treating human communication is altogether a dead end, an imposing signpost leading to a blank wall. Speech only "alters the probability pattern" of people's behavior if it is *understood*. That is what distinguishes it from infections, blows, flea bites and most of the other ways in which organisms interact.[3] That is what terms like *communication* are needed to mark. And the reason we pick out some interactions between animals as being communications is because that is true of them too. Threats, greetings, infantile appeals, submissive gestures have to be *understood* as such if they are to work. They are sometimes misinterpreted. And to explain the response to them one must always know what they are being understood *as*—what impression the receiver has formed of the signaler's motives and intentions. This is not true of infections, flea bites, or any sort of interaction between cells. And it is true of blows only insofar as blows are, among other things, a form of communication—they signal the striker's social attitude. Understanding communication therefore is interpreting it—whereas understanding causal interaction is no such thing. Communicating is conveying information and social attitudes. And this is something that it makes sense to talk only of conscious beings as doing.

Wilson's eccentric definition does not do him much harm, because he very sensibly ignores it, and goes on to decide what is and is not communication by common sense. But he can do this only by following, without acknowledgment, the everyday principle of confining it to beings which can reasonably be treated as conscious, and as *noticing* each other's signals. Far from being a retreat into mysticism, this provides the only context within which words like *communication* can have any kind of sense. Someone who genuinely thought of animals as unconscious automata would have to stop speaking of what they do as communicating. As that would lose us a most valuable distinction, nobody seems likely to try it.

[3] "The experimental animal which *eats* the dinner-bell has misinterpreted the sign."—Max Black. See pp. 225–227.

6

ALTRUISM AND EGOISM

DIFFERENT NOTIONS OF SELFISHNESS

I have been suggesting generally that if we are to understand the behavior of conscious beings, we must take their motives seriously and not try to reduce them to something else. I must now illustrate this in an important case: that of altruism and selfishness.

Many people have found altruism a problem, and Wilson is among them. In fact he speaks of it as "the central theoretical problem of sociobiology; how can altruism, which by definition reduces personal fitness, possibly evolve by natural selection?" (p. 3). But two totally distinct problems are combined here. There is, first, the traditional problem of selfishness in human life, which is a problem about motive. It runs "Can a conscious agent deliberately choose to do things that *he thinks* will not pay him?" This problem can be considered only by people willing to take motives seriously. Second, there is a problem about evolution, which runs "Can a trait survive if it leads its bearer to do things which do not (*in fact*) pay him?"

Both problems arise from the obvious point that people, and animals too, often do seem to do nonpaying things. The first inquirer is asking, can they know what they are doing? Do they perhaps understand their interest differently from the observer? He wants to make their acts *intelligible* in the sense of imaginable from the inside, from the angle of

another possible agent—to grasp their reasons, to see what makes them tick. The second, by contrast, is viewing action from outside, from a very remote standpoint indeed. The sense in which he wants to understand it is that of fitting it into the theoretical scheme of evolution, which means reconciling it with natural selection.

Taken separately, neither of these problems is too hard. Taken as a tangle, however, they become almost impenetrable.

They differ not only because they arise from different inquiries and use their terms in distinct senses, but because they point to quite distinct ranges of example. No doubt the acts that a conscious agent believes will pay him often actually do so. All the hard work of prudence is devoted to making this happen. But it is always an uncertain business. Gamblers may win fortunes while the most careful accountants overwork and die of thrombosis. And even if outer events turn out as expected, feelings may not. Both the successful (surviving) accountant and the successful gambler may find that they do not know what to do with their success. Things are liable to turn to ashes in people's mouths. There are therefore constant difficulties in deciding what "really pays." And what pays at one time of life or in one situation may be disastrous at another. When should the payoff be declared? People fall into Solon's paradox; they find themselves saying "call no man happy till he is dead," and then reflecting that that hardly seems the time to do it either.[1]

Let us now consider the two questions separately. First, can a conscious agent deliberately choose to do things that he thinks will not pay him?

The answer to this simply depends on what we mean by "pay" (or by advantage, benefit, or whatever similar word we fancy). If we use such words in a very wide sense, the answer must be no, but the point becomes rather trivial. What we deliberately choose must certainly in some sense please and satisfy us. But this is as true of suicide, martyrdom, alcoholism, and unprofitable vengeance as it is of eating and drinking. What we choose must somehow attract us. But it need not in any normal sense pay us. Saying that it does is trivial because it stretches such words as *pay* beyond their normal use for the sake of paradox.

[1] Aristotle, *Ethics* 1.10

The question has a much more ambitious and interesting point if we use words like *pay* in their direct and obvious sense, meaning "deliver a competitive, outward advantage." This is the way in which people making cost-benefit analyses use them. We now ask: can people deliberately do things that they believe will shorten their lives, lessen their income, damage their health, or diminish their power? Can they, that is, deliberately do such things for the sake of doing them, not trading off one advantage against another—say, losing income in order to increase health—but genuinely ignoring all such advantages? To keep this question clear of the earlier trivial one, we ought to have a list of publicly recognized advantages. It would probably be drawn from the list of our basic needs.

What we have now is a question of fact. And a brief glance around shows the answer to be yes. People do, sometimes, knowingly commit suicide. They do sometimes pursue vengeance, inquiry, or amusement that is sure to injure them. They smoke. They gamble, even against what they well know to be hopeless odds. (Think of Dostoevsky.) They indulge in futile competition. In fact, we all do some of these things sometimes. Prudence is not universal. Certainly no one is imprudent all the time. But even if imprudence were quite rare, its existence would still destroy the position of philosophic Egoism, for that position is that "expecting payment" in the full outward sense is essential to the notion of wanting or choosing, that to want something necessarily *means* "to believe that you personally will profit from it," means making a favorable cost-benefit analysis.

Now our selfishness does not work like this. We are far too lazy-minded to do so much calculation. Mostly we do either what we have to or what we feel like doing. And quite often what we feel like doing is helping or pleasing other people—though of course not so often as they, or bystanders, might hope. If we never felt like doing this, the word *selfish* could never have arisen, since it describes a person who does not have these feelings or is not moved by them to action, and distinguishes him from other people who do and are. But the first point to insist on is how little we actually calculate.

Even in cases that do lead to solid, visible benefits, such as eating, what moves us is not usually the calculation of the benefit, but direct desire. This becomes obvious when illness or distraction destroys the

appetite; we then find it really hard to eat. And people put on a diet for their health can have great difficulty in keeping to it. But the same holds for more complex desires, such as sociability, friendliness, curiosity, the taste for music, vengeance, or crime. Someone who tries to pursue these things without a genuine, spontaneous taste for them, merely for social advantage or because his doctor has told him to, will fail. His only hope of succeeding comes if he manages to *wake* a genuine taste.

It is true that many of the things we and other species like doing do eventually pay, and this fact is of great interest to the evolutionist. And it is true that we try to calculate our own interests. Calculation of this kind is even one of the things we have a taste for, as can be seen from the impractical lengths to which some people carry it. We are creatures naturally interested in the future, and it is a part of our rationality to be so. But neither our rationality nor our motivation can be limited to this tendency. The taste for future security is one taste among others, a taste that is often overborne.

The same thing is clear when we think how people choose their way of life. This choice is real—even in the most restricted life there is something worse that we try to avoid—but it is never undertaken by making a list of all conceivable modes of living and then doing an impartial cost-benefit analysis of their likely effects. Innumerable ways of life do not occur to us, and of those that do, many are marked "not for me." We use calculation of consequences only to choose among those that may seem initially possible and equal.[2] There is nothing irrational about this. We have to live in the present, not the future. The payoff pattern, the relation of means to end, is much less important in motivation than people suppose. What matters is the relation of part to whole. To give meaning to life, we want to see what we do as an element in something that, as a whole, satisfies us.

Naturally, this something does not have to be edifying. It may be vengeance or destruction quite as easily as friendship or discovery. But—and this is the point for my discussion of Egoism—because it needs to be something larger than our own life, something worth participating in, almost necessarily it involves other people. Even the most hell-bent financier usually needs somebody to share his wealth with. Failing to find someone, or finding only the wrong kind of people, he is frustrated. Even the most solitary artist or thinker hopes to find, eventually, a public that will understand him. Certainly some people are natural recluses, but—leaving aside the difficulty of finding out whether they are actually pleased with their lives—they do seem somewhat abnormal.[3] I am not criticizing them, but pointing out that they can hardly provide an escape for the rest of us from the dilemma endemic to a social species; awful though other people may be, most of the activities that we really care about must involve them. Solitude is necessary for many parts of our life, but it cannot be the climate of the whole. Nor do we need other people merely as scratching-posts, *means* to the adjusting of our own states of consciousness. Our nature demands for its fulfillment ends to aim at which lie outside ourselves.

[2] See a very interesting discussion by Bernard Williams in J. J. C. Smart and Williams, *Utilitarianism: For and Against* (Cambridge, 1973), pp. 82–92.

THE USE AND MISUSE OF EGOISM

At this point champions of philosophic Egoism tend to cry "humbug!" and stop listening. This is not very bright of them, since the point that I have made holds just as strongly for sinister motives such as revenge as for decent or heroic ones. (In fact the demand for real outward action is just what makes revenge such a dangerous motive; you cannot get people bent on it to accept psychotherapy or enlightened self-interest instead.) But the humbug difficulty must be met, because the whole trouble about evaluating Egoism is that people see the problem not as a real problem at all, but as a mere clash of styles and attitudes. William James pointed out how we often dramatize an argument as a clash between tough-minded and tender-minded attitudes, between partisans, as it were, of science and sympathy. This habit chronically infests and distorts certain philosophical controversies, particularly

[3] Proust, himself a thorough recluse, remarks, "In the case of the solitary, his seclusion, even where it is absolute and ends only with life itself, has often as its primary cause a disordered love of the crowd, which so far overrules every other feeling that, not being able to win, when he goes out, the admiration of the hall-porter, of the passers-by, of the cabman whom he hails, he prefers not to be seen by them at all, and with that object abandons every activity which would force him to go out of doors" (*Within a Budding Grove*, tr. C. Scott Moncrieff, Pt. 2 [London, 1924], p. 123).

about such tough-seeming but confused positions as determinism, hedonism, egoism and behaviorism. Role-playing of this kind paralyzes our thinking because it makes thought seem unnecessary; the positions are ready-made for us. Once they have imagined themselves to be tough-minded, people are quite liable to accept the loosest and most vacuous ideas uncritically, provided they are put forward in the right contemptuous tone of voice, while those who have cast themselves as tender-minded sometimes accept very brutal suggestions, provided that they sound familiar and traditional. But, as a minute's thought will show, science and sympathy cannot be alternatives, much less opponents. Anyone who treats them as such has forgotten the point of both. They are distinct aspects of life, and we all need both of them.

Looked at without this bias, Egoism is not a plausible account of motivation. And if we refuse to accept Egoism, the traditional problem about selfishness is not particularly difficult. We ask: can anyone deliberately act just for someone else's interest? This seems strange if we have already decided that the only reason he can act for is his own interest. But, as we have seen, people can and do constantly act for other reasons; they even act directly contrary to their own interests.

The point about hypocrisy, however, is important. For hypocritical distortion does *exaggerate* the claims others have on us, and play down the claims of self-fulfillment. That distortion is what philosophic Egoists—notably Hobbes, but also Aristotle, Spinoza, and Nietzsche—protest against.[*] They want to redress the balance. Hobbes put the position particularly strongly because he had a crucial political point to make. As he saw it, if people attended better to their own interests instead of showing off, they would avoid doing many of the terrible things they do; in particular, they would altogether give up waging wars of religion. The humbug of chivalry was therefore Hobbes's prime target, and the *Leviathan* is, among other things, the *Catch 22* of its day. Look, he says, at the realities. Death is not a joke. It is the end of you. Does not all other value, therefore, depend on, and perhaps reduce to, your power to avoid dying?

[*] Aristotle has an extremely interesting attempt to derive the love we feel for others from our love of ourselves (*Ethics* 9.4,8). This has been a very popular line with philosophers, but I think it is ultimately perverse. See the last two sections of chap. 13.

Now it is perfectly true that enlightened self-interest is often a far better guide in politics than posturing and *machismo*, both for our own interest and for other people's. Self-preservation is not only a strong general motive with us, but also a positive duty. What it cannot be is our only duty or our only motive. While we survive, we must *do* something. And further, the point of all we do cannot be (as the subtler "tough-minded" theories suggest) to manipulate our own state of consciousness. Even in politics, there are other things that concern us besides finding the best means either to staying alive or to self-stimulation. And in private life, this is still plainer. People do an enormous range of things "as ends in themselves"—that is, not as a means to anything else, and certainly not to adjusting their own state of consciousness (which is a delicate job, and perhaps could only be efficiently performed by the pharmacist or the hypnotist). Football players want to win their games, and supporters want them to win them. Neither will accept, even from the most skilled pharmacist or hypnotist, the offer of *the sensation of a won game* instead. Avengers do not want the sensation of avenging; they want people's blood. And similarly, rescuers and benefactors do not (if they are real ones) just want the sensation of rescuing and benefacting. They want to help people. This involves wanting the people actually to be helped. If somebody else does so first, they will usually be quite satisfied.

The question of hypocrisy must be considered here, because we have a kind of suspicion that all these people are pretending. This, however, can hardly be right. You cannot have fake Vermeers unless there are some real Vermeers. This is not because it is *difficult* to copy without a model, but because nobody could want to. The existence of fake rescuers and benefactors, far from disproving the existence of the real thing, actually establishes it.

Again, consider the arts. A real actor or musician is not trying to tinker with his own subjective state. He is trying to practice his art properly. If he is interested in altering anyone's state of mind, it is that of his audience. Feedback will please him, but only insofar as it shows that he has got across what he was trying to do. (If what he wanted was simply the sensation of being applauded, the simplest thing would be

to go through the motions of playing, while running a recording of an excellent performance by someone else. What's wrong with that?) The payoff pattern is quite unsuitable here. These actions point outward; they buy no return ticket.

Art and sport are good areas, too, for showing how much shallower and less useful the payoff pattern of means and ends is than the pattern of part and whole in understanding human motivation. Is the play in the first half of the game just a means to that in the second? Or is all the play just a means to goals? Why not shoot them right away and save time? If the positions in the league table are the point, why not determine them by calculating the probabilities, or with a randomizer?[5] Are the first bars of the *Appassionata* just a means to the last cadence? Again, within the life of a player, actor, or musician, at what point is the payoff to be declared? If the actors who play the bit parts in the first scene of *Hamlet* are each doing so merely as a means to playing Hamlet some day themselves, they will play them very badly. The bit parts may be a means to that, just as they are certainly a means to a paycheck. They may also be a means to diffusing a weird atmosphere that will produce the right emotional effect later in the play. But had the director treated that problem as someone editing a horror film would have to treat it—as an isolated practical problem in emotional manipulation—he would go wrong in much the same way as a Martian might if asked to organize a football game. Understanding that goals are the point, he takes twenty balls down to one end of the ground and fires them into the goal from a cannon. It is not going to be easy to explain to him—or to ourselves—just what is wrong with this. But activity for its own sake is a pattern that pervades human life. Researchers like Masters and Johnson have documented the sad effects on people's sex lives of treating the whole business of sex merely as a means to a quick orgasm; it is well known that things do not go much better if you treat it as a means to quick production of an heir. Payoffs are a part of our lives, but not at all such an important one as may appear. The misleading point is, perhaps, that where the whole extends through time, we naturally have our eye on the part we can still influence, and this is the future. But

[5] A possibility ingeniously exploited in Michael Frayn's novel, *The Tin Men* (London, 1972).

there is no moment when the payoff can be declared, and the search for one leads to confusion and despair. John Stuart Mill's adolescent breakdown was, it seems, a result of his entrapment in this feature of Utilitarianism.[6] Attempts to use the pattern consistently and exclusively must, as far as I can see, lead to this sort of barren desolation. The attraction of Egoism, in fact, vanishes if you are consistent about it. It has only the momentary impact of suggestions that demolish isolated pieces of humbug. Its most heady pronouncements are often combined with others pointing a very different way—as when Nietzsche, so often eloquent for Egoism, remarks all the same, "Man should be trained for war, and woman for the recreation of the warrior; all else is folly."[7] What's she supposed to get out of that?

HOW TO MISUNDERSTAND ALTRUISM

For great areas of human motivation, then, Egoists are forced not just to distort their account, but to make it meaningless.

Hobbes and Aristotle do this repeatedly. Two examples are all I have room for. Aristotle, considering the case of somebody who dies to save a friend, asks whether this must count as acting for the friend's sake, and answers, no, the man has merely secured for himself a benefit greater than his friend's, namely, glory. He "prefers a short period of intense pleasure to a long one of mild enjoyment."[8] But if we believed this, we should not be especially impressed with him for doing what he did, and the glory he is after could never have attached to his act. (Even if the fellow fails to think this out, of course, his cost-benefit analysis is miserably shaky; how can he be sure of enjoying himself enough in the time available?)

[6] "It occurred to me to put the question directly to myself, 'Suppose that all your objects in life were realized, that all the changes in institutions and opinions which you are looking forward to could be completely effected at this instant, would this be a great joy and happiness to you?' And an irrepressible self-consciousness distinctly answered, 'No!'" (*Autobiography*, chap. 5).

[7] *Thus Spake Zarathustra*, Pt. 1, "Of Womankind." He even adds "The man's happiness is, I will. The woman's happiness is, He will." Perhaps in fact egoism is always a sex-linked doctrine.

[8] *Ethics* 9.8.

Again—Hobbes defines pity as "Grief for the Calamity of another, arising from the imagination that the like calamity may befall himself" (that is, the person who feels it).[9] But if the only thing we ever felt on seeing the disaster of another was fear, our behavior would show only fear. There would then be no such word as pity, since there would be no such thing. And Hobbes could never have been called upon to define it.[10]

This is the traditional pattern of distortion. Anyone who doubts its power may be interested to watch Wilson falling into the very same hole. In *Sociobiology* he considers the case of someone who dives into water to save a drowning stranger, and remarks:

> The reaction is typical of what human beings regard as "pure" altruism. However, upon reflection, we can see that the good Samaritan has much to gain by his act. . . . If such episodes are rare, the Darwinist calculus would predict little or no gain. . . . But if the drowning man reciprocates at a future time, and the risks of drowning stay the same, it will have benefited both individuals to have played the role of rescuer. Each man will have traded a one-half chance of dying for about a one-tenth chance. [p. 120]

Now this passage might merely be saying that *in fact* they will, on average, have benefited themselves, or at least done themselves no harm. As we shall see, this is the most that the argument can possibly call for. But the language constantly implies more. It unmistakably refers to motivation, and analyzes it in the traditional distorting Egoistic style. The word that does so most plainly in the passage above is *traded*. *Trading* is a name for something that can only be done deliberately, and with appropriate motives. Traders must know what they are doing. To do something that merely turns out to benefit you is not trading, but if you make a deal, you have still traded if it turns out badly. Whatever the upshot, a rescuer was not "trading" unless his motive was a calculation of profit. Of course it could have been so. But people do in fact quite often rescue strangers *without* first checking that these strangers

[9] *Leviathan*, Pt. 1, chap. 6.

[10] See Butler, two long, very astute footnotes, on Sermon 1, sec. 6 and Sermon 5, sec. 1.

are husky, loyal, useful allies who will stay around until a counter-rescue becomes necessary. And a rescuer who finds that he has pulled out of the water a nonswimming old gentleman who lives at a distance does not necessarily react as he would if he had spent a lot of time salvaging a whiskey crate and then found it to be empty—that is, by throwing him back in again. He may still reckon that he has done what he set out to do.

The word "trade" does not stand alone. All Wilson's key terms are drawn from the language of conscious motivation. *Altruism*, *selfishness*, and *spite* are his names for activities that benefit (respectively) other people, the agent, and nobody. (It might have been better to talk of other-benefiting, self-benefiting, and purely injurious activities.) The dictionary, however, defines altruism as "regard for others, as a principle of action," which is a motive. And this is certainly its normal use. Selfishness and spite, still more obviously, are never names for what actually happens, but for the agent's state or habit of mind. A spiteful action does not cease to be spiteful because it is frustrated, nor is accidentally injuring somebody a piece of spite. Similarly, in opening the section on what he calls "reciprocal altruism," Wilson remarks, more in sorrow than in anger: "The theory of group selection has taken most of the good will out of altruism. When altruism is conceived as the mechanism by which DNA multiplies itself through a network of relatives, spirituality becomes just one more Darwinian enabling device" (p. 120).

The stern tone, suitable for one revealing unsuspected nepotism in the charities of a presidential candidate, suggests that we have examined the motives of all rescuers and have found that they are not what they seem. But no attention has been paid to these motives. In what sense can the theory of group selection "take the good will out of altruism"? It is supposed to do it by showing that traits will survive only if they benefit the group. But one cannot "take the good will out" by showing how good will evolves. *Good will* is a term describing a kind of motivation whose occurrence is what we start from. People do sometimes rescue those who cannot benefit them. The theory of group selection is there to show that such motives could not have been passed on if they had undermined the survival of the group—perhaps that they must have promoted it. This is an entirely different thing from

showing that they actually never evolved at all, that what looks like them is in fact only a cleverly disguised form of calculation. It seems important to notice that, if this were true, calculation of consequences would have to be immeasurably more strongly developed, commoner and more efficient than it actually is in man, let alone in other species. Hobbes's picture not only deprives us of our actual virtues; it credits us with others we do not have, at least in the necessary degree—a steady, unfailing prudence, a persistence in calculation, a stern intellectual honesty and consistency, a brisk readiness to act on extremely remote probabilities, such as we never actually see. (Since these really are important virtues, of which we are partly capable, the appeal of Egoism lies partly in celebrating them, as well as in attacking hypocrisy). If Hobbes's picture were accurate, the phenomena on the human scene would be totally different, and the "problem of altruism" in other less calculating species would remain quite insoluble.

This brings us back to the point with which I began this chapter—that we have two distinct problems before us. Wilson's kind of account—which is now extremely widespread—is weak in that it constantly carries into the evolutionary discussion an irrevelant and uncriticized Egoism. Officially it ignores motives, but in fact it makes constant reference to them, and because this reference is unacknowledged, its errors go uncorrected.

There is a real puzzle about the transmission of altruistic behavior. But that puzzle is solved by showing that it benefits one's kin and one's group. The further problem of inventing some way in which it can seem to benefit the agent himself arises only for Egoists, because only they have ruled that it has to do this to make him act. Wilson and his colleagues, however, take the Egoist stance, by pointing out that those benefited are often the agent's kin and suggesting that this makes it somehow possible for him to identify with them, since they will proliferate his genes for him if he does not live to do it himself.[11] Thus they give kin selection a much more prominent place in their argument

[11] Notice how this at once takes us away from the normal notion of selfishness. If someone said, "He is utterly selfish, he thinks of nothing but the prosperity of his relations," or (as we must shortly substitute), "He thinks of nothing but how many descendants he will have in five centuries' time," we could hardly fail to reply, "But why do you call that selfish? You surely mean that he is either (1) clannish or (2) crazy."

than group selection. The general point about identifying with one's kin I shall be discussing shortly. But it is worth remarking here that altruistic behavior can extend far beyond kin. It is true that those surrounding an animal very often are its kin, and that this makes a great difference to the actual transmission of its genes. But its behavior does not systematically alter when the surrounding animals are not kin at all, as in the case where a creature leaves its home group and joins another before breeding. And even if it stays put, much of its behavior benefits, and some is meant to benefit, non-relatives. The social insects are indeed a special case. For them, group and kin are effectively the same. But this is just what makes them a misleading model for studying the overall relationship of the two concepts.

Beyond this, however, the whole notion of identification is uncalled for and useless. Only a very intelligent and well-informed agent could *plan* successfully to act so as best to proliferate his genes. But we are dealing, mostly, with agents who do not plan at all, including, usually, the human rescuer. It is important that we understand such actions for what they are, as done with the motives that they actually are done with, rather than distorting them to fit a tidy theory. The motive has to be one that can function independently of planning.

There might, perhaps, have been an intelligent species somewhere which did not develop direct social impulses at all, but depended for all its social activity on calculation of consequences. We are not it. It would have done many of the things we do, but for quite different reasons. Others it not only would never have done, but could never have understood at all. But it is this alien species that is demanded by people who think of intelligence as the source of all social development. Thus Wilson writes: "a strong impelling force appears to generate social behavior in vertebrate evolution. . . . This force, I would like to suggest, is greater intelligence. The concomitants of intelligence are more complex and adaptable behavior and a refinement in social organization that are based on personalized individual relationships. Each member of the vertebrate society can *continue to behave selfishly* . . . but it can also *afford* to co-operate more" (p. 381, my italics). If intelligence had really been the only "impelling force," most of the concomitants would never have been found necessary. Why affection? Why time-consuming greeting procedures, mutual grooming, dominance and

submission displays, territorial boasting, and ritual conflict? Why play? Why (on the human scene) so much time spent in nonproductive communication of every kind—idle chatter, lovemaking, sport, laughter, song, dance, and storytelling, quarrels, ceremonial, mourning and weeping? Intelligence alone would not generate these ends. It would just calculate means. But these things are done for their own sake; they are a part of the activity that goes to make up the life proper to each species. Insofar as there is one "impelling force," it is sociability. From that comes increasing power of communication, which provides the matrix for intelligence.

THE MYSTERY OF THE UNCONSCIOUS ALTRUIST

Let us look now at our second question, the evolutionary one. Can a trait survive if it leads its bearer to do things that do not *in fact* pay him?

Here, too, the word "pay" must be taken in its normal sense, as referring to outward, competitive advantage, not to subjective satisfaction. And again, we should start from the phenomena. But which phenomena? The notion of what does pay will not be easy to delimit.

Animals do quite a number of things that can be called *altruistic* in the full and natural sense, because they are actually *aimed* at serving others. Many creatures take great trouble over rearing their young, and also defend them vigorously, sometimes getting killed in the process. Many also defend and rescue young that are not their own; some will adopt orphans.[12] Some (for example, dolphins and elephants) also help and rescue adults of their species in difficulties, and some, such as wild dogs, will also feed sick and injured adults. Many babysit, and some (such as wild dogs) will bring food to the babysitter. Beyond these things, however, there is a very wide and heterogeneous set of further activities that appear to pay others and not the agent: for instance, mobbing predators, giving warning cries, taking risks in finding new homes or food sources, and even leaving the nest when you are

[12] The widespread tendency of social creatures to help, favor, and delight in young that are not their own fits most awkwardly with the notion of "the selfish gene." If we were really concerned only with the competitive success of our descendants, ought we not to get busy suppressing the other competitors? In fact this habit is rare enough always to call for explanation when it happens. See Tinbergen, *The Herring Gull's World*, p. 169.

injured, so that (or rather, with the result that) your fellow worker-ants are not inconvenienced by your dead body.[13] Many of these things clearly are *not* done with any view to helping others, though there are borderline cases. But Wilson's method lumps them all together as altruism.

Now it is quite true that creatures take risks in doing all these things. The puzzle from the evolutionary angle is that dangerous practices should tend to eliminate themselves by natural selection. Inquisitive rabbits that tend to investigate predators do not leave many descendants. Ought "altruistic" tendencies to have been eliminated in the same way? Are they unaccountable?

Two things can make them seem so. One is a devotion to the Egoistic model. If you are convinced from the start that all motivation can aim only at the agent's own advantage, you certainly cannot make sense of action aimed at someone else's. But I have suggested that this model is wrong.

The other is an atomizing approach to impulses. If you believe that the tendency to each specific sort of action is inherited separately, then all tendencies carrying personal danger are surprising, because it should have been possible to eliminate them while keeping all the rest. But in fact there seems no reason to suppose that these tendencies are inherited in such small units, however convenient that arrangement might be to games theorists.

As for Egoism, the notions of group and kin selection make it quite unnecessary, and it ought to be dropped from the argument. Selection does not work by cutthroat competition between individuals, but by favoring whatever behavior is useful to the group. People with crude notions of "Darwinism" make an intriguing blunder here. They confuse the mere *fact of competing*, that is, of needing to share out a resource, with the *motive of competitiveness* or readiness to quarrel. Where creatures are competing (as a fact), their success will be decided by whatever tendencies they have that best help their predicament. These need not be quarrelsome tendencies at all. A species may prevail because it is better at finding food or turns to a food that is more plentiful, or because it grows protective coloration, or indeed because it becomes

[13] *Sociobiology*, p. 121.

less quarrelsome and more cooperative. Very often the best means (whether it counts as "prevailing" or not) is to move aside, finding a slightly different habitat, food, or mode of escape—inventing for oneself a new ecological niche. Now among the possible ways of prevailing, quarreling can certainly figure. But there are pretty sharp limits to its usefulness, and usually other methods are used. "Contests" between species are in general entirely figurative. Those involved need not even know what is going on. They come to the feeding place and find it bare; they find their eggs gone, or they unwittingly nurture a cuckoo. What they do is done for its own sake, not as part of a response to a challenge.

Competitiveness, as a motive, then, need not be involved at all in "competition" between species. And even where it is, it has to be limited sharply by prudence and common sense. Animals must avoid fighting that leads to serious injury; they have no doctor. Thus even a very hungry lion will abandon its kill to hyenas when the odds are against it, and vice versa. This is even more true of competitiveness within a species. While it has its uses, it can easily turn out very badly. The rival-fights of deer, for which they have evolved their antlers, are useful because they select powerful fathers for the young. But as the system develops, and the few strongest stags begin to monopolize the mating, things go less well. The gene pool is unnecessarily narrowed, and the job of protecting the young becomes harder. More serious, selection operates to produce good fighters rather than deer of good all-around capacity. This favors size, and particularly size of antlers. (Deer do not use their antlers against enemies outside their species; they drive them off with their forefeet.) Unless predators or other modes of selection intervene, stags then head towards the fate that seems to have overtaken the enormous Irish deer—a seven-foot spread of antlers and extinction.[14] Where the motive of competitiveness is

[14] For various views on this entrancing topic see Stephen Jay Gould's article, "The Origin and Function of 'Bizarre' Structures," *Evolution*, 28 (1974), 191–220. Gould quite rightly points out, (1) that the great creature prospered very well for a time—it was not some sort of Thurber animal which never made sense; (2) that the horns are not disproportionate to the body, but near the average proportion for deer. He adds the interesting fact that (3) they seem adapted for display rather than actual fighting. He concludes that they were not in fact impractical at all, but just an ordinary useful adaptation. Now he is clearly right to reject the view called *orthogenesis*, namely that such features develop in a straight line because of some mysterious force, regardless of value. Such forces are meaningless. But does it follow that sexual selection could not *continue* to favor larger forms at the point where (whether for reasons of fodder shortage or of thicker forests) they were becoming impractical? This sort of thing does occur in birds (see Wilson, p. 132, on grackles, Lorenz *On Aggression*, p. 40, on argus pheasants). It seems plausible that it could prevent the changes needed to delay extinction.

strong, it is hard for a species to get out of such a cul de sac. Strong fighting stags simply will not allow others to mate.

Thus the sort of "competition" involved in natural selection gives no support for an Egoist account of motives. All the same, some notion of possible motives is needed here. The question how a creature can come to rescue others of its species, or to give warning cries, is a genuine one, calling for an answer that makes genetic sense. We cannot treat these acts just as cultural phenomena, induced by precept or example. Could they just be imitated? As far as warning cries go, this is partly an empirical question about a particular species' power of mimicry. Some have none. But it is also about their kind of awareness, the sort of things they attend to, and the sort of reaction they are capable of. About rescuing, the point seems still plainer. Many species could never consider it, whatever examples might be shown them. We are dealing here with a difference in emotional *range* among species, with a capacity, found only in the most social species, to be so deeply concerned for others as to take drastic action. (A familiar example of this difference in emotional range is that between dogs and cats, which does not yield to cultural interference.) Now this gap has in fact somehow been bridged in evolution; this kind of active concern for others has—in several groups, separately—somehow become possible. And, since most of this bridging has been done well before the human stage, or quite independently of it, it has *not* been done by conscious calculation of the human type, but by adaptation. So it does seem likely that it carried a selective advantage. Of course, among human beings, example and training can help the reaction and make it much commoner. But they could not do this if the capacity were not there in the first place, just as it is in dolphins and elephants, although not in polar bears or hamsters. No purely cultural development could have generated it.

HOW TO MAKE THE WHOLE STUDY OF MOTIVES IMPOSSIBLE

The necessary complexity of such capacities points up the wrongness of an atomizing approach to impulses. It seems unrealistic to talk as though the tendency to rescue people were something that could be carried by a single gene. Something like a tendency to give warning cries might conceivably be so. But rescuing is plainly far too complex a phenomenon to be governed in this way. This is true partly because the kinds of danger, and therefore the kinds of rescue possible, do not follow a single pattern—but also because, in any fairly complex creature, the undertaking of dangerous actions *must* involve other traits in the character besides the impulse in question; the whole character has to be such as to permit them. Such behavior cannot stand alone.

Wilson in fact knows this perfectly well and sometimes says it. But he still does not give it enough weight, because he pays too much attention to the mass of speculation on these subjects generated recently by games theorists. (I do not say that this is all useless, but it is only useful so far as we can apply a binary model of winning and losing, pinned to a single, separately heritable characteristic while everything else is held constant, and this does not seem to be very far, with anything more complex than fruit flies.[15]) Thus, though he may not actually make the blunder of which his critics complain, of positing separate, single genes for spite, altruism, and so on, he much too readily accepts the language of those who do treat behavior as a mere mass of discrete units inherited separately.[16] This raises quite

[15] E. B. Ford (*Ecological Genetics* [London, 1964], pp. 9–10) spells out clearly the hopeless weakness of *a priori* genetics, conducted by "mathematicians who sit at home and deduce how evolution ought to work." He points out too the sharp limitations imposed on the work even of laboratory experimenters by what he calls Drosophilosophy—obsession with a few standardized lines of investigation concerning the black-bellied fruit fly. This animal (long the rat of genetics) does have certain advantages from the experimental angle, but Ford is clearly right to stress that no single species should ever be deemed typical in this way and studied in isolation both from comparison and from its own ecology in the wild.

[16] For the great complexity of this situation—the extent to which one gene can influence many characteristics and one characteristic depend on many genes—see, for example, Dobzhansky, *Mankind Evolving*, chap. 2. Failing to stress complexity invites incredulity

unnecessary problems. "Granted a mechanism for sustaining reciprocal altruism," he writes, "we are still left with the problem of how the behavior gets started. Imagine a population in which a Good Samaritan appears for the first time as a rare mutant" (p. 120). The passage goes on to cite calculations of the "critical frequency of the altruist gene, above which the gene will spread explosively. . . . How critical frequencies are attained from scratch remains unknown. Co-operative individuals must play a version of the game of Prisoner's Dilemma."

What is the problem? All the creatures that it makes sense to suppose could develop positive altruism are already caring for their young. And the first element of parental care which develops is defense and rescue from danger. (Even some fish and reptiles do this, as well as the social insects.) All that is needed is to extend the pattern to adults. Now the development of sociability proceeds in any case largely by this extension to other adults of behavior first developed between parents and young—grooming, mouth contact, embracing, protective and submissive gestures, giving food.[17] In fact, wider sociality in its original essence simply *is* the power of adults to treat one another, mutually, as honorary parents and children. It is enriched later with other patterns largely drawn from the interactions *between* infants: hence the enormous importance of play as a source of social sophistication. But quasi-parental interactions come first. They work well because they are adapted to soothe, to conciliate, to forge a bond. Once forged, why should this bond not carry its usual consequence of protectiveness? Besides this, in a stable group, many adults present will actually *be* descendants of the others, and others will be those whom they have

about the central point, which must still remain valid—namely that natural selection does manage to take place in spite of it, and sometimes to produce what appears the most refined adaptation of means to ends.

For the problems concerning "supergenes" (blocks of genetic material locked together by a twist in the DNA) and "switch mechanisms" (genetic factors making possible a kind of jump in evolutionary direction), see Ford, *Ecological Genetics*. The effect of such things can, as he points out, sometimes speed up evolution remarkably. In any case, *a priori* estimates of pace are useless, since, as he says, "it is completely unrealistic to consider the behavior of a large interbreeding community under constant conditions, since that situation, far from being realised, is not even approached in nature!" (p. 33).

[17] This is a dominant theme in both Eibl-Eibesfeldt's *Love and Hate*, and Wickler's *The Sexual Code*.

known as young, so that quasi-parental bonds are already formed. (Friendly behavior to young who are *not* one's own is very common.) Add, too, that those who, from whatever cause, are especially protective and good at rearing young, are likely to leave a disproportionate number of descendants in relation to those actually born. Again, the cry for help which a creature in danger gives is likely to resemble that of a young animal. Even if there has been no selection for the similarity, he suddenly finds himself in a position of infantile helplessness. Why should it not release the same behavior? What need is there of a drastic mutation? What compulsion to play at Prisoner's Dilemma?

That game would, no doubt, well represent the situation of an "altruistic" mutant in a nonsocial species, say a codfish or even a hamster. The effect of his distinctive actions would depend entirely on luck; he could not (for this is the point of the game) count on any understanding cooperation from other players. As Wilson says, "he rescues but is not rescued," so he is not likely to leave very many descendants.

But could there be such a mutant?

If there was, it seems that much more than just a single gene must have mutated. It is not clear how, starting cold, with the emotional equipment of a nonsocial species, he could become endowed at one leap with all the capacities for so complex a performance as a rescue. What is needed for such a change seems to be an alteration in the general programming system, a metamutation that would produce, not just one new behavior pattern (such as cries of warning) but a whole set, adjusted to a new way of life. If such changes are possible (as some have suspected they must be), the fact is of great importance and needs discussion; in particular, it would totally alter the rules of the games-theory game in this area.

What this supposition brings out is the absolute necessity of treating altruism as an aspect of motivation, and therefore as something that makes sense only in the context of a given emotional constitution. It cannot be dealt with if it is defined merely by results. The higher animals—unlike, perhaps, the social insects—have characters, general forms of life typical of their species. Their particular behavior patterns have to fit into these. When a particular species puts something high on its priority system—for instance, dominance, territory, or the care of the young—then it will face danger for that. Many species have at least

one such thing that they are willing to take risks for, though others specialize in running away and living to fight another day. Understanding a particular motive is seeing what set of motives it belongs to, and then what importance that set has in the life of the species. To see how each element can have developed, therefore, we need to think how the set as a whole developed. No doubt for a complex pattern like the care of mammalian young to be formed, many mutations have been necessary, but there are also other factors—for instance, the mere ordinary exchange of genes to form new combinations,[18] better feeding, or worse dangers, the stimulation of new environments, and the effect of other changes in the creature's life. And we have to ask about the effects on group and species survival of the whole set—not just of a particular action. It does not seem likely that the power to rescue others from severe danger makes much difference directly to a species's survival, though it can certainly do so to that of a small group. What does make an enormous difference, however, is the power to form the strong social bonds that make rescue possible. When we consider what is involved in the relation of parent and child in a slow-developing species, and remember that this bond shapes the early experience of every member of that species, further gradual progress to mutual help among adults does not seem too hard to imagine. And in that context, rescue is intelligible enough. It has to be explained not as an isolated gamble, but as part of a whole pattern of motives, which as a whole is advantageous. Explaining it is showing what company it keeps.

Wilson, however, would like to explain such habits in a much more simple and ambitious way—namely, by reconciling them somehow with the Egoist model, which he takes to be self-justifying. He feels it necessary to show altruistic behavior as somehow benefiting the individual who displays it. He goes about this by treating the behavior as contributing to that individual's *genetic fitness*—which means simply and solely his chance of leaving offspring.

[18] "The very lowest estimate of the number of variable genes in human populations will still be in the hundreds, and this is amply sufficient to make the potentially possible gene combinations far more numerous than can ever be realized in people now living or those to be born in the future. The mechanism of the Mendelian recombination of genes thus confers on a living species a capacity to produce a prodigious abundance of ever-new genetic endowments" (Dobzhansky, *Mankind Evolving*, p. 31).

This must be noticed at once as a disastrous move. The notion of fitness has long been a loose cog in arguments about evolution. Wilson greatly increases the strain upon it, without giving it the kind of support that that strain calls for. As it is such a central and widely invoked concept, it seems necessary to make clear the dangers of his treatment.

The trouble with words like "fit" in these discussions is that, if taken in a wide sense they are liable to become vacuous and if taken more narrowly they easily become tendentious. Thus the phrase "survival of the fittest" does not mean much if it means only "survival of those most likely to survive." If on the other hand it means "survival of those whom we should admire most" or the like, it describes a different state of affairs; we shall need different arguments to persuade us that this is happening. In just the same way, Wilson equivocates with the notion that to be "fit" is an advantage to anybody. If it means "healthy" or "able to do what he wants to do" then it usually is so. But if it only means "likely to have many descendants," there is no reason for treating it as an advantage at all. Yet for Wilson the latter meaning quickly becomes, not just *an* advantage, but the only advantage that matters. Thus he defines selfishness as "raising one's own fitness by lowering that of others" (p. 117), this fitness being genetic fitness. Again, of insects that discourage predators by tasting nasty, but can hardly do so without dying in the process, he writes, "it pays to stay around as long as possible to teach predators not to eat one's offspring. In contrast, cryptic saturniids have a short postreproductive life; it does *not* pay to teach predators that one's relatives are good to eat" (p. 125). *Whom* does it pay or not pay? The creature in question is no longer there to be paid. It is wrong to put the perfectly sensible point about the conditions under which a trait can survive in this highly metaphorical way. As we have seen, it can lead to personifying the gene itself, or the DNA, in order to find a candidate for the vacant position of plotter and payee. Just as often, however, the individual itself is nominated for the job of plotting, and that in some fairly surprising species:

> One of the most promising circumstances in which to search for *voluntary* population control is the evolutionary reduction of virulence in parasites. Virulence often (but not always) comes from the capacity to multiply rapidly. Thus the condition is likely to evolve by individual selection. But too high a level of virulence kills off the hosts, perhaps before infection of other hosts is achieved, so that virulence will be opposed by interdemic selection. *It may stretch credulity to think of an altruistic bacterium or self-sacrificing blood-fluke,* but in the sense that feeding ability or reproduction is curtailed in spite of competition from other genotypes, a parasite can be altruistic. [p. 116, my italics]

What is stretched is not credulity—the facts described are perfectly easy both to believe and to understand—but the use of words like "altruistic" and "voluntary."

Another awkward flaw in the scheme is that an individual shares genes with many others who are not his descendants. Since the proliferation of these genes too is supposed to benefit him, the notion of simple genetic fitness is expanded to that of *inclusive fitness*—"the sum of an individual's own fitness plus the sum of all the effects it causes to the related parts of the fitnesses of all its relatives" (p. 118). Thus, the care or rescue of any relative becomes compatible with egoism—provided that the arithmetic works out right: "If only first cousins were benefited ($r = \frac{1}{8}$) the altruist who leaves no offspring would have to multiply a cousin's fitness eightfold; an uncle ($r = \frac{1}{4}$) would have to be advanced fourfold, and so on" (p. 118). But if we are *really* to regard this "inclusive fitness" as representing our own interest, can these bargains be called altruistic at all? The answer seems to vary; thus, on the page before: "Self-sacrifice for the benefit of offspring is altruism in the conventional, but not in the strict genetic sense, because individual fitness is measured by the number of surviving offspring. But self-sacrifice on behalf of second cousins is true altruism at both levels." The mind reels. Genes identical to ours are scattered among all our relatives—indeed, if we move in a small stable group, over the whole population. What is special about the specimens passed through our own progeny? It is not as if we could keep them together in a *block* with which we can at all identify. Human inheritance systems obscure how fast these blocks break up. When that archdynast, Louis XIV, died and left the throne to his great-grandson, one-eighth of his genes remained to help rule his country. Seventy years later, Louis XVI carried about

one-thirty-second of them to the guillotine.[19] Anyone who tries to separate out his own interest as something persisting through these changes has forgotten how utterly dependent his own genes must be, at every stage, on the others that are incorporated with them. Thus, when Wilson proceeds from mentioning the second cousins to say, "and when directed at total strangers such abnegating behavior is so surprising (that is, "noble") as to demand some kind of theoretical explanation" (p. 117), he seems to make a rather arbitrary point. No man is an island; we can never tell whose life has been, or will be, bound up with ours. But apart from local kinship systems, which depend on culture, how could there be different *kinds* of motivation for rescuing first cousins, second cousins, and strangers? Cousins, after all, can *be* strangers; so, indeed, can uncles and brothers. (Wilson touches on the problem of unrecognized relatives on p. 247 of *Sociobiology*, but he does not seem to recognize its explosiveness.[20]) The demand for "some kind of theoretical explanation" arises only from the Egoist position. And from that position, offers of an increase in "inclusive genetic fitness" will be received with proper contempt. The agent himself is getting no sort of advantage at all.

In the second part of this book I have been discussing how we can best set about understanding human nature, granted that (as I argued in Part One) there actually is such a thing. One chief reason why humanely minded people have been unwilling to admit that we do have a nature has been the fear of a crude and unperceptive psychology.

If we want our psychology to be "scientific" in any sense that matters, we must not aim at a superficial likeness to the physical sciences, but at an intelligent adaptation of our concepts to a distinct subject

[19] Or possibly much less, since the half of his genes which a parent passes on to any particular child does not necessarily include a fair quarter from each grandparent, but varies indefinitely.

[20] It is worth noting, for instance, that in such species as the chimpanzee, multiple mating makes it impossible in principle for father and child *ever* to recognize each other as relatives. And in many species, an individual never knows as relatives anyone but his parents and the siblings of the same brood.

matter. And we should not expect to build a single monolithic science, but to use a number of different approaches, intelligently related.

Any such psychology ought indeed, as Wilson suggests, to make full use of the evolutionary perspective as a background. But it is equally necessary that it should be capable of dealing with the foreground—of abandoning the long perspective and looking directly at the motives of individuals. We must therefore take these motives seriously in their own right and not try to reduce them either to neurological phenomena or to behavior patterns. All such reduction must fail because it distorts essential characteristics of its subject matter. It only looks persuasive when it surreptitiously uses ordinary motive-concepts to flesh out its artificial schemes.

Motives have their importance in evolution and their own evolutionary history—but they have also each their own internal point, and it is virtually never a wish to bring about some evolutionary event, such as the maximization of one's own progeny. Confusion between the aims of individuals and the "aims" of evolution—if there can be said to be such things—is ruinous.

I shall move on to the question of how far evolution itself can be said to have aims in the next chapter.