effects on gene frequencies found in individuals in subsequent generations. The well-adapted character that survives from one generation to the next is not the same as the necessary conditions for its expression. Once these distinctions are made, we can ask: To what does that character belong? The character, which the metaphorical hand has supposedly selected, may be formed by more than one individual. The characteristics of the whole entity provide the adaptations to the environment. One assemblage of individuals, acting as an organised system can compete with another in the strict Darwinian sense of differential survival.

The possibility of group characters changing in Darwinian fashion is not in question among serious evolutionary biologists. However, the consensus in the last ten years has been that the conditions for such evolution were too stringent, since groups are usually much slower to die off than individuals and individuals can readily move from one group to another (see Maynard Smith, 1976). The consensus was

probably formed too readily and has been under attack in recent years¹.

The essential point is that the outcome of the joint action of individuals could have become a character in its own right. The nature of the argument may be perceived most clearly in the arrangements of different species that are obliged to live together in symbiotic partnership. A good example is provided by the lichens which are found on virtually every stable surface throughout the world from rocks and tree trunks to paving stones and old roofs. While they look like single organisms, lichens are composed of algae and fungi fused together in obligatory partnership. In Darwinian terms, though, overall features of a lichen might enable it survive better in a given environment than a lichen with other characteristics. Even though the character is replicated in an "offspring" lichen by the independent reproduction of the component algae and fungae, the mechanism of inheritance is irrelevant to the evolutionary process. So long as

¹ See for example Wilson (1980). Grafen (1984) argues that the "new group selection" stems from a paper by Hamilton (1975) and the logic applied to kin selection applies to the changes in the frequencies of genes within the successful groups. In one sense he is certainly right in that cooperating groups are likely to be much more closely related to each other than they are to members of groups that do not cooperate so effectively. In another sense, though, he misses the point that the character that makes one group more likely to survive than another is a property of the whole assemblage and not of the component individuals.

offspring characteristics are correlated with parental characteristics, it doesn't matter how they got like that.

To take a specific example, suppose that in one "individual" lichen, an algal mutation has made the lichen less tasty to reindeer (which are lichen specialists). The less palatable lichens will survive better than those without the mutant algae. This is not because of competition between algae, but because of the effects they have on the entity of which they are a part. In terms of my supermarket analogy, different recipes might be used for the biscuit and its chocolate coating, but the customers select the whole package and by doing so increase the numbers of copies of both types of recipes. The general point is that the methodology of focusing on the genetics of individuals merely serves to muddle the issue of what is necessary for differential survival with what is required for replication. Once liberated from the confusion, we can, with easier minds, examine the characters generated by the cooperative behaviour of social groups of animals.

The emergent properties of social life might have been important in evolution when cheating individuals were penalised by adversely affecting the group in which they lived (see Crook, 1980). Clearly, a cheat could sometimes obtain the benefits of the others' cooperation without joining in itself. However, such actions would not be evolutionarily stable if the cheat's social group was less likely to

survive than a group without a cheat and the cheat could not survive if it left its own social group.

If the conditions were right, the outcome of the joint actions of individuals in the social group would have changed as the result of Darwinian evolution. It is important to appreciate that this perfectly straightforward Darwinian argument does not undermine what we know about genetics or return to muddled good-for-the-species thinking. It merely draws attention to a higher level of adaptation. This requires acceptance that the characteristics of social groups are the emergent properties of the participating members and the logic of Darwinian theory applies as much to these characters as it does to those of individual organisms. Providing examples is never easy, because evolutionary history can't be replayed, but the conditions necessary for its occurrence are particularly likely to have operated in hominoid evolution. I should add, nonetheless, that such a view of differential survival of groups is still heterodox among biologists, mainly because they do not want to lose the ground that was won by examining the evolutionary effects of manipulation and reciprocity at the level of the individual.

4. Mutual Benefits

The third explanation for cooperation is sometimes known as "mutualism within a species" (e.g. West-Eberhardt, 1975; Wrangham, 1982). Two cooperating individuals are not necessarily related, but they are both more likely to survive and reproduce themselves if they help each other. This category includes examples of types like the iterated Prisoner's Dilemma, dealt with so interestingly by Axelrod (1984). For a specified set of pay-offs, everybody benefits by cooperating at the outset. I shall focus here on joint parental care of the offspring.

Table 1. The probabilities that offspring will be produced under four arrangements of bi-parental care (simplified from Maynard Smith, 1977).

		Female	
		cares	deserts
Male	cares	female gets P_2	male gets P_2
	deserts	female gets P_1	male gets $P_1(1+m)$
		$P_1(1+f)$	$P_0(1+f)$
		P_1	$P_0(1+m)$

Where P_0 , P_1 and P_2 are the probabilities of survival of young cared for by 0, 1 or 2 parents respectively, m is the probability that a

deserting male will acquire a new mate and f is the probability that a deserting female will acquire a new mate.

Every type of parental care is found in animals. Maynard Smith (1977) proposed a useful scheme which suggested how such diversity might have evolved. A simplified (and slightly modified) version of it is given in Table 1. Consider the bottom right-hand case where neither parent cares for the young; herrings are like this. If a mutant male entered the population and by caring for his young was able to have greater reproductive success than deserting males, male parental care should spread through the population. In many fish, such as the stickleback, males do care for the young and the females do not. Exactly the same argument would have applied if a mutant female had had a comparable advantage over other females and also applies when such a female enters a population where male care had been the usual practice. If such a female does better than the other females, females will eventually care for the young alongside the males. Biparental care is especially common in the birds.

Even when both parents care for young, their interests do not coincide. They certainly have a common interest in their offspring's survival, but they have diverging interests in as much as each one might be able to increase its reproductive success by spending time seeking extra mates elsewhere. Micro-economic models have been

borrowed to explain what happens when two animals cooperate but do not share identical interests. I shall describe a model originally proposed by Chase (1980) and which I have slightly modified.

Given that each individual can affect its reproductive success in more than one way, if it decreases the effort devoted to one form of reproduction, it must increase the effort devoted to the other by a specified amount in order not to reduce its overall reproductive success. Withdrawing care of offspring must be matched by stepping up the search for an additional mate. If the two forms of achieving reproductive success are taken as separate axes, it is possible to draw lines of equal success. These are comparable to the economists' lines of indifference, expressing the same level of satisfaction with different combinations of goods (see Fig. 1). Such lines have three features:

1. Higher levels of reproductive success are achieved as the joint amounts of the incompatible activities are raised.
2. The contours joining points of equal reproductive success never cross each other.
3. The shape of the contours will depend on how much help is received in caring for offspring¹.

¹ This is where I depart from Chase's (1980) formulation of the problem.

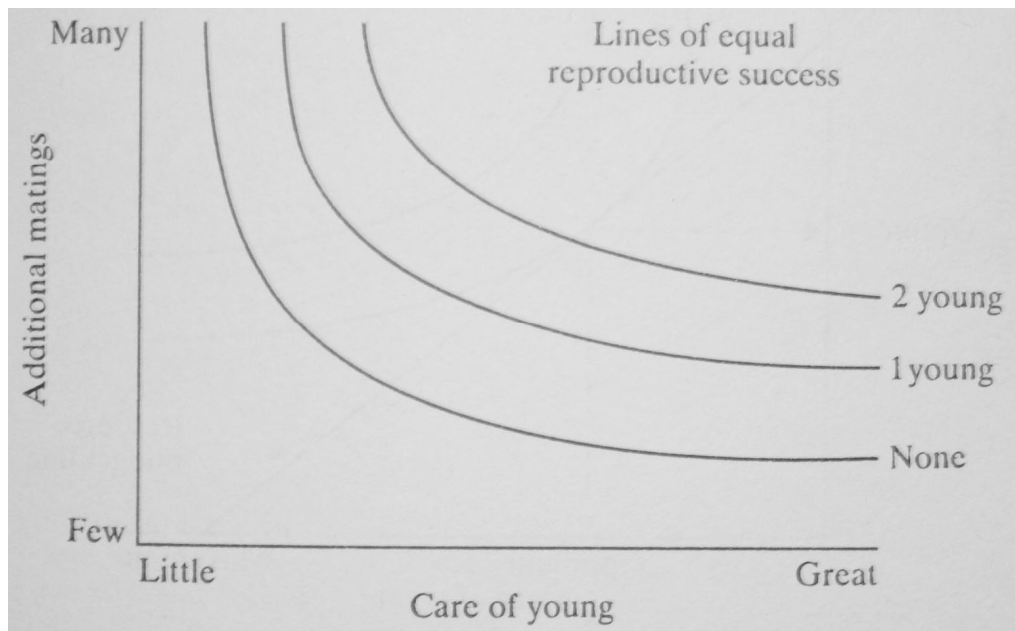


Figure 1. Hypothetical lines of equal reproductive success when an animal can increase its reproductive success by additional matings or by parental care but the two activities tend to be mutually exclusive (from Chase 1980).

In many species of birds, in which both sexes normally care for the young and one parent dies or disappears, the remaining mate increases the time and energy it devotes to caring for the young. This frequently observed event raises the question of the extent to which an animal can be a "free-rider" on the efforts of its mate.

We now need to calculate what economists would call a "Resource Budget Line". This is the line indicating the maximum amounts of all possible combinations of behaviour of which the animal is capable. If

the animal has a fixed amount of energy available, it might use all of that energy on caring for the young which it has had with a single mate. At the other extreme, it might spend all of it on looking for other mates. Those two points should be joined by a straight line (see Fig. 2), which may be drawn across the contours of equal reproductive success. It enables us to ask what is the best response in terms of producing the largest number of young. The best response, and therefore the one that is most likely to evolve, is where the resource budget line touches the contour of highest value in that set of conditions.

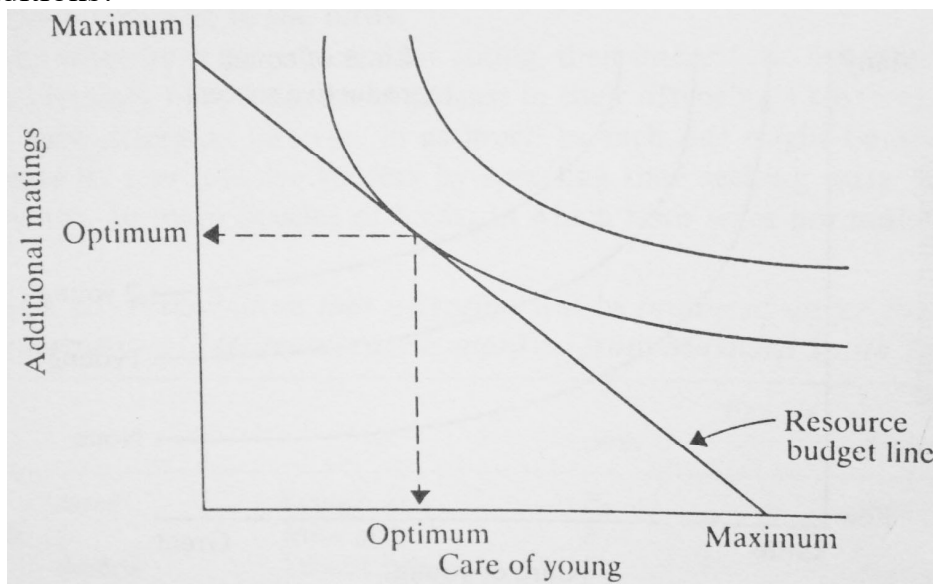


Figure 2. The optimal amounts of time devoted to seeking extra matings and to caring for young are found where the resource budget line touches a line of equal reproductive success at a tangent. The animal cannot exceed the maximum time spent searching for extra mates or the maximum amount of time caring for young because of energetic and other restraints.

Since the pattern of contours changes with the amount of help the animal gets in caring for its young, the best response would also be expected to depend on conditions. Therefore, animals that are capable of taking note of the conditions should have evolved so that they change the amount of time they allocate to care of the young in response to changes in conditions. The model is simple, but the expectation is matched by many observations of birds increasing the parental care devoted to their offspring when their mate deserts or disappears.

Free-riders who leave all parental care to their mates will not evolve if the respective amounts of care given by cooperating parents reaches an equilibrium when they both care for the young. Each animal involved in the cooperative care of young has an independent set of conditional rules about what to do if the help provided by its partner changes. These may not be the same for both sexes and will depend on the opportunities available for getting other matings. The rules will be the product of Darwinian evolution in the sense that the animals that had most offspring in the past would be those that most nearly found the optimum for a particular set of conditions.

First, consider the female's reaction to the male's contribution, basing it on the optimal response in a particular set of conditions (see Fig. 3). Next consider the male's reaction to the female's contribution.

Providing the slopes are right, the two reaction lines generate an equilibrium where they intersect. Note that a stable equilibrium is not an inevitable outcome for all combinations of lines (see Houston & Davies 1985). Similarly, if the two reaction lines do not intersect, the individual with the higher reaction line will do all the work.

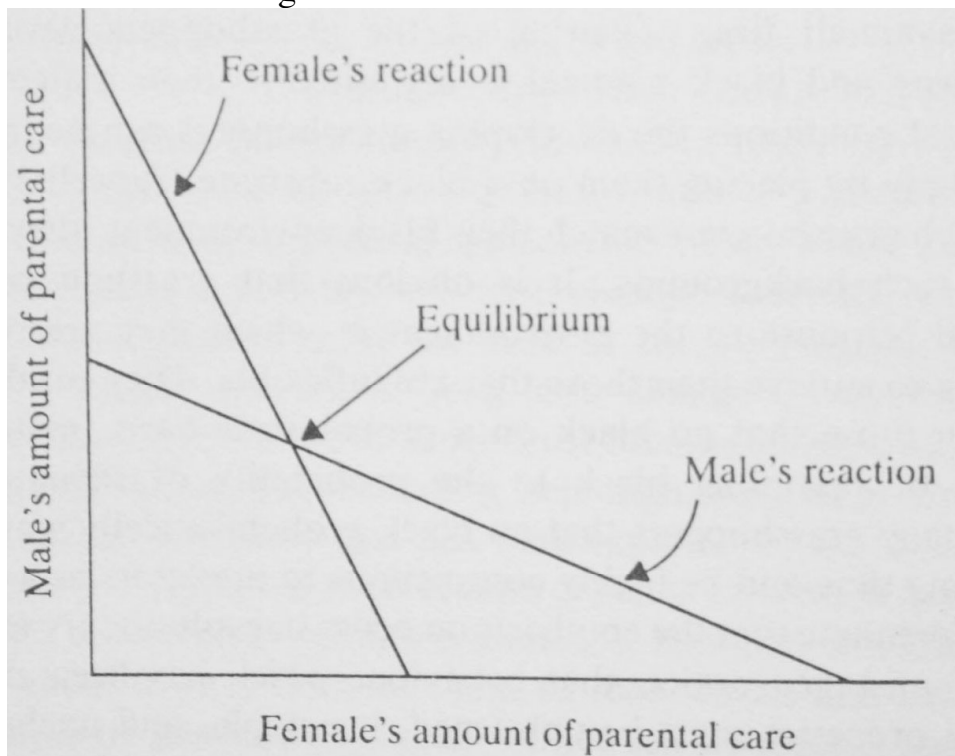


Figure 3. The female's reaction to a given amount of parental care by the male and the male's reaction to the female. Both sexes tend to reduce their parental care if the other increases its care, but the nature of the two individual's reactions to the other is such that if both sexes are present and healthy, the value for the amount of parental care given by each sex stabilises after interplay between the two of them. These values (which need not be the same for the two sexes) represent the separate optima for both of them.

Houston & Davies (1985) have provided an illustration of how such postulated rules might work in a common garden bird, the Dunnock. Birds are generally supposed to be monogamous, but in the Dunnock only some of the breeding arrangements involve a stable relationship between one male and one female. Some are polygynous, some are polyandrous and, even more remarkably, some breed in combinations of several males and several females (see Davies, 1985). In all breeding arrangements the amount of effort put into feeding the young increases with the number of young. Taking that into account, in monogamous pairs the female is responsible for slightly more than half of feeding. However, the female reduces the number of feeds to the brood when she is helped by two males. The feeding rate is about 7% less when she has two mates. Houston & Davies (1985) found that she does reduce her own rate of feeding the young when she has more help, but she certainly does not give up altogether - as might be naively expected if she operated on the principle of unenlightened self-interest.

5. Conditional expression

The example of the Dunnocks looking after their young emphasizes how the cooperative behaviour of animals may be exquisitely tuned to current conditions. Even in the plants and the invertebrate animals, examples of the induction of special responses to special environmental conditions are commonplace. Grasshoppers living on

African savannah are normally greenish-brown. They prefer backgrounds which are the same colour as themselves and they are difficult to see on such backgrounds. Their coloration and behaviour undoubtedly protect them from being eaten. After a savannah fire, offspring of the grasshoppers develop in a different way and black pigment is deposited in their cuticles. Under experimental conditions the developing grasshoppers can be induced to do this simply by placing them on a black substrate (Rowell, 1971). Not only do such grasshoppers match their black environment, they prefer to settle on such backgrounds. It is obvious that grasshoppers with a conditional response to the environment in which they were reared were more likely to survive than those that are inflexible. They would also have done better than those that became black on a probabilistic basis matching the likelihood of becoming black to the probability of savannah fires. Clearly, many grasshoppers that went black probabilistically would have done so at the wrong time and been highly conspicuous to predators as a result¹.

It was unfortunate that the emphasis on genes in evolutionary arguments led to the mistaken notion that behaviour which has been shaped by Darwinian processes must be unlearned, inevitable and unchangeable. One person who fostered this error was the dominant

¹ See Caro & Bateson (1986) for a general discussion.

promoter of sociobiology, E.O. Wilson (1975). When challenged about the way he had played down the interplay between the developing individual and its environment, he suggested that the role of developmental processes might usefully be "decoupled" in the study of social behaviour (Wilson, 1976). The strong implication was that the development of the individual is merely a complex process by which genes are decoded. Since developmental biologists dismissed this position as being naive (see Bateson, 1986), Wilson had to respond to the criticism, particularly when dealing with the obvious plasticity of human behaviour. His solution was to replace genes that prescribe the form of behaviour of an individual by genes that do the same for "epigenetic rules" (Lumsden & Wilson, 1981). The hypothetical rules are supposed to determine how development proceeds and how learning takes place. While the use of such explanatory devices had been commonplace in both biology and psychology, Lumsden & Wilson differed from other theorists of development in proposing that such rules are genetically determined, in the sense that the characteristics of the rules themselves develop independently of the state of external conditions. Such a position does not seem plausible *a priori* and the empirical grounds for doubting it are also substantial (see Bateson, 1976, 1983).

What an individual does is highly dependent on circumstances, in terms of both the short-term control of behaviour and in the

development of an individual's particular style (Hinde 1982; Huntingford 1984). No animal behaves in the same way irrespective of conditions. Nor, of course, do humans. Departures from the norm are not necessarily pathologies generated by abnormal conditions. They may well be highly adaptive responses to particular ecological conditions.

When, in the course of evolution, the form of behaviour became conditional on external circumstances, two things could have happened. First, novel environments were especially likely to generate surprising results. Removal of the buffering against fluctuations would have opened up the behavioural system to short-term change. Second, sensitivity to the social environment, as in the case of the cooperating Dunnocks, might have set the pattern of long-term change on an evolutionary pathway which led to a state in which trust became an important requirement for cooperation. The mediating step might have been the evolution of self-awareness, an issue that has been explored by Humphrey (1976, 1985) and Crook (1980). If what A does depends on what B has done and likewise B's behaviour is conditional on A's, then they are locked into a game in which it pays both of them to look ahead and calculate the consequences of particular actions¹. Once self-awareness had

¹ A similar argument was outlined by Jolly (1966) and played a part in the thinking of Trivers (1969, 1986) when he considered the evolutionary consequences of reciprocated aid.

evolved, trust might then have become one of the requirements for effective cooperation (see also Axelrod, 1984).

In conclusion, it is obvious that social life may sometimes involve conflict and intense competition. It may also involve real benefits and active cooperation. The balance between these conflicting pressures often change so that, if conditions become really difficult, the cooperative arrangements break down. Or if members of a group are not familiar with each other, no mutual aid occurs until they have been together for some time. As familiarity grows, individuals come to sense the reliability of each other. Furthermore, expectation of an indefinite number of future meetings means that deception is a much less attractive option. The major message is this. In cooperative behaviour, conditions matter a lot.

6. Conclusion

In this chapter I have described three explanations for the evolution of cooperation. My own view is that all three types of process have been important and they may all have been involved in the evolution of some forms of cooperation. Once evolutionary stability of cooperative behaviour was achieved by one or more of the Darwinian processes I have discussed, features that maintained and enhanced the coherence of the highly functional cooperative behaviour would then have tended to evolve. Signals that predicted what one individual was

about to do, and mechanisms for responding appropriately to them, would have become mutually beneficial. Furthermore, the maintenance of social systems that promoted quick interpretation of the actions of familiar individuals would have become important. Finally, when the quality or quantity of cooperation depended on social conditions, increasing sensitivity and self awareness would have become advantageous.

In ending, I want to reiterate four points. First, evolutionary process must be distinguished from its consequences; differential survival in the past does not necessarily mean social competition now. Second, in biology the outcome of differential survival is the "interest" and no motive is required for that interest to be maximised by a given course of action. Third, at least three explanations for the evolution of cooperation can be offered. Finally, the conditions in which cooperative behaviour occurs, and those in which it does not occur, need to be properly explored and understood. When that is done we may be able to explain the origins of trust.

Acknowledgements

I am grateful to Nick Davies, Diego Gambetta and Bernard Williams for their comments on an earlier version of this chapter.

References

Alexander, R.D. (1974) The evolution of social behavior. Annual Review of Ecology and Systematics, 4, 325-383.

Allee, W.C. (1951) Cooperation among Animals, with Human Implications. Schuman: New York.

Axelrod, R.(1984) The Evolution of Cooperation. Basic Books: New York.

Bateson, P.P.G.(1976) Rules and reciprocity in behavioural development. In Growing Points in Ethology, ed. by P.P.G. Bateson & R.A. Hinde, pp. 401-421. Cambridge University Press, Cambridge.

Bateson, P.[P.G.] (1983) Rules for changing the rules. In Evolution from Molecules to Men, ed. by D.S. Bendall, pp. 483-507. Cambridge University Press, Cambridge.

Bateson, P.[P.G.] (1986) Sociobiology and human politics. In Science and Beyond, ed. by S. Rose & L. Appignanesi, pp. 79-99. Blackwell, Oxford.

Caro, T.M. & Bateson, P.[P.G.] Organisation and ontogeny of alternative tactics. Animal Behaviour, 34,1483-1499.

Chase, I.D. (1980) Cooperative and non-cooperative behaviour in animals. American Naturalist, 115, 827-857.

Crook, J.H. (1980) The evolution of consciousness. Oxford University Press, Oxford.

Darwin, C. (1871) The Descent of Man, and Selection in Relation to Sex. Murray, London.

Davies, N.B. (1985) Cooperation and conflict among dunnocks, *Prunella modularis*, in a variable mating system. *Animal Behaviour*, 33, 628-648.

Dawkins, R. (1976) *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins, R. (1982) *The Extended Phenotype*. Freeman, Oxford.

Grafen, A. (1984) Natural selection, kin selection and group selection. In *Behavioural Ecology: An Evolutionary Approach*, second edition, ed. by J.R. Krebs & N.B. Davies, pp. 62-84. Blackwell: Oxford.

Hamilton, W.J. (1964) The genetical evolution of social behaviour. I & II. *Journal of Theoretical Biology*, 7, 1-52.

Hamilton, W.J. (1975) Innate social aptitudes of man: an approach from evolutionary genetics. In *Biosocial Anthropology*, ed. by R. Fox, pp. 133-155. Wiley, New York.

Harvey, P.H., Kavanagh, M. & Clutton-Brock, T.H. (1978) Sexual dimorphism in human teeth. *Journal of Zoology*, 186, 475-486.

Hinde, R.A. (1982) *Ethology*. Oxford Univ. Press, Oxford.

Houston, A.I. & Davies, N.B. (1985) The evolution of cooperation and life history in the dunnock, *Prunella modularis*. In *Behavioural Ecology* ed. by R.M. Sibly & R.H. Smith, pp. 471-487. Blackwells, Oxford.

Humphrey, N.K. (1976) The social function of intellect. In *Growing Points in Ethology* ed. by P.P.G. Bateson & R.A. Hinde, pp. 303-317. Cambridge University Press, Cambridge.

Humphrey, N.K. (1986) *The Inner Eye*. Faber & Faber, London.

Huntingford, F. (1984) The Study of Animal Behaviour. Chapman & Hall, London.

Jolly, A. (1966) Lemur social behavior and primate intelligence. Science, 153, 501-506.

Kitcher, P. (1985) Vaulting Ambition. MIT Press, Cambridge, Mass.

Kropotkin, P. (1902) Mutual aid: A factor of evolution. Heineman, London.

Lorenz, K. (1966) On Aggression. Methuen, London.

Lumsden, C.J. & Wilson, E.O. (1981) Genes, Mind, and Culture. Harvard University Press, Cambridge, Mass.

Maynard Smith, J. (1976) Group selection. Quarterly Review of Biology, 51, 277-283.

Maynard Smith, J. (1977) Parental investment: a prospective analysis. Animal Behaviour, 25, 1-9

Maynard Smith, J. (1982) Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Packer, C. (1977) Reciprocal altruism in *Papio anubis*. Nature, 265, 441-443.

Packer, C. (1986) The ecology of sociality in felids. In Ecological Aspects of Social Evolution: Birds and Mammals, ed. by D.I. Rubenstein & R.W. Wrangham, pp. 429-451. Princeton University Press, Princeton, New Jersey.

Rowell, C.H.F. (1971) The variable coloration of the Acridoid grasshoppers. Advances in Insect Physiology, 8, 145-198.

Trivers, R. (1971) The evolution of reciprocal altruism. Quarterly Review of Biology, 46, 35-57.

Trivers, R. (1985) Social Evolution. Benjamin/Cummings, Menlo Park, Calif.

West-Eberhard, M.J. (1975) The evolution of social behavior by kin-selection. Quarterly Review of Biology, 50, 1-33.

Wilson, D.S. (1980) The Natural Selection of Populations and Communities. Benjamin/Cummings, Menlo Park, Calif.

Wilson, E.O. (1975) Sociobiology: The New Synthesis. Harvard University Press, Cambridge, Mass.

Wilson, E.O. (1976) Author's reply to multiple review of "Sociobiology". Animal Behaviour, 24, 716-718.

Wrangham, R.W. (1982) Mutualism, kinship and social evolution. In Current Problems in Sociobiology, ed. by King's College Sociobiology Group, pp. 269-289. Cambridge University Press, Cambridge.