

The Australian National University



Centre for Resource
and Environmental Studies



The evolution of societal values
compatible with ecological sustainability

Fundamental Questions Paper No.3

A J D Bellett
John Curtin School of Medical Research
Australian National University

1990

FUNDAMENTAL
QUESTIONS PAPER

6QHS40

F86

BOUND SHELE



The evolution of societal values
compatible with ecological sustainability

Fundamental Questions Paper No.3

A J D Bellett
John Curtin School of Medical Research
Australian National University

1990



ISBN: 0 86740 381 0



This paper is the first to appear in the theme area of societal values and ecological sustainability of the Fundamental Questions Program.

Other forthcoming papers in this theme area include:

No.4. Ecological sustainability, deep environmental ethic and Tao:
a preliminary conjunction.

D. Bennett

No.5. Values for sustainability: the necessity of transcendence and sacred realms.

P. Cock

No.6. Environmental management and the political philosophy of trust.

G. McDonnell

No.7. Aborigines, values and the environment.

K. Palmer

No. 8. How green was my Mallee: changing Australian attitudes to their land.

B. Roberts

Work in this theme area was assisted by
financial support from the Howard Norman Trust.

Papers in the Fundamental Questions Paper series arise from Part 2 of the Fundamental Questions Program at CRES. The main outcome of Part 1 of the program is presented in *Our biosphere under threat: ecological realities and Australian opportunities* (S.Boyden, S.Dovers and M.Shirlow, Melbourne: Oxford University Press, 1990).

The Fundamental Questions Program seeks to engender and inform public discussion of the implications for Australian society of the need for long-term ecological sustainability.

The evolution of societal values compatible with ecological sustainability

A.J.D. Bellett

In this paper it is argued that morality began as a necessary component of cooperative social behaviour, during the evolution of humankind as social hunter-gatherers. Analysis of its biological roots suggests that morality is based primarily on reciprocity and mutual benefit in repeated social interactions, and on the innate desire of individuals for social approval and avoidance of disapproval. What is perceived to be of sufficient mutual benefit in a human population to require individual restraint or sacrifice, and thus the pattern of social approval and disapproval on offer, depends on cultural beliefs about reality and the nature of the good life. Although the moral cultures of modern societies are more complex and pluralistic than those of early hunter-gatherers, their fundamental basis remains unchanged. What morality can evolve in a given society is thus constrained both by its cultural history and by common innate human tendencies. Within these boundaries, there is constant moral debate. It is argued that values consistent with ecological sustainability must evolve gradually within a given culture, based on the development of belief that the quality and sustainability of human life depend on the integrity of the natural world. Norms of respect for nature would then evolve, which are perceived within the new cultural context to be in the broad long term interest of the people and their descendants. Abstract ethics that ignore biological and cultural reality are unlikely to be effective vehicles of change. Some environmental issues that need to be faced by our own culture if it is to achieve ecological sustainability are discussed using this framework.

'Enlightened self-interest is, of course, not the loftiest of motives, but those who decry it often substitute, by accident or design, motives which are much worse...' Bertrand Russell (1946).

1. Introduction

1.1 The problems addressed by the Fundamental Questions Program.

Human hunter-gatherer societies in most cases had a sustainable relationship with the very diverse ecosystems to which they belonged. Degradation of the complex relationships of in-

terdependence between land, water, air, microbes, plants and animals that sustains all life occurred in rare hunter-gatherer societies and some early farming societies. Ecological degradation was common in the hinterlands of early cities. However, it did not reach such magnitude that it began to threaten the stability of whole ecosystems and even the biosphere until the fourth, high energy phase of human societies began. Clearly the current levels of energy consumption, waste production and destruction of the very basis of all life including our own cannot continue indefinitely. Much of the problem is due to the sheer scale of human population and technology, but it is also one of attitudes and values.

The overall objective of the Fundamental Questions Program (FQP) is to contribute to community discussion of what the characteris-

tics of societies might be, that would allow them for an indefinite period to satisfy the health and wellbeing needs of people as well as those of the biosphere, and how existing societies could be changed to achieve such sustainability. There could be many different types of sustainable society, and each would be dynamic and changing, not static. However, in the outline of the FQP by the integrating group (Boyden, 1990; Boyden et al., 1990), four common biosocial imperatives have been identified: population must be stable or declining; the use of resources and energy and production of waste must be steady or decreasing (25% of present levels as an initial objective); human health, wellbeing and enjoyment of life must not depend on continuing increases in consumption; and neither must employment in the wider sense.

1.2 Issues to be considered in the values theme

The biosocial imperatives of the FQP imply that to achieve sustainability there have to be large changes in what is seen as of value in society, that is in social beliefs about the human good. In addition, at least two new categories have to be incorporated into the moral community as what Frankena (1985) calls moral patients: future generations of humankind, and other species with which we share the biosphere. Even soil, air, whole ecosystems and the biosphere have to enter our moral discourse at least indirectly, and some believe as moral patients in their own right. From none of these can we expect reciprocal consideration. The FQP includes a commitment to equity in achieving sustainability. We should therefore seek solutions to the problems of sustainability that maintain, and if possible improve, social and distributive justice within the present generations of humankind, while taking account of future generations of humankind, and the long-term stability of the biosphere. Acceptance of these changes is inconsistent with the dominant values and beliefs of contemporary high-energy societies. What the necessary changes in values and norms might be, and how they might be achieved, are the main questions to be tackled by the societal values and ecological sustainability theme area.

There has been much debate between philosophers as to whether the development of

ecologically responsible and sustainable societies requires a radically new ethic, or whether it can be achieved by altering and extending existing ethical or religious principles (Passmore, 1974; Goodpasture and Sayre, 1979; Routley and Routley, 1980; Attfield, 1983). Many have held that extension of ethics to other species requires the recognition of 'intrinsic' values, rather than instrumental values based on usefulness to people. The New Zealand Environment Protection law specifically calls for the recognition of intrinsic values in environmental assessments (Wright, 1988). However, a secure and plausible philosophical basis for intrinsic values has proved elusive in spite of strenuous and creative efforts (See for instance Calicott, 1982, 1985, 1986; Norton, 1986; Weston, 1985; Sober, 1986; Zimmerman, 1988). The most effective treatment is to my mind that of Rolston (1988) who postulates inherent values in other life-forms, that only become operative when they interact with people.

An alternative approach taken by some more radical ecophilosophers is to reject in its entirety mainstream Western philosophy and science, which is seen as the cause of the problem, and to base their biocentric or ecocentric values on Eastern philosophies (Sylvan and Bennett, 1988), pantheism (Zimmerman, 1988), mysticism, intuition and oneness with Gaia, the earth mother (Goldsmith, 1988), ecofeminism and even witchcraft (Salleh, 1989). Another common approach is to look to the values of hunter-gatherer societies for guidance (Calicott, 1983; Goldsmith, 1988). Some of these approaches will be considered by other contributors to the values theme of the FQP (Bennett, 1990; Cock, 1990; Palmer, 1990), and I shall not consider them in any depth unless their arguments impinge directly on those developed in this paper. For a good review of Australian environmental ethics see Bennett & Sylvan (1989). For a critique of ecofeminism from the standpoint of deep ecology see Fox (1989), and for a critique of Gaiaism from the standpoint of ecofeminism see Murphy (1989). My own contribution to the Fundamental Questions Program is in the Western rational/scientific tradition.

1.3 Rationale of this contribution

My first approach to the values necessary to

achieve sustainability, and to environmental ethics, was to use a social contract approach similar to that used by Rawls (1972) in his theory of justice, to develop a biocentric theory of inter-generational and interspecific justice. Several others have tried this approach (Manning, 1981; Elliot, 1984; Singer, 1988). I abandoned it for a number of reasons. Rawls (1972) himself argued that his theory was not applicable to relationships with future generations or other species. Moreover, I came to accept the philosophical objections of Hume (1740a) and others to all social contract theory (Sugden, 1984, 1986). Stronger Kantian versions of Rawlsian environmental ethics (Fuchs, 1981) ask too much of reason and are therefore vulnerable to the standard Humean objection that moral arguments must ultimately be grounded in moral premises that cannot be justified by further reasoning (Hume, 1740a; Harman, 1977). Moreover, there is little evidence that such ethics influence the behaviour of most people in real situations that involve conflict of interest.

Utilitarianism has produced powerful arguments for many conservation values; but it has well known difficulties dealing with human population issues and future generations, and no mechanism at all for dealing with distributive justice and conflicts of interest unless the ideal observer theory is invoked. The ideal observer theory itself, which some have proposed as the sole basis of environmental ethics (Taliaferro, 1988), is vulnerable to Humean objections similar to those that led me to reject social contract theory. Why should people alter their behaviour to conform with the opinions of an unreal hypothetical being; how are we to decide what the ideal observer would dictate in a real situation of moral choice; and how are we to choose between the opinions of different (real) people who recommend different courses of action, each based on what they think the ideal observer would have decided? The more I thought about human evolution, the origins and social functions of morality, and the enormous variation in the religions and ethics of different human cultures, the more doubtful I became of the validity and effectiveness of any formal and supposedly universal ethic. Religions are more effective than secular ethics, but I have no belief in the supernatural, and it is obvious that the beliefs of all of the 100,000 or so different human religions that anthropologists have identified cannot all be valid. It therefore seemed to

me likely that it is the social structure of religions and their use of ritual, dance and music that makes them more effective than formal secular ethics, rather than the validity of their beliefs.

A major objective of this paper is thus a review, and a limited synthesis and extension of, theories that regard morality as being essentially normative and social. It will be argued that what is seen as good, right and just in a human population is largely determined by a set of conventions and norms embedded in the culture of that population, which differ from those of other populations with different cultural histories and different paradigms of reality and the nature of the good life. This view of morality is essentially Humean, but I shall also draw on recent developments in evolutionary biology, economics and rational choice theory to try to define conditions under which moral norms can be expected to arise and be observed. These suggest that there are constraints, both cultural and biological, on what morality can evolve in a given population at a given time. This view of the nature of morality is then applied to the problems of the social evolution of values compatible with ecological sustainability, and of environmental ethics. I believe it suggests constraints on the extension of moral considerability to future generations of humankind, and to other species, which need to be considered if such extensions are to be possible and effective.

Another major theme of this paper is a reappraisal of the role of science in understanding and interacting with the environment. This was prompted by widespread criticism of science in the ecological movement, and by developments in the philosophy and sociology of science. My conclusion is again that there are no logically defensible external standards or methods that allow a population to arrive at empirical truth. Theories are essentially conventions of truth in the population that believes in them, related to observations of reality by members of that population by methods and standards that are normative rather than logically rigorous, and dependent on cultural history and an overall world view or paradigm of reality. Nevertheless methods and standards for judging facts and theories are essential for cooperative interactions with reality, and those that have evolved in Western science in the last 200 years are an integral part of our culture, and have no plausible

rivals within that culture. In this paper I shall argue that whatever philosophical or religious basis we have for our ecological values, in making decisions based on those values about our interactions with other living beings and the environment, we need to use the best available knowledge about ourselves and the biosphere, and about the predicted effects of our decisions. I shall further argue that such knowledge is best derived from empirical sciences, imperfect though they are. Because my approach to morality and values includes a consideration of the evolutionary origins of human social behaviour that is in part based on science, the nature of science will be discussed first.

2. The value of science in the change to sustainability

2.1 Critics of science in the ecological movement

There has been increasing hostility to science in the radical wing of the ecological movement, particularly amongst radical ecophilosophers, deep ecologists and ecofeminists. For instance, Goldsmith (1988) regards 'mainstream natural sciences (biology, ecology and anthropology) as being very seriously misguided'. Bowen (1985) says that 'the feminist and deep ecology movements join in condemning the tradition of patriarchal, exploitative science'; Salleh (1989) states that 'the current global crisis is a consequence of the traditional exclusion of women from patriarchal institutions; the most dangerous of these being science'; and Skolimowski (1988) sees a major goal of ecophilosophy and deep ecology as 'an attempt to replace today's vast scientific empiricist worldview'. Many such critics specifically repudiate major theories of evolutionary biology, scientific ecology and genetics.

There seem to be several reasons for this hostility:

- (i) A belief that science, and technologies derived from it, are principally to blame for the present ecological crisis, and can do nothing to help solve it.
- (ii) A belief that philosophers have discredited all theories of knowledge based on empiri-

cal observations.

- (iii) A belief that radical ecophilosophy or ecofeminism offer a superior way of knowing about and interacting with reality to that provided by empiricism.
- (iv) Transposed hostility irrelevant to either ecological sustainability or the nature and social functions of science.

2.2 The value of science in dealing with ecological problems

I have some sympathy with the first criticism of science and technology. There is no doubt that applied science and technology made possible many of the industrial and agricultural practices that now, combined with sheer human population pressure, threaten the long-term stability of the biosphere. Many scientists give little or no thought to the human value and social implications of their work. In part this is due to a professional ethos that all value considerations must be excluded from scientific work. I shall argue later that it is indeed important on purely pragmatic grounds that scientific hypotheses and theories are judged as far as possible solely on their goodness of fit with observation statements, but that what is investigated and what use is made of the knowledge inevitably involves value judgements, as does the standard of proof required by scientists for the acceptance of theories that have value implications (for instance, that a given technology has destructive effects on human health or the environment). The value assumptions, often unconscious, that underlie planning and applying scientific research therefore need much more thought, and open debate.

In spite of these criticisms, it is both inaccurate and unfair to blame science alone for our predicament. Science is merely a formalised and institutionalised version of methods people must of necessity use, and always have used, simply to live the best they can. The processes of imagination, reasoning, experimental trial, technological application of successful new concepts and cultural transmission that resulted in the replacement of the Acheulian stone tools of *Homo erectus* by the Mousterian tools of neanderthal *H. sapiens* and the Aurignacian of Cro-Magnon (Campbell, 1985) are no different in principle from those that resulted in the re-

placement of the steam engine by the internal combustion and jet engines. Improvements in stone age technology resulted in increased exploitation of the environment and increased population before humankind posed an ecological threat to the whole biosphere. The use of fire and stampede hunting techniques by the Aboriginal populations of Australia and North America probably contributed to the extinction of some species and changed the ecology, and certainly resulted in a release of greenhouse gases and toxic dioxins. Like all living beings, humans have three main objectives: survival, prosperity and reproduction; and we use our evolved capacities for observation, conceptualization and technological application to those ends. Our problem is we have been too successful for our own good, and the good of other species with which we share the biosphere. As well to blame all humankind as only scientists.

In fact many pure and applied scientists were to the forefront not only in recognising environmental problems, but also in the realisation that their solution requires changes in values as well as techniques. Such pioneers of ecological thought include Leopold, Dubos, Carlson, Ehrlich, Commoner, Boyden and Suzuki. Without science we would not even be aware of some of the more pressing global environmental problems, let alone trying to rethink our values as a consequence.

Planning for a more ecologically sound and sustainable future must also involve science. Knowledge based on science is necessary to know when it is safe not to intervene, as well as when and how to intervene if intervention is required; and science will also be needed to monitor and if necessary alter programs of action (or inaction). For instance, if we decide on ecological values to try to save any endangered species, and it seems that the numbers of that species are declining dangerously, we need to know whether the numbers are in fact declining, if so whether the decline threatens the species, if so what is the most likely cause of the decline, from this decide how to reverse it, and then monitor the results of our intervention to see if they are as predicted from our causal hypothesis. In the case of many species of large mammal, for instance, simply protecting the animals and their remaining habitat from human interference is not sufficient (Tudge, 1988). Wider investigations may be necessary to try to

detect patterns of species extinction and threatened extinction (for instance in arid habitats in central Australia), to identify possible causes, and to suggest possible changes in human impact on the region to reverse such trends.

The value of the best possible empirical knowledge in dealing with such situations should be obvious, for without it the best of environmental ethics can have little chance of helping its intended beneficiaries. Such is the criticism of science, however, that I feel it might be useful to illustrate the point with an example. In a National Park in Natal, the numbers of waterbuck recently declined to a dangerous level. The first hypothesis was that this was due to increased predation by lions. However, ecological field studies showed that this was not so. The alternative hypothesis suggested as a result of the investigation was that the decline in waterbuck numbers was due to competition with another antelope, nyala, which was displacing waterbuck on to poorer land that was tick-infested. Waterbuck were succumbing to competition, and to tick-transmitted infectious diseases. When some of the nyala were removed, the waterbuck population recovered (Tudge, 1988). Without a scientific ecological study, the appropriate remedial action would not have been taken.

2.3 Philosophical problems with empiricism

What is demonstratively false implies a contradiction; and what implies a contradiction cannot be conceived. But with regard to any matter of fact, however strong the proof may be from experience, I can always conceive the contrary, tho' I cannot always believe it. (Hume, 1740b).

A statement can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws (Quine, 1961).

Hume (1739) argued that the principle of induction is invalid, and theories cannot be inferred from observations by any logically rigorous process. Russell (1946) restated the Humean arguments very clearly, and Popper (1959) again argued against induction of theories from observations, and held that

theories are universal statements that are produced by an imaginative or intuitive process, and that they can be refuted, but not verified. Developments since Popper made the important step of identifying the potential refutation of theories as the key to understanding scientific method, have further complicated empiricism.

Kuhn (1970) and Lakatos (1974) argued that simple Popperian falsificationism does not adequately deal with the complexities of theory change in the history of science, and that since observations are theory dependent and fallible, conclusive refutation of a theory is not possible. These authors rightly saw science as a social activity, and thought that scientists worked within an accepted 'paradigm' or research program, which was not challenged during normal scientific research. Revolutionary change required a switch to a different program or paradigm. According to Lakatos, this occurs when a program becomes 'degenerate' and ceases to remain coherent and predictive compared with its successor. Kuhn's criteria of a successful paradigm include accuracy of prediction and simplicity, but these are judged by the scientific community; that is they are ultimately based on conventions or norms of specialist scientific populations.

Quine (1961) and Feyerabend (1975 a,b) have published arguments for an even more relativistic view of science. Quine maintained that the distinction between analytical and synthetic statements implied by Hume, Russell, the positivists and Popper could not be maintained logically. Quine's position, although powerful, appeared to be undermined by arguments of Lewis (1969) that analytical statements are true in all possible worlds. Since synthetic statements, because of their empirical content, are clearly not true in all possible worlds, the dichotomy epitomised by the quotation from Hume that heads this section appears to stand. However, I believe there are remaining doubts. Kuhn (1970), although not directly addressing the issue of the analytical/empirical dichotomy, pointed out that a change in scientific paradigm can make statements or concepts that appeared impossible, instead appear empirically possible or even quasi-tautological. Until the development of general relativity theory by Einstein, most people would probably have denied the possible existence of a world in which space was non-Euclidean.

Feyerabend (1975a) argued that no universal

rules or methods for choosing theories can be logically justified, and that if some of the rules proposed by empiricists were always observed, then some of the most important revolutions in science would never have happened. Like Kuhn, he argued that revolutions in science have resulted in replacement of one theory by another that is not commensurable with the first, but went on to claim that no rational means of preferring one to the other is possible. He therefore advocated a proliferation of theories without methodological restraints imposed by rules concerning empirical evidence. Scientists have to be 'lured away from a well-defined, sophisticated and empirically successful system' 'by irrational means such as propaganda, emotion, ad hoc hypotheses, and appeal to prejudices of all kinds' (Feyerabend 1975b). The trouble with this philosophy is that it collapses into the facile Dadaist view of 'anything goes' (Feyerabend 1975a, b), which as Krige (1980) and Chalmers (1982) have pointed out is impotent, because in practice it means 'everything stays'. Radical anti-science philosophers in fact want to reject some of the theories of modern science, and keep others. In a later section I shall consider whether the grounds on which they propose to choose between theories are as valid as those I suggest in the next section, pragmatic empiricism.

2.4 Pragmatic empiricism and the ethics of science

Observation and experience can and must drastically restrict the range of admissible scientific belief, else there would be no science. But they cannot alone determine a particular body of such belief. (Kuhn, 1970).

Each man is given a scientific heritage plus a continuing barrage of sensory stimulation, and the considerations that guide him in warping his scientific heritage to fit his continuing sensory promptings are, *where rational, pragmatic*. (Quine, 1961; my emphasis).

I first want to declare a realist position. I believe, and shall assume, that we are part of a real universe, and that certain things are true of the real universe whether or not anyone knows them, understands them or believes them. The object of science, and of the everyday use of any form of knowledge about reality, is to form con-

cepts that are sufficiently accurate models of some aspect of reality to enable us to predict future events in the light of experience, and to use those predictions to further the human good. The better the model of reality, the more useful the concept will be over a wide range of conditions. Our theories are human social constructs, and we cannot say that they are true; but they represent some aspect of reality well enough to be useful.

The philosophical arguments summarised in the previous section suggested that scientific theories can be refuted, but not proved; and further that refutation is a matter of social judgement rather than logical necessity. However, it is clear from these very arguments that to cling to a theory in the face of 'recalcitrant experience' involves a retreat from what is normally considered rational; it is unlikely that a theory that is inconsistent with reproducible observations is an adequate model of reality, and to believe in and act on such a theory rather than a rival that is consistent with the same set of observations is to lower our chances of understanding and dealing adequately with reality. There are, after all, some things it is more reasonable to believe than to doubt. The acceptance of the collapse of empiricism into Dadaism ('anything goes') faces extreme difficulties; it makes everyday life impossible as well as science. In the context of ecological sustainability it would forbid us, say, to reject on any conceivable evidence (let alone that presently available), the theory that technological progress will enable us to increase human population and energy use indefinitely without harm to the biosphere or ourselves.

The solution is, I believe, that our mental processes evolved like our senses to enable us to survive, prosper and reproduce in the real world; they are functional and pragmatic, not philosophically perfect. Indeed, there is evidence that other mammals share our abilities to recognise patterns of events and reason from them to predict future events in ways that are useful, and even to mentally combine two such sets of observations to predict whether their combined outcome will be beneficial or harmful (Tudge 1988). Be that as it may, we certainly did not invent our mental processes, they evolved, and imperfect as they are, their function is to deal with reality in ways that promote the human good. Empiricism is not perfect, but it usually works, and life without it would be impossible.

At least from the time of *Homo erectus* (over a million years ago), and probably earlier, our ancestors were using cooperative hunting techniques that required coordinated action on the basis of hypotheses inferred from previous experience, about the likely behaviour of the quarry in different situations. During the primaeval phase, and in early farming and urban phases, it is clear that because most of the activities on which human life depended were social and cooperative, and learned from the previous generation rather than discovered *de novo*, there would have been social pressures to accept some versions of reality rather than others. Moreover, the cultural traditions of a group would have had to deal with value and normative issues as well as facts, and these were usually combined, at least in part, in a religion. Since coordinated action was (and still is) of great importance, explanatory theories were probably tolerated in spite of inconsistency with some observations, provided they were reasonably predictive or at least not counterpredictive, were consistent with other concepts of the group's culture, and the culture as a whole promoted behaviour that was adaptive in the environment of the group to which it belonged.

Science emerged as a separate empirical discipline in Europe when social and political changes allowed a degree of cultural pluralism, so that new theories based on experimental observations were freed from the constraints of consistency with the religious culture of the day and appeals to authority. This freedom was not won without cost, and both the Copernican and Darwinian revolutions in science were accompanied by disputes with the religious and political establishments of their times. However, I shall argue later that perceptions of reality, values and norms evolve together, so that complete separation of theories from other aspects of the culture of the society that believes in them may be in the last analysis illusory.

One of the achievements of Kuhn (1970) was, I believe, to recognise that the standards and methods of science are not logical absolutes that lead to the discovery of empirical truth, but normative standards imposed by education and social pressure, that is science has its own ethic which is necessary if it is to be coordinated and able to fulfil its social functions. Modern science, like law, has to make social judgements based on evidence, and has developed methods and codes of behaviour to do so. As a social ac-

tivity, science involves coordination problems, and as argued in section 4.5, coordination problems are solved by the evolution of conventions. The conventions of science required for social choice of theories evolved because they usually work better than other ways of making the necessary judgements, not because they are philosophically rigorous. Experience has shown that errors in judging the validity of theories is seldom due to 'hallucination' or 'amending statements of the kind called logical laws'; but often to the human failings of wishful thinking, inaccurate observation, and protective denial of facts and theories inconsistent with values and beliefs deeply held and/or comfortably familiar. Professional ethics have developed to counter these tendencies.

This view of science is, I believe, consistent with the theory of knowledge and the use of arguments developed by Toulmin (1958). He suggested that science and ethics, together with other fields, have suffered from the unrealisable ambition to show that their conclusions are logically entailed by observations or axioms, by an analogy with mathematics. Once this impossible goal is seen as the holy grail it is, arguments for different points of view within each field can be seen to be supported to different degrees by evidence and principles, within the constraints of the overall paradigm and the methods and standards accepted within the discipline and culture concerned. The conclusions are not certain, but it is necessary to choose between different possible alternatives, and it is more reasonable to accept some than others, based on the quality of the evidence and arguments offered in their support.

Another intuition of Kuhn was that the increasing specialisation of science is necessary for its progression, but inevitably results in the development of esoteric language and concepts, and increasing inaccessibility to the lay population. The use of science in the social construction of models of reality, social decisionmaking about reality and development of new technologies (and I would argue that since it can no longer be claimed that it leads to truth, then science has no other unique social functions) must therefore be based on trust, as discussed by McDonnell (1990) in this series. What appears to be under stress is this trust, not science itself. What is needed is to develop mechanisms for open discussion of the value implications of what is investigated and how any new

knowledge may be used, and of making these responsive to the values required to achieve ecological sustainability. But to return to the situation where consistency of theories with societal values and norms is more important than consistency with careful observations and similarly tested theory is to deliberately prefer less predictive to more predictive models of reality. In the next section I shall argue that this approach should be rejected on both pragmatic and ethical grounds.

2.5 Do radical philosophies offer a better way of knowing?

Kuhn (1970) argued that a mature science can operate only within a socially accepted paradigm or body of theory, and that once having been developed, such a paradigm will not be discarded, in spite of accumulating anomalies, unless it is replaced by a rival that has the promise of resolving those anomalies. This seems also to be true of at least some other branches of enquiry. Radical ecophilosophers have criticised empiricism as a theory of knowledge. What alternatives do they offer, and are they superior to what remains of empiricism after the philosophical and sociological gestalt shifts discussed above?

Bowen (1985), after arguing against an inductivist objectivist view of science, concluded that it was not possible to define clearly the structure of an alternative conceptual framework, but that it must be holistic, democratic, and guided by a socially and environmentally responsible ethic. It is not clear exactly how theories are to be judged in this framework. An exposition of a radical ecofeminist epistemology has been made by Salleh (1989), who overtly rejects logic, and goes on to say that 'Our experience of knowing rests on fusion of consciousness with field, and sensitivity to the impermanence of both, as they shape each other'. The meaning of this statement is to say the least elusive, and how it could be applied to a real ecological situation unclear. Neither of these epistemologies appear to be serious or plausible rivals to empiricism, for all its faults.

Goldsmith gives a clearer statement of his alternative to empiricism. Ecology is, he says, holistic and emotional; 'the generalities of subjective ecological knowledge are subconscious', 'the most fundamental ecological

knowledge is acquired by intuition', and 'the 'truth' of an ecological proposition is the extent with which it fits in the world view of ecology' (Goldsmith 1988). I have no difficulty in acknowledging the role of intuition, for since Popper it has been accepted that theories arise by some such process. But how is their goodness of fit to reality to be judged? Different people have different intuitions, and not all of them approximate to any aspect of reality. Does Goldsmith agree with Feyerabend (whom he cites) that anything goes, and that therefore any statement about reality is as valid as any other? Apparently not, for like many radical ecologists he rejects some of the theories of natural science and accepts others. How is this judgement made? It seems that intuitive concepts must be accepted as a matter of faith, and facts and theories that are inconsistent with belief are rejected. Thus Goldsmith retains the normative aspects of science, but eliminates the requirement to model reality in a way that maximises consistency with observations within the current paradigm. The rejection by Goldsmith and many other radical ecophilosophers of evolution by natural selection and molecular genetics seems to be analogous to similar rejections by some fundamentalist religious sects, and rejection of the very same theories by Lysenko and his followers (who thought the theories were inconsistent with Marxism; Medvedev, 1969).

If, as I believe, empirical judgements are necessary to enable us to interact with reality in ways that are consistent with our values, the deliberate acceptance of a less coherent and predictive theory because its more realistic rival is inconsistent with preconceived ideas and values has unfortunate effects. The accepted theory must be less effective in predicting outcomes than its rival that takes account of observations inconsistent with the preferred theory. The deliberate choice on value grounds of the less realistic theory and its use to interact with reality will therefore have two ethically undesirable effects: some goods will be foregone, and some unexpected bads will occur that are inconsistent with the very values preference for the less realistic theory was designed to protect. A clear example of this is the harm done over three decades to Soviet agriculture and Soviet people by Lysenkoism and its doctrinaire rejection of 'bourgeois racist' genetic and evolutionary theory in favour of allegedly 'Marxist' alternatives that were unrealistic (Medvedev

1969). Of course scientific theories are not 'true', and therefore their predictions of reality are not perfect; but they represent the best that can be achieved with the current paradigm, technology and fabric of argument, and we have at the moment no alternative but to use them. 'It is possible for evidence to be sufficiently strong that it is unethical to ignore it in selecting a course of action' (Crawford-Brown & Pearce, 1989).

3. Evolution of cooperation and the origins of morality

Though Evolution gives no help in discovering what results of our efforts will be best, it does give some help in discovering what it is possible to attain and what are the means of its attainment (Moore, 1903).

One of the problems with the environmental movement has been that while in the main it has rejected the creationist Judaeo-Christian cosmology and myth of genesis, it has also been uncomfortable with its rational scientific Neo-Darwinian replacement, and has tried to function without a coherent and plausible alternative. It seems to me very difficult to develop an ethic of human relationships with the rest of nature, without a paradigm of how we came to exist as part of the natural world, and of our physical and biological relationships with other species. Some of the confusion in environmental ethics stems, I think, from unconscious retention of the creationist paradigm. For instance, rejection as much as acceptance of the 'dominion hypothesis' that the rest of nature was made for human use, makes sense only within a creationist paradigm. If the universe was not created by a moral arbiter for any purpose in which we have a duty to acquiesce, but simply is, and we simply evolved as part of it; then we can choose ourselves whether to use, or not to use, a given component of the natural world; and there is no reason to believe we should necessarily follow any general hypothesis, either permissive or prohibitory, supposed to cover all such choices. We know we have to use other species for food, clothing and shelter in order to survive; but it is becoming clear to us that uncontrolled use of the biosphere is also destructive of ourselves as well as of other

species we value. The solution has to be more complex than accepting or not accepting a general hypothesis of permissivity of use. The questions then arise: why did we evolve the capacity for moral choice, what is its basis, and how should we make wise judgements about our relationships with other species within that framework?

I start from a belief in the Neo-Darwinian account of genesis, and of our physical and biological relationships with the rest of the biosphere of this planet. This is, I believe, consistent with the conceptual basis of the Fundamental Questions program (Boyden, 1990; Boyden et al., 1990). Because of the philosophy of science developed in section 2, I do not claim that this leads to a unique 'true', 'objective' or 'scientific' view of human nature and human relationships of the kind assumed by sociobiologists (Wilson, 1975; Alexander, 1987), nor do I deny the importance of cultural factors and conscious thought in ethical aspects of human social behaviour. I am, however, suggesting that a biohistorical perspective (Boyden, 1987, 1990) on human sociality and moral culture, based on the Neo-Darwinian paradigm of natural and sexual selection of random mutations in the genes that are the basis of our inherited traits, may help us to understand our current predicament and the steps we need to take to improve our relationship with the rest of the biosphere. Indeed, without some sort of paradigm of the origin of humankind and our place in nature, it is probably going to be difficult to make progress in environmental ethics. The evidence in favour of Neo-Darwinian theory is very much stronger than that for its predecessor, the Lamarckian theory of inheritance of acquired characters, preferred by some deep ecologists (Goldsmith, 1988). Most professional biologists believe that Lamarckism has been conclusively refuted. Some, however, including my friend Ted Steele, believe that some phenomena that are difficult to account for by the Darwinian paradigm may be due to Lamarckian mechanisms coupled with natural selection (Steele, 1989, 1990), but most believe that they will be explained within the Neo-Darwinian paradigm by what Kuhn (1970) would call normal science, albeit with some minor changes in theory. Certainly the anomalies have produced no crisis in the dominant paradigm. The Darwinian paradigm is probably not, of course, complete and perfect, but neither is any

other theory. It is, in my opinion, the best theory of human genesis we have. Like Ted Steele (1990), I am conscious of wonder and awe when I watch and try to understand living creatures; but my belief that their evolution occurred essentially by natural selective forces acting on random variation in no way diminishes my wonder, rather the opposite.

Belief that biological evolution is somehow directed or goal seeking, which some deep ecologists think, appears to be an example of cryptic creationism; indeed its foremost proponent, Teilhard de Chardin, attempted to combine the Judaeo-Christian and Darwinian paradigms in an obscure and mystical, albeit poetic, philosophy (Teilhard de Chardin, 1965). The combination by Serafin (1988) of Teilhard de Chardin's belief in evolution towards a spiritual noosphere compatible with Christianity, with the hypothesis of Gaia the Earth Mother (Lovelock, 1979), further confuses cosmology. None of these rivals to the Neo-Darwinian paradigm seem to me to be supported by sufficient evidence to be taken seriously. I would claim, however, that evidence for the Neo-Darwinian view is ethically sufficient in the sense used by Crawford-Brown and Pearce (1989), that is it should be taken into account when planning ethically significant actions to which the theory may be relevant. The effects of application of Lamarckian 'Marxist' evolutionary and genetic theory by Lysenko (Medvedev, 1969), suggest that evidence for its failure, at least in that form, is also ethically sufficient. As it happens, the view of human nature and morality that is developing as a result of the application of the Neo-Darwinian scientific paradigm to hominid social evolution, is consistent with the school of thought in Western philosophy than began before Darwin with the work of Locke and particularly Hume and is now included under the umbrella term of rational choice theory. Evolutionary theory enriches understanding of, and explains the origin of, the human characteristics on which Hume based his moral theory; evolution basically suggests limits to what is feasible in morality, and defines conditions under which what is possible is most likely to occur. These uses of evolutionary theory are valid according to Moore (1903), unlike attempts to derive ethics directly from evolution in the style of Spencer, the social Darwinists, and the sociobiologists. But I believe Moore nevertheless underestimated the

relevance of understanding of human evolution in the development of theories of the good. Any theory of the human good that ignores the biologically evolved needs, desires, drives and limitations of humankind is bound to be unrealistic.

Ethical codes are required for life in cooperative groups consisting of individuals whose genetic inheritance and interests are not identical, and whose behaviour is not entirely pre-programmed and unconscious. All human behaviour results from a mixture of evolved innate foundations, cultural traditions, and varying individual and spontaneous components. Since Darwin (1859, 1883) it has become increasingly clear that humans evolved from primates that were already social. Our nearest contemporary relatives, chimpanzees, have a complex social life that includes exchanges of favours, incest taboos, limited cooperative hunting and sharing of meat, and sophisticated politics (Campbell, 1985; Tudge, 1988; Barnes, 1990). Of course human social behaviour will always be modified by culture, experience and conscious thought; but without an innate biological tendency towards cooperative group living, it is most unlikely that a moral culture would have developed (Alexander 1987). We do not each take a conscious and intentional decision to live in social groups, it is natural to humankind. Although I agree with deep ecologists and ecofeminists that one of the keys to changing environmental behaviour is cooperation, an assumption that cooperation is based solely on benign emotion and can be extended indiscriminately is simplistic and misleading. Much of this paper will be concerned with attempts to understand, within the Neo-Darwinian paradigm, how cooperation arose; and under what conditions, in which situations, and between which participants, cooperative interactions can flourish and be stable.

Since evolution is driven by competition between individuals for reproductive success, how cooperative behaviour could arise has been an important problem in evolutionary biology. Any individual that took account of and acted on the interests of others at cost to herself would promptly be taken advantage of, and lose evolutionary fitness. Of course, deep ecologists and social ecologists believe that cooperation rather than competition is the usual relationship in nature (Bookchin, 1987; Goldsmith 1988). I would refer the reader to any textbook of biology for a

contrary view (see also *Scientific American* 239 No. 3, September 1978; Darwin, 1859; Williams, 1966; Campbell, 1985; May, 1978; Tudge, 1988).

3.1 For the good of the group?

Cooperation was not at first seen as a problem in evolutionary theory, because it was thought that natural selection could operate at the level of the group or even species, and cooperative groups would have a selective advantage over non-cooperative groups. The theory of group selection is now thought to be wrong; its main problem is that it fails to account for how cooperation (or any other trait) could arise within a group in the first place. Williams (1966) compared the requirements for selection of groups that acquire a character that is advantageous to the group but disadvantageous to the individual by some random genetic process, with the usual process of genic selection, that is the natural selection of the more advantageous alternative form of a gene in a population of individuals. The conditions under which group selection could occur are exceedingly rigorous and improbable, and Williams (1966) argued that 'one should assume the adequacy of the simplest form of natural selection unless the evidence clearly shows that this theory does not suffice.' After considering a number of examples of apparent social adaptations that had previously been claimed to be explicable only by group selection, Williams concluded that they were either illusory, or could be explained by genic selection. This is consistent with the conclusion of molecular genetics that what selection acts on is the frequency of particular genes in a population of individuals (Dawkins 1976). One of Williams' examples is relevant to the Gaia hypothesis (Lovelock, 1979): the idea that ecosystems, and perhaps the whole biosphere, can be considered to be adaptive units that have goals like those of organisms. Williams (1966) concluded that they cannot. Cahen (1988) uses similar arguments against the grounding of the moral considerability of ecosystems in their supposed interests or goals, and an analogy with an organism. It seems, then, that we cannot explain human cooperation or ecosystem stability by selection for group interests or goods, but rather we have to explain them in terms of 'the interplay of biological relations that act to confer specific advantages or disadvantages on individual

organisms' (May, 1978).

3.2 Kin selection for altruism

One theory as to how cooperation could originate in a population of noncooperators is that cooperation consists of exchanges of altruism (acts that involve cost to the altruist and benefit to the recipient) between close relatives. Such a mechanism, known as kin selection, has been postulated by Hamilton, (1964), Williams (1966), Trivers (1971) and Dawkins (1976). Williams' (1966), analysis suggested that in all instances the altruist would be at a selective disadvantage, but the frequency of the gene for altruism could still increase in the group providing the behaviour were confined to siblings, and the cost to the altruist was much less than the benefit to the recipient. Evolutionary theory has increasingly taken a gene's eye view of selection (Dawkins, 1976), and the concept of Darwinian fitness has been extended to include not only the reproductive success of an individual and its offspring, but the reproductive success of all individuals in a social group that carry the same genes. This theory is particularly appropriate to evolution of cooperation in social insects, that have unusual genetic systems that result in very close relatedness of the individuals that cooperate, and there have been a number of experimental studies consistent with it. Most altruism and cooperation in nature occurs between closely related individuals; but there are exceptions such as symbiosis between unrelated species.

Can kin selection explain the evolution of cooperation in the ancestors of humans (known collectively as hominids)? It seems likely that love of a mother for her infant, caring behaviour towards her infant, and to a lesser extent altruism by women and men towards sisters and brothers, and similar but weaker relationships with more distant relatives, are all consistent with the theory. The theory explains the common tendency towards nepotism and preferential treatment of relatives in all human societies. But there are a number of reasons for thinking that it cannot alone explain human cooperativity. Most cooperative relationships in primaevial bands of hominid hunter-gatherers were probably between individuals not sufficiently closely related genetically for kin selection to be effective, and that is certainly the case in modern human societies. Incest taboos and

exogamy would make it unusual (although not impossible) for individuals to be closely related in the cooperative relationships of parenting, and food gathering. Some of the 'kinship' so prominent in hunter-gatherer societies is formal or even fictitious rather than genetic, and there is little evidence that humans can automatically and accurately recognise close genetic relationship independent of cultural context. Moreover, kin selection in the evolution of cooperative behaviour in birds and primates is now thought to be less important than originally supposed (Axelrod, 1984). While none of these difficulties for the kin selection theory is insuperable, they do suggest that it may not be the sole, or even the most important, factor in the evolution of human cooperativity. I shall assume that it can account for some human altruism, particularly of the nepotistic sort, and some weaker caring and benevolence towards more distantly related individuals.

3.3 Selection for altruism by sexual choice

It has been argued that evolution from promiscuity through one-male groups (polygyny) towards monogamy was necessary to reduce between-male sexual competition in the evolution of hominid cooperativity (Campbell, 1985; Alexander, 1987). In many primates males show little altruism towards females and their infants. Amongst others, however, friendly and altruistic behaviour towards a particular female and her infant is a better predictor of sexual access than position in the male dominance hierarchy (Campbell, 1985). Altruistic behaviour by a male towards a female would in such a situation increase his reproductive fitness, while the reproductive fitness of the female would be increased by male help in provisioning and care of her infants. Cooperative behaviour between parents could therefore be selected by sexual choice, that is choice by females of altruistic males and vice versa (Darwin, 1883; Parker, 1987). Two of the earliest events in hominid evolution, before the expansion of the brain, were the evolution of bipedalism and reduction of canine teeth, probably concomitant with the evolution of the carrying of nuptial gifts of scavenged meat by males (Parker, 1987) and of infants by females (Lovejoy, 1981; Campbell, 1985), and of the smile as a friendly sig-

nal(Parker, 1987). Foley and Lee (1989) presented arguments that only a limited number of evolutionary pathways were possible in hominid evolution, based on the main biological parameters of primate social organisation: solitary life or association of each sex with kin, lineage or non-kin of like sex; and the stability of association between males and females. Evolutionary constraints on pathways between different social states suggest a trend towards longer and more stable relationships between males and females. In this evolution, the increased reproductive fitness of each partner due to their cooperation would be threatened by the temptation to cheat by desertion or adultery. There is a similar problem in understanding the evolution of any cooperative relationship, that will be explored in the next section.

3.4 Conditions for the evolution of in-group cooperation based on reciprocity: the two-person prisoner's dilemma

An alternative theory that does not assume (or exclude) genetic relatedness or a sexual relationship between the interacting individuals, is based on the evolution of a stable pattern of reciprocity in a series of interactions between them. This theory does not require that there be specific genes for specific behaviours; the patterns can be learned or culturally determined. It assumes that the individuals that develop a cooperative relationship are pursuing their own interests, but not that those interests are necessarily or exclusively reproductive.

Initially, when considering the biological origins of cooperation, it is clear that cooperative behaviour would not have begun, spread, and been able to resist challenge by cheaters unless it contributed directly or indirectly to the reproductive fitness of individuals. But once a tendency towards reciprocal cooperation had evolved biologically, further cooperative interactions could arise from innate, spontaneous, or culturally determined behaviour patterns, emotional drives, or from a mixture of all as seems most likely for any human behaviour pattern. The advantage of the theory is that since it is compatible with both biological and cultural mechanisms, it is not necessary to draw a

precise line between them.

I shall assume that a tendency towards increased cooperativity began in bands of hominids numbering 30 to 40 individuals, that moved out into the savannah as the climate and ecology of Africa changed about 5 million years ago, and became omnivorous, and therefore more dependent on gathering seeds and nuts and scavenging meat than their forest ancestors (Campbell, 1985). Such a group will be called an in group. The requirements for the evolution of cooperation required by the theory are spontaneous exploratory social and environmental behaviour, the ability to recognise and remember other individuals, to distinguish between beneficial and harmful outcomes of interactions with them, to remember the outcomes of different interactions with each other individual, the ability to learn to preferentially repeat those interactions with those individuals that resulted more often in favourable outcomes, and to imitate successful interactions by others. Many of these abilities are also required for the recognition of patterns of events in the environment, and their use to predict future events in a way that furthers the human good, which I suggested were the original basis of understanding and exploitation of the environment.

In theories of cooperation based on reciprocity, the participants (usually called players because the formal representations of situations is based on game theory) are presumed to be acting solely out of self-interest, and altruism and conscience are ignored. This therefore represents a worst-case scenario for the evolution of morality, but if it is possible to define conditions for stable cooperation on this assumption, then altruism, emotional bonds and conscience may then make cooperation more certain under those conditions, and possible in more demanding situations. Most rational choice theorists however, appear to believe that self interest is the only human ethical motive, while others take this merely as a simplifying assumption.

The basic tensions in simple cooperative interactions are exemplified by the two person prisoner's dilemma shown in Table 1 (Axelrod and Hamilton, 1981; Axelrod, 1984).

Each player does better if both cooperate rather than both defect, since the reward for cooperating ($R=3$) is better than the punishment for mutual defection ($P=1$). The dilemma arises because each could do even better by defecting

Table 1. The two person prisoner's dilemma.

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	R=3, R=3	S=0, T=5
	Defect	T=5, S=0	P=1, P=1

(T > R > P > S)

The score for player 1 (row) is shown to the left of that for player 2 (column)

when the other cooperates, scoring the prize of the temptation to defect (T=5) and forcing the other player to accept the sucker's payoff (S=0). Many potential cooperative problems are of this type (Axelrod, 1984). The payoff matrix can be varied, but to be a prisoner's dilemma $T > R > P > S$. If a game is played only once, to defect is the preferred option for both players, for if the opponent cooperates the score will be T=5 rather than R=3, and if the opponent defects, the score will be P=1 rather than S=0. In this case, defection (D) is said to be dominant to cooperation (C).

Deep ecology, social ecology and ecofeminism all seem to accept that evolution occurs, but reject natural selection as the mechanism without proposing a viable alternative, and claim that the natural state of interactions between individuals is cooperation, that is that C is dominant to D. This would be true only in a world in which there was no temptation, or one in which everyone always resisted temptation, or one peopled by masochists who preferred other people to win at their expense rather than to win themselves. In a world such as our own, even if universal, unconditional and unprovokable cooperation were operative, it could easily be invaded by players who always defected (ALL D can invade ALL C), since in every encounter ALL D would score 5 and ALL C 0. ALL C is an exploitable strategy, and is not evolutionarily stable.

If a prisoner's dilemma is repeated a known number of times, each player's dominant strategy is again to defect (Axelrod, 1984; May, 1987). However, if the same two players are to meet an indefinite number of times, so that they cannot be sure when their last interaction will occur, it is possible for a pattern of cooperation to evolve under some conditions. The critical factor is the weight (w) of the next move rela-

tive to the current move, which is less than one for two reasons: the probability that the two players will meet again is less than one, and the value of a future payoff is normally perceived as less than that of one immediately enjoyed (for this reason w is also known as the discount factor). It can be shown that if w is sufficiently high, there is no best strategy independent of the strategy used by the other player. To get over this difficulty, Axelrod (Axelrod and Hamilton, 1981; Axelrod, 1984) organised two series of computer tournaments in which strategies submitted by people from several relevant disciplines were played against each other. Strategies that were successful in eliciting cooperation had some of the following characteristics: they were conditional in that they were dependent on the moves made by the opponent, brave in that they cooperated without waiting for the opponent to cooperate; nice in that they were never the first to defect; provokable in that they punished defection by the opponent by defecting themselves on subsequent move(s); and forgiving in that they cooperated again after punishing defection by the opponent. The most successful strategy, tit-for-tat, has all these characteristics and was also one of the simplest submitted; it cooperates on the first move, and then repeats each move made by the opponent. In computer simulations in which the representation of each cooperative strategy in each generation was proportional to its success in the previous generation, tit-for-tat again emerged as the most successful, and maintained that position (Axelrod, 1984). There is evidence that the evolution of cooperation in interactions between individual fish and birds follows the tit-for-tat pattern, and that it often evolves in laboratory non-zero sum games between people (Axelrod, 1984; May, 1987). Its most convincing example in real human interactions, the evolution of *de facto* truces in trench warfare during world war one (Axelrod, 1984) is a special case of an N person prisoner's dilemma, and will be considered in section 4.1. However, it is interesting that the social evolution of reciprocal cooperation in this case was followed by that of concern for the opponent's welfare, and cultural transmission of the successful pattern of behaviour. Even tit-for-tat would be unable to establish cooperation in a population of ALL D players if it were confined to lone individuals. However, if it is considered to arise as spontaneous social behaviour amongst individuals with the charac-

teristics I suggested above as necessary, and which were probably common in early hominids, rather than as a specific behaviour determined by a specific gene, then clusters of individuals using the tit-for-tat strategy could arise by chance and have a selective advantage. Cultural transmission and broadly based genetic evolution of the required characteristics could then follow. May (1987) has calculated what value of w is required to make sure that a population using tit-for-tat as the basis of cooperativity could not be invaded and exploited by individuals using ALL D. The requirement is:

$$w > (T - R) / (T - P).$$

This result shows that the evolution of cooperation based on reciprocity depends on the value of w as well as on the payoff matrix. Cooperation is favoured by conditions where w is large, that is when players are likely to meet each other time after time for an indefinite period, and build up a cooperative personal relationship. These conditions are exactly those in a hominid in-group, in which altruistic relationships are also likely to have evolved biologically by kinship and sexual selection. Once hominid evolution was dependent to an important extent on in-group cooperation, the necessary adaptive social behaviour would be reinforced by evolution of the emotional responses, under appropriate conditions, of trust and suspicion, gratitude and resentment, and friendship and enmity. Further key events in the development of in-group morality would be the evolution of an increased emotional requirement for human company, and the desire for social approval, and avoidance of disapproval (Boyden, 1987).

4. Social evolution of moral behaviour

In this section I shall consider the evolution of collective action and the provision of public goods once agriculture and urban life began to replace life in hunter-gatherer in-groups. For the first time, conflicts of interest over the ownership of property and surplus material goods arose, and it was necessary to interact daily with strangers. Some of the problems were intractable and required the development of institutions and governments to impose order when it could not be achieved spontaneously. Most of the social

structures that evolved were, of course, inequitable and exploitative. I shall assume that the most efficient and equitable system for maintaining non-spontaneous order is representational democracy, and that order cannot be achieved and maintained without a government, that is that anarchism does not work. No government, and particularly no democratically elected government, can indefinitely and without threat to its power impose a policy that is against important interests and wishes of the majority of the governed. Whether desirable changes in its policies are government initiatives, or result from reaction to public pressure, they are unlikely to occur unless a substantial body of the governed are united in advocating and cooperating to implement such changes. Understanding change in a democracy therefore requires amongst other things an understanding of how spontaneous changes in values and conduct arise in small groups and spread through society.

The view of human nature and of the origins of cooperative and moral behaviour suggested by evolutionary biology is remarkably consistent with the philosophy of David Hume (1740a). Hume argued that people are motivated primarily by self-interest, and prefer an immediate to a delayed satisfaction of their desires. He showed that moral principles cannot be derived solely from reason, or from empirical knowledge, but must ultimately be grounded in 'moral sentiment' that cannot be justified by further moral arguments. Hume believed that people have altruistic feelings and can behave altruistically, but that this is mainly confined to family and friends, and further extension is capricious and unreliable. The order and moral behaviour required for wider social life is based on conventions that gradually evolve, under conditions where personal share in the common advantage is weighed against personal cost. Hume identified the key problem in collective action based on such a balance; while cooperation can emerge between two people or in a small group, it becomes increasingly difficult to achieve as the size of the group increases, because there is an increasing temptation to attempt to share in the benefits without contributing to the costs.

4.1 The N person prisoner's dilemma as a model of collective action

In an N person prisoners' dilemma, each of N players has to choose between contributing to some public good (say, by not using some ecologically damaging but desirable product), or defecting (by using the product). As in the two-person dilemma, defection is dominant to cooperation. This requirement may be too strong, as there is empirical evidence that about 50% of modern humans in laboratory situations fail to exploit an opponent playing unconditional cooperation (Hardin, 1971), but there are difficulties about the interpretation of this result, and for the moment I shall assume the more conservative restriction that $D_i > C_i$ for each player i . An equilibrium in such a situation is a combination in which no one would have been better off had s/he alone acted otherwise, given the actions of the others. The outcome where all cooperate is better for each player than the outcome where all defect, but as in the two person case, unconditional ALL D is an equilibrium, but unconditional ALL C is not; if cooperation is to be stable, some players must use strategies of conditional cooperation. Suppose T_n is a tit-for-tat strategy, in which a player using T_n chooses C in the first round, and thereafter chooses C if, and only if, at least n players chose C in the previous round. If $n = N - 1$ the strategy is conditional on cooperation by all other players. If $f(v)$ is the payoff to a player who chooses C when v others also choose C, and $g(v)$ the payoff to a player who chooses D when v others choose C, universal T_n is an equilibrium if, and only if,

$$w_i > g(N-1) - f(N-1) / g(N-1) - g(0)$$

for each player i (Taylor, 1987). This equilibrium is weak if N is large. Olson (1965) has argued that if N is such that no individual contribution makes a perceptible difference to that of the group as a whole, then 'a collective good will not be provided unless there is coercion or some outside inducements'. An assumption in this equilibrium is that all players are perfect in the execution of their strategy; on more realistic assumptions, a mistake by any player in any round (either accidental defection, or the mistaken interpretation of a cooperative move by another player as a defection) will destroy cooperation permanently.

Moreover, invasion by a single individual using any strategy that includes unprovoked defection in any round will also destroy cooperation.

In spite of this, cooperation in a real N person prisoners' dilemma can be stable. Axelrod (1984) cites the case of the 'live and let live' system in trench warfare in world war 1, in which soldiers shot to avoid causing casualties or damage, and did not shoot at all at certain times unless the enemy did so, in which case the fire was returned. What evolved was a tit-for-tat strategy to reduce casualties on both sides, although either side could have exploited the de facto truce to obtain an advantage in the war of attrition in the trenches. The reason cooperation was stable, although thousands of men were involved, was because small groups faced each other for long periods of time at any particular point on the front, that is the N person dilemma was reduced because of the situation into constituent local dilemmas in which w was large and N small. Generalising from the two person case (Axelrod, 1984), cooperation in an N person prisoners' dilemma can be promoted by 'enlarging the shadow of the future' (increasing w), changing the payoff matrix so that cooperation is more rewarding and temptation less so, and breaking the problem down to subunits in which w is large and N small.

It has been suggested that an N person collective action problem can be considered as a social choice problem that is in some conditions reduces to a prisoners dilemma (Hardin, 1971). If the problem of choice facing N participants is viewed as a matrix of 2^N possible outcomes (for instance, there are 1024 for 10 participants), then only those outcomes in which the ratio of benefits to costs is more than 1 for each player are realizable. Hardin showed that there is a Condorcet choice (by definition unique) amongst the realizable outcomes, that in which everyone contributes. He further argued that it would therefore be rational to use sanctions against those who did not contribute. This begs the question in two ways: firstly in assuming that the situation would be accepted by the participants as a social rather than individual choice, and secondly in assuming that the mere existence of a Condorcet choice makes the use of external sanctions rational. Indeed, one of the problems in working towards sustainability, for instance, seems to me one of getting the issues on to the democratic social choice agenda, and I believe that premature and unconsidered use

of sanctions is likely to be counterproductive. Moreover, there is a logical problem with the use of sanctions in such a situation; it involves costs to the sanction imposers (both material and social), which they would each prefer to avoid while enjoying the benefits of the provided public good, and therefore causes a nested infinite regress of prisoners' dilemmas within the first. There are, of course, situations in which sanctions are necessary, justified and effective, but they are not a simple, universal resolution of every N person prisoners' dilemma. What other ways are there of coping with those who fail to cooperate in the provision of a public good?

4.2 Responsible defectors, free riders and foul dealers

Are there stable equilibria in which a public good can be provided by a subgroup of cooperators in the presence of another group who defect? Obviously the first requirement is that the amount of good necessary can be provided by less than N players. It is instructive first to consider the situation in which interactions between individual players take the form of iterated two person prisoners' dilemmas. If the probability, w that two players who cooperated with each other will play each other again, is higher than the probability, u , that two players who failed to cooperate with each other will play each other again, then for certain values of w and u there are equilibria in which cooperation by a group of cooperators can be stable in the presence of a group of non-cooperators (May 1987). We then face a question with normative overtones; is it rational for the cooperators to share the good they have produced by cooperation, with the non-cooperators? In some cases, of which many environmental goods are examples, it may be impossible for the cooperators to prevent the non-cooperators from enjoying the good, so they have no choice (the good is said to be nonexcludable). If they can exclude the non-cooperators, under what circumstances should they allow access? I assume from arguments in previous sections that cooperators will likely share the goods with non-cooperators with whom they have an altruistic relationship based on kinship or sexual partnership. However, there is a more interesting case. Assuming that people are variable in interests and skills, and that people like contemporary *H. sapiens* have multiple rational life goals, then

each individual may cooperate with a preferred group to produce good G_1 , but defect when another group of which s/he is not usually a member cooperates to produce good G_2 . If goods G_1 and G_2 are complementary in some way, and those in the first group can produce more G_1 than G_2 , and those in the second more G_2 than G_1 , then exchange of G_1 and G_2 between the subgroups will produce a Pareto superior outcome to universal cooperation of all to produce $G_1 + G_2$. The temptation is to make unfair exchanges. The obvious solution is to evolve a cooperative tit-for-tat strategy in exchanges of G_1 for G_2 , but an alternative that will become clear later is to evolve a convention as to what the exchange rate should be. Free markets involve both. I raise these issues because in the literature on N party collective action problems, whether considered as prisoners' dilemmas or social choice problems, it seems to be assumed that defection is always socially undesirable. It may be in many instances, and in others if a problem is considered in isolation, but if defection from one cooperative activity can allow a person to provide a different public good of greater, or of equal but complementary, worth, then the first defection may be socially desirable. This is particularly important when sanctions are considered. I shall call a defector a responsible defector if, as a result of defection in provision of one public good, s/he is able to contribute to the provision of a different but greater public good, or a complementary good of equal worth, and it is difficult or impossible to contribute to both. In general, responsible defectors will choose a pattern of public good contributions and defections that minimise the cost of public goods provision to themselves, and maximise the provision of those public goods they desire, and in which they can share.

Generally, however, in N person prisoners' dilemmas those that profit from a public good without contributing to it are considered to be free riders. The type of dilemma in which free riders can exist, is one in which no cooperator is made worse off by a lone defector than they would be under universal defection (Pettit, 1986), and is characterised by a number of conditions: the good is nonexcludable, it can be provided by less than N members, but not by one alone. It is generally assumed that $D_i > C_i$ for each player i . However, in view of the empirical evidence that some people play unconditional ALL C in laboratory situations in spite of the fact that rational choice theorists say it is irra-

tional, it may be more realistic to consider situations in which some players have an ALL C strategy some ALL D, and some Tn. If the numbers of players using these strategies are n_c , n_d and n_t which together sum to N , and the total number of cooperators $n_c + n_t = m$, it can be shown that there are no equilibria if $n < m-1$ or $n > m-1$ for every Tn player (Taylor 1987). However if $n = m-1$ for every Tn player (that is if their cooperation is conditional on cooperation by all other cooperators) there can be a cooperative equilibrium if, and only if, for each Tn player i ,

$$w_i > g(m-1) - f(m-1) / g(m-1) - g(n_c) \text{ (Taylor, 1987).}$$

Again this equilibrium is weak, and poses another fallibility problem; it is necessary for each Tn player to reliably identify every other Tn player and ALL D player, and respond appropriately to them. If a Tn player mistakes an ALL D player for a Tn player, the defection will be interpreted as a defection amongst the cooperators, the mistaken Tn player will defect in the next round in retaliation, and cooperation will be destroyed. This vulnerability to mistaken identity is compounded in an apparently more forgiving tit-for-tat strategy that has some currency: play C on the first round, and in each subsequent round play C unless another player defects who is neither an ALL D player, nor a Tn player punishing a defection. These difficulties in the stability of cooperation in the presence of a group of free riders suggest that the free riders must be licensed and clearly identifiable by some prominent convention. It may be rational in some circumstances to license responsible defectors as free riders.

A less desirable parasite on the cooperators providing a public good is the foul dealer (Pettit, 1986). Foul dealers can occur only in a situation where one of the players can cause the payoff to one or more others to fall below that received under universal ALL D. The obvious example is disarmament: all countries are better off if they all disarm, but disarmament exposes them to the threat of a worse situation than that of universal armament, being the helpless victim of another's bombs; and defection while others disarm carries the temptation of remaining the only one armed. Although the effects of the foul dealer are devastating, while those of the free rider are merely irritating, Pettit (1986) has argued that tit-for-tat is a more credible strategy against a foul dealer, because retaliation

against defection is an obviously rational, legitimate and therefore believable response to defection by a foul dealer. The two main problems are having the courage to make the first cooperative move, and making sure that none of the opponents is defecting. Fortunately, Mr. Gorbachev had the courage and rationality to start this process amongst the nuclear powers.

4.3 Chicken and assurance

As well as increasing the shadow of the future, cooperation in a prisoners' dilemma can be promoted by changing the payoff matrix. Before considering how this might be done, it would be useful to consider two games with different logic that sometimes occur in collective action problems, and into which the prisoners' dilemma is transformed by appropriate changes to the matrix. Table 2 shows the relationships between the games.

Table 2

Prisoner's dilemma		
	Cooperate	Defect
Cooperate	R=3, R=3	S=1, T=4
Defect	T=4, S=1	P=2, P=2
(T > R > P > S)		
Chicken		
	Cooperate	Defect
Cooperate	R=3, R=3	S=2, T=4
Defect	T=4, S=2	P=1, P=1
(T > R > S > P)		
Assurance		
	Cooperate	Defect
Cooperate	R=4, R=4	S=1, T=2
Defect	T=2, S=1	P=3, P=3
(R > P > T > S)		

Provision of a number of environmental public goods that are 'lumpy' and can be provided by a subgroup of players, can be modelled by the chicken game, for instance restriction of catch of a fish to a level that is indefinitely sustainable, or restriction of pollution of a river or lake to a level that is absorbable, beyond which there is catastrophic ecological damage. Each of two players would prefer to take a free ride on the cooperation of the other rather than mutual cooperation if that were possible, but the

threatened ecological catastrophe is so bad that each would rather refrain from overfishing or polluting if the other is doing so. There are two equilibria: mutual cooperation or cooperator with free rider. The outcome of a repeated N person game of chicken is not predictable; Taylor (1987) has argued that cooperation is more likely than in the prisoners' dilemma, but others have predicted a 'scramble' to defect in order to force others to provide them with a free-ride. Risk-averse players are more likely to succeed in providing the public good. If we again assume, based on limited empirical evidence, that at least some players will pre-commit to contribute, then if the number that do so is sufficient the good will be provided. The problem is that if others know some players are committed to contribution, the temptation to defect is greater; and if too many defect, catastrophe will result.

In two-person Assurance, neither player can provide the public good alone, but if both contribute, the good is provided. There are two equilibria, mutual cooperation and defection, but as each prefers mutual cooperation to defection, cooperation is the likely outcome. The N person equivalent would be one in which a public good would not be provided unless all contributed. In a game of chicken, if it is known at some stage that further defections will result in failure to supply the public good, the remaining game becomes one of assurance.

4.4 Altruistic egoists

Altruism can have beneficial effects in achieving cooperation, and it is therefore often assumed that it is inherently desirable and should always be encouraged. In this section, as well as discussing the beneficial effects of altruism, I want to sound a cautionary note. Just as defection in a particular activity can, in a wider social context, be responsible, there are also situations in which altruism is at best a mixed blessing.

In a famous music hall act, Claude and Cecil, two pure altruists, both approach a doorway. Each wishes to defer to the other: 'After you, Claude' 'No, after you, Cecil'.... 'NO, after you Claude' 'NO, after YOU, Cecil'. This shows two things, that altruists can have conflicts of interest, and that altruism does not automatically solve problems of coordination and conflict of interest. If Claude and Cecil were trying to leave a burning building they would be in trouble, and

if Cecil were an egoist, he would have done Claude a good turn, and the outcome for both would have been better. Pure indiscriminate altruism would be of doubtful social value, but is probably nonexistent as expected from biological evolutionary theory. A more realistic view of human nature is that we are egoists with altruistic tendencies, and that the altruism is weak unless its object is a sexual partner, relative, friend, or someone who is likely to reciprocate. This on the whole works well, which is not surprising since it is the product of the pragmatic forces of natural selection. There is a tension between our egoistic and altruistic tendencies, and the balance point is variable.

In a formal treatment of mixtures of egoism and altruism, Taylor (1987) assumed that a player seeks to maximise the sum of his own payoff and that of another player. If the payoffs to the two players are p_1 and p_2 , the altruistic egoist, i , maximises a utility u_i which is a weighted sum of p_1 and p_2

$$u_i = a_i p_i + b_i p_j$$

When $b_i = 0$, player i is a pure egoist and I assume $a_i = 1$; when $a_i = 0$, i is a pure altruist, and I assume $b_i = 1$. The case I want to consider is that of egoism with positive altruism. Unlike Taylor (1987), I assume that $a_i + b_i = 1$, that is one cannot be simultaneously highly altruistic and highly egoistic; and that $b_i > 0.5$, that is that no one cares more about anyone else than about him or herself. If the game between pure egoists is a prisoners' dilemma, then the utilities that characterise the game transformed by altruism are shown in Table 3.

Fig. 1 shows how different degrees of altruism affect the severity of prisoners' dilemmas ($R=2.5$, $P=1$, $S=0$ and $T=5, 7.5$, or 10 as indicated), expressed as the value of the discount factor, w required to make cooperation stable, if both players have the same degree of altruism. As expected, the model predicts that increasing altruism makes cooperation more stable. In the example shown, for a low initial value of T , there is a zone in which $T > R > S > P$ (chicken) before cooperation is assured by the domination of C over D for both players. However, in more severe prisoners' dilemmas, altruism reduces the discount factor required for stable cooperation, but does not convert the matrix into chicken or allow stable cooperation in a one-shot encounter.

A more realistic assumption would be that

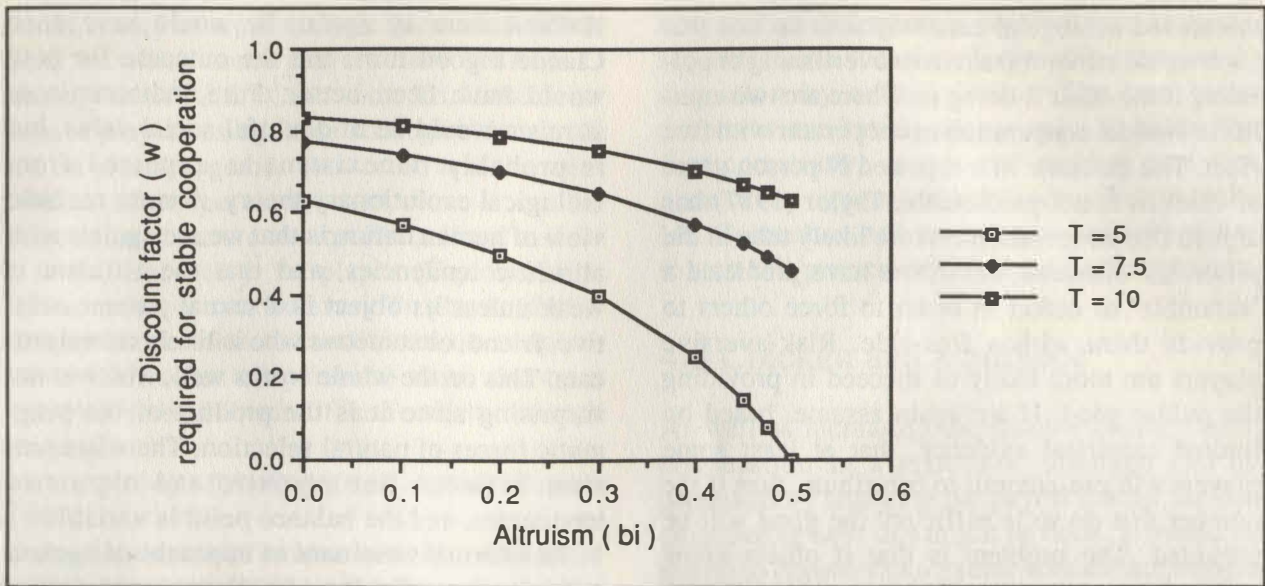


Figure 1.

Table 3 Prisoner's dilemma between altruistic egoists.

	Player 2	
	Cooperate	Defect
Player 1	Cooperate R, R	S'1 = a1S+b1T, T'2=a2T+b2S
	Defect	P,P
	T'1 = a1T+b1S, S'2 = a2S+b2T	

people vary in their degrees of altruism. It also seems likely that although a person may have a characteristic average level of altruism, their altruism towards others will vary; i's altruism towards j will be high if j is his/her child, permanent sexual partner, relative or friend, and low if j is an unrelated stranger, and very low or even negative if j belongs to a group with which i's group is competing. In an N person dilemma, i's altruism might be a weighted average of his/her altruism towards the others involved; if many of the group are his preferred recipients of altruism, the average may be high, if not it may be low. Although the arguments above suggest that reciprocal and equal altruism may make spontaneous cooperation more stable, some combinations of different degrees of altruism may make it less stable in mixed groups. The obvious danger is in situations where the original or transformed game is chicken. If it becomes known in an repeated game that some players are highly altruistic, more egoistic players will be even more tempted to pre-commit themselves to D, to force the altruistic players to provide a public good on which the egoists can take a free ride. This works only if there are sufficient

altruists to provide the good; altruists contribute and egoists take a free ride. But if there are not enough altruists, so that one or more egoists is forced to contribute to avert disaster, cooperation may actually be less probable and stable than if nobody in the group were altruistic.

Another problem with altruism arises if it is considered to be sophisticated, as defined by Taylor (1987), that is if each person's utility depends on the utility of the others. If I cannot decide my utility until you have decided yours, and you cannot decide yours until you know mine, we have an infinite regress of indecision (M. Common, personal communication).

The conclusion, then, is that altruism can make cooperation easier in certain circumstances. It is particularly beneficial in groups where it is strong and universal, but these groups are likely to have compositions that already favour in-group cooperation. In mixed groups of egoists and altruists, altruism may also be beneficial, but there is a danger that it may increase the temptation of the egoists to free ride. This is particularly important if the public good is 'lumpy', as it often is in environmental problems, and the original or transformed

matrix that of chicken.

4.5 Moral conventions, obligations and rights

'And this may properly enough be call'd a convention or agreement betwixt us, tho' without the interposition of a promise; since the actions of each of us have a reference to those of the other, and are performed on the supposition, that something is to be perform'd on the other part. Nor is the rule the less deriv'd from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it.' (Hume, 1740a).

Many aspects of social life require the evolution of the regularities in behaviour called conventions. They may be reinforced by more formal regularities such as rules and laws, but these usually rely on convention for their regular observance. Some conventions involve little or no conflict of interest (for instance, Hume's examples of two men rowing a boat, currency and language); but problems that involve conflicts of interest as well as coordination can also be solved by conventions. Coordination of individual choices to produce a socially desirable outcome against the grain of individual preferences seems to be essential to any ordered social life, and exactly what ethics is all about. Lewis (1969) defined a coordination equilibrium as a combination in which no one would have been better off had any one agent alone acted otherwise. Equilibria in situations of pure coordination are always coordination equilibria. Coordination equilibria also occur in situations of mixed conflict of interest and coincidence of interest like those we have been considering, but not all equilibria in such situations are coordination equilibria. Tit-for-tat is a coordination equilibrium in a repeated N person prisoners' dilemma if and only if it is of indefinite duration (Hardin 1982).

In his classic study of convention, Lewis (1969) defined a convention in terms of coordination equilibria as follows:

A regularity, R, in the behavior of members of a population P when they are in a recurrent situation S is a convention if and only if it is true that, and is common knowledge in P that, in any instance of S among members of P,

- (1) Everyone conforms to R;
- (2) Everyone expects everyone else to conform to R;
- (3) Everyone prefers to conform to R on condition that the others do, since S is a coordination problem, and uniform conformity to R is a coordination equilibrium in S.

Sugden (1986) used a more relaxed and general definition of a convention as 'Any stable equilibrium in a game that has two or more stable equilibria'; that is it is a uniform conformity of choice of one possible equilibrium over another.

Conventions can evolve in a population in a recurrent situation without deliberate and conscious social choice, and be self-enforcing. Lewis focussed on precedence in the evolution of a convention, that is common knowledge in the population of the way in which past instances of the coordination problem were solved. Schelling (1960), and Sugden (1986) reported that people are often able to solve coordination problems that are novel, even if they are unable to communicate with each other. Solutions in such situations tend to be prominent or salient, that is based on the most obvious and easily observed asymmetry in a situation, on the assumption that the other people involved will tend to pick the same salient feature of the situation to solve the coordination problem. If the situation is repeated, this then acts as a precedent. Another common method of solving novel coordination problems is the use of a precedent from an analogous, but not identical, situation.

It is clear from the analyses of Hume and Lewis that crucial features of the self-enforcing nature of conventions are the expectation that if I conform, others will also, and the fact that mutual conformity is to everyone's advantage. It has been suggested that these features form a basis for reciprocal obligations and rights; that is that conformity to conventions is a large part, or even the whole, of morality (Hume, 1740a, Sugden, 1986; see also Chapter 9 of Harman, 1977). The argument is that others' expectations that I will conform providing they do also is the basis of my obligation in the situation, and my expectation that they will reciprocate providing I conform is the basis of my right. A similar theory of rights can be derived from Rawls' theory of justice (Martin, 1985). In the rest of this paper I shall assume that the Humean tacit

convention theory of morality is correct in principle, and explore some of its complications, and its extension to problems of environmental ethics and sustainability. Aspects of the theory that I shall briefly address here are what distinguishes moral conventions from conventions that have no moral overtones, how they acquire moral force, and pluralism.

What is considered immoral behaviour in one society may be acceptable in another, and definition of what is the subject of morality in a given society is part of the social grammar of that society, itself established by the evolution of a set of conventions. With this background, it seems unlikely that a precise distinction between conventions that carry moral force and those that do not is possible. However, most of the conventions that are perceived as moral in our society are those where the good to be produced is believed to be of great social benefit, while conformity requires individual sacrifice or restraint. Failure to conform to a moral convention attracts more social disapproval than failure to conform to a convention not perceived in the society as moral, and this social pressure is how the convention acquires moral force, that is moral conventions are norms (Pettit, 1990). Moral conventions have also been characterised as conventions of respect (Harmann 1977).

The main advantage of the tacit convention theory is that morality is seen to arise by interactions between real people in a real society, taking account of the structure and realities of that society. It is consistent with the intuition that morality has both external and internal standards: the external standards are the judgements of other people of our social behaviour, the internal standards are our own judgements of ourselves and others. It is not necessary to appeal to external hypothetical or supernatural arbiters of timeless, universal and absolute standards, such as gods, impartial spectators, perfectly rational beings, or parties to a hypothetical social contract in a state of nature. But the theory is inherently pluralistic and conditional. Rights and obligations have limits; they are legitimate expectations only within the population, situation and time in which the convention that defines them is active; they are conditional on reciprocation. Again these conclusions are consistent with one interpretation of Rawls' theory of justice, at least so far as rights are concerned (Martin, 1985). There may

be a natural moral law of conventions common to all human societies, necessary for any ordered social life; but if so it appears to be confined to conventions of truth, promise keeping, restraint from inflicting harm, mutual aid, and restraint from theft, within each moral community; and even these 'natural duties' are defined in each society by a set of conventions concerning the people, situations and times to which they apply.

The convention theory as commonly conceived is also inherently conservative (Sugden, 1986). It implies that one should comply with the established conventions of one's own society, which depend on its cultural history. But conventions are not immutable; like language they change gradually, both by a random process of cultural drift, and in response to new situations and changed perceptions of the population. Although a convention can attract social approval once the regularity in behaviour is established, it is often assumed that a new convention cannot emerge as a result of a new perception or attitude. The argument is that to do so it would be necessary to impose sanctions, and sanctions carry two costs to those who impose them: costs of detection and execution. These costs would occur before the benefits of the new convention flowed from established conformity, so that imposition of sanctions would itself be a public good, and each would try to free ride on sanctions imposed by others, while failing to impose sanctions her or himself. To resolve this paradox it is necessary to consider the role of conscience in the development of norms.

4.6 Conscience and the psychology of morality

'When the happiness or misery of others depends in any respect upon our conduct, we dare not, as self-love might suggest to us, prefer the interest of one to that of many. The man within immediately calls to us that by doing so, we render ourselves the proper object of the contempt and indignation of our brethren' (Adam Smith, 1759).

At some stage during human evolution, we developed the capacity to monitor the behaviour of ourselves, as well as that of other people, with respect to a set of moral conventions or norms characteristic of the society into which we were born. Adam Smith (1759) suggested that the way that these rules are maintained is by social approval of those who conform, and disapproval

of those who do not; and that approval of our own compliance, and disapproval of our own non-compliance, constitute conscience. This is consistent with the evolved tendency of members of all human societies to seek social approval and avoid disapproval, while what is approved and disapproved of in a given population at a given time depends on its cultural history (Boyden, 1987). I shall argue that moral norms evolve within societies as solutions to the problems inherent in the cooperative life of a population in a given situation, and that the reason we developed the capacity for self-monitoring (conscience) was that it was of selective advantage in an increasingly social species, because it maximised social approval and minimised disapproval, without incurring the costs to all parties that result from overt sanctions.

Freudian theory suggested that a moral code was learned early in life from a child's parents, was internalised as a 'superego' without modification, complete by the age of 5, and often led to psychological problems. In fact moral development is much more gradual, flexible and complex, and continues into the late teens (Brown, 1965). The moral grammar of a society has to be learned like the grammar of its language by generalising from many, and often inconsistent, individual instances as well as by direct instruction. The moralization of an individual involves not only cognitive learning, but also development of the moral emotions of guilt, shame, and self-respect, and learning of moral conduct by imitation, selective reinforcement and conditioning (Brown, 1965).

Since Freud it has often been assumed that conscience is a disadvantageous and even pathological thing; for instance, Hardin (1968) talked of the 'pathogenic effects of conscience', and argued that in environmental issues it should be replaced by 'mutual coercion mutually agreed on'. Quite how this can be established given the cost of sanctions and the punishment paradox is not clear, but I want to challenge two implicit assumptions of the argument: that conscience is disadvantageous to the individual, and that coercion is preferable. Conscience, when we do or contemplate something perceived as against the moral norms of our society, gives rise to the unpleasant moral emotion of guilt. If we do something against those norms and are discovered, we receive the social disapproval of our peers, and experience the more in-

tense unpleasant emotion of shame. There is no doubt that to seek social approval, and to avoid social disapproval, ridicule and shame are universal innate human characteristics (Boyden, 1987), which probably therefore have an evolved biological basis. Guilt is the social equivalent of pain; it is an unpleasant signal that alerts us to the fact that something is happening that is dangerous to our welfare. It is a signal that if we deviate from a norm, our behaviour is likely to result in social disapproval, which we dislike, and which may be accompanied by sanctions that will harm us socially and physically. Conscience allows us to coordinate our social behaviour with that of our peers with the minimum of overt disapproval and sanctions that are costly to all parties. In neurotic illness guilt may be excessive, just as pain may be in physical illness, but in both cases the signal shows that there is something wrong and identifies the area of the problem. It does not follow that the signal is of no value to the individual in normal life, rather the opposite.

4.7 Feasible norms

Not all conventions are moral conventions. Moral conventions may have a number of features that distinguish them from other conventions: they are usually necessary for social life, or for the production of important public goods, they require individual cost to produce the socially beneficial outcome, and they imply respect of the other parties to the convention (Harman, 1977; a point I shall return to later). But above all, moral conventions are enforced by serious social pressure, that is they are norms.

Pettit (1990) considered what makes norms feasible, that is able to emerge, persist and be usually complied with. Norms that are not feasible, however desirable we may think them, will be paid only lip service, and seldom observed. A proposed moral norm may even be impossible given the principle established by Kant that 'ought' implies 'can'. For a norm to be feasible, Pettit suggested, its observance must be in the broad economic and social self-interest of those who comply with it. I have indicated above that this assumption of rational choice theory may be too pessimistic, but again it can be taken as a worst case scenario for the evolution of moral behaviour.

By analogy with Lewis' definition of conven-

tions, Pettit (1990) defined a norm as follows:

'A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a norm if and only if it is true that, and it is a matter of common belief that, in any instance of S among members of P,

1. Nearly everyone conforms to R
2. Nearly everyone approves of nearly everyone else he finds conforming and disapproves of nearly everyone else he finds deviating: and
3. The fact that nearly everyone approves and disapproves on this pattern helps to explain why nearly everyone conforms'.

It was mentioned earlier that tit-for-tat is a coordination equilibrium in an N person prisoners dilemma providing it is indefinitely repeated (Hardin, 1982). With the definition above, it can be seen how a norm could evolve to reinforce conformity to tit-for-tat:

1. Universal tit-for-tat is a Pareto-optimal equilibrium
2. Therefore nearly everyone evolves this strategy
3. Therefore nearly everyone cooperates
4. Universal tit-for-tat is also a coordination equilibrium
5. Therefore everyone disapproves of anyone else unilaterally defecting, and approves of anyone else conforming
6. 5 gives everyone an extra motive not to defect unilaterally.

In this example, everyone's economic interests are met because tit-for-tat is Pareto optimal, and everyone's social interests are met because conformation to the norm wins them social approval and avoids disapproval. The norm of conformity to tit-for-tat is therefore feasible.

In this derivation the behaviour of the norm evolves first, and then acquires moral force by attracting social approval. For norms that are more environmentally sound to evolve, it is likely that attitudes would have to change first. Is it possible for a norm to evolve as the result of a change in attitude? According to standard rational choice theory it is not, because it would be necessary to use sanctions that are costly.

However, Pettit (1990) has produced an attitude derivation of norms providing 5 assumptions are met:

1. Interaction. The collective action situation is such that nearly everyone is better off if everyone else takes one particular option than if they do not. Moreover, nearly everyone is made better off if anyone else takes the beneficial option.
2. Publicity. Some people will know, or are likely to notice, whenever someone else acts in a way that promotes the collective benefit, and when they act in a way that does not.
3. Perception. Anyone noticing a relevant act, will realise that that act is either beneficial to everyone and to the observer in particular, or harmful to everyone and to the observer in particular.
4. Sanction. Nearly everyone who observes a relevant act and understands its likely effects will approve of an act that is collectively and personally beneficial, and disapprove of one that is not. This need not be expressed in the form of deliberate and costly punishment. The person who disapproves may never deliberately and voluntarily express disapproval, but the person acting in a way that is collectively and personally harmful will know that the observer disapproves, because s/he would disapprove if the situation were reversed. Moreover, there are innate non-verbal communication signals between humans that express approval and disapproval, and are at least in part involuntary.
5. Motivation. The last assumption is that people are motivated in part by a desire that others not think badly of them, and if possible think well. This appears to be the case (Boyden, 1987). Assumptions 4 and 5 are the key to the attitude derivation of norms; they represent important aspects of human social behaviour that have generally been ignored in rational choice theory.

If these assumptions are met, then norms can emerge and persist as a result of attitudes in a society. First, the situations in which such norms arise have to be those in which one option is col-

lectively better than alternatives, and everyone is better off if anyone else chooses the collectively beneficial option, and worse off if they do not. The situation also has to be such that the actions of others in the situation will be noticed, and their implications understood. The second stage of the derivation is that a pattern of approval and disapproval will develop, so that nearly everyone will approve of nearly everyone he finds acting in the beneficial way, and disapprove of nearly everyone he finds acting otherwise. Thirdly, recognising the pattern of approval and disapproval on offer, people will be motivated to conform to it. Fourthly, a norm will be established and persist providing the motive to receive the approval of others and particularly to avoid their disapproval outweighs other considerations in the situation. We may expect attitude derived norms to emerge in situations where the individual benefit of conformation to the norm is high and its cost low, and where those who do not conform are free-riders. We should not expect such norms to be feasible in many party collective action problems where the individual cost is high and the benefit low, and where those who do not conform are foul dealers.

4.8 Interactions in the evolution of values, norms and theories

What is perceived as of value with respect to an ultimate end (the Good) is usually seen as separate from norms that regulate behaviour (the Right, the Just) (Moore, 1903). Some philosophers believe that definition of the Good is the foundation of ethics. Values are seen by others as reducing ultimately to individual preferences with no moral content, and they are affected by experience, and lack the permanence and universality commonly (and I believe mistakenly) associated with ethics. Yet in our society some values are seen as moral values, and every moral norm, if we accept the derivations above, is based on the perception that some course of action is better, collectively and individually, than another; that is the existence of a moral norm implies a shared value judgement and a shared conception of the good. In this section I shall argue that values and norms evolve together, and both are influenced by changed perceptions of reality and of the good life.

Values change as a result of experience. We

may value something highly until we experience it, and then change our minds. We may not value something until threatened by its loss, when we realise that far from being valueless, it was a good taken for granted. Repeated failure to achieve a goal seen as valuable may reduce our perception of its value; feasible goals are more satisfying. We may change our valuation of something as a result of a change in understanding or beliefs about it. Thus although, as originally argued by Hume (1740a), values cannot be derived solely from facts because they require the prior existence of some independent motive, nevertheless beliefs and theories about reality alter how we value different things and actions with respect to that motive.

My main argument here is that perception of values may change as a result of social, as well as other, experience, so that values, norms and theories interact as they evolve. I will illustrate the plausibility of this hypothesis by two examples. At one time smoking was thought not to be harmful to health, although a nuisance to non-smokers, while it was perceived by many as being enjoyable as well as somehow conferring social status and sexual attractiveness on those who smoked. Smokers valued smoking for both the pleasure and social status it gave them, and these perceived benefits outweighed any inconvenience it might cause to non-smokers. There was a norm of smoking, at least among the majority of men; they approved of others smoking, and assumed that others approved of their smoking. Once the theory that smoking caused lung cancer was proposed, and accepted, and followed by theories that smoking contributed to other cancers and cardiovascular disease, many former smokers reduced their valuation of smoking, and stopped smoking. This was followed by a growing perception amongst the population that smoking is both collectively and individually harmful; people increasingly disapproved of anyone else they saw smoking, and approved of those who refrained from smoking in a public place. New norms evolved against smoking in various situations, and spread. In many public situations these were eventually codified in rules and laws, but it is important to realise that the main reason that smokers refrain from smoking in such situations is the social experience of disapproval, not formal penalties. Moreover, the social norm contributes to the decreasing value ascribed to smoking amongst smokers, and is a motive in addition to protec-

tion of health to refrain from smoking. We then come to a further development, my interpretation of which may surprise scientists. Scientific tests of the extended hypothesis that smoking may be dangerous to the health not only of smokers, but also those breathing the same air, could be seen as merely a scientific development. But could it not also be seen as a value judgement that smoking is collectively as well as individually bad, influencing the development of a science? Is not the norm that resulted from a changed value judgement that resulted from a changed perception of reality, now contributing to a further change in the perception of reality?

The second example of the effect of social experience on values is the norm of telling the truth reliably, and not just expediently. In standard sociobiological theory, an individual is expected to tell the truth only when it serves his or her interests. Some sociobiologists have even suggested that deception and the detection of deception were important factors in the evolution of human intelligence, and Alexander (1987) argued that unless a speaker and audience shared extensive interests, communication would consist of a minimal core of truth with embellishments limited only by the credulity of the audience. This could have been true in a hypothetical 'state of nature' like Hobbes' war of every man against every man, but it is not true of any real human society. The reason is that people have social as well as economic interests, and even their economic interests, broadly conceived, are seldom served by lying if as a result others do not believe them and routinely lie to them in return. Lewis (1969) has shown that the essential foundation convention in language is a convention of truthfulness within a population. Without it, what would evolve would not be a language, but a tower of Babel. Pettit's attitude derivation of norms can readily explain the norm of truthfulness in communication that exists in our society (and probably all human societies); there is certainly no doubt that we disapprove of anyone else we find lying and approve of anyone else we find telling the truth, nor that nearly everyone feels guilt when lying for short term selfish reasons, and shame when detected in a lie. My argument is that as a result of repeated social experiences of this sort, we come to value honesty more, and the personal gains that can be made by lying less; and that similar changes in perceived

values, towards those socially approved, result from the evolution of all norms. The modified values that we develop as a result of moral norms, are what we intuitively think of as moral values.

It is probably impossible to model such subtle changes with much accuracy, but as a first approximation it seems reasonable to assume that in deciding whether to conform to an evolving norm, individuals might construct a modified matrix of values consisting of the sum of their original personal preference (broadly economic) matrix, and a perceived social approval matrix. For instance, if the original preference matrix represented a prisoners' dilemma between two people, i and j , and an evolving norm required conformity to C_i , C_j we might symbolise i 's approval of j 's conformity as $a(i,j)$, and i 's disapproval of j 's defection as $d(i,j)$. The original preference and social approval matrices, and the transformed matrix of their sum, can then be written as shown in Table 4.

I assume that $a(i,j) = -d(i,j)$, and for simplicity that approval and disapproval are symmetrical, that is $a(i,j) = a(j,i)$. Table 5 shows calculation of transformed value matrices for two different individual preference situations, taking $a(i,j)$ as 1.

If this simplified model is at all valid, then the values on which people act in situations to which a norm applies are not merely personal preferences, but a combination of such preferences and social values derived from the norm. If the norm is a moral norm in the population, the transformed matrix will be perceived as a matrix of moral values, and if the norm is strong and stable in the population, people's individual preferences will tend to shift towards those of the transformed matrix with repeated experience of the situation.

From the point of view of a population of moral agents P rather than an individual, if the achievement of some purpose p in some recurrent situation S requires cooperative action, then its achievement involves two coordination problems in P : coordination of purpose or ends, and of values required to achieve that purpose, that together constitute the good of P in S . The coordination of purpose could be solved by a coincidence of interests, or by a convention or norm that p is a desirable end. If this is achieved, then possible equilibria in the value problems are that a given behaviour B is valuable for p , or

Table 4

		Original preference matrix	
		j	
		C	D
i	C	R(i), R(j)	S(i), T(j)
	D	T(i), S(j)	P(i), P(j)

		Social approval matrix	
		j	
		C	D
i	C	a (j,i), a (i,j)	a (j,i), d (i,j)
	D	d (j,i), a (i,j)	d (j,i), d (i,j)

		Transformed value matrix	
		j	
		C	D
i	C	R(i) + a (j,i), R(j) + a (i,j)	S(i) + a (j,i), T(j) + d (i,j)
	D	T(i) + d (j,i), S(j) + a (i,j)	P(i) + d (j,i), P(j) + d (i,j)

Table 5

		Original preferences		Social approval		Transformed values	
		j		j		j	
		C	D	C	D	C	D
i	C	3, 3	0, 4	1, 1	1, -1	4, 4	1, 3
	D	4, 0	1, 1	-1, 1	-1, -1	3, 1	0, 0

		Prisoners' dilemma		Norm C(i), C(j)	
		j		j	
		A	B	A	B
i	A	1, 1	0, 0	-1, -1	-1, 1
	B	0, 0	1, 1	1, -1	1, 1

		Pure coordination		Norm B(i), B(j)	
		j		j	
		A	B	A	B
i	A	1, 1	0, 0	-1, -1	-1, 1
	B	0, 0	1, 1	1, -1	1, 1

that B is not valuable for *p*. If most individuals in a series of S observe and/or believe that if nearly everyone does B, *p* nearly always occurs; whereas if nearly everyone does not do B, *p* nearly always does not occur; then the solution B is valuable for *p* in S is salient, and this convention will evolve. Such a convention could, I think, be called a value theory of P, and the way it evolves and its social nature would be similar to a factual theory. Similar theories would evolve concerning the value of components of the physical and biological environment for the

achievement of *p*. Individual value statements would be fallible and dependent on theories about B and *p* in the same way that observation statements are fallible and theory laden. Indeed, a value theory in the narrow sense would be a factual theory that happened to be relevant to some cooperative purpose of the population that believed in it. I have already argued that scientific theories are human social constructs that are sufficiently accurate approximations of reality to be used to promote the human good. The overall conclusion, then, is that values and

facts are both theory dependent, and all three are instrumental to cooperative ends of the population of agents that believes in them. This does not mean that the values are necessarily selfish or short term; the agents might have altruistic and socially desirable long term collective ends. But biological evolutionary theory and rational choice theory suggest that collective ends that evolve are likely to be perceived to be in the long term enlightened self-interest, broadly conceived within the cultural context of the population, of the individuals that make up the population.

It seems likely from this analysis that values and norms evolve together, and that both are related to beliefs about reality, environmental and social, and the nature of the good life. A given combination of beliefs about reality, values and norms is characteristic of the society in which those beliefs, values and norms obtain. In *primaeval* hunter-gatherer societies, changes in the cultural mix were probably very slow. In modern secular Western societies such changes are rapid; nevertheless they occur by evolution within a given cultural background. Changes in values cannot occur in isolation; they must start from the cultural status quo, and be compatible with the changed understanding about the human situation, and new social norms of behaviour, with which the new values must mesh. In European Australian society, that means the changed values must evolve in one, or preferably both, of the major cultural traditions: the rational secular tradition, or the Judaeo-Christian religious tradition.

4.9 The problem of moral considerability: norms of respect

'Now it appears, that in the original frame of our mind, our strongest attention is confin'd to ourselves; our next is extended to our relations and acquaintance; and 'tis only the weakest which reaches to strangers and indifferent persons' (Hume, 1740a).

The type of morality that has been considered so far is confined to interactions between different members of a population of human moral agents, and is based on reciprocity between those agents: justice in the sense used by Hume and Rawls. But some types of morality do not fit

this pattern, at least at first glance, and many of the moral problems to be faced in working towards sustainability are not confined to the population of agents, for instance relationships with future generations and other species. Extension of morality to other species, ecosystems, and even the whole biosphere has been suggested by many environmental philosophers. However, I believe it is necessary to again consider the origins and nature of human morality before trying to decide whether, and under what conditions, such extensions of moral considerability might be possible and effective. If it is possible, we might ask whether morality that is not based on reciprocity, and for which people are not the moral patients, is necessarily of the same type as morality between people, and carries the same weight. Stone (1988) has argued on philosophical rather than evolutionary grounds that ethical monism, the view that there is one ethic that should guide all our moral interactions with the biosphere as well as each other, can no longer be assumed.

The processes discussed in section 3 help to explain how altruism and cooperation may have arisen in hominid in-groups, because of their contribution to the improved survival, prosperity and reproduction of individuals within the group. The whole of human evolution until very recent times took place in in-groups living as hunter-gatherers. Thus cooperation and morality evolved within such groups, and there has been insufficient time for their fundamental innate basis to have changed since. All of the conditions that favoured the evolution of altruism and reciprocal cooperation within in-groups change dramatically when we consider interactions between members of different out-groups. Such individuals are unlikely to be close kin or to have a continuing sexual relationship leading to cooperative parenting, or a long term friendly relationship. Moreover, the value of *w* would be low in such an interaction, making ALL D rather than tit-for-tat the more prudent and successful strategy, even more so if the groups to which two interacting individuals belonged were in competition, as they often were.

This puts a different emphasis on the problem of moral considerability. Many environmental philosophers have asked what reasons we may have to justify not including other species in our moral considerations. Theories of the evolution of cooperation suggest we should rather ask why

we include people with whom we do not share kinship, an ongoing sexual relationship or a longstanding cooperative personal relationship, let alone why we should include members of other species. Alexander (1987) has even suggested that competition between human groups drove the evolution of cooperation within groups. The invention of agriculture, and the development of urban life, forced us to extend moral considerability beyond the simple in-group in our own enlightened self-interest, but it has been a hard-won battle. To assume we can further extend considerability indiscriminately to other species is to beg the question. There is evidence that we can, under the right conditions, extend some moral values to humans outside our interlocking personal networks of in-groups. Unfortunately, there is equally convincing evidence that we can also treat out-groups with hostility, violence and a tendency to genocide. The examples of the cannibalistic neanderthals of Krapina and Hortus (Campbell, 1985), and the regimes of Adolf Hitler, Idi Amin and Pol Pot make it doubtful that such tendencies are confined to one race or generation of humankind. I shall argue that extension of moral considerability to people outside our own in-groups was necessary to solve practical problems of civilized living, and drew on the already established strategies of reciprocity. But extension of morality outside the natural barrier is of necessity a difficult and uncertain process, and it is necessary to consider why, and under what conditions, extended cooperation is possible. Moreover, we have no choice but to eat other living beings in order to live, and like most species we have to use other species in various ways to serve our own ends. I am not arguing that what is natural is necessarily right, but that it puts limits on what can reasonably be expected.

This point was in part anticipated by Goodpasture (1978) in his objection 6 to his own principle of moral considerability: 'The clearest and most decisive refutation of the principle of respect for life is that one cannot live according to it, nor is there any indication in nature that we were intended to'. His reply was that the life principle asks for 'sensitivity and awareness, not for suicide'. Thus Goodpasture's principle is not so strict and universal that it prohibits the use of other species by humankind to fulfil legitimate ends, that is unlike some later radical biocentric and frankly misanthropic ethics such

as 'Earth first' it does not demand that every other being be treated like a member of the human in-group. Some deep ecologists also recognise limits to biocentric moral considerability (Fox, 1989; Bennett, 1990), but most appear to fear the 'slippery slope', and assume that any use of other species by humankind implies every use is justified. Since we have no choice but to exploit other species to live, we cannot live with a principle that forbids all use of other species. We have to evolve an ethic that defines for a given population what use of what other living things by humankind is justified under what circumstances, and what might make such norms feasible, given the evolutionary origin of morality in the human in-group.

It may be instructive first to examine cases of morality where reciprocity is limited, but not ruled out. As examples, the Australian blood banks, coastal patrols, surf life-saving clubs and bushfire brigades are all based on voluntary contribution to a public good, in situations where it is unlikely (although not impossible) that the contributor will receive reciprocal benefits. Altruism may be the motive of many contributors, although in most cases they can expect increased self-esteem and social approval as a result of their contributions. I suggest that in these situations a public good can be provided by a small number of the population, and there is a norm of free-riding on committed contributors; everyone has a right to the good, but no-one has an obligation to provide it, unless they do so as the result of a promise made in joining the organisations. Thus although these systems appear to be running on pure altruism, they are actually made stable by two strong norms: that goods that are actually excludable be made available to all in these particular recurrent situations, and that promises are kept. A committed contributor who failed to contribute, or attempted to restrict access to the good, would be treated with strong disapproval; but someone who did not enter into a commitment to contribute would not.

Suppose we next consider voluntary contributions to a completely separate population with almost no possibility of direct reciprocity, say contributions to famine relief in Africa. Such schemes at first sight appear to run solely on weak altruism, or what Hume called moral benevolence or sympathy. My argument is that no-one has an obligation to contribute, and the suffering have no right to expect relief, if no

norm defining those rights and obligations exists in the population of agents. Although some relief may be supplied by the more altruistic agents, it is likely to be weak and capricious. If, on the other hand, there is a norm in the population of agents that contributions should be made, that norm defines obligations on the part of the agents to contribute, and rights on the part of the recipients for relief. But according to Pettit's theory, such a norm is feasible only if it is perceived to be in the economic and social interests, broadly conceived, of the population of agents. The paradoxical conclusion is that the morality of contributing to relief is likely to be stronger, more effective and reliable if it is instrumental, and based on perceived indirect self-interest, than if it is based on the supposedly nobler morality of altruism.

What circumstances might result in the evolution of a feasible norm within a population of agents, of which the beneficiary, at first glance, is outside the population of moral agents? I suggest such a norm would be feasible if and only if there is a common belief within the population of agents that the moral patient can influence the welfare of the population of agents in some way, either for good or for bad. Based on the suggestion of Harman (1977) that moral conventions based on reciprocity are characterised by respect for each other, I shall call norms that identify moral patients within or outside the population of agents norms of respect.

Over periods of 40,000 to 60,000 years, humankind (neanderthal and modern) in Europe and north America, has observed rituals of respect for bears with which human populations were in competition (Campbell, 1985). These were based on respect for a powerful and dangerous adversary. But other norms of respect are based on a perception that their subjects are beneficial to humankind, but vulnerable. Aboriginal cultures in Australia had (and often still have) respect for a wide range of animals, plants, waterholes, rocks, places, names and mythical beings; some of these were perceived as powerful and frightening, others as beneficial and vulnerable, but respect for all of them was maintained by strong norms. We also have norms of respect for entities that have no ends or goals of their own, for instance our national flag. Harman (1977) considered the hypothetical example of whether invading martians would develop moral conventions towards humankind. His conclusion was that martians

would only see us as moral patients if we were able to influence their welfare in some way.

The result of this analysis is that the tortuous attempts to find some common basis for the extension of moral considerability to other lifeforms in interspecies and environmental ethics has probably been based on a misconception. The assumption has always been that moral patients have some common property that they share with us (such as sentience, interests or goals) that defines them as worthy of our consideration. I would now suggest that moral patients are defined not by any common property, but by a relationship to the population of moral agents. This might be formalised as follows:

An entity E is a moral patient of a population P at time t in a recurrent situation S if there exists within P at t a norm of respect towards E in S.

Some might object that moral patients could also be defined by an altruistic relationship between members of P, and E; but this is still a relationship rather than a property. Even a relationship with a very vulnerable entity seems better based on respect than altruism to me, although there are undoubtedly norms that enforce altruistic behaviour (for instance, towards infants). These could also be seen as norms of respect (for instance, respect for human life). To take account of arguments that altruism might nevertheless characterise a moral patient perhaps it should be included in the definition, but only if specified by a norm. Otherwise, as I have argued, it does not determine rights and obligations. But a norm of altruism must be active in a population to identify a patient, and it will not become active unless it is feasible as defined above. Taking these considerations into account, the definition could be extended as follows:

An entity E is a moral patient of a population P at time t in a recurrent situation S, if and only if a norm of respect or altruism towards E is in force in S, in P at t.

This does not mean that the moral patients of a population lack common properties, for regularities are to be expected because of salience, precedence, and the human tendency to minimise logical inconsistency. But whether or not a given entity is a moral patient of the population of agents cannot be established simply by rational argument about the properties of the supposed patient, without reference to the

norms actually operating in the population at the time.

R. Goodin (personal communication) suggested to me that moral considerability should deal with possible moral patients of P, and not just those currently the subject of norms. Considering the entities that historically have qualified for moral considerability by human populations, I doubt that the class of possible patients can easily be defined. However, there is clearly evolution of moral considerability within the culture of each human population, and at a given time, it may be possible to define a class of potential moral patients that is the subject of current moral argument:

An entity E is a potential moral patient of a population P at time t in a recurrent situation S if either;

1. There is an emerging belief amongst members of P at t that E can influence their welfare in S individually and collectively, for good or bad, and that respect for E in S is in their biological, economic or social interests, broadly conceived; or
2. Some members of P are motivated to behave altruistically towards E, and there is an emerging belief amongst members of P at t that altruistic behaviour towards E in S is in their biological, economic or social interests, broadly conceived.

5. Evolution of values compatible with ecological sustainability

If the biosocial imperatives of the fundamental questions program are taken as a starting point, they represent a greatly changed perception of the reality of the human situation compared with the dominant perception in phase 4, high energy urban society. This will in turn require equally large changes in values and norms of behaviour, if society is to achieve a sustainable relationship with the biosphere in the long term. It means the end of increasing population and energy use, and of the perception that humankind can have an indefinitely increasing standard of living based on more and more technological exploitation of the

environment. It will require a deliberate restraint on consumption, and abandonment of the belief that human happiness depends mainly on material possessions and consumption of material goods. In essence, it will require the replacement of standard of living by quality of life as the major human value. Humankind can no longer see themselves as masters of the universe, but rather as they were as hunter-gatherers, just one part of a complex ecological system of interdependent life forms. The expansion of human populations that began with the invention of agriculture has reached its limits. The changes we must undergo may well be as great as those in adapting to settled urban life based on agriculture.

In this section the analysis of the origins and nature of human morality already undertaken will be used to suggest how the necessary changes in values and norms could occur. Although there are enormous problems, and there will undoubtedly be mistakes and setbacks, I believe there may be some cause for optimism, at least so far as Australian society is concerned. We still have wilderness areas, and other areas relatively unaffected by human activity and of immense significance globally as well as nationally. Awareness of environmental problems and participation in organisations trying to address them is high. We have the biological and economic resources to respond to the challenge, and relatively peaceful, ordered, equitable and democratic social institutions. It has been argued that if the environmental cause is lost in Australia, it will be lost everywhere (Bennett and Sylvan, 1989). Although the required decreases in material and energy consumption and generation of wastes may at first sight seem formidable, the reduced levels can nevertheless support a rich and varied lifestyle. I shall argue that many of the necessary changes can occur by mechanisms that are already established and understood, providing perception of the reality of the human situation is achieved throughout society. The required changes are understanding and acceptance of the new human situation, revision of ends and values in line with this revised perception, collective action and the evolution of new conventions and norms to achieve stabilisation of population and reduction of consumption and waste production, and protection of the remaining natural environment, and transfer to the political agenda of problems that cannot be solved without government action. We also need to develop new norms

respecting the needs of future generations, other species with which we share the world, and the ecology of the biosphere. There are signs that these changes are already underway both in Australia (particularly in Tasmania) and overseas (for instance, in Sweden).

Changes compatible with sustainability must occur by a slow evolution of perceptions of reality, values and norms, at the grassroots level in the first instance, and only when this is secured, transferred to the social choice agenda and codified in laws with coercive sanctions. There is likely to be strong opposition from those with interests and values best served by the maintenance of the status quo. No government can enforce laws that are not in the perceived interests, individual and collective, of the majority of citizens without cost, and if a democratically elected government attempts to do so, it simply loses power. For ethical as well as pragmatic reasons, human values also need to be protected while developing new ecological values.

5.1 Cultural pluralism and ecological values

Arguments presented in sections 3 and 4 suggest that what morality can evolve in a given human population will depend on its cultural history, which will include paradigms of reality and the nature of the good life, and conventions, norms, rules and laws that regulate behaviour. It was suggested that the biological roots of social behaviour also put limits on what morality can reasonably be expected to evolve, resulting in a number of trends that are likely to be found in most human societies, in spite of their cultural diversity. People will tend to be motivated in their social, political and moral behaviour by what they perceive, within their cultural context, to be in the long term biological, economic and social interests of themselves, their children, relatives, sexual partners, and those with whom they have long term cooperative personal relationships and shared interests. An important determinant of social behaviour in all human societies is likely to be the desire for social approval and the avoidance of disapproval, but what behaviour receives approval and what disapproval is greatly influenced by cultural variables (Boyden, 1987).

Two points related to this view of morality

need to be developed here. The first is the hypothesis that because there has been insufficient time since the development of farming and urban settlement for much change due to biological evolution, it is unlikely that the biological foundation of human social and moral behaviour has been modified to any extent since then (Boyden, 1987), and there is certainly no time to wait for biological evolution to respond to the realities of post-expansion life. It follows that changes required for ecological sustainability must be cultural. However, that does not mean that our biological heritage can be ignored in planning for the future; biological restraints on what morality can reasonably be expected to emerge and be stable have to be taken into account, and if they are not the consequences of planned actions are likely to be different from those predicted.

The second point is that in a complex society such as exists in Australia, there will always be a degree of cultural pluralism, and therefore moral pluralism. Any large change in values, such as that required to achieve ecological sustainability, will need to occur in several different moral cultures. To achieve a change throughout the society will require cooperation between people with different perceptions and cultures. In such alliances, the interests of different groups will vary, and their understanding and trust of each other will have limits. It will be necessary to take this into account, as well as the biological foundations of cooperative behaviour. Successful cooperation is likely to be built up in repeated interactions involving personal relationships, probably with each party using a tit-for-tat strategy.

As an example of this type of interaction, it might be instructive to ask why local Aborigines supported the Daintree-Bloomfield road, although its route went through a number of sacred sites as well as ecologically sensitive coastal rainforest. Although ecologically conscious Australians who are ethnically European may tend to support (as I do) Aboriginal land rights and the relationship between traditional owners and their land, the romantic assumption of some ecologists that Aborigines and other people with 'vernacular' cultures (Goldsmith, 1988) have 'natural' conservation values has not always been borne out in practice, and could even be considered mildly paternalistic and racist (Anderson, 1989; Palmer, 1990). Aboriginal groups, like ethnically European

groups, sometimes decide in favour of conservation, and sometimes of development. Pro-development Aboriginal decisions are often blamed on European corruption of 'natural' Aboriginal values. In fact, *conscious* ecological values are as European as conscious exploitative 'progress' values, and ethnically European people were trying to influence the Aborigines on both sides of the Daintree road debate. The question is, was the eventual pro-road Aboriginal decision influenced by European opinion, and if so, why did the European conservationists lose?

According to Anderson (1989) the Bloomfield Kuku-Yalanji, and specifically their leaders, had the major say in Aboriginal opinion about the Bloomfield-Daintree road because the route included their sacred sites; these sites decided whose voice was most influential, not what the decision should be. The Kuku-Yalanji leaders decided in favour of the road because it increased their importance in the Aboriginal community, it would improve their access to European goods and services, and because it was supported by local Europeans with whom the Kuku-Yalanji had built up long term, personal cooperative relationships (Anderson, 1989). In other words, they saw the road as being in their long term economic and social interests within the local cultural context. Local anti-road Europeans had failed to develop personal relationships with the Kuku-Yalanji or understanding of their culture or politics, and had in some cases restricted their access to hunting and gathering territory. Outside conservationists assumed local Aborigines would be anti-road without, in most cases, consulting them (Anderson, 1989). In pre-contact society, interactions with strangers with no on-going social contact had almost always been hostile, and the Kuku-Yalanji therefore mistrusted the anti-road Europeans (cf section 3.4, and 4.9 paragraph 2). The lesson seems clear; Aborigines, whatever their exposure to Europeans, are people like any other people, and their morality is subject to the same biological and cultural constraints as anyone else's. Assumed coincidence of interests without communication, understanding or personal relationships is unlikely to be a good basis for ecological cooperation between people from different subcultures.

5.2 Changing the perception of reality

According to the thesis developed in this paper, the first step in the evolution of values compatible with ecological sustainability is an individual and collective change in perception of reality, that is in beliefs and theories about the place of humankind in the biosphere. The new perception that evolves in post-expansion secular societies must, if it is to be effective, be based on scientific understanding rather than the explanatory myths of primaevial societies. I gave a number of pragmatic and ethical reasons why this should be so in section 2, but in fact we cannot go back. We have eaten of the fruit of the tree of knowledge and lost our innocence; we cannot return to the garden of Eden. The explanatory myths of hunter-gatherer societies were deeply embedded in their cultures. Our new perception of our place in nature can only evolve within our own cultural heritage.

The overall perception of the relationship between humankind and the biosphere that is necessary as a basis for the evolution of values and norms compatible with ecological sustainability, has been outlined in previous publications of the Fundamental Questions program (Boyden, 1990; Boyden et al., 1990). These state the origins of *H. sapiens* by biological evolution, the biohistorical processes that have led to the present position, the inescapable dependence of humankind on the health and integrity of the biosphere, and the threats to that integrity that will result if the expansion of human population and technometabolism is allowed to continue. There is a great need for further research on the state of the environment, and our dependence on, and effects on the biosphere. While there will be need to develop new technologies that consume less energy and resources, and cause less pollution and danger to the health of the biosphere, I believe the major value of science will be in increasing understanding of nature and our place in it, rather than attempting to control and exploit it even more. It is only by research for understanding that we shall complete the change in perception necessary, and more and more people will be convinced that this changed perception is right.

In using science to understand our relationship with the biosphere, there are two caveats. One is the conclusion from the philosophy of

science that the theories we use are human social constructs that are approximate models of some aspects of reality, and are likely to be replaced in time by different and better theoretical constructs. Although we use theories that are the best available when we need to act, we should do so with caution and the realisation that further research may again change our perception of reality. The other caveat is that we always have to act when our knowledge is incomplete, and to some extent uncertain, and that some relevant aspects of reality may in fact be chaotic, and not amenable to scientific study or technological control. The understanding of the biosphere and our place in it based on science must always be incomplete and imperfect, and we need also to acknowledge that it is accompanied by intractable uncertainty and ignorance. The combination of this understanding and acknowledged ignorance should, I believe, engender an attitude to nature similar to that which existed for different reasons (and rarely consciously) in the cultures of many hunter-gatherer societies: understanding of our dependence on other life-forms, and of the complexity of the myriad relationships on which all life depends; concern for the welfare of all living things on which we depend indirectly as well as directly; and humility because of limits to our power to understand and control other living things and natural phenomena; in a word, respect.

The understanding of life and natural forces that has flowed from science has demystified our perception of reality. We no longer believe in supernatural beings with strange and terrible powers that shape and control our destiny. There is no doubt that this demystification has been accompanied in some people by a loss of respect. I argued in section 4.9 that respect results from the acknowledgement of other entities that are not entirely in our control, and which can influence our welfare. Some writers on environmental ethics have suggested that we need to return to the mysticism and supernatural beliefs of what Goldsmith (1988) has called 'vernacular' societies. But it seems to me that it is respect for nature that must be regained in Western secular culture, not belief in the supernatural. To support this argument, I cite the Aztec and Easter Island societies that had strong beliefs in the supernatural, faced or believed they faced environmental crises that threatened their society, and in which increasingly obses-

sive placatory religious rituals were counterproductive. On Easter Island, the felling of trees to act as sledges for the transport of increasingly large religious statues contributed to the environmental degradation that, together with warfare between rival communities, eventually resulted in a catastrophic decline in population and social organisation. These societies were acting on what Boyden (1987) calls 'cultural delusions', beliefs that human sacrifice or the building of large statues would cause supernatural beings that controlled their destiny to ensure the continued prosperity of their societies. I argued in section 2.5 that failing to use the best available empirical knowledge results in poor ethical as well as practical decisionmaking. A further problem about perceptions of reality that are based almost entirely on mysticism, intuition and dogmatism without empirical content is that when things start to go wrong, as from time to time they must, there is a tendency to become fanatical in a vain attempt at correction. It seems to me that respect for nature based on scientific understanding of our dependence on it, and acknowledgement of limits to our knowledge and control, is a much better basis for our values, norms and actions, than a return to superstition and mysticism.

5.3 Enlarging the shadow of the future

It was shown in section 3.4 that an important determinant of the stability of cooperation based on reciprocity was w , the discount factor, which has two components: the probability that the interacting individuals will meet again, and the weight given to the payoff in future encounters compared with that from a current encounter. Theoretically, cooperation can be fostered by increasing w , that is 'enlarging the shadow of the future' (Axelrod, 1984). There is a sense in which this is central to the values required for sustainability, for if we took no heed for future generations, we should not worry about ecological sustainability. The solution seems at first sight simple, we should lift our horizons to several hundred generations hence, and this would assure cooperation amongst our contemporaries to achieve sustainability. However, there are several complications that make this solution problematic.

There are good reasons why we discount the future. Humankind has an evolved preference for immediate to delayed satisfaction of desires. A delayed but anticipated payoff may never occur, for many reasons. To sacrifice an immediate payoff for a later one of equal value is therefore to risk a reduction in total utility over time. Moreover, predictions of the future are necessarily more uncertain than knowledge of the present, so that the consequences of present actions on future outcomes are less certain than consequences on immediate outcomes. All of these reasons make it rational to discount the future; the size of the discount factor, however, depends on many things. The most important of these are the anticipated value and priority of a possible future gain, relative to its immediate cost, and the risks that the future gain may never be realised, or may depreciate. In environmental questions and thinking about sustainability, the balance between egoism and altruism also affects the outcome. The more distant the time horizon, the greater is the component of altruism involved, for actions that may benefit very distant generations will surely not benefit ourselves, and for various reasons may also not benefit our descendants, in whom we have evolutionary interests. Since universal and indiscriminate altruism is unlikely amongst human populations, it seems doubtful that time horizons can easily be extended indefinitely. To promote reliable consideration of future generations, norms of altruism towards them need to evolve; and to be feasible, these norms must be based on a perception that they are in the biological, economic and social interests (broadly conceived) of the present generation. The only plausible basis for such a perception, it seems to me, is our innate evolved concern for the welfare of our own children and grandchildren, and to a lesser extent for all the children of the in-groups to which we belong. Such consideration also draws on the other component in w , the probability that we shall again interact with the other agents in the situation. These arguments suggest that far from concentrating on possible altruism towards very distant generations in the evolution of values compatible with sustainability, we should rather concentrate on the biological, economic and social advantages of ecologically responsible behaviour for the future welfare of those in the current generation with whom we interact, and particularly their children and grandchildren.

There is another way in which over-detailed

concern for the welfare of very distant future generations may be counterproductive. The theories, technologies, perception of reality, values and norms of distant generations will be very different from ours, and our predictions of their preferences are as a result likely to be extremely inaccurate. They will not belong to the same culture. To make irreversible decisions on the basis of what we expect to be their preferences, may prevent them from achieving their own ends, by closing off their options. Decisions we take for such distant generations should be designed to ensure that their needs are met, insofar as common human needs can be defined; beyond that we should strive merely to be responsible stewards of the earth, leaving as many options as possible open. We should walk lightly, socially as well as ecologically.

5.4 Changing the values

The first biosocial imperative in the transition to ecological sustainability is that human populations should be stable or falling. Since it is not possible to maximise simultaneously two variables, the higher the population in a sustainable society, the lower must be the consumption per capita, other things being equal (Hardin, 1968). The relationship between food production and human population is a critical one, both for humans and other species. If human population is not controlled, human and environmental catastrophe is inevitable, which is a bad in anyone's ethics. Because much of the present population growth is in the third world, the magnitude and seriousness of the problem is often not appreciated. I am saddened to hear some of my colleagues argue, for instance, that a 20% increase in the yield of rice, milk, bacon or trout as a result of genetic engineering will solve the problem. Unchecked, population increase is exponential, and surely even the most optimistic should realise that there cannot be an indefinite exponential increase in agricultural production, whatever technology is used. Present human populations are already putting pressure on Australian water and soil, and population control is as important here as anywhere else.

Fortunately for human and non-human life, scientific study of human reproduction has allowed technological separation of sexual intercourse from reproduction. It is in this area that even the most ardent believers in religious or

Kantian ethics and noble motives must surely concede that to be effective, a norm must be feasible. If we relied on the establishment of a norm of controlling population by abstinence from sexual intercourse, humankind and the biosphere would be in deep trouble. Contraception is probably the single most crucial factor in achieving sustainability, unless we intend to control population by starvation and war, or return to a marginal diet and four years of breast feeding per child. The problem is twofold: to change values so that reproduction is less valued, and norms disapproving of reproduction above replacement level emerge; and to make contraceptive technology available, culturally acceptable and affordable where it is needed.

Like all living creatures, humankind has in-built biological desires, drives and emotions that result in behaviour that in the normal course of events results in reproduction. Evolutionary theory, indeed, suggests that all behaviour is in the end directly or indirectly adapted to maximising reproductive fitness, for it is competition in reproduction together with genetic variation that drives the evolutionary process. But this does not mean that we instinctively maximise the number of children we have; like many mammals, humans have small numbers of children and invest large biological effort in rearing them to sexual maturity. The goal is to have the optimum number of children for the prevailing capacity of the environment to support. What this optimum is depends not only on the physical and biological capacity of the environment, but also on the expected per capita consumption per child. There is therefore an inevitable trade-off between economic and reproductive goals and rewards. The decreased birthrate in affluent societies over the last few decades is due not only to technological improvements in contraception, but also to two perceptions: that the probability that each child will survive to reach sexual maturity is high, and that the combined economic capacity of parents and children is inversely proportional to the number of children. These perceptions decrease the expected value of further reproduction as each child is born. The technology of contraception makes reproductive limitation possible, but it would not be used unless there were a motive to use it. One of the most important ways of working towards values compatible with sustainability is therefore to emphasise the necessary trade-offs between reproductive and

economic goals and rewards, and to put environmental costs of reproduction on the agenda. At the national level, policies that reward and encourage reproduction need to be reversed, and the environmental impact of migration needs to be considered.

The next most difficult step in the evolution of morality compatible with sustainability is likely to be the replacement of broadly economic values, by longer term biological and social values. Some changes that involve comparatively trivial costs (for instance, reducing the valuation of chlorofluorocarbons as spray propellants) have already taken place. But the magnitude of the changes that have to occur is only slowly becoming apparent, even among those who accept their inevitability if sustainable human societies are to evolve. To a society that even has difficulty in accepting a trivial and temporary reduction in standard of living for longer term economic gain, the prospect of a 75% reduction in standard of living, as it is currently perceived, with no tangible economic benefit, is almost impossible to imagine. We know it is possible for human societies to exist at this lower level of economic activity, and for life within them to be enjoyable, for they existed in the past. What we do not know is whether it is possible without great social disruption to achieve such a state from one in which a higher level of consumption is enjoyed, and economic values are perceived as preeminent. Such a change could occur only gradually over several generations, and there would undoubtedly be false starts and backtracking. Although some individuals may be able to slowly change their own values because, say, they realise their own continued consumption at the current rate threatens more basic needs of their children and grandchildren, and the complex biological relationships on which their lives will depend, others will want to protect their existing economic values by denying the reality of the changed perception of the human situation. To be sure, life at the reduced energy and resource consumption and pollution generation levels compatible with sustainability can still be attractive and varied (Boyden, 1990; Boyden et al., 1990). But I very much doubt, until new norms and moral values have evolved over a long period of time, whether more than a tiny minority of people will actually prefer to consume less, any more than most people would prefer to have less sex or a smaller number of

friends. Of course, small groups of people have set up communities over the past three decades that have to some extent withdrawn from the mainstream economic life of capitalist countries; but the vast majority of these have always in fact been dependent on mainstream society for information, resources, markets, public services and supplementary or even principal income. They have therefore consumed more than they could have if their economies were truly independent of the rest.

If this argument is at all correct, and the sacrifices required to achieve sustainability will not initially be seen by most people as desirable in their own right, then norms compatible with sustainability will, if they are to be feasible, require some compensatory perceived benefit which outweighs the perceived cost. The individual sacrifices to be made must be seen as contributions to a public good that is individually and collectively beneficial; that is the revised behaviour will only emerge if it is seen as yielding a biological and social benefit, individually and collectively, that will in the long term exceed the immediate economic cost.

As an example, we might take contributions to the cost of leaving an area as wilderness, rather than exploiting it for some economic purpose: the economic sacrifice is not valued in itself, but suffered in order to achieve the objective of preserving wilderness. This can only happen if wilderness is seen as having some value that outweighs the economic cost. But cost there must be, and another consideration is who bears the cost, and who gets the benefits. Those who want to preserve wilderness without themselves suffering economically are attempting to take a free ride on those whose livelihood is directly affected. Eventually, because of changed values, it may be possible to obtain an economic benefit from wilderness, in which case the sacrifices may be seen as savings rather than costs; but an economic return may itself threaten the wilderness value of the area.

It seems to me inescapable to me that if wilderness, species, ecosystems and the biosphere are to be preserved, they must have a value that is higher in priority than short term economic value, or any value requiring consumption, to the individuals in the society that preserves them, and preservation must carry net economic and human reproductive costs. The values are human values, and instrumental; the individuals in the society that preserves (say)

wilderness think and feel that preservation is in the long term biological and social interests, broadly conceived, of themselves, their relatives, friends, children and grandchildren. There is evidence that individuals in at least some human societies have such values, and so long as the values are acted on, I believe they need no further justification; they need not be 'intrinsic'. It is important to recognise that people have, and are prepared to pay for, such values. There appear to be two broad categories of non-consumption values: values that require experience but not consumption, and existence values that do not require experience (at least of any direct sort by the current generation). Amongst experience values of wilderness, for instance, are aesthetic values. Existence values include biological diversity and heritage values, and values related to the stability of the biosphere. Species preservation likewise requires existence value for the species.

5.5 Against intrinsic values

There have been many who think that there must be intrinsic values, to justify say, preservation of wilderness or species. I do not think this is the case, and because so much time, energy and paper has been spent in the search for intrinsic values, I will say why I believe the search is doomed to failure. There is also a need to point out that extreme biocentric or ecocentric ethics based on supposedly equal intrinsic value of all species are not feasible in the sense used in section 4.7, and in many situations are even physically impossible if immense human suffering and death are to be avoided. This point was realised by Goodpasture (1978), and is conceded by some deep ecologists (Fox, 1989; Bennett, 1990) as discussed in section 4.9. Weston (1985) nevertheless argues that the search for a philosophical basis for intrinsic values has weakened environmentalism, and diverted it from more important practical issues.

The theory of intrinsic values is associated mainly with the intuitionist ethics of G.E. Moore (1903), who regarded the question of what is good in itself or has intrinsic value as the most important in ethics. However, Moore claimed that whether or not a thing had intrinsic value was an intuitive truth unrelated either to other truths or to experience. A thing was held to be intrinsically valuable if, on reflection, its existence was held to be good in complete isolation.

An important part of Moore's theory was the 'principle of organic unities', which states of intrinsic values that '*the value of a whole must not be assumed to be the same as the sum of the values of its parts.*', or even proportional to that sum (Moore, 1903). Moore's principle does not say that intrinsic value can belong only to a whole, nor that all things that have intrinsic value are of equal intrinsic value (he believed in degrees of intrinsic value); as I understand it his principle says only that intrinsic values do not obey the rules of arithmetic. Nevertheless, he finally concluded that things that are intrinsically good are 'highly complex wholes, composed of parts which have little or no value in themselves'.

At first sight the principle of organic unities seems to apply also to ordinary instrumental values. For instance, the instrumental value to me of the computer on which I am writing this paper is apparently higher than that of the sum of its parts, because the sum of the parts is not functional, and I do not have the knowledge or skill to assemble the parts into a functional whole. To resolve the paradox, I suggest we have to think about whether something was not left out of the original concepts of both the whole and the list of parts; the whole is not simply a random heap of the physical parts. To make a functional whole the parts have to be assembled in a particular way which involves a decrease in entropy, and inputs of information and energy. These are not mysterious components of the whole, we pay for similar things every day. When these components are included as parts in the concept of the whole computer, the sum of the values of the parts no longer falls short of the value of the whole. I do not claim this disproves the principle of organic unities; for if intrinsic values do exist, it is possible they may behave non-linearly, unlike instrumental values. I do, however, suggest that careful thought may show that an apparent difference between the value of a whole and that of the sum of its parts may be due to something missing from the equation rather than a deviation from linearity.

Moore's intrinsically valuable wholes consisted of some function of a valuer (such as consciousness or love; never apparently the valuer per se), and some property of a valued object (such as beauty). He used the principle of organic unities to justify one intuition, that the whole has intrinsic value, against the counter-

intuition that none of the identified parts has intrinsic value. However, if the principle of organic unities is taken seriously, its use to justify the intrinsic value of wholes against counter-intuitions about their parts is bought at the price of impotence in other respects; for unless it specifies some rules other than those of arithmetic for relating the value of a whole to that of its parts, the principle cannot be used to derive the value of a part or a whole even if the value of everything else is known; the value of each part and of the whole can be decided only in isolation, and by intuition. In the end, the intuitionist argument for intrinsic value is justified solely by intuition. As I argued for factual theories in section 2.5, intuition is undoubtedly a valuable creative and synthetic human capability; but it needs to be coordinated, and tested by experience and reasoning, if it is to be of use to a human population in interacting with the real world. What moral intuitions people may have about the good can, and must, be affected by personal experience and social pressure within a particular culture, and judged by predicted and actual consequences. Intuition alone is not enough.

It might be easier to understand what Moore meant by intrinsic value, by considering examples of things he thought had this property, in spite of his contention that they have nothing besides intrinsic value in common. Things he thought were intrinsically valuable included consciousness of a beautiful object, true knowledge, love of a good person, and Beethoven's 5th Symphony. What is extraordinary to me, is that he did not realise these all involve subjective human values that would certainly vary from one human culture to another. It seems that what Moore held to be intrinsic values are very similar to what I have called non-use experience values (particularly aesthetic values) and existence values. However, they seem to me to be obviously human instrumental values, greatly influenced by culture, rather than intrinsic values.

Another theory of intrinsic values was proposed by Brentano (1902), who rejected arguments for a principle like Moore's principle of organic unities. Brentano's theory has recently been extended by Chisholm (1986), with the reintroduction of Moore's principle. According to this theory, a thing is intrinsically valuable if it is right to love it, and has no intrinsic value if it is right to hate it. This immediately raises the

questions of how emotions can be right or wrong, and what defines the rightness of an emotion. In spite of much discussion, no answers to these questions that I find plausible are provided; some of those given involve the principle of organic unities, and others the assertion that aesthetic values are objective (which I deny).

Values are held by a valuer, and have no point unless they affect her or his behaviour. In the usual broader sense values refer both to an end thought to be good by the valuer, and to values relative to that end (values in the narrower sense). Intrinsic value, if such exists, can therefore apply to the end of the valuer, to something valued relative to that end, or to the combination of both. The theories of intrinsic value already discussed appeared to be concerned with the third possibility ('organic wholes'). The use of the concept in environmental philosophy has usually been different (Routley and Routley, 1980); it has been used to suggest that valued entities can have value that is intrinsic to themselves, and not dependent on valuation by a person (or any other valuer) for any purpose. This concept of intrinsic value seems to me to be defective in that it fails to specify for what purpose the object is valued, and by what valuer. It was in fact the implausibility of the location of intrinsic value solely in the valued object that led Moore to postulate the principle of organic unities in order to support his intuition of intrinsic value in complex wholes. For instance, uranium is valued by some people because it can be used indirectly to produce energy, treat cancer, or kill other people. If I am not motivated to kill other people, that value does not exist for me; and if I have a prior motive not to harm people, and I believe I can only use uranium to generate energy by harming others, then uranium also ceases to have that value for me. The goodness or otherwise of the ends of treating cancer or killing people, or the primacy of the end of not harming people of future generations over that of economic energy production in this generation, could be decided even in ignorance of the existence of uranium. But the value of uranium to humankind or to any other hypothetical valuer can only be specified relative to those, or other, ends. What is the intrinsic value of uranium independent of the use to which it will be put? The question seems to me unanswerable.

Some say that intrinsic values are values of

God, or gods in the case of those who think that environmental ethics require a return to pantheism. Then the values are still instrumental, but instrumental to some motive or purpose of the god or gods. In what way are such values intrinsic to the valued object? If people believe that a god will be angry and punish them, or disapprove of them, if they do not act according to the god's values, and the people act in a way that they believe will avoid the disapproval or punishment of the god, then the values that motivate their behaviour are still instrumental.

It is said that to value something intrinsically is to value it in and for itself. In the way this has developed in environmental ethics, it is assumed that a thing can be valued in and for itself only if it has its own ends. This is an extension to environmental ethics of the Kantian principle of treating every man as an end in himself, but its relevance to intrinsic value is dubious; most of the things thought by Moore to be intrinsically good did not have their own ends; Hitler had his own ends, but they hardly show him to be intrinsically good. Should I value uranium in and for itself? The answer on the 'own ends' theory is no, because it lacks its own ends or interests. However, individuals of *Plasmodium falciparum*, which cause malaria, have their own ends: to invade, prosper and reproduce in my bloodstream. Should I value each individual *Plasmodium* for its own ends? If the ends of humans and gods do not define intrinsic value, why should the interests of individuals of *P. falciparum*? I believe the whole approach is based on a misapprehension; ends or interests define a valuer, not a valued object. There is an additional difficulty in defining moral considerability by intrinsic value and then intrinsic value in terms of interests or ends; there are strong arguments against abstract classes such as species, ecosystems or wilderness having real ends or interests (Williams, 1966; Cahen 1988), so argument for their moral considerability on this basis is vulnerable to the same criticism.

Holistic rationalists, on the other hand, say only the whole is intrinsically valuable, and parts have no intrinsic value. This requires once again Moore's principle of organic unities, with all its difficulties. Moreover, it is a dangerous argument to use in environmental ethics. It could imply, for instance, that a species is valuable, but not the individuals of that species; or that an ecosystem is valuable, but none of the

members of its biotic community. Some say that the whole is Gaia, who can replace or change her parts, and thus heal herself. If she is indestructible, why should we care what we do to her; we cannot destroy her. If we care because of what she may do to us, we are back to human instrumental values. As Calicott (1986) has said, with friends like holistic rationalists, species preservation needs no enemies.

Some say there is no distinction between the whole and the part, or between the subject and the object, so if I am intrinsically valuable, everything is. This argument has been based on quantum theory (Calicott, 1985; Zimmerman 1988) which is a fallible empirical theory, refers to a particular level of organisation of matter in the universe, and is not necessarily applicable to other levels of organisation. But if the reality and separateness of biological entities and their conflicting interests is to be ignored at the level of organisation appropriate to environmental ethics, then this and all other forms of ethics are impossible. How can we define rights and duties if the interacting parties are held to be unreal and indistinguishable? Use of this doubtful lack of discrimination between beings is compounded when it is used to defend the thesis that all beings are intrinsically valuable, for it depends on the assumption that at least one of them is intrinsically valuable, for which no argument is offered, other than that it has ends or interests. As I have argued earlier, ends or interests define a valuer, not an object of value.

Some say intrinsic value is objective, while instrumental values are subjective. If we mean by 'objective' true of the real world independently of human judgement, human values are clearly not objective. To propose an hypothesis that intrinsic value independent of human value is an observable property of members of a given class of entities in the real world, I have to have some means of distinguishing empirically between those entities that carry intrinsic value, and those that do not. I know of no such test, and Moore argued that such a test is impossible. Even if an empirical test were possible, the arguments in section 2 suggest that the hypothesis that a certain class of entities are intrinsically valuable cannot be held to be objectively true; it still depends on human judgement. If it is held to be true by intuition, then that intuition is surely not objective.

The most persuasive advocate for intrinsic values, in my opinion, is Rolston (1988) who

suggests if I understand him rightly, that humans are the only moral valuers in our world, but that intrinsic value is somehow carried by the valued object, and revealed or activated when there is an interaction between a person and the valuable object. However, for a critique of an earlier version of this theory, see Weston (1985). Rolston's theory, by retreating from the concept which has crept into environmental ethics of intrinsic value independent of a valuer, avoids the difficulties discussed above. It represents a return to something like the intuitive concept of value in 'organic wholes' current 90 years ago, but in the absence of something equivalent to the principle of organic unities, allows some of the value to be assigned to the valued object. Rolston's concept of value inherent in the valued object and perceived as a result of an interaction with the valuer expresses an important intuition, but because it does not allow that valuation has to be with respect to some end of the valuer, it is incomplete. When motive of the valuer is included, I think it becomes clear that it is some property that is relevant to the motive of the valuer that is intrinsic in the valued entity, not the value per se. Even this is too strong a statement, for to value an entity, a valuer has only to believe that it has a property that will enable him or her to achieve some purpose, it need not be true that the entity actually has such a property. Uranium, then, is valuable only to valuers who believe that uranium has some property that will enable them to achieve some end that they desire, such as treating cancer, killing people or generating energy. We seem to have come back once more to human instrumental values.

5.6 Ecologically responsible norms

In his classic paper on the tragedy of the commons, Hardin (1968) argued that as maintenance of the commons is an N person prisoners' dilemma, spontaneous cooperation is impossible, and coercion the only solution. However, tit-for-tat is a coordination equilibrium in an indefinite N party dilemma, and conventions and norms are possible solutions to the dilemma. In fact, the commons in Europe and England were not run as open access systems; grazing rights were defined by conventions, as were access rights and maintenance obligations in common forests (Cox, 1985; Taylor, 1987; J. Dargavel, personal com-

munication). Pettit (1990) has argued that under these conditions, the interaction assumption of his attitude derivation of norms is satisfied, for not overgrazing is collectively beneficial, and everyone benefits to some extent by anyone else not overgrazing. Anyone overgrazing is likely to be noticed, the harm to the person noticing overgrazing would be perceived, and nearly everyone noticing anyone else overgrazing would disapprove; the publicity, perception and sanction assumptions are therefore satisfied. Since not overgrazing is a small cost if nearly everyone else does the same, the motivation to avoid disapproval and seek approval would be sufficient to make nearly everyone conform to a norm of not overgrazing, and the commons would be maintained. History suggests that Pettit rather than Hardin is right. Commons were well maintained and norms of not overgrazing were stable for hundreds of years (Cox, 1985; Taylor, 1987; J. Dargavel personal communication). Some commons are still running. Those that no longer continue did not become degraded because of the commons system; the land was appropriated by more powerful individuals or groups and used for other purposes, or its use was converted to more 'efficient' agricultural technologies that were not sustainable (Cox, 1985).

Most problems in environmental conservation and the achievement of ecological sustainability are also collective action problems, where a public good is to be produced at individual cost by refraining from or limiting consumption, pollution or reproduction in a recurrent situation. In most of the situations, Pettit's action type assumption is satisfied, for nearly everyone is better off if everyone takes one option than if they reject it, at least according to the perception of reality, ends and values shared by most participants in the FQ program, and a growing number of others. But this does not mean that evolution of ecologically responsible norms will be easy. The attitude derivation of norms works well if the problem is an A type dilemma, and those who try not to conform are potential free-riders, rather than foul dealers. It is also dependent on the publicity assumption, which is easy to satisfy in a small community whose members interact daily, but more difficult to satisfy in larger communities. Even in cities, or in international relations, however, there are some environmental and sustainability problems where nearly everyone, or every country, taking the collectively undesirable op-

tion is easily detected and exposed to disapproval, and where nearly everyone is motivated to avoid disapproval. Bicchieri (1990) has also shown that under certain conditions, norms of cooperation can begin in small groups and spread by an evolutionary process to the rest of the population. Providing these conditions are met, ecologically responsible norms are likely to emerge relatively spontaneously, and be stable.

There are, however, many situations where spontaneous evolution of appropriate norms will be difficult, and transfer of the problem to the social choice agenda will be necessary to develop government backed formal incentives or sanctions to reinforce spontaneous processes. If Pettit is right, this may paradoxically be most readily accepted in type B dilemmas whose solution is threatened by foul dealers, and where behaviour-based norms are appropriate. The enormous individual as well as collective harm that foul dealers can do in such situations is such that strong sanctions against potential defectors are clearly rational and collectively and individually desirable, and therefore likely to be approved by cooperators and believed by potential defectors, and effective with little need to actually apply sanctions in most cases. An example of such a situation might be, say, a proposal to build a factory likely to discharge highly toxic waste into a creek that drains into a lake in a National Park, or one used by professional fishermen observing a norm of fishing to sustainable limits.

More difficult problems are likely to be those in which the dilemma is severe because the sacrifice required is high, and individual share in the public good small. For instance, it has been relatively easy to establish norms of not using sprays propelled by CFCs and translate them into government legislation and international agreement, because the individual and national cost is small; if we did not use sprays at all, it would not inconvenience us much. It is much more difficult to take similar action about refrigerants. Such action requires development of adequate alternatives, because the cost of giving up refrigeration is too large. Some have argued against developing alternative refrigerants, and suggest giving up refrigeration. Such proposals take no account of the dependence of human food supply and distribution on refrigeration in the cities of developed nations, or of the improvements in health that have

resulted from refrigeration. It may eventually be necessary to stop using refrigeration, but it could only be done after a large change in cultural perception of reality, ends and values, and massive physical reorganisation of food production, distribution and storage in developed countries. In the present circumstances, a norm of not refrigerating food would simply not be feasible.

Another factor that is often forgotten is the distribution of costs and benefits. It is much easier to develop a norm of sustainability if the costs and perceived benefits are distributed equitably in the population. If a required public good can be provided by a small number of contributors, the cost to the contributors is low and they have some altruistic motives or compensating egoistic rewards, then a public good may be provided as in the example of blood banks. But if the cost to contributors is high while the benefits are enjoyed by others, it is inequitable to expect them to bear the cost without compensation, and they are unlikely to do so without coercion and protest. This applies, for instance, to restraint on logging in National Estate forests in southern NSW. I believe that such forests should not be logged by methods that involve actions that are incompatible with the sustainable maintenance of the present level of biological diversity, and renewable resource potential for future generations. But I also believe it is inequitable and unrealistic to expect local people whose livelihood depends on timber harvesting to regard such a norm as feasible, when they believe they are being asked to sacrifice their livelihood for what they see as our values, which they do not share because their perception of reality is different. It is all very well to occupy the high moral ground and say their values are wrong, but it is they, not we, who are being asked to pay the cost. The solutions in such cases are some form of short-term compensation, negotiation and compromise, the development of alternative, ecologically responsible local employment, patient discussion and education to change the perception of reality, and the eventual evolution of norms of respect for the forest based on long term enlightened self-interest. This type of approach seems to be working in a number of areas in Tasmania. It is less likely to work if harvesting and employment are controlled by a large multinational company that does not have a stake in our future.

Another situation where equity is involved is where the public good required to achieve sustainability is 'lumpy', not excludable, generally perceived to be essential, and can be provided by less than all those involved. This applies to many environmental situations where there is a threshold such as a necessary replacement population of plants or animals, or a level of pollution that can be absorbed by the system, and exploitation beyond the threshold causes ecological catastrophe. The situation is equivalent to the game of chicken, and although it should intuitively be easier to solve than a prisoners' dilemma because the possible catastrophe is everyone's least favoured option, it is often intractable. This is because tough individuals can trade on the others' fear of catastrophe to force them to concede a free ride, but one too many risk tolerant individuals can realise everybody's worst fear (Taylor, 1987). The situation can be even worse if recalcitrant, risk-tolerant individuals or nations commit themselves to higher and higher consumption, forcing the others to reduce theirs, until a single individual or nation is taking all the maximum sustainable yield, and the sustainability is being provided by the restraint of others under duress. This sort of behaviour would not be tolerated within a community, because sanctions against it are rational and believable; but it has occurred in situations where the individuals involved are unlikely to meet each other, and belong to different communities, states or nations, for instance fishing in international waters. Not unexpectedly, such exploitation is more likely off someone else's coast than one's own, for if catastrophe does occur, it is likely to affect the victims of exploitation more than the exploiter. Another situation where some can exploit others in a chicken-like situation is pollution of a river, where the exploiter does not even have to be brave, merely upstream. There are a number of possible defences against such exploiters, but none is guaranteed of success. Appeals to higher values and morality are unlikely to be effective with someone already acting in this way. Publicity in a forum where the reputation of the exploiter is both valuable and vulnerable is the possible basis of a norm of equitable restraint. Negotiation of a bargain of restraint in exchange for something else of value is another possibility. If the worst comes to the worst, however, providing the situation is sustainable,

restraint by the exploited under duress is preferable to retaliatory consumption or pollution; at least it maximises the options of future generations of humankind, and whatever other species are involved.

Relations with other species must, if my arguments about moral considerability are accepted, be based on norms of respect and/or altruism. This immediately, to my mind, solves one problem, for while norms of altruism require that the moral patient has interests or ends, norms of respect do not. We can therefore accept norms of respect for other species, wilderness, ecosystems, rivers, places and the biosphere, which I believe are essential for sustainability, but which are presently implausibly based on supposed class interests or goals and intrinsic values of the patients. Norms of altruism are worthwhile and defensible, but I would again mention their limitations. People are likely to feel and act altruistically towards individual large-bodied mammals that are intelligent, cuddly or cute, and to a lesser extent towards birds that have attractive plumage or song. The number of species currently thought to be facing extinction as a result of human activities is horrendous, and most of these do not have members likely to inspire altruism in people. For instance, the most numerous class of threatened species is probably that of insects that inhabit the canopies of rain forests (May, 1988), and few people feel altruism towards insects, or towards other invertebrates, reptiles, amphibians, plants (except perhaps trees), protozoa, bacteria or viruses. The only hope for most of the non-human species in the world is the evolution of norms of respect for other lifeforms and their habitats, based on long-term enlightened self interest, by human moral communities.

5.7 Pluralistic moral reasoning and practical environmental decisionmaking

In this paper I have emphasised the evolutionary origins and the social and cultural aspects of morality, because they give a richer and more realistic account of human morality than most *a priori* ethics, and because it seems to me that neglect of biological and cultural constraints on what morality can evolve in a given society at a given time has led a number of authors to devise

abstract environmental ethics that ask too much of human nature, or to advocate cultural transplants that are unlikely to be effective vehicles of change. On the other hand, I have probably neglected the role of moral reasoning and argument. Within any given cultural context at any given time a number of different opinions about a given moral issue can generally be defended by rational argument. These arguments will have different degrees of coherence, cogency and social support, which may well change with time. The functions of moral arguments are, I suggest, communication between members of a moral community; persuasion of others to a given moral opinion; coordination and the forging of consensus; and the justification of moral change or defence of the status quo. But above all, moral reasoning serves to guide individual choice and practical social decisionmaking by determining to what extent a given action can be justified given individual conscience, and the set of moral conventions, norms and rules current within the population. The forms ethical arguments can take are diverse and context dependent, and their structure has much in common with those used in jurisprudence (Toulmin, 1958).

Stone (1987) also recognised the close relationship between jurisprudence and ethics, and argued that any substantial claim to legal considerability must be based on an underlying moral claim. But the grounds accepted for legal considerability are wider and more varied than those currently accepted as grounds for moral considerability. Stone opposed what he called moral monism, the view that a single ethical system that recognises only one criterion of moral considerability can proceed by analytical arguments from some universal premise to produce for every moral dilemma a uniquely correct solution. Toulmin (1958) showed that such a system is bound to produce paradoxes when confronting the real moral world, which inevitably requires judgements that balance competing moral principles and conflicts of duty. Stone's alternative is moral pluralism; moral analysis and argument that recognises a number of different planes. Each plane can have different criteria for moral considerability and therefore can deal with different classes of entities (for instance, persons, future persons, non-persons (including subclasses of sentient animals and nonsentient living creatures), things, and membership entities such as species, wilderness or ecosystems). Importantly, these

different planes can have different governances; different moral rules, moods, textures, standards and force. This has, to my mind, two great advantages: it allows us to ascribe moral considerability for various reasons to nonhuman entities without at the same time demanding that they be treated morally in exactly the same way as people; and it allows us to recognise rational choice, utilitarian or Rawlsian/Kantian arguments as appropriate and valid for some situations on some planes without assuming that any one of these represents the only valid moral governance that must be applied indiscriminately and with equal force to every dilemma on every plane.

Stone's moral pluralism at present represents only a sketch of a possible way forward. Nevertheless it appears to me to be a great advance in concepts in ecological ethics; it is certainly more plausible, flexible and practical than previous attempts to extend monolithic and supposedly universal theories of human morality indiscriminately and without change into new areas we need to address if we are to work towards ecological sustainability. Stone (1987, 1988) was at great pains to distinguish pluralism from relativism, which he rejected, and also opposed the convention theory of morality, and accepted intrinsic good as one possible basis for moral considerability on non-utility planes. Clearly these views are inconsistent with some of those advanced here. Stone illustrated his pluralistic approach by analysis of the problems raised by proposals to drill for oil off the coast of Alaska in the path of migration of bowhead whales, which belong to an endangered species and are hunted by Inupiat Indians. What I shall attempt here is a similar, but necessarily brief and preliminary, analysis of an Australian controversy of equal complexity: whether or not to allow mining of Coronation Hill in the Conservation Zone within Stage 3 of Kakadu National Park. What I hope to show is that a decision not to mine can be defended both by pluralistic arguments of the type considered by Stone, and can also be supported within many of the planes by the type of evolving social interaction arguments developed in this paper. With Stone, I shall maintain that if a given action can be supported by arguments of more than one type on more than one plane, then it should count in favour rather than against the action. This approach has something in common with the shared platform of different branches of deep

ecology developed in David Bennett's paper (Bennett 1990).

I shall take a number of things as given in the Coronation Hill controversy, which have been fairly widely (but not universally) accepted as a result of the Woodward Commission, the Ranger Uranium Environmental Inquiry, the Australian National Parks and Wildlife Service plan of management, the Senate report on the Potential of the Kakadu National Park Region (Senate Standing Committee on Environment, Recreation and the Arts (SSCERA), 1988), The World Heritage status of the Park, and various Acts of Parliament. I shall assume that mining is out of the question in the Park proper, but was until recently expected to occur for a 5 year period within the Conservation Zone before its incorporation in Stage 3 of the Park, that is I assume that mining at Coronation Hill is not excluded a priori. Another assumption is that the Park proper has very high conservation, heritage and aesthetic value, but that these values are lower in the Conservation Zone, because of its history of pastoral use and small scale mineral exploration and mining. Proper management of the Conservation Zone is nevertheless vital to the Park, because it is the habitat of a number of species rare or endangered in the region, and because it forms part of the catchment area of the South Alligator river, which flows through it. I shall assume that Aboriginal people (collectively Gagudju within stages 1 and 2 the park, but locally speaking the Jawoyn language) have established moral rights to the land, whether or not this is in all areas legally established; that the 50,000 year history of Aboriginal culture in the area is an important part of the heritage value of the Park; that the pre-contact Aboriginal relationship to the land was demonstratively sustainable; and that proposals for mining require consultation with the traditional owners of the land. Legitimate arguments in the consultation process include Jawoyn sacred sites and beliefs in the presence in the area of routes and resting places of 'bula' beings from the dream-time, who if disturbed may destroy the earth and the people (sickness country). Advice on the presence of such sites in an area is sought from the Aboriginal Sacred Sites Authority, but requires consultation with traditional owners. Because of loss of local culture and connection with the land which is only slowly being regained, and of the difficulty of communication across cultural barriers, the consultation process

is inherently slow and uncertain (K. Palmer, evidence to SSCERA, 1988). If mining is allowed, the Jawoyn people would be paid royalties, probably administered by a Jawoyn association. I assume that the proposed mining process includes the production of alkaline cyanide waste that if released would cause severe ecological damage to the Alligator river system, but that the waste would be contained within a tailings dam with a backup retention pond. The risk of damage to the Park ecosystem would thus be small, but not negligible, since release of toxic waste could occur by damage to the dam and pond, or by seepage, and the hydrology of the region is not well understood. Other damage would occur by earthmoving, road-building and accommodation of workers at the site, and the introduction of weeds. If mining at Coronation Hill were permitted, it is likely that as a result of this precedent there would be pressure to allow much more extensive exploration and mining within the Conservation Zone, for there are 13 potential uranium mines and 5 gold mines in the area; mining on this scale could pose a serious threat to the Park in both the short and long term, despite strict environmental controls (SSCERA 1988).

My basic viewpoint is that mining and other types of non-Aboriginal economic exploitation of the region originally ignored the interests of Aboriginal people and of the living creatures in the region, but that over a period of about 30 years, norms of respect for the indigenous people, fauna and flora gradually evolved, and were formalised in the ways outlined above. A cooperative interaction between the Gagudju and non-Aboriginal people, particularly the NPWS, developed. Cooperative management of the Park and recognition of Aboriginal ownership of the land were the main goals of this collective action. These conflicted with existing norms and laws, that for economic reasons frequently gave mining priority over other potential uses of land irrespective of the wishes of owners. The solution that developed in the Kakadu region was to isolate small enclaves where particularly lucrative mineral deposits were located and allow mining there, but not elsewhere. Even within these areas, mining was subjected to increasingly thorough appraisal and control to ensure, so far as possible, that it did not compromise the other values guiding use of the Park. The root of the controversy over the mining of Coronation Hill is the inherent am-

biguity of the function of the Conservation Zone within Stage 3 of the Park with respect to this framework; is it a mining enclave or part of a National Park?

To begin a preliminary pluralistic analysis of the type advocated by Stone, the assumptions listed above would be used to prepare a series of action-influence overlays on a map of the Park, to show the influence of different degrees of mining exploitation, and of damage to or leakage of the toxic waste containment, on the Park. These maps indicate different possible degrees of damage to: the local region; to Aboriginal sacred sites, camp sites and hunting grounds; to regions important to the recreational, aesthetic, conservation and heritage functions of the Park proper; and to the flora and fauna. These could be used to identify zones and scenarios where a given level of mining might cause a high degree of risk to an area already identified as important to one or more functions of the Park, or important to a particular group of people, nonpersons, or things. Stone's moral pluralism attempts to value possible actions in terms of these risks and of possible benefits on a number of moral planes.

The first of Stone's utility planes, U1, considers only people as morally considerate and their preferences as fixed and determinate. Aboriginal traditional owners, miners, conservationists, tourists or archaeologists are treated merely as people with different preferences. Animals, plants, hills and rivers are considered only insofar as they enter into the utility calculations of people, and these utilities are summed, with the future heavily discounted. Considerations of this sort are also expected to be important from arguments developed in sections 3 and 4 of this paper, for one of the prime drives in the evolution of cooperation, conventions, individual conscience and social norms was held to be the achievement of pareto-optimal outcomes. Precise analysis on this plane of the Coronation Hill issue would require much more work, but my expectation is that the benefits would outweigh the costs if only Coronation Hill were to be mined. The economic gain would be high, local damage would be in areas not important to most people's utilities, and the risk of damage to more important areas in the near future would be low providing containment precautions were adequate. This might change if further mining were allowed, for local damage would increase, country near more important

bula sites would be disturbed, and the risk of contamination of the South Alligator river and damage to important areas of the Park higher.

The second utilitarian plane, U2, considers all sentient beings as morally considerate. On the basis of arguments developed in Section 4.9 of this paper, it would be expected that within a given culture, norms of altruism or respect towards some sentient animals might evolve, but would be unlikely to be identical to similar norms concerning people. Moral considerability of all sentient creatures would be unlikely. The obvious problems in utilitarian analysis on this plane, with or without universal considerateness, are how to determine the preferences of nonpersons, what weight to give them, and how to combine them with the human preferences of U1; that is how to avoid the traps of simply extending the governance of reciprocal human morality to moral patients that are nonpersons. Stone concentrated on the first problem, and made little comment on the others. In fact, in this case I think preference determination is relatively simple. We are not considering here the development of a complete utilitarian ethic for all sentient creatures in the Park, but merely asking how the utilities of those nonpersons we hold to be morally considerable would be affected by our possible decisions about mining. However ignorant we may be of what makes archerfish, crocodiles, lotus birds or Leichardt's grasshoppers happy, we can be very sure that all of them have a strong preference not to be exposed to alkaline cyanide. On the other hand, we can be equally sure that their utilities will be little advanced, if at all, by the economic return from mining. The preferences of those sentient creatures we hold to be morally considerable that are located in high risk zones would therefore count strongly and unequivocally against mining.

The approach to considering the utilities of nonpersons together with human utilities that I suggest, is based on Taylor's treatment of altruism, which I discussed in Section 4.4 in a rational choice context. I suggest that the utilities to be summed consist of the human preferences, plus a series of preferences of those nonpersons we think are morally considerable, to which we assign different altruistic weights. This will not satisfy those who favour biocentric egalitarianism, but again I emphasise we are not attempting the implausible task of constructing a biocentric ethic in which nonpersons are as-

sumed to have the same status as people within an extended moral community, but rather trying to sketch how we might reasonably begin to incorporate some nonpersons into human morality as moral patients. The assignment of different weights to the preferences of different creatures does not mean they are to be included merely according to their utility to people; rather their own preferences are to be considered, but assigned weights according to our degree of altruism towards them. If we stick strictly to sentience as the criterion of considerability, we might assign weights according to the likely degree of sentience of different creatures, based on the level of organization of their nervous systems. Certainly the attitudes of ecologically aware people to dolphins and whales, for instance, suggests that we are in fact more altruistic towards them, than towards fish or krill, that is we intuitively assume their preferences have a higher weight than those of less advanced creatures in human morality. Moreover, it seems obvious that such weights can be context dependent. We may assign a high weight to the utilities of crocodiles in a National Park, but this would be quite inappropriate in (say) a creek flowing through the grounds of a suburban primary school. The extent to which the preferences of nonpersons in U2 would outweigh the presumption in U1 that mining at Coronation Hill itself should go ahead, therefore depends on how many nonpersons are thought to be morally considerable, and the weights applied to their preferences. It is possible that in some circumstances the weight of nonperson preferences might be higher than those of persons (for instance, when considering harvesting of an endangered species of elephants for ivory). In general it seems reasonable that the preferences of nonpersons should be discounted by some factor. Nevertheless, within the context of the Kakadu National Park, many people might suggest that at least some mammals and birds are morally considerable, and that their utilities should not be heavily discounted. If this is accepted, I believe the balance in analysis on planes 1 and 2 combined would be against mining.

Stone's utility planes U3 and U4 consider people remote in space and time as morally considerable. I gave reasons in sections 4.9 and 5.3 for thinking on the basis of biological evolution and extended rational choice theory, that we might expect some concern for future genera-

tions and remote people, but that the more remote and unconnected to our own lives those people were, the less might be our concern for them. The difficulties here for extended consequentialist ethics are that the more remote people are in space and time, the more inaccurate is our knowledge of their preferences, and the more uncertain is our prediction of how our actions may affect them (if at all), and even of whether or not they exist or will come to exist. For these reasons, even if we are in principle spatiotemporally egalitarian, I believe it is rational to discount what we believe to be the preferences of remote people by a factor proportional to their remoteness. This is no doubt not acceptable to strict extended utilitarians such as Singer, but Stone (1987) has produced two defences that I find convincing. In any case, I think it can be argued that the preferences of the spatiotemporally remote would have little effect on the outcome already suggested for the Coronation Hill mine decision. The impact of both the potential benefits and risks on planes U3 and 4 must decrease in proportion to remoteness. There is no reason, a priori, to suggest that the pattern of preferences of remote people will be different from those already considered in any predictable way. If we consider spatiotemporally remote nonpersons in the same way, as I believe we must, then we have merely an increasingly diffuse zone shading away from the immediate, that will not predictably alter the balance of utilities. There is only one circumstance in which this would not be the case; that in which the predicted benefits and risks diffuse at different rates as spatiotemporal remoteness increases. This would, for instance, be the case for a uranium mine; but not, I suggest, for a gold mine. Analysis on planes U3 and U4, therefore, weakly reinforces the conclusions from the combination of U1 and U2.

The foundation assumption in all Stone's nonutility planes is a principle I argued on extended rational choice grounds, that moral values are not exclusively utilitarian, but that on occasion obligations and rights may evolve that take priority over summed welfare. For instance, in our society we have an obligation not to kill an intellectually or physically handicapped child, or a frail old person, even though killing them may increase total welfare by reducing the cost of supporting nonproductive members of the community. I suggested such obligations evolve socially as norms of respect

or altruism towards the moral patients, but they are, of course, readily derived from Rawlsian or Kantian principles.

The first of Stone's nonutility planes, N1, is based on respect for (or inherent value of) all persons. In the Coronation Hill controversy, it appears to me that this requires some degree of relativism as well as pluralism, and recognition of conflicting interests. It requires, in particular, that mutual respect between Aboriginal and non-Aboriginal Australians within a combined moral community, means that each must respect as legitimate within the other subculture arguments based on beliefs and values that are not valid within their own subculture. There are at least three legitimate competing moral claims, based on different constituencies, interests and values. Miners had a legitimate expectation to be allowed to mine, based on the original plan for mineral exploration within the Conservation Zone and an original government intention that mining be permitted within a defined period if important finds were made, and on the precedent of the previous decision to allow mining at Ranger. The Jawoyn had legitimate expectations based on the outcomes of the Woodward Commission, the Ranger Uranium Enquiry, the Senate enquiry, and the administration of previous stages of the Park, that mining would not be allowed to interfere with their privacy, hunting rights or land they considered sacred or part of the bula cycle. Conservationists had legitimate expectations on the same basis that nothing would be allowed to significantly threaten the conservation of the Park or the South Alligator river catchment, or further degrade the Conservation Zone in a way that would interfere with its rapid incorporation into Stage 3 of the Park. It should be clear that no decision would be able to simultaneously fulfil all these expectations to the satisfaction of those holding them, all of whom regarded their own expectations as presumptive rights.

I would argue that the Jawoyn have the strongest moral case, based on traditional ownership, the importance of the decision to what they have managed to preserve or rebuild of their traditional way of life, and the extreme vulnerability of that way of life. The next strongest case is, I suggest, that of non-Aboriginal conservationists, based on the high conservation and heritage value of the Park, its uniqueness, and vulnerability to damage and contamination caused by mining. The weakness

of the case is the limited direct conservation value of the region in which mining would actually occur, and the fact that there is a risk rather than a certainty of damage to more important regions. There is in the overall Kakadu region, unlike the Bloomfield region discussed in Section 5.1, considerable cooperation between the Gagudju and non-Aboriginal conservationists and the NPWS, at least as reported by Breeden and Wright (1989), but it is unclear to what extent this is true of the Jawoyn (SSCERA, 1988). I suggest that arguments on Aboriginal ownership and conservation grounds are certainly consistent with each other and may be partially supportive, strengthening the case against mining. The case for mining is based mainly on legitimate expectations flowing from the historical priority of mining over other forms of land use, legislation and government reports and statements. However, it is, I suggest, morally weaker. While the obligations to respect Jawoyn traditional ownership and conserve the Park are of a strength that is best expressed by saying that we ought to do these things, the strongest case for mining establishes only that there was an expectation that it might be permitted, not that we ought to mine. For these reasons, I suggest that analysis on plane N1 comes down firmly against mining. However, the mining joint venturers have a strong enough moral claim that a decision against their interests ought to be accompanied by some compensation for the loss of their investment.

Stone's second nonutility plane, N2, considers people remote in time and space. In this case we ask are there grounds for a moral obligation to remote people either to mine or not to mine, irrespective of the utility arguments on planes U2 and U3? I can see no grounds for an obligation to mine. It could be argued that we have obligations to future generations of Jawoyn not to interfere with the land they will inherit in a way that will limit their enjoyment of it as a birthright, but our best guide to whether or not such an obligation includes an obligation not to mine must be contemporary Jawoyn. Likewise, it could be argued that obligations to spatiotemporally remote people under our heritage and conservation responsibilities for the Park indicate that perhaps we should not mine. However, it seems to me that these obligations are much weaker than those considered under N1, and Stone came to a similar conclusion in the Beaufort context.

On plane N3, we ask if we have moral obligations to living nonpersons (animals and birds, assumed to have some sentience and preferences) that might affect our mining decision, and if so, how strong they might be. I am sympathetic to arguments of Stone (1987) that at least higher animals are morally considerate, but should be treated on a different moral plane by a different moral governance, because they differ from humans in morally salient ways that cannot be captured by a single monistic principle. The lower the form of life, the more we might question whether a given creature should be morally considerate at all, rather than whether we are justified in treating it differently from a person. This obviously requires debate that will continue long into the future. I shall in the meantime base my analysis on the theory of moral considerability developed in 4.9, and the empirical observation, which is consistent with the theory, that both in Aboriginal and non-Aboriginal cultures that are relevant to the Coronation Hill controversy, some animals are held to be morally considerate in the sense that all or some people have obligations not to harm them in particular ways, but that these obligations are highly context dependent and always weaker than equivalent obligations not to harm people. It might be that future philosophical inquiry will establish universal obligations to animals of exactly the same kind and strength as obligations to humans; but I remain sceptical, and believe the onus is on those who propose such extensive obligations to convince others, rather than vice versa.

I assume then, that many people within non-Aboriginal as well as Aboriginal communities relevant to the Coronation Hill decision hold that there are obligations not to harm some of the indigenous animals and birds within the Park, although these obligations are more context dependent and have less moral force than similar obligations with respect to people. Obligations along these lines are in fact incorporated in some Acts of Parliament, the NPWS plan of management and the report of the Senate Committee (SSCERA 1988). These imply that the obligations are of sufficient weight that some people have a duty to ensure, so far as possible, that any actions they take within the Park area will not cause unintentional harm to wildlife. The main question with respect to proposed mining at Coronation Hill is whether or not this duty would be discharged by careful

management and the construction of a tailings dam and retention pond. My argument on plane N3 would be that the duty would be better discharged by not mining, than by mining with these precautions, because there is no possibility of harm in the first case, but a risk in the second. However, I think this establishes only that not mining is morally preferable on N3 to mining with the best possible precautions; it does not establish that we should not or must not mine. The obligation not to endanger wildlife would argue much more strongly against any further mining, even if the Coronation Hill mine were allowed, for the probability of harm to wildlife is proportional to the number and extent of mining activities.

There are further complications on N3. The inquiries and Acts mentioned above recognise Aboriginal rights to hunt within the Park; clearly within that culture, as in non-Aboriginal culture, the moral considerability of sentient creatures is context dependent and guided by a complex set of conventions that does not preclude killing some of them for food. Moreover, in both subcultures the killing of feral animals is seen as not only permissible, but desirable, and a cooperative campaign to reduce buffalo numbers has been carried out. However, there is pressure from the Gagudju Association to maintain a herd of disease-free buffalo for their hunting (SSCERA). This again suggests that indigenous people, like other people, recognise limits to the applicability and moral force of obligations to sentient creatures.

On plane N4, we consider possible obligations to preferenceless and nonsentient entities. Many people would hold that plants, for instance, although they have interests lack sentience, and cannot be morally considerable except in so far as they are useful to higher forms of life. Considered with respect to traditional moral theories, this may be so. However, I believe there are circumstances in which norms of respect or altruism towards plants evolve, and they become morally considerable in their own right. For instance, Green Bans in Sydney prevented the destruction of a number of trees to make way for a car park for the Opera House. Plants, then, only become morally considerable under unusual circumstances, and the force of any obligations towards them is weak. But although the crushing of a flower is not as wrong as the crushing of a child, there are circumstances in which it is clearly morally

preferable to make choices that avoid damage to plants (Stone, 1987). It seems to me plausible to argue that we carry such an obligation to the indigenous plants of the Kakadu National Park. Again, I would point out the context dependence of any such obligation; as with feral animals, most people would accept that not only is it permissible to destroy weeds within the Park, but that we have an obligation to do so in order to protect indigenous flora from indirect consequences of our actions. The outcome of analysis of possible obligations to plants on N4 is thus again that it is morally preferable not to mine, than to mine with precautions, but the moral force of the conclusion is weak.

The evolution of norms of respect for things, in my opinion produces stronger support for their possible moral considerability than other moral theories. Stone (1987) based such considerateness on intrinsic good, and on 'the beautiful, the majestic, the rare, the untouched, the intricately complex and the profoundly simple'. Certainly parts of the landscape of Kakadu National Park may be all of these things, but again I suggest that what is considered to be beautiful or majestic, for instance, is highly influenced by culture. Stone argues that all these criteria may be above mere convention. I would point to the power in Jawoyn tradition of the belief in bula sites and of the presence of three major bula sites, and five lesser sites associated with the bula cycle, in the Conservation Zone and adjacent areas of Stage 3 (SSCERA, 1988). One of these sites is located on Coronation Hill, but it is probably not one of the most important (SSCERA, 1988). I doubt if any of the sites fit Stone's criteria, and Coronation Hill certainly does not. Their considerability is not based on a belief that they are intrinsically good (or bad), but rather on a belief that if they are not respected great harm will result. Arguably, the Alligator Rivers have a similar, but less important, status in the non-Aboriginal belief/value system, which fits Stone's criteria better. In this case the catchment of the South Alligator River, which flows through the Conservation Zone, is perceived as valuable but vulnerable, and if it is not respected, harm to the river system will result. The moral considerability of each of these things evolved because of a pattern of beliefs and values within one of the two different subcultures, but the things have come to be respected in part for themselves, and obligations towards them are now recognised within

both subcultures. Moreover, it seems to me that these obligations have more moral force than those already recognised under N2, 3 and 4; not only is it morally preferable that we do not disturb a bula site, or contaminate the South Alligator River, but we ought not to do these things. Some mining advocates have pointed out that a few Jawoyn people have ventured onto Coronation Hill, and some have worked on the site. I suggest this is entirely consistent with the rational choice account of the nature and evolution of norms (Pettit, 1990) within the context of Aboriginal culture in a process of rapid change.

There is to my mind a gradation between the entities Stone considers on N4 as things; and those held considerate on N5, membership entities such as species or ecosystems; and qualities considered on N6, such as ecosystem stability or biodiversity. I have already subscribed to the view that any moral considerability of such entities cannot depend on their supposed possession of interests, or analogy with an organism (Section 3.1; Cahen, 1988). Stone recognised their moral considerability largely because some similar entities and qualities are recognised in law. I would suggest that moral norms about things, classes and qualities evolve when there is a perceived need for their consideration within a given ecological situation and cultural context. Stone, however, believes their considerability is more than a matter of convention. Clearly this area needs clarification. I am content to observe that things, membership entities and qualities often appear in everyday moral discussions, and suggest that they are also relevant in the example considered here. It would be an incomplete account of value issues in the Coronation Hill controversy that took no account of (say) ecosystem stability, endangered species or biodiversity. The big advantage of considering as morally considerable on a separate plane membership entities such as species and ecosystems is that it allows value to be located in, and obligations to be held towards, both wholes, and the individuals of which they are composed. Moreover, in law and moral life, membership entities may be subject to a different governance from that which guides interactions of individuals (Stone 1987). What is not clear in the Coronation Hill example, is how many such entities and qualities should be considered, and how. As a start, we could locate on action-in-

fluence maps of the Park all known species and ecosystems, and ask if we might plausibly recognise obligations to any of them with respect to the mining decision. The most vulnerable natural ecosystems would be those bordering the South Alligator River, and a number of species might be lost from the Conservation zone itself if mining occurred. I suggest that it is plausible to recognise their moral considerability, and argue that this suggests that on this basis not mining is morally preferable to mining, and perhaps that there is a consequent obligation not to mine stronger than that derived from the considerability of individual creatures or plants. Since everything is connected to everything else, we might suggest similar obligations with respect to other ecosystems and species; but as in the case of earlier planes, these must be weaker than obligations to ecosystems and species in zones of high hazard, and such obligations must decrease with spatiotemporal distance.

There is much to be done in clarifying and improving pluralistic moral analysis of the complex decisions necessary in the evolution of ecological sustainability. Nevertheless, I think its application to this example has been instructive. The most obvious result is that in my opinion, the only plane on which a decision to mine is indicated is U1, the utility plane in which only (contemporary) people are morally considerate, and their morality governed only by aggregate welfare. On all other planes, there are either strong arguments for not mining, or not mining is morally preferable to mining, and a strong overall case for not mining can be made. It also appears to me that analysis in terms of evolving social norms has complemented the application of utilitarian and deontological theories. Analysis on some planes seems to have been more useful than others. Further work on, and experience with, this type of analysis may allow the development of tools in ecological decisionmaking that go beyond the narrow confines of traditional cost-benefit analysis on plane U1

6. Coda

The position I have come to is roughly this: values, facts and the theories on which they depend are all in essence conventions of truth and

value in a given human population at a given time. These are expressed in the language of the population, which is itself a set of conventions based on a metaconvention of truthfulness in that language. All of these are instrumental to the Good of the population. In so far as it is individual, the Good consists of a complex set of often conflicting motives ultimately grounded in the biologically evolved needs, desires and drives of the individual, which have to be balanced against each other, and against collective goods. In so far as it is collective, the Good consists of limited altruistic motives of the individuals, and a complex set of conventions of goodness that coordinate collective goals that are perceived as being in the long term enlightened self-interest of the individuals, and which depend on cultural history. There are no standards or sources of empirical truth, morality or value external to the population, whether or not there is a convention within the population that there are. In spite of the enormous diversity in human moralities that exist or have existed, there are constraints on what morality can evolve within a given population at a given time. Some of these constraints have biological roots, while others are culturally determined. Likewise, although it is also a social construct, what is considered valid empirical knowledge by a given population at a given time must be constrained by observations of reality by members of the population, if the knowledge is to be of any use to the population in its interactions with the real world.

If this analysis is at all right, environmental ethics is, and can only be, based on a paradigm of reality and observations of nature by the population, and a set of norms of respect for nature, which are perceived at a given time and in a given cultural context to be in the broad long term enlightened self-interest of the individuals in a moral population, and which evolve slowly within that cultural context. Although morality is a social construct, we cannot simply construct whatever we like and expect it to work, without taking account of the biological and cultural heritage of the human population in which it is to apply. Although many would prefer to believe that an environmental ethic can be derived *de novo* from some unassailable premise by some logically rigorous process, and demonstrated to be valid, there seems little likelihood that this will happen. Rather such an ethic will evolve by a gradual social process, and be grounded in existing cultural traditions. The key events in such

an evolution will be, I suggest, a change in social perception of the human situation in the real world to one that recognises we are part of nature and inescapably dependent on it for the quality and sustainability of human life, and a concomitant change in the social vision of what constitutes the good life. Appropriate dependent values, norms, rules and laws will follow.

Acknowledgements

I thank the Australian National University for support for this project; particularly Professor Ian Ross when Deputy Vice Chancellor for being persuaded that such a risky investment was worthwhile, and for suggesting that I talked to Philip Pettit. I thank Henry Nix and the staff of CRES for hospitality, and Stephen Boyden for support, encouragement, advice, ideas and discussions, and for inviting me to participate in the FQ program. I thank the following for discussions that influenced my viewpoint, in some cases by forcing me to face issues I would otherwise have ignored: Frances Barnes, Margarita Bowen, Peter Cock, Mick Common, Stephen Dovers, David Dumaresq, Robert Goodin, Paul Harvey, Robert May, Kingsley Palmer, Phillip Pettit, Vernon Reynolds, Helen Ross and Ted Steele.

References

- Alexander, R.D. 1987. *The biology of moral systems*. New York, Aldine de Gruyter.
- Anderson, C. 1989. Aborigines and conservationism: the Daintree-Bloomfield road. *Australian Journal of Social Issues* 24: 214-227.
- Attfield, R. 1983. *The ethics of environmental concern*. New York, Columbia University Press.
- Axelrod, R.M. 1984. *The evolution of cooperation*. New York, Basic Books.
- Axelrod, R.M. and Hamilton, W.D. 1981. The evolution of cooperation. *Science* 211: 1390-1396.
- Bennett, D.H. 1990. *Ecological sustainability, deep environmental ethics, and Tao*. Fundamental Questions Paper. Canberra, Centre for Resource and Environmental Studies, Australian National University (In press).

- Barnes, F. 1990. *Human nature and an ecologically sustainable world*. Fundamental Questions Paper. Canberra, Centre for Resource and Environmental Studies, Australian National University (In press).
- Bennett, D.H. & Sylvan, R. 1989. *Australian perspectives on environmental ethics*. UNESCO Project.
- Bicchieri, C. 1990. Norm of cooperation. *Ethics* 10: 838-861.
- Bookchin, M. 1987. *The modern crisis*. Montreal, Black Rose.
- Bowen, M. 1985. The ecology of knowledge: linking the natural and social sciences. *Geoforum* 16: 213-225.
- Boyden, S.B. 1987. *Western civilization in biological perspective: patterns in biohistory*. Oxford, Oxford University Press.
- Boyden, S.B. 1990. *The Fundamental Questions program and its conceptual basis*. Fundamental Questions Paper No. 1. Canberra, Centre for Resource and Environmental Studies, Australian National University.
- Boyden, S., Dovers, S. & Shirlow, M. 1990. *Our biosphere under threat: ecological realities and Australian opportunities*. Melbourne, Oxford University Press (in press).
- Breeden, S. and Wright, B. 1989. *Kakadu - looking after the country the Gagudju way*. Brookvale, Simon and Schuster.
- Brentano, F. 1902. *The origin of the knowledge of right and wrong*. (C. Hague, trans.) London, Constable.
- Brown, R. 1965. *Social Psychology*. London, The Free Press.
- Cahen, H. 1988. Against the moral considerability of ecosystems. *Environmental Ethics* 10: 195 - 216.
- Calicott, J.B. 1982. Hume's is/ought dichotomy and the relation of ecology to Leopold's land ethic. *Environmental Ethics* 5: 163-174.
- Calicott, J.B. 1983. *Traditional American Indian and traditional Western European attitudes towards nature: an overview*. In Elliot, R. and Gare, A. (eds.) *Environmental philosophy. A collection of readings*. Philadelphia, Pennsylvania State University Press.
- Calicott, J.B. 1985. Intrinsic value, quantum theory and environmental ethics. *Environmental Ethics* 7: 257-275.
- Calicott, J.B. 1986. *On the intrinsic value of non-human species*. In: Norton, B.G. (ed.) *The preservation of species*. Princeton, Princeton University Press.
- Campbell, B.G. 1985. *Humankind emerging* (fourth edition). Boston, Little-Brown.
- Chalmers, A.F. 1982. *What is this thing called science?* (second edition). St. Lucia, University of Queensland Press.
- Chisholm, R.M. 1986. *Brentano and intrinsic value*. Cambridge, Cambridge University Press.
- Cock, P. 1990. *Values for sustainability: the necessity of transcendence and sacred realms*. Fundamental Questions Paper, Canberra, Centre for Resource and Environmental Studies, Australian National University (In press).
- Cox, S.J.B. 1985. No tragedy of the commons. *Environmental Ethics* 7: 49-61.
- Crawford-Brown, D.J. & Pearce, N.E. 1989. Sufficient proof in the scientific justification of environmental actions. *Environmental Ethics* 11: 153-167.
- Darwin, C. 1859. *The origin of species; by means of natural selection or the preservation of favoured races in the struggle for life*. (1962 edition) New York, Collier.
- Darwin, C. 1883. *The descent of man, and selection in relation to sex*. (second edition). London, John Murray.
- Dawkins, R. 1976. *The selfish gene*. Oxford, Oxford University Press.
- Elliot, R. 1984. Rawlsian justice and non-human animals. *Journal of Applied Philosophy* 1 : 95-106.

- Feyerabend, P.K. 1975a. *Against method: outline of an anarchistic theory of knowledge*. London, New Left Books.
- Feyerabend, P.K. 1975b. How to defend society against science. *Radical Philosophy* 11: 3-8.
- Foley, R.A. and Lee, P.C. 1989. Finite social space, evolutionary pathways, and reconstructing hominid behaviour. *Science* 243: 901-906.
- Fox, W. 1989. The deep ecology-ecofeminist debate and its parallels. *Environmental Ethics* 11: 5-25.
- Frankena, W. 1979. *Ethics and the environment*. In: Goodpature, K.E. and Sayre, K.M. (eds.) *Ethics and problems of the twenty-first century*. Notre Dame, University of Notre Dame Press.
- Fuchs, A.E. 1981. Duties to animals: Rawls' alleged dilemma. *Ethics and Animals* 2: 83-87.
- Goldsmith, E. 1988. The way: an ecological world view. *The Ecologist* 18: 160-185.
- Goodpasture, K.E. 1978. On being morally considerable. *Journal of Philosophy* 75: 308-325.
- Goodpasture, K.E. and Sayre, K.M. (eds.) 1979. *Ethics and problems of the twenty first century*. Notre Dame, University of Notre Dame Press.
- Hamilton, W.D. 1964. The genetical evolution of social behaviour. *Journal of theoretical Biology* 7: 1-52.
- Hardin, G. 1968. The tragedy of the commons. *Science* 162: 1243-1248.
- Hardin, R. 1971. Collective action as an agreeable n-prisoners' dilemma. *Behavioural Science* 16: 472-481.
- Hardin, R. 1982. *Collective action*. Baltimore, Johns Hopkins University Press.
- Harman, G. 1977. *The nature of morality: an introduction to ethics*. New York. Oxford University Press.
- Hume, D. 1739. *A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects. Book I, Of the understanding*. (Selby-Bigge, L.A. ed., second edition 1978) Oxford, Clarendon Press.
- Hume, D. 1740a. Ibid. Book III. *Of morals*.
- Hume, D. 1740b. Ibid. *Abstract of the treatise: a book lately published entitled a Treatise of Human Nature & C.*
- Krige, J. 1980. *Science, revolution and discontinuity*. Brighton, Harvester.
- Kuhn, T.S. 1970. *The structure of scientific revolutions*. Chicago, University of Chicago Press.
- Lakatos, I. 1974. *Falsification and the methodology of scientific research programmes*. In Lakatos, I. and Musgrave, A. (eds.) *Criticism and the growth of knowledge*. Cambridge, Cambridge University Press.
- Lewis, D.K. 1969. *Convention. A philosophical study*. Cambridge, Mass., Harvard University Press.
- Lovejoy, C.O. 1981. The origin of man. *Science* 211: 341-350.
- Lovelock, J.E. 1979. *Gaia: a new look at life on earth*. New York, Oxford University Press.
- McDonnell, G. 1990. *Environmental management and the political philosophy of trust*. Fundamental Questions Paper, Canberra, Centre for Resource and Environmental Studies, Australian National University (In press).
- Manning, R. 1981. Environmental ethics and Rawls' theory of justice. *Environmental Ethics* 3: 155-166.
- Martin, R. 1985. *Rawls and rights*. Lawrence, Kansas, University Press of Kansas.
- May, R.M. 1978. The evolution of ecological systems. *Scientific American* 239: 118-133.
- May, R.M. 1987. More evolution of cooperation. *Nature* 327: 15-17.
- May, R.M. 1988. How many species are there on earth? *Science* 241 : 1441-1449.
- Medvedev, Z.A. 1969. *The rise and fall of T.D. Lysenko*. New York, Columbia University Press.
- Moore, G.E. 1903. *Principia ethica*. Cambridge, The University Press.
- Murphy, P.D. 1988. Sex-typing the planet: Gaia imagery and the problem of subverting patriarchy. *Environmental Ethics* 10: 155-168.

- Norton, B. G. (ed). 1986. *The preservation of species: the value of biological diversity*. Princeton, Princeton University Press.
- Palmer, K. 1990. *Aboriginies, values and the environment*. Fundamental Questions Paper, Canberra, Centre for Resource and Environmental Studies, Australian National University (In press).
- Parker, S.T. 1987. A sexual selection model for hominid evolution. *Human evolution* 2: 235-253.
- Passmore, J. 1974. *Man's responsibility for nature*. New York, Scribner.
- Pettit, P. 1986. Free riding and foul dealing. *Journal of Philosophy* 83: 361-379.
- Pettit, P. 1990. Virtus normativa: rational choice perspectives. *Ethics* 100: 725-755.
- Popper, K.R. 1959. *The logic of scientific discovery*. London, Hutchinson.
- Olson, M. 1965. *The logic of collective action*. Cambridge, Mass., Harvard University Press.
- Quine, W.V.O. 1961. *Two dogmas of empiricism*. In From a logical point of view. New York, Harper & Row.
- Rawls, J. 1972. *A theory of justice*. Cambridge, Mass., Harvard University Press.
- Rolston, H. 1988. *Environmental ethics. Duties to and values in the natural world*. Philadelphia, Temple University Press.
- Routley, R. and Routley, V. 1980. *Human chauvinism and environmental ethics*. In Marrison, D. McRobbie, M., and Routley, R.(eds). *Environmental Philosophy*. Canberra, Research School of Social Sciences, Australian National University.
- Russell, B. 1946. *History of Western philosophy and its connection with political and social circumstances from the earliest times to the present day*. London, George Allen & Unwin.
- Salleh, A. 1989. Stirrings of a new renaissance. *Island Magazine* 38: 26-31.
- Schelling, C. 1960. *The strategy of conflict*. Cambridge, Mass., Harvard University Press.
- Senate Standing Committee on Environment, Recreation and the Arts, The Parliament of the Commonwealth of Australia. 1988. *The potential of the Kakadu National Park region*. Canberra, Australian Government Publishing Service.
- Serafin, R. 1988. Noosphere, Gaia and the science of the biosphere. *Environmental Ethics* 10: 121-137.
- Singer, B.A. 1988. An extension of Rawls' theory of justice to environmental ethics. *Environmental Ethics* 10: 217-231.
- Skolimowski, H. 1988. Eco-philosophy and deep ecology. *The Ecologist* 18: 124 -127.
- Smith, A. 1759. *The theory of moral sentiments*. In: Schneider, H.W. (ed.). 1948. *Adam Smith's moral and political philosophy*. New York, Hafner.
- Sober, E. 1986. *Philosophical problems with environmentalism*. In: Norton, B. G. (ed.) *The preservation of species*. Princeton, Princeton University Press.
- Steele, E.J. 1989. Mechanism of directional mutations? *Mol. Reprod. Dev.* 1, 231-232.
- Steele, E.J. 1990. Scientific forward. In Green, D. (Ed.) *Descent of spirit : writings of E. L. Grant Watson*. Sydney, Primavera Press.
- Stone, C.D. 1987. *Earth and other ethics: the case for moral pluralism*. New York, Harper and Row.
- Stone, C.D. 1988. Moral pluralism and the course of environmental ethics. *Environmental Ethics* 10: 139-154.
- Sugden, R. 1984. *Social contract theory*. Oxford, Blackwell's.
- Sugden, R. 1986. *The economics of rights, cooperation and welfare*. Oxford, Blackwell's.
- Sylvan, R. and Bennett, D. 1988. Taoism and deep ecology. *The Ecologist*. 18: 148-159.
- Taliaferro, C. 1988. The environmental ethics of an ideal observer. *Environmental Ethics* 10: 233 - 250.
- Taylor, M. 1987. *The possibility of cooperation*. Cambridge, Cambridge University Press.
- Teilhard de Chardin, P. 1965. *Building the earth*. Wilkes-Barre, Pennsylvania, Dimension Books.

- Tokar, B. 1988. Social ecology, deep ecology and the future of green politics. *The Ecologist* 18: 132-141.
- Toulmin, S.E. 1958. *The uses of argument*. Cambridge, The University Press.
- Trivers, R.L. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology*. 46: 35-37.
- Tudge, C. (ed.) 1988. *The encyclopedia of the environment*. London, Christopher Helm.
- Weston, A. 1985. Beyond intrinsic value: pragmatism in environmental ethics. *Environmental Ethics* 7: 321-339.
- Williams, G.C. 1966. *Adaptation and natural selection: a critique of some current evolutionary thought*. Princeton, Princeton University Press.
- Wilson, E.O. 1975. *Sociobiology; the new synthesis*. Cambridge, Massachusetts, Belknap.
- Wright, J.C. 1988. *Future generations and the environment*. Studies in resource management No. 6. Canterbury, New Zealand, Centre for Resource Management.
- Zimmerman, M.E. 1988. Quantum theory, intrinsic value, and pantheism. *Environmental Ethics* 10: 3-30.

1745079



A.N.U. LIBRARY

THE AUSTRALIAN NATIONAL UNIVERSITY
The Library

This book is due on:

CANCELLED
22 AUG 1991

15 SEP 1994
CANCELLED

CANCELLED
4 APR 1991

CANCELLED
6 MAY 1991

CANCELLED
27 JUN 1991

CANCELLED
3 AUG 1991

CANCELLED

CANCELLED
20 APR 1991

18 AUG 1994
CANCELLED