THE ENVIRONMENT FOUNDATION

THE VALUES PLATFORM FOR SUSTAINABILITY

INAUGURAL LECTURE

Wednesday, 14 November 2001, at the Royal Institution
with Sir Geoffrey Chandler, CBE, Trustee, the Environment Foundation, in the Chair

PROFESSOR RICHARD DAWKINS, FRS
Charles Simonyi Professor of the Public Understanding of Science, University of Oxford

Sustainability doesn't come naturally: a Darwinian Perspective on Values

Sir Geoffrey Chandler

The Environment Foundation is a small charity, which has been working quietly with others to put sustainability on the business agenda. We are now taking a major step forward in developing a five year programme, examining the values that will be necessary if we are to have a sustainable world. It will bring together people from different sectors of society and different generations, and with differing views, to probe and question and identify those values. It is a particular pleasure that so many are with us tonight representing so broad a spectrum of the world in which we live – from education, research, business, government, non governmental organisations and the media. Indeed we have teenagers present and at least one highly distinguished nonagenarian.

This occasion is intended as the point of departure for the Foundation's new programme. It is a programme which tackles one of the most fundamental problems of our time – the conflict between the values we hold and the manner in which we behave and, where business is concerned, the conflict between values and financial wealth creation. We believe that action in pursuit of sustainability must essentially address values and attitudes, not simply technology and legislation. We live in a world of our own making, but it is a world that we cannot much like. It suffers from social inequity, economic inequality and accelerating degradation of the physical environment and is manifestly unsustainable, which is why sustainability is now at the top of the agenda. But is what we have our inescapable inheritance? Was the 17th century poet, Fulke Greville, correct when he wrote 'Oh miserable condition of humanity, Born under one law, to another bound'? Are we stuck with what we have got or do we in fact have a choice?

We could have no more appropriate speaker this evening to help us begin to answer those questions than the evolutionary biologist, Professor Richard Dawkins. He is the first holder of the Charles Simonyi Professorship of the Public Understanding of Science at the University of Oxford, but perhaps is more widely known for his books, such as *The Selfish Gene* and *The Blind Watchmaker,* and as one of the most challenging intellects of our time.

Professor Richard Dawkins

"What comes naturally" is a topic which Darwinism might be expected to illuminate. Darwinian natural selection gives us just about everything else in our nature – our bones, our organs, our instincts. If there is a reason to exclude our *values*, it had better be a good one. The values of sustainability are important to all of us here and I enthusiastically include myself. We therefore might hope that these too are built into us by natural selection. I shall tell you today that this is not so. On the contrary, there is something profoundly anti-Darwinian about the very idea of sustainability. But this is not as pessimistic as it sounds. Although we are products of Darwinism, we are not slaves to it. Using the large brains that Darwinian natural selection has given us, it is possible to fashion new values that contradict Darwinian values and that is the policy that I shall urge upon you.

Our starting point must be the fundamental logic of Darwinism itself. Simply stated, everybody has ancestors but not everybody has descendants. We have all inherited genes for being good at becoming an ancestor. Ancestry is the ultimate Darwinian value. In a purely Darwinian world, all other values are subsidiary and, synonymously, gene survival is the ultimate Darwinian value. So, as a first expectation, all animals and plants can be expected to work ceaselessly for the long-term survival of the genes that ride inside them. The world is divided into those for whom the simple logic of this is as clear as daylight and those who, no matter how many times it is explained to them, just don't get it.

Alfred Russell Wallace put the problem in a letter to his co-discoverer of natural selection, "My dear Darwin, I have been so repeatedly struck by the utter inability of numbers of intelligent persons to see clearly, or at all, the self-acting and necessary effects of natural selection". Those who don't get it either assume that there must be some kind of personal agent in the background to do natural selection's choosing; or they wonder why individuals should value the survival of their own genes, rather than, for instance, the survival of their species, or the survival of the ecosystem of which they are a part. After all, say the second group of people, if the species or the ecosystem don't survive, nor will the individual. So it must be in their interest to value the species and the ecosystem. As we shall see, this is faulty reasoning. If only it were true, the values of sustainability would simply be built into us by natural selection. What an appealing thought that would be.

Who decides, then, that gene survival is the ultimate value? Nobody decides, and there is no personal agent doing the choosing. It all just follows automatically from the fact that genes reside in the bodies that they build and are the only things (in the form of coded copies) that reliably persist down the generations. This is the modern version of the point Wallace was making with his apt phrase 'self-acting'. Individuals are not miraculously nor cognitively inspired with values and goals that will guide them in the paths of gene survival. Only the past can have an influence, not the future. Animals behave as if striving for the future values of the genes simply and solely because they bear and are influenced by genes that survived through ancestral generations in the past. Those ancestors that, in their own time, behaved as if they valued whatever was conducive to the future survival of their genes, have bequeathed those very genes to their descendants. So their descendants behave as if they, in their turn, value the future survival of their genes. It is an entirely unpremeditated, self-acting process, which works as long as future conditions are tolerably similar to past conditions. If they are not, the result is often extinction of the species. But it is not differential species extinction itself which constitutes the process of natural selection. If you understand that, then you understand Darwinism in my view. The word Darwinism, by the way, was coined by the ever-generous Wallace.

I am going to continue my Darwinian analysis of values using bones as my example for the moment because they are unlikely to ruffle political or other human hackles. It is not that I mind ruffling hackles per se but, in this case, it would be a distraction and I do mind distractions that get in the way of clarity.

Bones are not perfect. They sometimes break. A wild animal that breaks its leg is unlikely to survive in the harsh competitive world of nature. It will be vulnerable to predators, unable to catch prey, whatever it might be. So, why doesn't natural selection thicken bones so that they never break? We humans, by artificial selection, could breed a race of dogs whose leg bones were so stout that they literally never broke. So why doesn't nature do the same? The answer of course is costs – economic costs – and this implies a system of Darwinian values.

Engineers and architects are never asked to build unbreakable structures, impregnable walls, bridges that can't fall down, trains that can't come off the rails. Instead, the engineer is given a monetary budget and asked to do the best he can according to certain criteria within that constraint. Or he may be told the bridge must bear a weight of some number of tons and must withstand gales three times more forceful than the worst ever recorded in this area. Now go ahead and design the cheapest bridge you can that meets these specifications. Safety factors in engineering imply that we put a monetary value on human life. If we don't like that, that's tough – there's no other way.

Designers of civil airliners are more risk-averse than designers of military aircraft. All aircraft and ground control facilities could be made safer if more money was spent More redundancy could be built into control systems. The number of flying hours demanded of pilots could be increased, and so on. Recent events may make us wish that safety checks on aircraft and security checks on passengers were more stringent and time-consuming, and they have recently become so. The balance has shifted slightly, but there will always be cost constraints on how stringent they can become. We are prepared to pay a lot of money for human safety, but not infinite amounts. Like it or not, we are forced to put monetary value on human life. People who think it somehow wicked to talk about putting monetary value on human life, people who emotionally declare that a single human life has infinite value, are living in cloud cuckoo land.

Darwinian selection, too, optimises within economic limits and can be said to have values in the same sense. My colleague Nicholas Humphrey, continues this argument with another analogy from engineering. Henry Ford, it is said, commissioned a survey of the scrap yards of America to find out if there were parts of the Model T which never failed. His inspectors came back with reports of almost every kind of breakdown – axles, brakes, pistons – all were liable to go wrong. But they drew attention to one notable exception – the king pins of the scrapped cars invariably had years of life left in them. With ruthless logic, Ford concluded that the king pins on the Model T were too good for their job and ordered that, in future, they should be made to an inferior specification. Nature, Humphrey concludes, is surely at least as careful an economist as Henry Ford. Humphrey applied his lesson to the evolution of intelligence, but we can apply it to bones or anything else.

Imagine that we commissioned a survey of the corpses of gibbons, and looked to see whether there are any bones that never break. Suppose we found that every bone in the body breaks at some time or another, with one exception – let's say it's the thigh bone, the femur, which has never been known to break? Henry Ford would be in no doubt – in future, the femur must be made to an inferior specification – and natural selection would agree. Individuals with slightly thinner femurs, who have diverted the material saved into some other purpose, say building up other bones, would survive, or at least reproduce, more successfully. In a machine or an animal, the simplified ideal is that all the parts

should wear out simultaneously. If there is one part that consistently has years of life left in it after the others have worn out, then it is over-designed. Material that went into building it should instead have been diverted to other parts. If there is one part that consistently wears out before everything else, then it is under-designed and should be reinforced, using materials taken away from the over-designed parts. Natural selection will tend to uphold an equilibration rule – rob from strong bones to pay weak ones until all are of equal strength.

I said that that's an over simplification, the reason being that not all the bits of an animal or a machine are equally important. That's why in-flight entertainment systems in airliners go wrong, thankfully, more often than rudders or jet engines. A gibbon, unlike a human, might be able to afford a broken leg better than a broken arm, because its way of life depends on its swinging through the trees. So a gibbon with a broken leg might just survive to have another child, whilst a gibbon with a broken arm probably wouldn't. So the equilibration rule I mentioned has to be tempered – rob from strong bones to pay weak ones, until you have equalised the risks to your survival accruing from breakages in all parts of your skeleton.

But who is the 'you' that we are talking about in giving these instructions? Obviously it is not an individual gibbon. The 'you' is an abstraction. You can think of it as a lineage of gibbons in ancestor-descendant relation to one another, represented by the genes that they share. As the lineage progresses, ancestors whose genes make the right adjustments survive to leave descendants who inherit those correctly equilibrated genes. The genes that we see in the world tend to be the ones that get the balance right, because they have survived through a long line of successful ancestors, who have not suffered the breakage of under-designed bones, nor the waste of over-designed bones.

So much for bones. Now values. We need to establish in Darwinian terms what values are doing for living things. Where bones stiffen limbs, what do values do for their possessors? Having established that the ultimate Darwinian value is gene survival, we are now going to mean something closer to what humans ordinarily mean by values. By values I am going to mean the criteria in the brain by which animals choose how to behave. What are the proximal values in the brain for which animals can be expected to strive, given that the ultimate value is gene survival?

The majority of things in the universe don't actively strive for anything. They just are. I am concerned with the minority that do strive for things, and this minority I shall call value-driven. Some of them are animals and plants, and some are man-made machines – thermostats, heat-seeking missiles. Numerous physiological systems in animals and plants are controlled by negative feedback. There is a target value which is defined in the system. Discrepancies from the target value are sensed and fed back into the system, causing it to change its state in the direction of reducing the discrepancy, until the discrepancy becomes ideally zero. Other value-seeking systems improve with experience. From the point of view of defining values in learning systems, the key concept is reinforcement. Reinforcers are either positive, in which case we call them rewards, or negative punishments. Rewards are states of the world which, when encountered, cause an animal to repeat whatever it recently did; and punishments are the opposite: states of the world which, when encountered, cause an animal to avoid repeating whatever it recently did. The stimuli that animals treat as rewards and punishments are primitive *values*.

Psychologists make a further distinction in primary and secondary reinforcers. Chimpanzees, for example, can learn to work for food as a primary reward, but they will also learn to work for the equivalent of money, which they can then put into slot machines to get food. Some scientists, such as Konrad Lorenz, the grand old man of ethology, have argued that Darwinian natural selection has built

in specific rewarding mechanisms, specified differently and in detail for each species to fit its unique way of life. Lorenz believed, for instance, that squirrels had an appetite not just for food, but an appetite to perform the motor patterns of getting food – of cracking nuts in this case – quite independently of the desire to eat them. He would have said that, for a beaver, the act of building a dam has a rewarding value in itself. The nervous system is pre-equipped with the *value* of liking building dams.

Perhaps the most elaborately surprising examples of primary values of this kind come from bird song. Different species of bird develop their songs in different ways, of course. The American Song Sparrow is a fascinating mixture. Young Song Sparrows brought up completely alone end up singing normal Song Sparrow song. So unlike, say Bullfinches, they don't learn by imitation of other birds, but they do learn. Young Song Sparrows teach themselves to sing by babbling at random and repeating those fragments that sound as a Song Sparrow song ought to sound. There is a template built in of Song Sparrow song genetically specified. You could say that the information of what a Song Sparrow song sounds like is built in by the genes, but note that it is not built in on the motor side. It is not built in as a set of instructions, "Sing like this". It is built in on the sensory side. The instructions are, "Sing at random, until you hear a fragment that sounds like this and then repeat that fragment". So it's like the rat in the skinner box but, unlike the rat, this reward is highly elaborate and highly specific.

It is examples like this that stimulated Lorenz to use the colourful phrase 'innate schoolmarm', or innate teaching mechanism, in his various lengthy attempts to resolve the ancient dispute over nativism versus environmentalism. His point was that, however important learning is, there has to be innate guidance of what to learn. In particular, each species needs to be supplied with its own specifications, its own values, specifying what to treat as rewarding and what punishing. "Primary values", Lorenz was saying, "have to come ultimately from Darwinian natural selection". It should follow that, given enough time, we should be able to breed changed values, by artificial selection of the kind we used to breed, say, bulldogs from wolves. We should be able to breed a race of animals that enjoy pain and hate pleasure. Of course, by the animals' newly evolved definition, this statement is an oxymoron and I have to re-phrase it – by artificial selection, we could reverse the previous definitions of pleasure and pain. The animals so modified would be, of course, less well equipped to survive in the wild than their wild ancestors, just as bulldogs incidentally are for many other reasons. Bulldog puppies can't be born – they need a caesarean section.

Wild ancestors have been naturally selected to enjoy those stimuli most likely to improve their survival. They have been naturally selected to have the right values, the right proximal values to promote their ultimate value of gene survival and, of course, to treat as painful those stimuli most likely to injure them and prevent their surviving. So injury to the body – puncturing the skin, breaking bones – are all perceived as painful, not for arbitrary reasons, but for good Darwinian reasons. Our artificially selected animals in this hypothetical experiment will enjoy having their skin pierced, will actively seek to break their own bones and will bask in a temperature so hot or so cold as to endanger their survival. And similar artificial selection, I venture, would work with humans. Not only could you breed humans with changed tastes, changed primary values, but you could breed for all sorts of things like callousness, sympathy, loyalty, slothfulness, petty meanness or the protestant work ethic. This is a less radical claim then it sounds, because genes don't fix behaviour deterministically. They only contribute quantitatively to statistical tendencies, which are already influenced by many other things. Nor does it imply a single gene for each of these complicated things, any more than the feasibility of breeding race horses implies a single gene for speed. In the absence of artificial breeding, our own values are presumably influenced by natural selection under conditions that prevailed in the Pleistocene of Africa and before.

Humans are unique in many ways and perhaps the most obviously unique feature is language. Whereas eyes have evolved between 40 and 60 times independently around the animal kingdom, language, as far as we know, has evolved only once. Superficially, language seems to be purely learned, but there is strong genetic supervision of the learning process. The particular language we speak is of course learned, but the tendency to learn language, rather than just any old thing, is inherited and evolved specifically in our human line. We inherit evolved rules for grammar. The exact readout of these rules varies from language to language, but their deep structure is laid down by the genes and presumably evolved by natural selection, just like our bones.

Evidence is good that the brain contains a language module, a computational mechanism that actively seeks to learn language, and actively uses grammatical rules to structure it. According to the young and thriving discipline of evolutionary psychology, the language learning module is just an example of a whole set of inherited special-purpose computational modules in the brain – perhaps modules for sex and reproduction; for analysing kinship, which is important for doling out altruism and avoiding incest; for counting debts and policing obligations; for judging fairness and natural justice; perhaps for throwing projectiles accurately towards a target; and for classifying animals and plants. These modules will presumably be mediated by specific built-in values.

If we turn our Darwinian eyes on our modern civilised selves and our predilections – our aesthetic values, our capacity for pleasure, our arts, our philosophies – it is important to wear sophisticated spectacles. Don't ask how a middle manager's ambitions for a bigger desk and a softer office carpet benefit his selfish genes. Ask instead how these urban partialities might stem from a mental module which was selected to do something else in a very different place and time. For office carpet perhaps (and I really mean perhaps) read soft and warm animal skins whose possession betokened hunting success.

A little parable here. We might, on seeing moths flying into candle flames ask, "What is the Darwinian survival value for moths of making burnt offerings of themselves in candle flames?" My point will be that that's the wrong question to ask. Instead we should be asking, "What's the survival value of the kind of nervous mechanism which, when there are candles about, has the effect of guiding moths into them?" A possible solution is this. Lots of insects use rays from distant celestial objects as a compass. You can see why this works because the rays from, say, the moon, the stars, or the sun, are hitting us from infinity. They are therefore parallel, and the rule of thumb in the nervous system that maintains a fixed angle relative to these rays will work, and cause the animal to maintain a fixed compass direction. That presupposes that the object is a celestial object, or at least is at optical infinity. A candle is not at optical infinity. The rays radiate out from a central point and, if you follow that same rule of thumb while maintaining an acute angle to the rays, you will describe a neat logarithmic spiral into the candle flame. So the right way to express the story of the moth and the candle flame is not to ask why they kill themselves, but to ask why they maintain a fixed angle relative to light rays. If you put it like that, and think your way back to a time before candles were invented; before artificial close sources of light at night were invented; back to where any source of light had to be at optical infinity, then you will get the right answer. That's the kind of thing we have to do when asking questions about the evolution of human values.

Why do men want to be rich and powerful? Remember the parable of the moth and the candle. In our society wealth tends, on the whole, not to be translated into genetic success. We have to think our way back to a time when society might have been more like the West African pop singer who has been married 80 times and is married to his entire backing group. In our society wealth more usually buys

things like Rolls Royces, although occasionally it can buy what it primitively used to, which would have been a harem and therefore reproductive success. It's just another illustration of the parable of the moth and the candle.

Evolutionary psychologists have coined the term environment of evolutionary adaptedness, or EEA, for the set of conditions in which our wild ancestors evolved. There is a lot we do not know about the EEA. The fossil records are limited and some of what we guess about it comes from a kind of reverse engineering, from examining ourselves and trying to work out the sort of environment to which our attributes would have been well adapted. We know from fossil evidence that the EEA was located in Africa, probably but not certainly scrubby savannah land. It is plausible that our ancestors lived in these conditions as hunter gatherers, perhaps in something like the way modern hunter gatherer tribes live but, at least in earlier periods, with a less developed technology. We know that fire was tamed more than a million years ago by *Homo erectus.* We know various other things, but not a great deal. Whenever the exodus from Africa happened, and that is controversial, there has evidently been time for humans to adapt to local conditions. Arctic humans are very different from tropical ones, physically as well as culturally. There has been time for biochemistries to diverge in response to diet. Some peoples, perhaps those with herding traditions, retain into adulthood the ability to digest lactose, a sugar found in milk. In other peoples, only children can digest milk, and the adults suffer from an unpleasant condition – lactose intolerance. Presumably such differences have evolved by natural selection quite rapidly in different cultural environments. If natural selection has had time to shape our bodies and our biochemistries since some of us left Africa, it should also have had time to shape our brains and our values over the rather longer time that we consider our ancestors to have lived in the EEA.

Various researchers, notably Gordon Orians of the University of Washington, have been round the world on rather a cushy research assignment, looking at gardens – at what sort of gardens people like – to test the hypothesis that there is some sort of innate specification of the kind of world we like to live in, which is reflected in the gardens we cultivate. Is it something like the EEA? You might guess that an important virtue of a site for our ancestors to live in might have been the presence of water. Maybe this is why everybody loves a stream or pond in their garden, and why so many of us claim to be lulled to sleep by the reassuring sound of running water. There have been studies in which children have been asked to judge which kind of landscape they find most attractive, and Orians at least claimed that very young children are most drawn to East African savannah. I must say that I am a little bit sceptical of the inferences drawn from this, but you can see that it is an interesting kind of approach.

Fear of heights, which is not shown by steeplejacks building skyscrapers in New York, is shown by virtually all of the rest of us. Vertigo and the common dreams of falling might well be natural in species that spend a good deal of their time up trees, as our ancestors did. Fear of snakes and spiders and scorpions might, with benefit, be built into any African species. If you have a nightmare about snakes, it is just possible that you are actually dreaming about snakes, rather than symbolic phalluses. Biologists have often noted that phobias against snakes and scorpions and heights are a lot more common than phobias against electric light sockets, motor cars and guns. Yet, in our temperate and urban world, snakes and spiders on the whole no longer constitute a source of ever-present danger, whereas electric sockets, guns and cars are potentially lethal.

It is notoriously hard to persuade drivers to slow down in a fog, or refrain from tailgating at high speed. The economist Armen Alchian has ingeniously suggested that we should abolish seat belts and instead compulsorily fix a sharp spear to all cars in the middle of the steering wheel, pointing straight at the driver's heart. I think I would find it persuasive, whether or not for atavistic reasons. I also find

intellectually persuasive the following calculation: if a car travelling at 80 miles per hour is abruptly slammed to a complete halt, this is equivalent to hitting the ground after falling from a New York skyscraper. In other words, when you are driving fast, it's exactly as if you were hanging from the top of the Empire State Building by a rope, sufficiently thin that its probability of breaking is equal to the probability that the driver in front of you will do something really stupid. I know almost nobody who could happily sit on a window sill up a skyscraper, and very few who do things like bungy jumping willingly. Yet almost everybody happily drives at high speed along motorways, even when they clearly understand in a cerebral way that the dangers are precisely equivalent. I think it quite plausible that we are genetically programmed to be afraid of heights, but not to be afraid of travelling at high speeds horizontally in wheeled vehicles, because our ancestors would never have met them.

Continuing our guesswork about our ancestors' world – the EEA – there is reason to think that we lived in stable bands, either roving and foraging like modern baboons or, perhaps, more settled in villages like present day hunter gatherers, such as the Yanomami of the Amazon jungle. In either of these cases, stability of grouping in villages or roving bands means that individuals would tend to encounter the same other individuals repeatedly through their lives. Seen through Darwinian eyes, this could have had important consequences for the evolution of our values. In particular it might help us to understand why, from the point of view of our genes, we are so absurdly nice to each other and I shall be referring back to that in a moment.

I now finally want to come to sustainability itself and the values that might encourage it. From a Darwinian point of view, the problem with sustainability is this: sustainability is all about long-term benefits of the world or of the ecosystem at the expense of short-term benefits. Darwinism encourages precisely the opposite values. Short-term genetic benefit is all that matters in a Darwinian world. Superficially, the values that will have been built into us will have been short-term values not long-term ones.

People of goodwill such as, I suspect, everybody in this room, are rightly preoccupied with sustainability, with renewable resources, with taking the side of the future against short-term private gain. Not surprisingly, the rhetoric of such people tends to place nature on a pedestal, where every prospect pleases and only man is vile. For reasons we have just seen, alas, it is not like that, quite the contrary. But as I said at the outset, this is not a reason for despair, nor does it mean that we should cynically abandon the long-term future, gleefully scrap the Kyoto Accords and similar agreements, and get our noses down in the trough of short-term greed. What it does mean is that we must work all the harder for the long-term future, in spite of getting no help from nature, precisely because nature is not on our side.

There is a confusion here with another strand of rhetoric – that of the noble savage. Tribal, so-called primitive, peoples have been thought to be in tune with nature, conserving stocks for the future, taking only what they need, living in harmony with the land, respecting their prey even as they kill them. This rhetoric falls foul of the facts. Unfashionable though it may be to say so, it is looking more and more likely, for example, that the magnificent Pleistocene megafauna of North America died out as a direct consequence of the arrival, perhaps some 13,000 years ago, of hunter gatherers, who had walked across what was then the Bering land bridge. Primitive agriculture too tends to be of the slash and burn variety, which is the very opposite of sustainable, the very opposite of forward-looking.

Humans are no worse than the rest of the animal kingdom. We are no more selfish than any other animals, just rather more effective in our selfishness and therefore more devastating. All animals do what natural selection programmed their ancestors to do, which is to look after the short-term interest

of themselves and their close family, cronies and allies. If any species in the history of life has the possibility of breaking away from short-term Darwinian selfishness and of planning for the distant future, it is our species. We are earth's last best hope, even if we are simultaneously the species most capable in practice of destroying life on the planet. When it comes to taking the long view we are literally unique. No other species is remotely capable of it. If we do not plan for the future, no other species will.

In the 1950s when it was becoming fashionable to worry about over-population and pollution, ecologists talked about prudent predators. Human fisheries, whale fisheries and so on, would ideally protect future stocks by banning, say, small-mesh nets. Wholesale slaughter of whales, at least theoretically, was supposed to be replaced by carefully managed cropping. Those 1950's ecologists thought that wild predators were equally prudent conservationists. They thought wild predators didn't over-hunt their prey. They called them 'prudent predators.' Nobody was suggesting that these prudent predators were consciously or deliberately foresightful, in the way that human conservationists are, or can be. So it had to be done by some kind of natural selection, and the name 'group selection' was used. Those groups or species whose individuals single-mindedly pursued prey stocks to extinction would themselves go extinct. The world would be left with those groups or species whose individuals behaved, albeit unconsciously, in a prudent, conservationist, far sighted, sustainable way. It is a pity, and I wish it were otherwise, but group selection models don't work. Differential group survival obviously happens, in the trivial sense that some groups go extinct and others survive, but there is no evidence that any form of group selection drives evolution. Group selection is based on no coherent theory. The only coherent theory of adaptation we have is the neo-Darwinian theory of differential survival of replicators, usually genes in gene pools. Any other kind of Darwinism, if it is to work at all, must substitute a true replicator for the gene. The 'meme', for example, has been suggested as the cultural analogue of the gene. There could, at least in theory be a meme-based version of Darwinism. Memes, like genes, are true replicators. But I shall say no more about memes today. Groups and species are not replicators.

To see why the idea of prudent predators is theoretically unsound, imagine that a race of prudent predators somehow managed to come into existence. Each individual in the population restrains itself from over-hunting the food supply. It sacrifices its own short-term gain in the interests of a sustainable long-term supply for the species. Now imagine what will happen if a single mutant arises who ignores long-term sustainability and instead goes all out for short-term gain. Whose genes will spread through the population – the genes of the selfish exploiter or the genes of a typical member of the prudent majority? You can see the answer, and mathematical models confirm it. The majority will soon cease to be a majority. In the jargon of our subject, prudent predator is not an evolutionarily stable strategy.

I suppose I should mention here that there is a workable modern theory which calls itself group selection, but it isn't true group selection at all. It is something very different, masquerading under the name group selection. The so-called 'new' group selection is a hamfisted way of re-expressing the well established Darwinian theories of kin selection and reciprocation, which we have had for a long time. We have long understood that natural selection can favour genes that make individuals look after their close kin, who statistically share the same genes, or will look after unrelated individuals with whom they can build up relationships of mutual back scratching. That is not group selection, and it certainly does not provide a satisfactory theory of prudent predators.

There is a tension between short-term individual welfare and long term group welfare or world welfare. If it were left to Darwinism alone there would be no hope. Short-term greed is bound to win. The only hope lies in the unique human capacity to use our big brains with our massive communal database and

our forward simulating imaginations. This is what the Kyoto Accords and similar initiatives are all about. To a Darwinist it is not surprising that it is so hard to get agreement in support. It is not good enough, of course, to just write down a prescription for the future of the world as though we were a benevolent dictator, with the power to make things happen. Alas, we are not a benevolent dictator, and even dictators who start off benevolent seldom remain so. We seem to be stuck with some sort of democracy and we had better make the most of it.

To resolve the tension between short-term and long-term interests is hard. How do you get people – millions and millions of mostly nice (but not overwhelmingly nice), people, somewhat altruistic (but not very altruistic), people – to agree to forgo some of their own short-term gains and do something about the long-term future of the world? As a leader, assuming you do not have dictatorial powers, how do you persuade people and still get elected next time around? Two connected theoretical frameworks are often invoked, known as 'The Tragedy of the Commons' and 'The Prisoner's Dilemma', from which lessons can be drawn. I haven't time to explain these, so must hope that they are sufficiently well known under these names.

The optimal, or more strictly called evolutionarily stable or collectively stable, individual strategies for prospering in a many person Prisoner's Dilemma game have been much studied. Under some conditions a limited form of altruism can prosper in a fundamentally selfish world. It is also interesting to think of mega-strategies that a government might employ for engineering the rules of the game in the right direction – engineering the game in such a way that individual players are more likely to prosper from their own forward looking altruism.

Taxes are a good example. Nobody likes paying taxes, but most of us recognise them as necessary. We pay them as a necessary evil – a tithe on our own short-term selfish gains in the interests of society as a whole, and, we hope the long-term future. Even if we have no children we recognise, as a purely Darwinian machine would not, the long-term desirability of educating the children of our society. We want to live in a nation that educates its young and cares for its old, so we pay our taxes even though we may grumble as we do so. What we find much harder – I speak for myself, but I have never heard anyone dissent – is the thought that we are paying our taxes and somebody else is not. We are deeply indignant at what we perceive as unfairness. I think this sensitivity to unfairness is probably another of the fundamental values built into us primitively. Most of us do not too much mind giving up some selfish benefit for the future benefit of the community, so long as we can be reassured that the system is fair and is being properly enforced, so that others are not getting away with failing to play their part.

The same is true of the Tragedy of the Commons. All the cattle owners know that if too many cattle are placed on the common land, overgrazing will lead to erosion and starvation. All individuals can see that it would be better if they all showed restraint, and rationed their use of the common land. The Tragedy of the Commons is that the benefit of cheating accrues to the individual who does the cheating and him alone, but the cost of cheating is borne by everybody equally, not just the cheat but everybody else too. So, in a world of voluntary restraint and no policing, cheating unfortunately makes sense. If you rely on voluntary rationing somebody will break the convention and in this case put too many cattle out on the common. What honest participants in the tragedy of their commons crave is strong policing to punish cheats. The only alternative is fencing. Divide the land up, so that each individual farmer has his own small plot and that way the costs of overgrazing are borne by the individual overgrazer just as exclusively as the benefits of grazing. This is ultimately why the majority of farmland is fenced and it is, incidentally, why territoriality is so common in the animal kingdom as well. It is the tragedy of the sea and of the atmosphere that they cannot be fenced in this way. So

whales are hunted to extinction. Greenhouse gases are spewed out, to the immediate benefit of the industries doing the spewing, but the costs are shared equally by everybody.

I began by saying that Darwinism was not friendly to the values of sustainability. To the extent that our values stem from the Darwinian selection of our ancestors, this sounds like a pessimistic conclusion. The only solution to the problem of sustainability is long-term foresight, and long-term foresight is something that Darwinian natural selection does not have. I have said that hope lies in a uniquely human capacity for foresight. But how, you might ask, do we manage to have foresight given that we ourselves are products of Darwinian natural selection, which favours only short-term gain? Some people have even complained at what they see as an inconsistency in my position. How can I on the one hand say that we are the products of Darwinian selection of selfish genes, which is incorrigibly shortsighted, yet at the same time say that salvation lies in humanity's capacity for looking far ahead?

The answer lies in the fact that brains, although they are the products of natural selection, follow their own rules, which are different from the rules of natural selection. This is obvious in the case, for example, of contraception. Contraception is clearly anti-Darwinian. It would be hard to imagine anything more anti-Darwinian than contraception. Yet we do it. The brain is big enough to over-ride the genes in this case. The brain exists originally as a device to aid gene survival. The ultimate rationale for the brain's existence, and for its large size in our own species, is like everything else in the natural world, gene survival. As part of this, the brain has been equipped by the natural selection of genes with the power to take its own decisions – decisions based not directly upon the ultimate Darwinian value of gene survival, but upon other more proximal values, such as hedonistic pleasure or something more noble. It was Darwinian selection of genes that built into our brains values such as hedonistic pleasure, orgasm, enjoyment of a sweet taste, or determination to kill oneself in a Jihad – also obviously an un-Darwinian act. It is a manifest fact that the brain – especially the human brain – is well able to over-ride its ultimate programming; well able to dispense with the ultimate value of gene survival and substitute other values. I have used hedonistic pleasure as just an example, but I could also mention more noble values, like a love of poetry, or music, and of course the long-term survival of the planet – and sustainability.

Discussion

Giles Chitty, Independent Financial Adviser: Are there examples of groups of predators who behave in a prudent manner in a way that, perhaps, if they have a maverick among them who is imprudent, they eliminate him? I am thinking strongly of the parallel around Kyoto.

Professor Dawkins: I don't know of any direct examples of that. Something a little like it can happen, at least theoretically, in groups of individuals who know each other as individuals. I think, for example, of pack hunting animals, like wolves, or lions, who are in a certain amount of danger when they attack large prey. One could imagine that some individuals might selfishly hang back and allow the others to bear the brunt of the buffalo's horns. One could also imagine theoretical models in which those who hang back are punished by the other members of the group, noticing that this is going on and driving the shirkers off the kill. But this does not really get to the problem of prudent predation, because we are now talking about a group of animals which know each other and, once you do that, then you are immediately into standard Darwinian theory of reciprocation, which is not to do with long-term altruistic considerations, but can all be handled in short-term language. So, although you probably could find examples which might look superficially like prudence, there are strong theoretical reasons for doubting it.

Phil Clothier, CorpTools UK: What is your definition of a value?

Professor Dawkins: A value is something which is maximised. So, in the case of the ultimate value, what is maximised by all animal and plant behaviour is gene survival. Animals and plants behave as if they had undertaken extremely sophisticated mathematical calculations, in which they are striving to maximise the survival of their genes. That is the ultimate value. More proximally, what animals are maximising are things like a full stomach or an orgasm – something which the nervous system values. The nervous system is pre-equipped with a tendency to value this sensation, whatever it might be, or the sensation of having a nice warm nest – something like that – and this is, from a Darwinian point of view, a proximal value in the service of the ultimate value of gene survival. But the general definition I suppose is something which is maximised.

Jess Kingsford: You say that large brains have been selected through evolution. Do you think that our ability to over-ride our more primal impulses was possibly what was favoured, or what was selected, or is it that what was favoured was the capacity to take the long view?

Professor Dawkins: I think that it was not the capacity to take the long view, in the sense in which all of us here would wish it to be, that was favoured by natural selection. I think that is an emergent property. In the same sort of way, electronic computers were originally built as mathematical calculating engines, and then it was an emergent property that they turned out to be very good at word processing and playing chess and things like that. So the capacity to see into the future would have been a useful thing for the short-term gain, the short-term benefit, of the individual – the capacity to plan a hunt, the capacity to take provision for a drought that's coming, the capacity for storing food for the winter. These are all forward-looking enterprises, but they are all for the selfish gain of the individual. It is that that built into our brains the ability to plan for the future and the ability to plan for the world's future, as opposed to just our own selfish future. That is the emergent property which would never have been directly selected, as such, by Darwinian selection.

Kate Rawles, Philosopher: You talked about the ultimate values that come from Darwinism and then proximal values. What is the relationship between them? What room for manoeuvre have we as humans got and, in particular, are there any constraints on our secondary values that we just can't get?

Professor Dawkins: I speculated that one might breed animals that enjoyed pain, and that would be a fanciful example of changing values. I think your intuition on final constraints is as good as mine. I imagine there are pretty severe limits to what could be achieved by, not necessarily artificial selection, but by training. Could you imagine teaching children to completely reverse normal values? Could you train a group of children to grow up valuing things which are very, very far from what Darwinian selection would have built into them? Imagine bringing up children to kill themselves. That is pushing pretty far away from what Darwinian selection would allow. You are asking how tied are the teachable values to the primary Darwinian values. So, teaching people to kill themselves is pushing it about as far as it can go. In the last couple of months, we have seen disturbing examples where this apparently has been done, so it looks as though, rarely, something like that can happen. I suppose maybe you were asking because of the hope that one might be able to teach people to forgo short-term selfish gain in the interests of long-term world benefit. I am more optimistic about that. There are an awful lot of people who, either for cultural reasons or educational reasons or I don't know quite what, do seem to be capable of subjugating their selfish desires for the good of humanity as a whole, or even living creatures as a whole. The fact that some people seem to manage to do this gives me hope that more people might.

Bob Boote, BTCV: There is so much that is going on today which adds up to evil. What is the value of evil in your context?

Professor Dawkins: I suppose I felt that I did not really need to stress evil because, in a way, many of the things that we call evil do seem to follow more naturally from the Darwinian background. One does not really need to stress that one expects that selfishness, ruthlessness, aggression, riding roughshod over the needs of others weaker than ourselves, are likely to follow from Darwinian natural selection. I suppose I ought to say that, as a passionate Darwinian in the academic sense that I believe Darwinism is the explanation for all of life, I am also a passionate anti-Darwinian when it comes to deriving values for our own life. A pretty good definition of the kind of society in which I don't want to live is a society founded in the principles of Darwinism. That is, in a way, the central message of my lecture.

Michael Quint: You have mentioned the importance of strong policing. Does that not suggest our only hope of going forward is with the much maligned United Nations?

Professor Dawkins: I am naïve about such political matters. I suppose that governments within countries are at least capable, theoretically capable, of the kind of policing I am talking about – making people pay their taxes and suppressing too much manifestation of self interest. When it comes to international interactions, where you do not have world government, organisations such as the one you are citing are the nearest approach we have. It is clear that they are teetering on the edge of being workable, but they do not have the sort of teeth, the sort of powers, that strong governments within countries do.

Dominic Scholfield, People & Planet: I know you have written in the past about the possibility that ideas might develop a Darwinian pattern, using the meme as the unit of cultural transmission that can be replicated. If that is the case – if there is one idea that will survive, is it likely to be sustainability?

Professor Dawkins: The point about memes is that there is nothing special about genes. Darwinism can work with anything which has the property of being a self-replicating entity, which DNA molecules undoubtedly do have. One can theoretically imagine some other things having that property, like computer viruses and perhaps like ideas in a culture. Ideas in a culture may survive in the culture because they have survival value. They have what it takes to survive and if you look around our culture, you see trivial examples, like epidemics, crazes of fashion, games that children play in playgrounds. You are raising the hope that an idea like sustainability might be a good meme and might have a high survival value, in the sense that it would survive, perhaps because a world in which all individuals are imbued with sustainability is a world which is going to continue. Unfortunately that sounds perilously like the group selection argument that I mentioned earlier. One could say the same thing about a species and gene survival. A species, all of whose individuals work for the long-term survival of the species, is more likely to survive than a species whose individuals work for their own selfish good. But it is of the nature of Darwinism that short-term survival is what counts and, if the striving for short-term survival drives the species extinct, that's just too bad. It is too late for natural selection among species, if there were such a thing to come along and save the situation because, by then, the species has already gone extinct. I rather fear the same thing is likely to arise for the meme analogy that you are proposing, but you might come up with an ingenious mathematical model to make it work.

Questioner: How do you work out if you have got to the right level of question rather than the right question?

Professor Dawkins: In the case of the moth and the candle flame, you could first of all check that, given a light source of optical infinity, the moth really does maintain a fixed angle to it. You could then experiment by systematically changing the position of the light to see if the moth changes its own. In other words, see if you can steer the moth just by switching lights on and off. So that would be a test that the moth is actually following that rule. Then I suppose you could test whether the trajectory of the moth in the vicinity of the candle really is a logarithmic spiral, by taking high speed films and analysing that. Let me weaken my position by saying that I am not necessarily saying that any particular 'moth and candle flame' kind of explanation is the right one, but you should be eternally alert to the possibility that the question you are asking is the wrong question. That does not mean you know when you have got the right question. But when somebody challenges you as a Darwinian to explain why people fight over shopping trolleys in Sainsbury's or something, you don't give them a naive answer at the wrong level. You say to yourself, "Moth in candle flame". It is a kind of self-warning.

Questioner: You talked about Darwinism as a framework by which you might be able to understand what conflicts with sustainability. I wonder whether the framework of economics makes better sense? Your example of contraception being rather anti-Darwinian might be very sensible from an economic point of view.

Professor Dawkins: Contraception makes economic sense even from an individual economic point of view. An individual impoverishes himself or herself by having too many children. Yes, economic values – just maximising one's own wealth or any of the other things that economists call utility, whether it is personal wealth or sum of human happiness or whatever it is – all these are values which economists consider might be maximised. Economists, in a way, have an easier time because they are allowed to postulate any kind of utility function, any kind of value that might be maximised, and then look at the consequences. Darwinism is more constrained, in that we know what the fundamental utility function of nature is. It is gene survival. All other utility functions which are not gene survival have to come about as a kind of liberation from the deep Darwinian utility function. But having established that, we can liberate ourselves – and that was one of my central points tonight – we are left with the economists' way of looking at things. What other kinds of utility functions do people maximise and how do they maximise them?

Richard Wilson, Environment Council: You said that people were afraid to enter into international agreements such as the Kyoto Protocol for fear of somehow damaging their own individual position. However, often by entering into dialogue, you actually expose yourself to complex facts and knowledge which will allow you to improve your position. How do you know that you've got the right answer when you don't have all the information?

Professor Dawkins: Uncertainties abound in nature as well. What turns out, as a matter of fact, to be the optimal proximal decision for ultimately maximising the gene survival is never obvious. Animals frequently get it wrong. But the assumption we make is that, in effect, an indefinitely complicated piece of mathematics goes on unconsciously inside the animal. The animal behaves as if it were a very powerful computer which has been programmed by generations of natural selection. Complexity exists in wild nature exactly as it does in the human economic situation. What is more simple in the Darwinian case is that the utility function is known. It is not known how it is maximised. That is extremely complicated, but the utility function is known. In the case of human economic decisions, we don't even know what the utility function is. Different people could be maximising different ones. People could change their minds about what they are maximising. They could have some kind of

curious weighted sum of different utility functions. In Darwinism the practice is just as complicated, but the fundamental value is known.

Rt Rev John Oliver, Bishop of Hereford: I speak on environmental issues on behalf of the Church of England. I am wondering whether there is really such a contrast between altruism and self-interest. Is it not possible to say that because we do have very big brains we can understand that it is essential for gene survival that we must have a sustainable future? And that is actually a very hopeful sign?

Professor Dawkins: I agree with that. In different words that is what I was trying to say. Big brains allow you to take a long distance view of your own self-interest and allow you to take actions which natural selection per se could never have allowed you to do. I would resist any suggestion that that is why natural selection gave us the big brains in the first place. I think it is an emergent spin off from the fact that we have big brains for other reasons. But, as a result, we can actually say my long-term self-interest is different from what a naive Darwinian computer would say it is. My long-term self-interest is to forgo short-term benefits in the interests of long-term benefit and that is a hopeful sign, I agree. However, I would shrink from calling it a legitimate evolution of the Darwinian process, because it might be misunderstood as suggesting that natural selection put it there for that reason – in the same way as natural selection put wings on birds so that they could fly. It is a bit of a different thing. I used the word spin-off just now and I think that is about right. There are precedents for that too, of course. The swim bladders of fish, which are used as flotation devices, started out as lungs, and it was a spin-off benefit that they could be used as flotation devices as well. Nature is rife with such cases. They are called pre-adaptations and I think you could say that what we have here is a pre-adaptation.

Anthony Forsyth: Your perspective on sustainability is obviously founded upon the Darwinian beliefs that you hold. However, in environmental circles, there has historically been a great association with different forms of spirituality – not only people like the Bishop, but a vast range of spiritual beliefs. Yet it is apparent from what you are saying that your Darwinian perspective allows very little room for spiritual beliefs. Do you feel that your Darwinian approach to sustainability in practice would be significantly different from an approach to sustainability allowing for a spiritual perspective? Assuming there is a significant difference, is your Darwinian approach to sustainability likely to be broadly adopted within the environmental community?

Professor Dawkins: I don't think it was possible for you to tell from my lecture what my attitude to spiritual beliefs might be.

Anthony Forsyth: It clearly came across as viewing man as a dichotomy of mind and body, with no possibility of a soul. Whatever spiritual beliefs one might have, much of what you were attributing purely to genes would normally be attributed to genes in conjunction with what one might call a soul.

Professor Dawkins: I don't find that a helpful way of looking at the world and so I am not the right person to answer that question. When I say I don't find it a helpful way of looking at the world, that is putting it very mildly indeed.

Sir Geoffrey Chandler: We are enormously grateful to Professor Dawkins for his admirable lecture and generously thoughtful responses to the questions. I am particularly grateful for his confession that he is passionately anti-Darwinian in the context of what we are trying to do. It gives the Foundation courage in our values programme. This is the beginning of a dialogue and not the end of it. Over the next five years, you will be invited to participate in a series of values assessments. The results will be

posted on our website. We aim to show how our values change through the use of our thankfully large brains during this period. Finally, we work in partnership with others and if there are those present, individually or institutionally, who would like to participate in what we are trying to do, please do make contact.