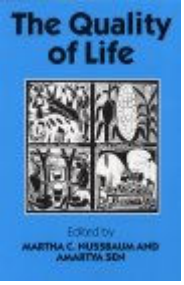


**Nussbaum, Martha** (Editor) , Professor of Law and Ethics , University of Chicago
**Sen, Amartya** (Editor) , Master of Trinity College, Cambridge

# The Quality of Life

Print ISBN 0198287976, 1993

## Contents

List of contributors ix
Introduction
*Martha Nussbaum and Amartya Sen* 1
PART I LIVES AND CAPABILITIES
Equality of What? On Welfare, Goods, and Capabilities
*G. A. Cohen* 9
Capability and Well-Being
*Amartya Sen* 30
Commentators on Cohen and Sen: *Christine M. Korsgaard and Wulf Gaertner* 54
Descriptions of Inequality: The Swedish Approach to Welfare Research
*Robert Erikson* 67
Commentator: *Bengt-Christer Ysander* 84
Having, Loving, Being: An Alternative to the Swedish Model of Welfare Research
*Erik Allardt* 88
Quality of Life Measures in Health Care and Medical Ethics
*Dan Brock* 95
Commentator: *James Griffin* 133
PART IITRADITIONS, RELATIVISM, AND OBJECTIVITY

Objectivity and the Science-Ethics Distinction
*Hilary Putnam* 143
Commentator: *Lorenz Krüger* 158
Objectivity and Social Meaning
*Michael Walzer* 165
Commentator: *Ruth Anna Putnam* 178
Value, Desire, and Quality of Life
*Thomas Scanlon* 185

end p.vii

Commentator: *Sissela Bok* 201
Explanation and Practical Reason
*Charles Taylor* 208
Commentator: *Martha Nussbaum* 232
Non-Relative Virtues: An Aristotelian Approach
*Martha Nussbaum* 242
Commentator: *Susan Hurley* 270
PART III WOMEN'S LIVES AND GENDER JUSTICE
Women and the Quality of Life: Two Norms or One?
*Julia Annas* 279
Commentator: *Margarita M. Valdés* 297
Justice, Gender, and International Boundaries
*Onora O'Neill* 303
Commentator: *Martha Nussbaum* 324

Commentator: *Sissela Bok* 201
PART IV POLICY ASSESSMENT AND WELFARE ECONOMICS
Distributing Health: The Allocation of Resources by an International Agency
*John E. Roemer* 339
Commentator: *Paul Seabright* 358
The Relativity of the Welfare Concept
*B. M. S. van Praag* 362
Commentator: *Siddiq Osmani* 386
Pluralism and the Standard of Living
*Paul Seabright* 393
Commentator: *Derek Parfit* 410
Life-Style and the Standard of Living
*Christopher Bliss* 417
Commentator: *Amartya Sen* 437
Index 445

end p.viii

# Introduction

show chapter abstract and keywords
hide chapter abstract and keywords

**Martha Nussbaum**
**Amartya Sen**

'And he said, Now, this schoolroom is a Nation. And in this nation, there are fifty millions of money. Isn't this a prosperous nation? Girl number twenty, isn't this a prosperous nation, and a'n't you in a thriving state?'
'What did you say?' asked Louisa.
'Miss Louisa, I said I didn't know. I thought I couldn't know whether it was a prosperous nation or not, and whether I was in a thriving state or not, unless I knew who had got the money, and whether any of it was mine. But that had nothing to do with it. It was not in the figures at all,' said Sissy, wiping her eyes.
'That was a great mistake of yours,' observed Louisa.
(Charles Dickens, *Hard Times*)

When we inquire about the prosperity of a nation or a region of the world, and about the quality of life of its inhabitants, Sissy Jupe's problem still arises: How do we determine this? What information do we require? Which criteria are truly relevant to human 'thriving'? Girl number twenty quickly discerns that just knowing how much money is available for a given number of people (the analogue of GNP per capita, still widely used as a measure of quality of life) will not take us very far. For we also need, at the very least, to ask about the distribution of these resources, and what they do to people's lives.

The problem is actually more complex still. For if we are really to know much about the 'thriving' of Sissy Jupe and her fellow citizens, we need to know not only about the money they do or do not have, but a great deal about how they are able to conduct their lives. We need surely to know about their life expectancy (think of the miners of Coketown in Dickens's novel, who keep their families from 'want and hunger', but go to a premature death). We need to know about their health care and their medical services. We need to know about education-and not only about its availability, but about its nature and its quality (it seems likely that Mr Gradgrind's school may actually diminish the 'thriving' of its pupils). We need to know about labour-whether it is rewarding or grindingly monotonous, whether the workers enjoy

any measure of dignity and control, whether relations between employers and 'hands' are human or debased. We need to know what political and legal privileges the citizens enjoy, what freedoms they have in the conduct of social and personal relations. We need to know how family relations and relations between the sexes are structured, and how these structures foster or impede other aspects of human activity. We need, perhaps above all, to know how people are enabled by the society in question to imagine, to wonder, to feel emotions such as love and gratitude, that presuppose that life is more than a set of commercial relations, and that the human being-unlike the steam engines of Coketown-is an 'unfathomable mystery',

end p.1

not to be completely 'set forth in tabular form'. In short, to think well about Sissy's problem, we seem to need a kind of rich and complex description of what people are able to do and to be-a description that may be more readily available to the reader of Dickens's novel than to those who confine their reading to the narrowly technical and financial documents favoured by Sissy's teachers.

Economists, policy-makers, social scientists, and philosophers are still faced with this problem of measurement and assessment. They need to know how people are doing in many different parts of the world, and they need to know what is really involved in asking that question. When they face the problem well, they face it, so to speak, with wonder (to use, deliberately, the word least tolerated in Mr Gradgrind's school); with a sense, that is, of the profound complexity of assessing a human life, and with a desire to admit, at least initially, the widest possible range of accounts of how one might go about this, of what indicators one might trust.

Of course, it is possible to wonder not at all-to stick to a mechanical formula that is easy to use and which has been used before. The unasked question does not have to be answered. This volume is an attempt to ask questions and to propose and examine some possible answers. By examining the arguments for and against a variety of different accounts of how to measure quality of life, it aims to generate a more complex understanding of alternative positions and their respective merits. The original motivation for this project lay in our perception that these issues were being debated in several different fields whose communication with one another was unfortunately slight. Most social scientists and economists would agree with Sissy Jupe that GNP per capita is a crude and incomplete measure of quality of life; and yet such measures continue to be widely used when public policy is made. Again, philosophers have for some time been debating the merits of measuring the quality of human life in terms of utility (whether understood as happiness or as the satisfaction of desires or preferences). Some philosophers continue to defend this general approach-though usually with considerable qualification, producing utilitarianisms with complex and subtle restrictions on the nature of the preferences that may be taken into account. Others have concluded that the whole utilitarian approach should be rejected-to be replaced, perhaps, by an account of the many different kinds of activity that actually make up a 'thriving' human life. (Such a programme can take various forms and some of these are explored in the papers by Cohen, Sen, Brock, Scanlon, and Nussbaum.) The philosophical debates have not had much impact on the making of public policy in much of the world; nor have they been particularly noticed in the standard works in economics. Our hope was that by getting the participants in these debates together and encouraging further debate among them, we might advance the state of the question, encourage further co-operative inquiries, and present the debate in a form accessible not only to professionals in these academic disciplines but also to policy-makers and the general public.

end p.2

## Lives and Capabilities

The papers in Part I address the general questions we have already described, examining how we may try to find adequate criteria for assessing the quality of life. Cohen's and Sen's papers both discuss the 'capability approach' presented by Sen, but Cohen also examines a number of other approaches, including, among others, utilitarian calculations, the Rawlsian focus on primary goods, and Dworkin's use of resources. After criticizing these approaches, Cohen argues for concentrating on what he calls 'midfare', which is fairly close to the idea of functionings used in the 'capability' literature.

The life that a person leads can be seen as a combination of various doings and beings, which can be generically called functionings. These functionings vary from such elementary matters as being well nourished and disease-free to more complex doings or beings, such as having self-respect, preserving human dignity, taking part in the life of the community, and so on. The capability of a person refers to the various alternative combinations of functionings, any one of which (any combination, that is) the person can choose to have. In this sense, the capability of a person corresponds to the *freedom* that a person has to lead one kind of life or another. Cohen's paper discusses why one has to go in this direction (rather than staying with the traditional concerns: incomes, utilities, resources, primary goods), but he ends by questioning whether the freedom-type idea of capability is the precisely correct alternative. Sen presents the capability approach and its rationale, and also attempts to answer Cohen's critique (along with a few other criticisms presented elsewhere).

The papers by Erikson and Allardt discuss some methods and strategies for measuring the quality of life that have been used by Scandinavian social scientists for a long time. Their approach has much in common with that of focusing on functionings and capabilities, and the actual measurement techniques used have obvious relevance for the use of the capability approach. The philosophical underpinnings of the capability approach, on the other hand, provide some defence of the Scandinavian practices and also some suggestions for extension. While the Swedish and Finnish approaches differ in some significant respects, they agree in focusing neither on opulence nor on utility, but on the ways in which people are actually able to function, in a variety of areas. And they insist that any adequate measure of quality of life must be a plural measure, recognizing a number of distinct components that are irreducible to one another.

Brock's comprehensive study of measures of quality of life in the area of health care shows, among other things, still another convergence, as doctors and philosophers, looking for the best way to assess the quality of patients' lives, have increasingly turned to a list of functional capabilities, not unlike those proposed in the capability literature and in the theory and practice of Scandinavian social scientists. The field of health care provides a rich ground for comparing, contrasting, and assessing different approaches.

end p.3

## Traditions, Relativism, and Objectivity

When standards are chosen to assess the quality of life of people in different parts of the world, one has to ask whose views as to the criteria should be decisive. Should we, for example, look to the local traditions of the country or region with which we are concerned, asking what these traditions have regarded as most essential to thriving, or should we, instead, seek some more universal account of good human living, assessing the various local traditions against it? This question needs to be approached with considerable sensitivity, and there appear to be serious problems

whichever route we take. If we stick to local traditions, this seems to have the advantage of giving us something definite to point to and a clear way of knowing what we want to know (though the plurality and complexity of traditions should not be underestimated, as frequently happens in cultural-relativist accounts). It seems, as well, to promise the advantage of respect for difference: instead of telling people in distant parts of the world what they ought to do and to be, the choice is left to them. On the other hand, most traditions contain elements of injustice and oppression, often deeply rooted; and it is frequently hard to find a basis for criticism of these inequities without thinking about human functioning in a more critical and universal way. (For example, Dickens's criticism of British materialism and class injustice relies centrally on some extended reflections about what a human being is—reflections that, in their socially radical character, would not have been easy to extract from the traditions of English class relations.)

The search for a universally applicable account of the quality of human life has, on its side, the promise of a greater power to stand up for the lives of those whom tradition has oppressed or marginalized. But it faces the epistemological difficulty of grounding such an account in an adequate way, saying where the norms come from and how they can be known to be the best. It faces, too, the ethical danger of paternalism, for it is obvious that all too often such accounts have been insensitive to much that is of worth and value in the lives of people in other parts of the world and have served as an excuse for not looking very deeply into these lives.

The papers in Part II confront this difficult set of questions, exploring arguments about cultural relativism from a variety of viewpoints and showing the bearing of these arguments on questions about the quality of life. Putnam's paper attacks the widely accepted idea that questions of science are matters of fact, and ethical questions, by contrast, are questions in resolving which we can appeal to nothing beyond subjective preferences. He argues that 'fact' in science is a more anthropocentric and 'internal' matter than is frequently realized, and that the very same sort of truth and objectivity that is available in science is available in ethics as well. Walzer, by contrast, defends a qualified form of relativism, in which, however, much attention is paid to the various forms of criticism and discontent that may arise within a society. However inarticulate these dissenting voices may be, they are, he argues, to be taken as important

end p.4

data in understanding what a good life would be for the society as a whole. Scanlon's paper, closely related to the debate about utility in Part I, questions the adequacy of desire as a measure of quality of life, arguing instead for exploring an approach based on a critical scrutiny of a 'substantive list' of elements that make human life valuable. Taylor examines the forms of reasoning that people use when they argue that one way of life is better for human beings than another. Although such arguments always take place in a particular historical context and do not have the deductive structure that we have come to associate with good scientific argument, they can none the less, he claims, be perfectly reasonable, and can succeed in showing that some ways of doing things are indeed better than others. Nussbaum examines one major account of the quality of life in terms of a list of basic human functions-Aristotle's-that does claim to have validity for all human beings. She argues that although Aristotle himself did not confront the subtle arguments cultural relativists today might use against his account-for example, the argument that even such basic human experiences as reasoning and desiring are constructed and experienced differently in different societies-his account can still be made to reply

convincingly to such objections. (The Aristotelian list of functions converges surprisingly, once again, with the proposals of Sen, Erikson, Allardt, and Brock in Part I, although they emanate from different intellectual traditions.)

## Women's Lives and Gender Justice

In no area are there greater problems about measuring quality of life than in the area of women's lives and capabilities. The question of whether utility is an adequate measure and the question of cultural relativism take on a special urgency here. For in most parts of the world women do not have the same opportunities as men. These inequalities-and the deficiencies in education and experience often associated with them-tend to affect women's expectations and desires, since it is difficult to desire what one cannot imagine as a possibility. Desire-based approaches to measuring quality of life frequently end up, for this reason, affirming the status quo, informing us, for example, that the women of country *Q* have no need of literacy because, when investigated by the authorities of *Q*, they do not express an unsatisfied desire for literacy. An approach based on a substantive account of human capabilities would ask different questions here, and would probably arrive at a different recommendation. Again, our solution to the problem of cultural relativism will have especially clear implications where women's lives are concerned, for most local traditions oppress women. A universal account of human functioning seems to have greater critical potential here. (Current work in progress at WIDER will press these issues further, attempting to construct a non-relative account of human capabilities in the context of questions about women.)

end p.5

But as we try to answer such questions, we also need to decide whether the quality of a female life has the same constituents as the quality of a male life. Some respected philosophical answers to this question (for example, that of Jean-Jacques Rousseau), while universalist rather than relativist about the good, have divided humanity into two distinct 'natures', with different norms and goals for each. If Rousseau's Emile were found leading the life that is judged best for his consort Sophie, tending the house and caring for children, Rousseau would judge his quality of life to be low indeed; so too for Sophie, were she to be discovered (*per impossible*, as Rousseau would have it) exercising the virtues of citizen autonomy. Some contemporary feminist writing tends in a similar direction. In Part III, Annas's paper examines this question, defending a single account of human functioning. O'Neill partly concurs, and describes a strategy we might use to judge that a particular practice is unjust to women. She links the question of gender justice to related questions about justice across international boundaries.

## Policy Assessment and Welfare Economics

Part IV explores a number of ways in which questions about quality of life arise in welfare economics (broadly defined) and the formation of public policy. Roemer studies the criteria of resource allocation used by the World Health Organization, comparing them with some criteria proposed by economists and philosophers. The axioms implicit in practical policies can be made explicit and then illuminatingly contrasted with rival demands. Van Praag analyses the results of some empirical studies of how people perceive their well-being. He shows that perceptions are highly relative to surrounding social issues.

Seabright examines the problem of living standards in terms of the plural influences that call for accommodation. He explores the connection between contractarian justifications of social theories and pluralist conceptions of the social good.

Finally, Bliss discusses the measurement of the standard of living, arguing (against Sen) that this question should be separated from the broader question of quality of life. He presents, partly by implication, a partial defence of more orthodox measures of living standards, in particular real income, but develops his own line of departure in terms of the concept of 'life-style', the contents and implications of which he investigates extensively.

The contributions of the commentators at the original WIDER conference were, in general, of such a high order that it was decided to publish them along with the papers. Although this introduction has not commented on their contents, they are an integral part of the philosophical and practical debate of which this volume is the outcome.

end p.6

## Part I Lives and Capabilities

end p.7
end p.8

# 1 Equality of What? on Welfare, Goods, and Capabilities

show chapter abstract and keywords
hide chapter abstract and keywords

**G. A. Cohen**

## 1 Introduction

In his Tanner Lecture of 1979 called 'Equality of What?' Amartya Sen asked what metric egalitarians should use to establish the extent to which their ideal is realized in a given society. What aspect(s) of a person's condition should count in a *fundamental* way for egalitarians, and not merely as cause of or evidence of or proxy for what they regard as fundamental?

In this study I comment on answers to Sen's question in recent philosophical literature. I take for granted that there is something which justice requires people to have equal amounts of, not no matter what, but to whatever extent is allowed by values which compete with distributive equality; and I describe and criticize what a number of authors who share that egalitarian view have said about the dimension(s) or respect(s) in which people should be made more equal, when the cost in other values of moving towards greater equality is not intolerable.[1]

[1] I discuss conflict between equality and other values in section 1 of 'On the Currency of Egalitarian Justice', *Ethics*, 99 (July 1989). That article represents part of the contribution which I prepared for the July 1988 Helsinki WIDER symposium on 'The Quality of Life', and the present article is also drawn from that contribution. There is overlap between the articles in that a similar critique of Rawls appears in both. Discussions of Dworkin and Scanlon which appear in the *Ethics* article are not reproduced here.

The publication of John Rawls's *A Theory of Justice* in 1971 was a watershed in discussion bearing on the question, derived from Sen, which forms my title. Before *A Theory of Justice* appeared, political philosophy was dominated by utilitarianism, the theory that sound social policy aims at the maximization of welfare. Rawls found two features of utilitarianism repugnant. He objected, first, to its aggregative character, its unconcern about the pattern of distribution of welfare, which means that inequality in its distribution calls for no justification. But, more pertinently to the present exercise, Rawls also objected to the utilitarian assumption that welfare is the aspect of a person's condition which commands normative attention. Rawls replaced aggregation by equality and welfare by primary goods. He recommended normative evaluation with new arguments (goods instead of welfare quanta) and a new function (equality[2]

[2] Or, strictly, the maximin function, which enjoins departures from equality when the worst off benefit as a result. But that complication is of no significance here.

instead of aggregation) from those arguments to values.

Rawls's critique of the welfare metric was undoubtedly powerful, but, as I



shall argue, his motivation for replacing it by attention to primary goods was not correspondingly cogent. He did not consider, as an alternative to equality of welfare, the claims of equality of opportunity for welfare, which his criticisms of equality of welfare do not touch. What is more, those criticisms positively favour equality of opportunity for welfare as a remedy for the defects in the rejected doctrine.

But while equality of opportunity for welfare survives Rawls's criticisms of equality of welfare, arguments against the welfare metric which were later advanced by Sen also apply against its opportunity-defined cousin. Sen called for attention to something like opportunity (under the title 'capability' [3]

[3] Immediately after introducing the notion of '*capability* to function' in the Dewey Lectures, Sen shifts to the alternative language of 'opportunity' to express the same idea. See 'Well-Being, Agency and Freedom: The Dewey Lectures 1984', *Journal of Philosophy*, 82 (April 1985), 200-1. Cf. *Commodities and Capabilities* (Amsterdam, 1985), 59; *The Standard of Living* (Cambridge, 1987), 36.

), but it was not welfare, or not, at any rate, welfare alone, which Sen thought people should have the opportunity to achieve. Instead, he drew attention to the condition of a person (e.g. his level of nutrition) in a central sense captured neither by his stock of goods (e.g. his food supply) nor by his welfare level (e.g. the pleasure or desire satisfaction he obtains from consuming food). In advancing beyond Rawls, Sen therefore proposed two large changes of view: from actual state to opportunity, and from goods (and welfare) to what he sometimes called 'functionings'.

In my view, Sen's answer to his own question was a great leap forward in contemporary reflection on the subject. But often a thinker who achieves a revolution misdescribes his own achievement, and I shall argue, at some length, that Sen's work is a case in point. He moved away from Rawlsian and other views in two directions which were orthogonal to each other. If Rawls and welfarists fixed on what a person gets in welfare or goods, Sen fixed on what he gets in a space between welfare and goods (nutrition is delivered by goods supply and it generates welfare), but he also emphasized what a person *can* get, as opposed to (just) what he *does*. Sen's misdescription of his achievement lay in his appropriation of the word 'capability' to describe both of his moves, so that his position, as he presented it, is disfigured by ambiguity. I shall here expose the ambiguity in Sen's use of 'capability' (and cognate terms), and I shall also propose an answer to his (my title) question which departs from his own in a modest way.

## 2 Rawlsian Criticism of Equality of Welfare

Before examining Rawls's critique of equality of welfare, a word about what will here be meant by 'welfare'. Of the many readings of 'welfare' alive (if not well) in economics and philosophy, the present inquiry recruits two: welfare as enjoyment, or, more broadly, as desirable or agreeable state of consciousness, which



I shall call *hedonic welfare*; and welfare as *preference satisfaction*, where preferences order states of the world, and where a person's preference is satisfied if a relevant state of the world obtains, whether or not he knows that it does. [4]

[4] These two readings of welfare correspond to Sen's 'happiness' and 'desire fulfilment' readings and exclude his 'choice' reading: see 'Well-Being, Agency and Freedom', 187 ff. It is reasonable to ignore the 'choice' reading, since, as Sen shows, it reflects confusion about the relationship between preference and choice. My two readings also correspond to Ronald Dworkin's 'conscious state' and

'relative success' conceptions: see 'Equality of Welfare', *Philosophy and Public Affairs*, 10 (Summer 1981), 191-4, 204-9, 220-1. I do not consider welfare as 'overall success' (ibid., 209 ff.) because it is very hard to handle, and in any case it is, arguably, undermotivated: see my 'On the Currency of Egalitarian Justice', n. 34. I also set aside so-called 'objective theories of welfare' (Dworkin, 'Equality of Welfare', 224-6), since most philosophers would consider them alternatives to any sort of welfare theory: Thomas Scanlon, for whom welfare is preference satisfaction, would describe his theory as anti-welfarist, yet it is an objective theory of welfare in Dworkin's sense. Finally, to mention an author whose work is salient in this study, Richard Arneson has the same understanding of welfare as Scanlon and Rawls has not specified a particular conception-which is not to say that he should have done.

It will sometimes be necessary to say which of those two I mean by 'welfare', but not always. Often the debates on which I comment have a similar shape under either interpretation of welfare, so that I shall have both in mind (by which I do not mean some amalgam of the two) at once. Unless I indicate otherwise, my contentions are meant to hold under either of the two readings of 'welfare' that I have just distinguished.

Rawls advances two objections to equality of welfare, and they apply against both its hedonic and its preference interpretations. I shall call his objections the 'offensive tastes' and 'expensive tastes' criticisms. I believe that each criticism can be accommodated by a welfare egalitarian through a natural modification of his original view. In the case of the offensive tastes criticism, that would probably be conceded by Rawls (and by Ronald Dworkin, who develops the criticism more systematically and at some length [5]

[5] See 'Equality of Welfare', 197-201.

). But the second criticism is supposed by Rawls and Dworkin to justify an abandonment of the terrain of welfare altogether; as I shall indicate, I do not think that it does. The second criticism also creates a problem for Rawls's system, which I shall describe in a brief digression.

Rawls adverts to offensive tastes in the course of his critique of utilitarianism, but, as Amartya Sen notes, [6]

[6] 'Equality of What?' in S. McMurrin (ed.), *Tanner Lectures on Human Values* i (Cambridge, 1980), 211.

he is at that point really criticizing welfarism as such, where welfarism is the view that just distribution is some or other function of nothing but the welfares of individuals. It follows logically that the offensive tastes criticism also applies against a conception of justice in which equality of welfare is the only principle. And although a believer in equality of welfare who also affirms further principles need not be a welfarist other than with respect to the metric of equality in particular, it is extremely unlikely that a good criticism of welfarism proper will not also apply to that restricted welfarism which acknowledges the relevance of no information but welfare in the context of

end p.11

equality, even if its proponent admits non-welfare information elsewhere. In any case, the offensive tastes criticism strikes me as powerful against the welfare-egalitarian claim, whether or not other claims are conjoined with it.

The offensive tastes criticism of welfarism is that the pleasure a person takes in discriminating against other people or in subjecting others to a lesser liberty should not count equally with other satisfactions in the calculus of justice. [7]

[7] *A Theory of Justice* (Cambridge, Mass., 1971), 30-1.

From the point of view of justice, such pleasures deserve condemnation, and the corresponding preferences have no claim to be satisfied, even if they would have to be satisfied for welfare equality to prevail. I believe that this objection defeats welfarism, and hence equality of welfare. But the natural course for a welfare

egalitarian to take in response to the offensive tastes criticism is to shift his favour to something like equality of *inoffensive* welfare. The criticism does not seem to necessitate abandoning equality of welfare in a more fundamental way.[8]

[8] In fairness to Rawls, one should recall that he presented the offensive tastes criticism as an objection not to equality of welfare but to utilitarianism, and for utilitarians a move to 'inoffensive welfare' would no doubt constitute a pretty fundamental shift. From the fact that the same criticism applies against both views, and that each should be revised in the same way in the face of it, it does not follow that the distance between the original and the revised view is the same in both cases.

The 'expensive tastes' criticism *is* thought to necessitate such an abandonment. It occurs in the context of Rawls's advocacy of primary goods as the appropriate thing to equalize. Now the phrase 'primary goods' covers a set of different things in Rawls,[9]

[9] See *A Theory of Justice*, 62.

but for our purposes, we can, as Rawls does in the texts I shall quote, take it narrowly, as referring to goods in the economist's sense, or the power to purchase them.

Prosecuting his case against welfare and for primary goods, Rawls asks us to 'imagine two persons, one satisfied with a diet of milk, bread and beans, while the other is distraught without expensive wines and exotic dishes. In short one has expensive tastes, the other does not.' A welfare egalitarian must, *ceteris paribus*, provide the epicure with a higher income than the person of modest tastes, since otherwise the latter might be satisfied while the former is distraught. But Rawls argues powerfully against this implication of the welfare egalitarian principle:

> as moral persons citizens have some part in forming and cultivating their final ends and preferences. It is not by itself an objection to the use of primary goods that it does not accommodate those with expensive tastes. One must argue in addition that it is unreasonable, if not unjust, to hold such persons responsible for their preferences and to require them to make out as best they can. But to argue this seems to presuppose that 'citizens' preferences are beyond their control as propensities or cravings which simply happen. Citizens seem to be regarded as passive carriers of desires. The use of primary goods . . . relies on a capacity to assume responsibility for our ends.

People with expensive tastes could have chosen otherwise, and if and when they

end p.12

press for compensation, then others are entitled to insist that they themselves bear the cost of 'their lack of foresight or self-discipline'.[10]

[10] 'Social Unity and Primary Goods', in A. K. Sen and B. Williams (eds.), *Utilitarianism and Beyond* (Cambridge, 1982), 168-9. Cf. 'Fairness to Goodness', *Philosophical Review*, 84 (April 1975), 553; 'Justice as Fairness, Political not Metaphysical', *Philosophy and Public Affairs*, 14 (Summer 1985), 243-4. For a somewhat different explanation of why justice ignores expensive tastes, with less (not no) emphasis on the idea that they are subject to the agent's control and more on the idea that it is appropriate to hold him accountable for them, see the reply to Arrow in Rawls's unpublished 'Citizens' Needs and Primary Goods'.

For interesting comment on and sympathetic development of Rawls's views on responsibility for preference, see Bruce Landesman, 'Egalitarianism', *Canadian Journal of Philosophy*, 13 (March 1983), 37.

I believe that this objection defeats welfare egalitarianism but that it does not, as Rawls supposes, also vindicate the claims of the primary goods metric. The right way for an erstwhile welfare egalitarian to respond to the objection seems to me to be the following: 'To the extent that people are indeed responsible for their tastes, the relevant welfare deficits do not command the attention of justice. We should therefore compensate only for those welfare deficits which are not appropriately traceable to

the individual's choices. We should replace equality of welfare by equality of opportunity for welfare. It would be utterly unjustified to adopt a primary goods metric because of the expensive tastes counter-example.'

Equality of opportunity for welfare,[11]

[11] For an elegant exposition and defence of equality of opportunity for welfare, in its preference interpretation, see Richard Arneson, 'Equality and Equality of Opportunity for Welfare', *Philosophical Studies*, 56 (1989).

unlike equality of welfare, permits and indeed enjoins departures from welfare equality when they reflect the choices of relevant agents, as opposed to deficient *opportunity* for welfare. If a person's welfare is low because he freely risked a welfare loss in gambling for a welfare gain, then, under the opportunity form of the principle, he has no claim to compensation. Nor has a person who frittered away welfare opportunities which others seized. Nor, to take a different kind of example, has a person who chose to forgo welfare out of devotion to an ideal which (expressly, or merely as it happened) required self-denial. A person like Rawls's epicure who, *ex hypothesi*, could have controlled himself and maintained more austere tastes, has, in consequence, no grievance at the bar of equality of opportunity for welfare if his actual welfare is substandard because his tastes are expensive. The eligibility of equality of opportunity for welfare as a reply to Rawls shows that his conviction that the reasons he gives for rejecting equality of welfare are also reasons for affirming equality of primary goods is misplaced.

Before leaving Rawls, I want to indicate a serious problem for his political philosophy which his remarks about expensive tastes raise. The problem would be less interesting than it is in fact were it merely a problem in the Rawlsian architectonic and not also expressive of a general tension in left-egalitarian thinking.

The problem, as it appears in Rawls, is that the picture of the individual as

end p.13

responsibly guiding his own taste formation is hard to reconcile with claims Rawls uses elsewhere in a fundamental way to support his egalitarianism. These claims, which are quite standard on the left, express scepticism about special efforts as a ground for rewarding people unequally. Here is how Rawls formulates this fairly familiar scepticism:

the effort a person is willing to make is influenced by his natural abilities and skills and the alternatives open to him. The better endowed are more likely, other things equal, to strive conscientiously, and there seems to be no way to discount for their greater good fortune. The idea of rewarding desert is impracticable.[12]

[12] *A Theory of Justice*, 312.

Now there are two ways of taking this passage. One way is as I think Rawls intended it, and the other is as Robert Nozick took it, and on the basis of which he directed strong criticism against Rawls. Nozick, I am sure, misread the passage, but his misreading of it constitutes a correct reading of what many socialists and egalitarians say about effort, so it will be worth our while to pause here, to attend to Nozick's criticism. On either reading of the passage, it is hard to reconcile with what Rawls says about foresight, self-discipline, and expensive tastes. But I shall come to that point in a moment, for the passage, can also be criticized independently, and I want to do that first.

The two readings of the passage divide with respect to how they take the word 'influenced' in Rawls's use of it here. On my reading of it, it means 'influenced', that is, 'affected, but not necessarily wholly determined'. On Nozick's reading of the passage, 'influenced' means something like 'wholly determined'. There is a difficulty

for Rawls whichever way we take the passage, but not the same difficulty in each case.

On my reading of Rawls, in which he means 'influenced' by 'influenced', he does not say that those who make a lot of effort have no control over, and therefore deserve no credit for, the amount of effort they put in. His different point is that we cannot reckon the extent to which their above-par effort is attributable not to admirable striving but to 'greater good fortune': there is no 'way to discount' for the latter. That is a practical objection to trying to reward effort that deserves reward, not a claim that there is no such effort: see the final sentence of the passage.

If Rawls is right that not all effort is deserving, then, we might agree, not all effort deserves reward. But why should it follow that effort deserves no reward at all? The practical difficulty of telling how much of it merits reward hardly justifies rewarding it at a rate of 0 per cent, as opposed to at a rate of somewhere between 0 and 100 per cent, for example through a taxation scheme whose shape and justification escapes, because of its deference to effort, the writ of the Difference Principle. (To be sure, it is hard to tell how much effort a person supplies, but Rawls's claim is not that total effort is indiscernible, but that we cannot identify the contribution to effort of fortunate endowment.)

end p.14

But that criticism of Rawls is mild by comparison with the one to which he is exposed on Nozick's reading of his remarks. The plausibility of that reading is enhanced by Nozick's careless or mischievous omission of what follows 'conscientiously' when he exhibits the *Theory of Justice* passage quoted above. Nozick thereby creates the impression that Rawls is presenting a familiar egalitarian determinist doctrine. Nozick's response to that doctrine is very powerful:

So denigrating a person's autonomy and prime responsibility for his actions is a risky line to take for a theory that otherwise wishes to buttress the dignity and self-respect of autonomous beings . . . One doubts that the unexalted picture of human beings Rawls' theory presupposes and rests upon can be made to fit together with the view of human dignity it is designed to lead to and embody.[13]

[13] *Anarchy, State and Utopia* (New York, 1974), 214.

Nozick is pressing a dilemma: either people have real freedom of choice, in which case they may be credited (at least to some extent) with the fruits of their labours; or there is no such thing as free choice, in which case liberals should take the purple out of the passages in which they set forth their conception of humanity, and — we can add — socialists should stop painting inspiring pictures of the human future (unless they believe that people lack free will under capitalism but will get it after the revolution).

On Nozick's reading of the 'effort' passage, it is clearly inconsistent with the responsibility for taste formation with which Rawls credits citizens. That does not matter too much, since Nozick's reading is a misreading. But it is not easy to reconcile what Rawls says about effort with what he says about tastes, even on my less creative reading of his text. On my reading of it, effort is partly praiseworthy and partly not, but we cannot separate the parts, and the indicated policy consequence is to ignore effort as a claim to reward. Now, the passage about tastes begins with the thought that 'citizens have *some* part in forming and cultivating their final ends and preferences', though it ends by assigning a more wholesale responsibility for preferences to citizens. If we stay with the opening thought, then we can wonder why partial responsibility for effort attracts no reward at all while (merely) partial responsibility for expensive taste formation attracts a full penalty (and those who

keep their tastes modest reap a welfare reward). And if we shift to the wholesale responsibility motif, then we can wonder why beings who are only in a limited way responsible for the effort they put in may be held wholly responsible for how their tastes develop.

The most natural view about these matters is that people are partly responsible and partly not, both for how much effort they put into production and for whether or not they develop expensive tastes. [14]

[14] I press that natural view against Dworkin's development of the theme of expensive tastes (see 'Equality of Welfare', 228-40) in section IV of 'On the Currency of Egalitarian Justice'.

This does not mean that it is easy to tell where responsibility begins and ends or, consequently, what the dictate of distributive justice is in particular instances — but there is no antecedent

end p.15

reason whatsoever for supposing that judgements about justice, at a fine-grained degree of resolution, are easy.

# 3 Sen and Capability

*a*. The foregoing critique of Rawls does not prove that quantity of primary goods is the wrong metric for egalitarian evaluation; it just proves that a major reason Rawls offered for favouring primary goods points, instead, to equality of opportunity for welfare. For a more thorough refutation of the primary goods proposal, I turn to Sen's 'Equality of What?' That seminal article also argues persuasively against the welfare metric, and, while Sen did not address equality of opportunity for welfare, his argument against equality of welfare readily extends itself to the former view. After presenting and endorsing Sen's negative arguments, I shall argue that his positive replacing proposal, capability equality, suffers from a severe expositional obscurity.

Sen's argument against the primary goods metric was simple but powerful. It was that differently constructed and situated people require different amounts of primary goods to satisfy the same needs, so that 'judging advantage in terms of primary goods leads to a partially blind morality'. [15]

[15] 'Equality of What?' 216.

It is, Sen rightly said, a 'fetishist handicap' to be concerned with goods as such, to the exclusion of what goods '*do* to human beings'. [16]

[16] Ibid., 218; see also 'Ethical Issues in Income Distribution', in *Resources, Values and Development* (Oxford, 1984), 294; *Commodities and Capabilities, 23; The Standard of Living*, 15-16, 22.

Or, as Sen later expressed the point: 'what people get out of goods depends on a variety of factors, and judging personal advantage just by the size of personal ownership of goods and services can be very misleading . . . It seems reasonable to move away from a focus on goods as such to what goods do to human beings'. [17]

[17] 'Introduction', in *Choice, Welfare and Measurement* (Oxford, 1982), 29-30.

The principle of equality condemns equal goods provision to a sound-limbed person and a paraplegic, because greater resources are necessary to enable the latter to achieve mobility, a desideratum to which a metric of stock of wealth is blind. [18]

[18] 'Equality of What?' 218.

Sen also used the example of the needy cripple to good effect against the welfare alternative to primary goods. For the egalitarian response to his plight is not determined by a judgement that he suffers a welfare deficiency. Perhaps, indeed, he suffers no such thing, 'because he has a jolly disposition. Or because he has a low aspiration level and his heart leaps up whenever he sees a rainbow in the sky.' [19]

[19] Ibid., 217.

So while both hedonic and preference-satisfaction welfarists are free of the goods theorist's fetishistic neglect of what goods do to human beings, Sen criticized them for their too-narrow view of what people get from goods, for focusing 'not on the person's capabilities but on his mental reaction', not, for example, on how much nourishment a person gets from food, but on how much

end p.16

utility, which is a matter of mental reaction or attitude, he derives from such nourishment. [20]

[20] Ibid., 218. 'Mental reaction' must here cover not only a kind of experience but also a subjective valuation, to cater for the preference form of welfarism.

Utility is an unsuitable guide to policy, if only because a person may adjust his expectations to his condition. The fact that a person has learned to live with adversity, and to smile courageously in the face of it, should not nullify his claim to compensation. [21]

[21] See, for further development of this point, *Commodities and Capabilities*, 21-2, 29; 'Introduction' in *Resources, Values and Development*, 34; 'Rights and Capabilities', in ibid., 308-9; 'Goods and People', in ibid., 512; *The Standard of Living*, 8-11.

His high welfare score is thus not a decisive reason for not assisting someone who labours under a severe disadvantage, which is recognizable as such from an objective point of view. His equanimity may, after all, reflect admirable and reward-worthy striving to overcome a natural reaction of misery. But even if no such striving is necessary, because the person is blessed from birth with an extra-sunny disposition, the requirement of compensation retains intuitive force. And that means that not only equality of welfare but also equality of opportunity for welfare falls before the case of the cripple. Consider the poor and lame but sunny-spirited Tiny Tim. Tiny Tim is *actually* happy, by any welfarist standard. And we may also suppose that, because of a fortunate innate disposition, he is blessed with abundant *opportunity* for happiness, that he need not do much to get a lot of it. Yet egalitarians would not on that account strike him off the list of free wheelchair receivers. Hence they do not think that wheelchair distribution should be controlled exclusively by the welfare opportunity requirements of those who need them. They need wheelchairs to be adequately resourced, whether or not they also need them to be, or to be capable of being, happy.

*b*. In the course of making the critical points reported and endorsed above, Sen used the term 'capability', and he appropriated that term to denote his own positive counter-proposal. I shall now argue that, in 'Equality of What?', Sen brought two distinct aspects of a person's condition under that single name, and that this unnoticed duality has persisted in his subsequent writings. Both aspects, or dimensions of assessment, should attract egalitarian interest, but one of them is not felicitously described as 'capability'. The identification of that latter dimension constitutes a particularly striking contribution to normative understanding, but just that dimension is hard to perceive in Sen's exposition, because of the unfortunate and ambiguous nomenclature.

As we have seen, Sen arrived at what he called 'capability' through reflection on the main candidates for assessment of well-being that were in the field when he gave his 1979 lecture, to wit, utility, or welfare, and Rawlsian primary goods. [22]

[22] A notable further candidate not yet then in print is Dworkin's equality of resources. It would be a worthwhile-and difficult-exercise to distinguish each of the two Sen dimensions I shall describe from the Dworkin resources dimension. (For pertinent remarks, see Sen's excellent rebuttal, all of which strikes me as correct, of Dworkin's criticism of Sen's view: 'Rights and Capabilities', 321-3.)

Sen pleaded for a metric of well-being which measured something

falling *between* primary goods and utility, in a sense that will presently be explained-a something which had, amazingly, been largely neglected in previous literature. He called that something 'capability': 'what is missing in all this framework [23]

[23] That is, the framework of discussion restricted to the rival claims of primary goods and utility as measures of well-being, and, within 'primary goods', to goods in the ordinary sense. It is that subset of primary goods which is pertinent here.

is some notion of "basic capabilities": a person being able to do certain basic things'. [24]

[24] 'Equality of What?' 218.

But that characterization of the missing dimension was different from another which Sen offered in the same text, and which was more in keeping with his *argument* for the new perspective.

According to that argument, as we have seen, it is necessary to attend to what goods do to (or for) human beings, in abstraction from the utility they confer on them. But to call what goods supply to human beings 'capability' was a mistake. For even when utility has been set aside, it is not true that all that goods do for people is confer capability on them-provide them, that is, with the capacity to do things-or that that is the uniquely important thing they do for them, or that that is the one thing they do for them that matters from an egalitarian point of view. In naming his view 'Basic Capability Equality' Sen failed to delineate the true shape and size of one of the dimensions he had uncovered, and which I shall now try to describe.

It is indeed false that the whole relevant effect on a person of his bundle of primary goods is on, or in virtue of, his mental reaction to what they do for him. There is also what welfarists ignore: what they do for him, what he gets out of them, apart from his mental reaction to or personal evaluation of that service. I shall call that non-utility effect of goods *midfare*, because it is in a certain sense midway between goods and utility. Midfare is constituted of states of the person produced by goods, states in virtue of which utility levels take the values they do. It is 'posterior' to 'having goods' and 'prior' to 'having utility'. [25]

[25] *Commodities and Capabilities*, 11. Midfare is the '*state* of a person' in the sense of ibid., 23.

Midfare is a heterogeneous collocation, because goods do categorially various things for people: (1) they endow them with capabilities properly so called, which they may or may not use; (2) through people's exercise of those capabilities, goods contribute to the performance of valuable activities *and* the achievement of desirable states; and (3) goods cause further desirable states directly, without any exercise of capability on the part of their beneficiary: an example would be the goods which destroy the insects that cause malaria. Capability (properly so called) is, then, a part of midfare, for it certainly cannot be excluded from the range of things that goods confer on people, yet, equally certainly, it does not exhaust that range.

Each terminus of the goods-midfare-utility sequence has seemed to some the right focus for assessment of a person's situation from an egalitarian point of view. Rawlsians look at the beginning of the sequence and welfarists look at


its end. Welfarists think that the Rawlsian measure is too objective, that it takes too little account of distinguishing facts about individuals. Rawlsians think that the welfare measure is too subjective, that it takes too much account of just such facts. The reasons each side gives for disparaging its opponent's dimension suggest that each should prefer midfare to the dimension favoured by its opponent. Welfarists draw attention to utility because, so they say, people do not care about goods as such but about the utility they provide. But, since people also care more about midfare than

about goods as such (save where they are *themselves* being fetishistic), the welfarist reason for preferring welfare to goods is also a reason for preferring midfare to the latter. Advocates of goods oppose the welfare metric because, they say, the welfare consequences of goods consumption are (1) too subject to volition (Rawls, sometimes [26]
[26] See pp. 12-13 above.
), (2) too much a matter of people's (not necessarily chosen) identifications (Rawls at other times, [27]
[27] See the reference to 'Citizens' Needs and Primary Goods' in n. 10 above.
Dworkin [28]
[28] See 'Equality of Welfare', 228-40; 'Equality of Resources', *Philosophy and Public Affairs*, 10 (Fall 1981), 302-3, and, for criticism of the latter, see section III *c* of my 'On the Currency of Egalitarian Justice'.
), or (3) too idiosyncratic (Scanlon [29]
[29] See his 'Preference and Urgency', *Journal of Philosophy*, 72 (Nov. 1975), 659-66, and see section V *b* of my 'On the Currency of Egalitarian Justice' for criticism of Scanlon's idiosyncrasy claim.
). On all three grounds midfare arguably scores better than utility does.

Given that each side in the foregoing division has reason to prefer the midfare dimension to the one favoured by its opponents, it is extraordinary that midfare had not been uncovered, and Sen's reorienting proposal was consequently profound and liberating, albeit remarkably simple. For it simply says that, in the enterprise of assessing a person's well-being, we must look to her condition in abstraction from its utility for her. We must look, for example, at her nutrition level, and not just, as Rawlsians do, at her food supply, or, as welfarists do, at the utility she gets out of eating food. [30]
[30] 'Introduction' in *Choice, Welfare and Measurement*, 30.

But this significant and illuminating reorientation is not equivalent to focusing on a person's capability, in any ordinary sense. Capability, and exercises of capability, form only one part of the intermediate midfare state. *What goods do to people is identical neither with what people are able to do with them nor with what they actually do with them* (and it is also not identical with all or part of the combination of these two things). To be sure, it is usually true that a person must do something with a good (take it in, put it on, go inside it, etc.) in order to be benefited by it, but that is not always true, and, even where it is true, one must distinguish what the good does *for* the person from what he does *with* it. The colloquial question 'How are you doing?' can be used to ask after a person's midfare (especially when this pedantic rider is attached to it: 'by which I do not mean how do you feel about how you are doing'), but the usual answer will not be (just) a list of capacities, activities, and results of activities, because not all midfare is capability or an exercise of capability or a result of exercising capability. And many midfare states which are indeed a result of

end p.19

exercising capability have a (non-utility) value which is unconnected with their status as effects of exercising capability, and which is not clearly exhibited in its true independence of capability by Sen.

The case of food, which has, of course, exercised Sen a great deal, illustrates my claims. The main good thing that food does for people is that it nourishes them. Typically, of course, people become nourished by nourishing or feeding themselves, by exercising the capability of nourishing themselves which ownership of food confers on them. But the fact that food gives a person the capability to nourish himself is not the same fact as (and is usually less important than) the fact that it

enables him to be nourished. To say that food enables him to be nourished is to say that it makes it possible for him to be nourished. That he characteristically actualizes that possibility himself is a further (and usually less important) fact. When, moreover, we ask how well nourished a person is, we are not asking how well he has nourished himself, even though the answer to the two questions will usually be the same; and we are usually primarily interested in the answer to the first question.[31]

[31] In one place ('Goods and People', 510), Sen makes one of the distinctions on which I am insisting: 'while goods and services are valuable, they are not valuable in themselves. Their value rests on what they can do for people, or rather, what people can do with these goods and services.' But why does Sen here reject his first and, in my view, superior suggestion in favour of the second? Because, I suggest, of an interest in advocating freedom, which is a desideratum different from midfare: see subsection *d* below.

The difference between midfare and capability (properly so called) will perhaps become more evident if we reflect a little about small babies. Small babies do not sustain themselves through exercises of capability. But it is false that, in the case of babies, goods generate utility and nothing else worth mentioning. When food is assigned for the consumption of either a baby or an adult, each is enabled to be nourished. The fact that only the adult is able to nourish himself does not mean that he alone gets midfare. The baby gets it too. Hence midfare, the product of goods which, in turn, generates utility, is not co-extensive with capability, and 'capability' is therefore a bad name for midfare.

If food does not make my case strongly enough, since babies do suck and chew, think instead of clothes. No collaboration on the baby's part is needed when its parent confers the midfare of warmth and protection on it by dressing it. Or consider the midfare supplied by the nutriment in a hospital drip, to baby and adult alike, or, for that matter, by the rays of the sun. There is no relevant exercise of capability by benefited agents in these instances, but there is an important benefit to be described in non-welfarist-midfare-terms. Hence the concept of capability is insufficiently general to capture one of the things that Sen wants to identify.

There are two powerful motivations for pointing to something other than either goods or utility when concerning oneself with egalitarian policy, but the motivations point at different things. There is good reason to look at what a

end p.20

person *can* achieve, independently of his actual state; *and* there is good reason not to reduce the evaluation of that *actual* state either to an examination of his stock of resources or to an assessment of his utility level. But these are distinct points, and the language of capability felicitously covers the first one only.

The ambiguity I have tried to expose appears in a number of Sen's dictions, including the apparently harmless phrase 'what people get out of goods'.[32]

[32] See p. 16 above, text to n. 17.

On one reading of it, under which 'get out of' means (roughly) 'extract', getting things out of goods represents an exercise of capability. But 'get out of' can also mean, more passively, 'receive from', and it does not require capability to get things out of goods in that sense. Goods (and welfare) theorists ignore (some of) what people get out of goods in both senses of the phrase, but while only the first sense relates to capability, the second denotes something at least as important.

*c.* In Sen's discourse, to have a capability is to be capable of achieving a range of what he calls 'functionings'. But Sen characterizes functionings differently at different times, and thereby adds further imprecision to the presentation of his view. Sometimes, in keeping with the ordinary meaning of 'functioning', and in line with Sen's original gloss on 'capability' as 'being able to *do* certain basic things',[33]

[33] 'Equality of What?', 218 (italics added).

a functioning is by definition an activity, something that a person does. [34]

[34] ' "Functionings" are what the person succeeds in *doing* with the commodities . . . at his or her command' (*Commodities and Capabilities*, 10).

The questions 'Can they read and write? Can they take part in the life of the community?' [35]

[35] 'The Living Standard', *Oxford Economic Papers*, 6 (1984), 84.

inquire into people's functionings in this familiar sense of the term. But at other times, functionings are not by definition activities but all (desirable) states of persons, and 'being well nourished', 'being free from malaria', and 'being free from avoidable morbidity' [36]

[36] Ibid., 84, and see 'Well-Being, Agency and Freedom', 197. These examples fall under the characterization of functionings as 'activities . . . or states of existence or being' (ibid., 197). (In one place, Sen describes 'being well nourished' not as a functioning but as a capability, but that is probably a slip: see *The Standard of Living*, 18).

are consequently entered as examples of functionings, although, not being activities, they are not functionings in the ordinary sense of the term. (Even though 'I am free from malaria now' can be part of the answer to the question 'How are you doing?' in its colloquial use.)

When Sen writes that 'Functionings are . . . personal features; they tell us what a person is doing', [37]

[37] 'Rights and Capabilities', 317.

he places his incompatible broad and narrow definitions of 'functioning' on either side of the semi-colon. For not all personal features, and not all of the personal features that Sen wishes to encompass, are things that a person is doing. Unlike reading and writing, being free from malaria is not something that one does. Elsewhere, a broader definition of 'functionings'

end p.21

is offered, under which 'they tell us what the person is doing or achieving', [38]

[38] 'The Living Standard', 84.

and it is true that being free from malaria is something that one *may* achieve. But it is surely of supreme (midfare) importance even when one cannot be credited with achieving it.

Sen himself notes that being free from malaria may be entirely due to 'anti-epidemic public policy'. [39]

[39] *Commodities and Capabilities*, 16.

What he fails to note is the consequent impropriety of regarding it, in that instance, as something the person achieves, as the exercise of a capability of any kind. Yet Sen would surely not want to exclude heteronomously obtained freedom from malaria from the balance sheet of how a person 'is doing'. *And that proves that he has a concern to promote forms of midfare which does not derive from his concern to promote the claims of capability as such*. Indeed, one may go further: the lacks in people's lives which Sen is *most* concerned to draw to our attention are midfare lacks which are not lacks in capability proper, and the alleviation of which need not always proceed through an enhancement of the sufferer's capability. He is concerned with people who are 'ill-fed, undernourished, unsheltered and ill', [40]

[40] Ibid., 21

who lack 'basic clothing, ability to be housed and sheltered, etc.' [41]

[41] Ibid., 73.

Being able to be housed is not the same thing as being able to house oneself. Entitlements to goods make desirable states possible for people. They generally realize these possibilities themselves, by exercising a capability to do so. But, with respect to the lacks which most exercise Sen, it is the possibilities that matter, and the corresponding capabilities matter only derivatively.

At one point [42]

[42] I am embarrassed to say that I have lost my record of where these phrases occur.

Sen extols the importance of 'a person's ability to function without nutritional deficiency' and of 'the capability of avoiding nutritional deficiency'. Such functionings and avoidings are genuine activities, [43]

[43] The *ability* to function without nutritional deficiency is trickier. I have that ability if and only if there is something I am able to do and I lack nutritional deficiency. But, in general, the characterization of the result as an ability is owing to the first clause in that statement of necessary and sufficient conditions, rather than to my nutritional good order.

but the generative desideratum here is not activity but simply lack of nutritional deficiency, the fundamental desirability of which is lost in these athletic phrasings. It is not hitting the nail on the head to say that food is desirable because it enables a person to avoid nutritional deficiency, as though performing that activity is the (one) important thing here. Decent living space, to change the example, is a primary good which helps to maintain a person in good health, and it often does that when it would be false to say that it helps him to maintain himself in good health. Whether he is doing that is an exquisitely subtle question, a negative answer to which is consistent with decent living space delivering its hygienic boon. More generally, the 'kind of life I am living' cannot be identified

end p.22

with what I am 'succeeding in "doing" or "being" ', [44]

[44] *Commodities and Capabilities*, 28. Cf. ibid., 51, where 'well-being' is described as depending on 'the particular achievements of the person-the kind of "being" he or she succeeds in having'. This is entirely implausible when 'achievements' and 'succeeds' are taken literally, and (see subsection *d* below) Sen has a reason to want their literal meanings to resonate.

unless we put scare-quotes around 'succeeding' as well. There are many benefits I get which I do not literally succeed in getting.

It is true that the better nourished I am, the larger is the number of valuable activities of which I shall be capable. But that capability, the important capability which food confers, is a *result* of eating it. It is not the Sen capability associated with food, which is a capability to *use* food to achieve various 'functionings': being nourished, conducting a ceremonial, entertaining friends. [45]

[45] At one point Sen writes that 'the essence of the capabilities approach is to see commodity *consumption* as no more than a means to generating capabilities' ('Goods and People', 522 (italics added), but that is either an aberration or an unsignalled change of doctrine. For elsewhere commodity *entitlement* generates capability, which is the power to use or consume the commodity in a variety of ways, each of which uses is a functioning. In the different conceptualization just quoted, capability is a consequence of what elsewhere is called a 'functioning'.

One cannot infer from the central place in life-and midfare-of action and capacity that capability spreads across the entire space of midfare. And, as we have seen, not everything which merits attention under the broad midfare construal of Sen's contention is an activity or achievement.

We may conclude that, while Sen's focus on what goods do for people apart from the mental reaction they induce is original and illuminating, it is unnecessarily narrowed when the object of the focus is described in functioning/ capability language. Comprehending as it does everything which 'goods do for people', [46]

[46] See the catalogue characterization of midfare at p. 18 above.

midfare cannot be identified either with capability or with what Sen calls 'functioning', nor can it be factored into the two without a confusing stretching of the meanings of words.

*d*. Why did Sen use the language of capability and functioning to express claims which that language fits quite imperfectly? Because, I hypothesize, he had something in addition to midfare in mind, to wit, freedom, and he wrongly thought that attending to a person's midfare-to what he gets from goods apart from the utility upshot of getting it-is attending to how much freedom he has in the world. Both the misrepresentation of all desirable states as a result of the exercise of capability and the tendency to represent all desirable states as activities reflect an interest in freedom distinct from, but not clearly distinguished by Sen from, the move from both utility and goods to midfare.

There is a case for installing the notion of freedom within egalitarian discourse, but that is a different exercise from vindicating the claims of midfare as such. There are *two* powerful motivations for pointing to something other than either goods or utility in a comprehensive characterization of well-being, but the motivations justify two distinguishable deviations from each of those metrics: possession of goods and enjoyment or utility are not the only actual states that matter,

end p.23

and — here is the freedom motivation — it is not only actual states, but the range of states the agent can attain, that matter.

According to Sen, 'the category of capabilities is the natural candidate for reflecting the idea of freedom to do', [47]

[47] 'Rights and Capabilities', 316. Cf. 'Economics and the Family', in *Resources, Values and Development*, 376.

since 'capability to function reflects what a person *can* do'. [48]

[48] 'Rights and Capabilities', 317.

Hence 'the concept of capabilities is a "freedom" type notion', [49]

[49] *Commodities and Capabilities*, 14.

and the functioning vectors accessible to a person determine her *'well-being freedom'*. [50]

[50] 'Well-Being, Agency and Freedom', 201.

All that may be true of capability (more or less) strictly so called, but it is not true of 'capability' where the term is used to denote the entire midfare dimension between goods and utility. Sen intends capability to have an athletic character. He associates it with the Marxist idea of a person fulfilling his potential through activity, which is to be contrasted with the idea of a person finding his *summum bonum* in passive consumption. [51]

[51] For relevant citations of Marx, see 'Development: Which Way Now?', in *Resources, Values and Development*, 497; 'Goods and People', 512; *The Standard of Living*, 37.

But, in Sen's wider construal of it, as midfare, capability covers too much to provide 'the perspective of "freedom" in the positive sense'. [52]

[52] 'Economics and the Family', 376. The sentence continues: 'who can *do* what, rather than who has what bundle of *commodities*, or who gets how much *utility*.' My point is the simple one that what people can do with their commodities is not identical with what their commodities (can) do for them.

The ambiguity between capability as a form of freedom and capability as midfare was not resolved in Sen's contribution to the July 1988 WIDER symposium ('Capability and Well-Being'), to which, with characteristic generosity, he has allowed me to refer. Instead, and as I shall explain, his ambiguous use of 'capability' was matched by an ambiguous use of 'freedom'.

At p. 5 of his 1988 typescript, Sen says that 'capability reflects a person's freedom to choose between different ways of living'. That formulation more or less identifies capability with freedom of choice (how much it does so depends on what 'reflects' means here: it might mean 'is'). In line with that characterization of capability is Sen's description of the rich faster, who 'has the capability to be well nourished, but chooses not to [be]'. [53]

[53] 'Capability and Well-Being', 38.

Elsewhere, however, something very different from the freedom to choose whether or not to eat, namely, freedom *from* hunger, is denominated a 'capability'. [54]

[54] Ibid., 41.

In fact, though, freedom from hunger is *being* well nourished. It is not the ability to choose which the rich faster has: it is what he chooses not to have. Freedom from hunger is a desirable absence or privation, the sort of freedom which even beings that are not agents can have. Healthy plants have freedom from greenfly, and sound houses are free from dry rot. (Note that a person might even be described, in a special context, as free from nourishment, for example, when he wants to fast, or by captors who want him to starve.)

end p.24

Unlike the freedom to choose whether or not to eat, freedom from hunger is not constitutively freedom to *do* anything. Sen speaks of *exercising* such 'capabilities' as freedom from hunger and freedom from malaria. [55]

[55] Ibid.

But they are not freedoms that are *exercised*. Sen's application of the term 'capability' *both* to the freedom to *avoid* morbidity [56]

[56] Ibid., 6.

*and* to freedom *from* morbidity [57]

[57] Ibid., 41.

shows that, in the attempt to bring the very different issues with which he is concerned under the single rubric of 'capability', he is led to make equivocal use of the term 'freedom'. [58]

[58] There is further evidence of the persisting ambiguity at pp. 17-19 of 'Capability and Well-Being'. At pp. 17-18, and in line with the definition of capability on p. 5, a capability set is characterized by the 'various alternative combinations of beings and doings any one (combination) of which the person can *choose*' (italics added). But at pp. 18-19, it is allowed, in seeming contradiction of that characterization, that the realized combination in a capability set may or may not be chosen. What is 'achieved' might not be 'achieved on the basis of choice' (p. 19).

When Sen introduced capability equality in 'Equality of What?', he was modest about its claims. It was 'a partial guide to the part of moral goodness that is associated with the idea of equality'. [59]

[59] 'Equality of What?' 220.

Five years later, his claim for the new perspective was much stronger. For, in the Dewey Lectures, Sen said that 'the *primary* feature of well-being can be seen in terms of how a person can "function", taking that term in a very broad sense', and that 'the accounting of functioning vectors' provides 'a more plausible view of well-being' than competing conceptions do. [60]

[60] 'Well-Being, Agency and Freedom', 197 (italics added), 226. See further ibid., 195, where there is an implied identification of 'having "well-being" ' with functionings. Cf. *Commodities and Capabilities*, 25, 51; *The Standard of Living*, 16.

Elsewhere, we are advised that, in assessing 'well-being and advantage', we should focus 'on the capability to function, i.e. what a person can *do* or *be*'. His utility is only *evidence* of a person's advantage in that central sense, [61]

[61] And often it is rather unreliable evidence, since people tend to adjust to adverse conditions: see pp. 16-17 above.

and the goods at his disposal (here called his 'opulence') are only *causes* of that advantage. [62]

[62] *Commodities and Capabilities*, Preface. Cf. ibid., 52. (Strictly, opulence is a magnitude which *supervenes* on command of goods: see ibid., 58).

The position of midfare between primary goods and utility, thus construed, is given as a reason for treating it as the central dimension of value.

These are strong claims, but they are easier to accept in that functionings are now explicitly described as 'doings *and* beings' so that both 'activities' and 'states of existence or being' come under the 'functioning' rubric. [63]

[63] 'Well-Being, Agency and Freedom', 197 (italics added).

What I cannot accept is the associated athleticism, which comes when Sen adds that 'the central feature of well-being is the ability to achieve valuable functionings'. [64]

[64] Ibid., 200.

That overestimates the place of freedom and activity in well-being. As Sen

end p.25

writes elsewhere, 'freedom is concerned with what one *can* do' and 'with what one can *do*': [65]

[65] 'Rights and Capabilities', 318.

midfare fails, on both counts, as a representation of freedom.

*e*. I said earlier that there are two powerful motivations for pointing to something other than either goods or utility when concerning oneself with egalitarian policy: there are other actual states that count, and it is not only actual states that count. In the last section I have shown how confusion of those two points is visible in the attempt to express both in the language of freedom, which is appropriate to the second point only.

Under one exegetically plausible disambiguation of Sen's formulations, they recommend equality of capability to achieve functionings, where 'capability' carries something like its ordinary sense (and is therefore not confused with midfare), and where 'functionings' denote all desirable states, and not desirable activities only. So disambiguated, Sen's theory displays two departures from equality of welfare: there is a change of modality, in that capability or opportunity, rather than final achievement, is key; and there is an enrichment of the conception of what opportunities are *for*-not welfare alone, but more broadly conceived good states of the person. In this reconstruction, the error of forcing the concept of capability to denote both the element of opportunity and the move to a broader conception of advantage is eliminated.

When Sen first invoked capability, it was in the context of a proposal that we attend to '*basic* capability equality'. [66]

[66] That was the title of section 4 of 'Equality of What?'

The relevant capability was of a fundamental sort, capability whose absence disables the person from satisfying his basic needs. Such need satisfaction is, while clearly related to the achievement of welfare, also irreducible to the latter: one may need something for which one has no desire and one may desire something which does not constitute a need. At the basic level, we can, with some confidence, rank capabilities in importance without paying attention to people's tastes. But, as Sen

points out, capability rankings are more moot once we pass beyond the basic desiderata of a normal human life:

when there is diversity of taste, it becomes harder to surmise about capability by simply observing achievement. For extreme poverty this problem is less serious. Valuing better nourishment, less illness and long life tend to be fairly universal, and also largely consistent with each other despite being distinct objectives. But in other cases of greater relevance to the richer countries-the informational problems with the capability approach can be quite serious.[67]

[67] 'The Living Standard', 87.

For capabilities which go beyond need satisfaction, it is hard to see how rankings are possible without recourse to utility valuations of the relevant states. In a critical comment on Sen, Richard Arneson tries to exploit the dependence on preference of the value of 'higher' capabilities:

end p.26

I doubt that the full set of my functioning capabilities [matters] for the assessment of my position. Whether or not my capabilities include the capability to trek to the South Pole, eat a meal at the most expensive restaurant in Omsk . . . matters not one bit to me, because I neither have nor have the slightest reason to anticipate I ever will have any desire to do any of these and myriad other things.[68]

[68] 'Equality and Equality of Opportunity for Welfare', 93.

Arneson infers that, in so far as the capability approach claims our attention, it is only as a different way of presenting the idea of equality of opportunity for welfare. But that conclusion is hasty. For one might hold that objective (non-welfare) assessment of capability is possible at the basic level, even though, beyond that level, we evaluate capability according to the range of desires which it enables a person to satisfy. The capability which matters as such (that is, independently of its welfare consequences) is capability definitive of a normal human existence, capability whose absence spells non-satisfaction of need. This answer to Arneson is anticipated by Sen:

The index of capabilities can be sensitive to the strength of desires without converting everything into the metric of desires. The welfarist picture drops out everything other than desires. A non-welfarist over-all index of capabilities may not drop out desires and may well be sensitive to the strength of desires *without ignoring* other influences on the indexing.[69]

[69] 'Rights and Capabilities', 319. Not only does Sen allow strength of desire to condition capability evaluation, but, somewhat curiously, he is also disposed to classify happiness itself as a functioning. 'Being happy' is described as a 'major functioning' at p. 13 of 'Well-Being and Agency' (unpublished, 1987) and as a 'momentous' one at p. 200 of 'Well-Being, Agency and Freedom'. See also 'Goods and People', 512; *The Standard of Living*, 8, 11, 14. See, too, *Commodities and Capabilities* (15, 52) for more tentative statements of happiness's credentials as a functioning.

And, one might add, the sensitivity of the capability index to desire is inversely related to the degree of 'basicness' of the region of capability space under exploration.

Still, if capability in its higher reaches waits on utility for its significance, it is in its more basic reaches that it makes its distinctive normative contribution, as Sen acknowledges: 'The issue of capabilities-specifically "material" capabilities-is particularly important in judging the standard of living of people in poor countries-it is also important in dealing with poverty in rich countries.'[70]

[70] 'The Living Standard', 85.

And even if utility and opulence offer more general, non-dependent (on other metrics) assessments of people's conditions, because of not being restricted to the basic, the

notion of basic capability equality may provide an apter reading of the egalitarian impulse than they do. The problem of characterizing well-being in general is not the same as the problem of the priorities of egalitarian justice, and basic midfare, if not basic capability as such, rather than goods bundles or utility quanta, surely is the first priority of justice.

*f*. In the last sentence of subsection *e*, I reintroduced the equivocation between capability and midfare. Here I shall explain why I did so, and why, more

end p.27

particularly, capability as such is not, in my view, the right thing for an egalitarian to focus on.

I have elsewhere proposed that the right thing to equalize is what I called 'access to advantage'. [71]

[71] See 'On the Currency of Egalitarian Justice', 920-1.

In that proposal, 'advantage' is, like Sen's 'functioning' in its wider construal, a heterogeneous collection of desirable states of the person reducible neither to his resources bundle nor to his welfare level. And, while 'access' includes what the term normally covers, I extend its meaning under a proviso that anything which a person actually has counts as something to which he has access, no matter how he came to have it, and, hence, even if his coming to have it involved no exploitation of access in the ordinary sense (nor, therefore, any exercise of capability). If, for example, one enjoys freedom from malaria because others have destroyed the malaria-causing insects, then, in my special sense, such freedom from malaria is something to which one has access. That special construal of access is motivated by the thought that egalitarians have to consider states of a person which he neither brought about nor ever was in a position to bring about, states which fall within category (3) of midfare, as it was sub-classified above (desirable states caused directly, without any exercise of capability by the beneficiary). Under the disambiguation of Sen's position articulated in subsection *e* above, such states go unconsidered in the egalitarian reckoning (though Sen himself is, of course, supremely concerned about them).

Under equality of access to advantage, the normative accent is not on capability as such, but on a person not lacking an urgent desideratum through no fault of his own: capability to achieve the desideratum is a sufficient but not a necessary condition of not suffering such a lack. My own proposal strikes me as better attuned than capability equality to the true shape of the egalitarian concern with such things as health, nourishment, and housing.

Equality of access to advantage is motivated by the idea that differential advantage is unjust save where it reflects differences in genuine choice (or, more or less, capability) on the part of relevant agents, but it is not genuine choice as such (or capability) which the view proposes to equalize. The idea motivating equality of access to advantage does not even imply that there actually is such a thing as genuine choice. Instead, it implies that if there is no such thing-because, for example, 'hard determinism' is true-then all differential advantage is unjust. The fact that my view tolerates the possibility that genuine choice is a chimera makes salient its difference from Sen's. In my view, Sen has exaggerated the indispensability of the idea of freedom in the correct articulation of the egalitarian norm. No serious inequality obtains when everyone has everything she needs, even if she did not have to lift a finger to get it. Such a condition may be woeful in other ways, but it is not criticizable at the bar of egalitarian justice.

end p.28

## Bibliography

Arneson, R. (1989). 'Equality and Equality of Opportunity for Welfare', *Philosophical Studies*, 56.
Cohen, G. A. (1989). 'On the Currency of Egalitarian Justice', *Ethics*, 99.
Dworkin, R. (1981*a*). 'Equality of Welfare', *Philosophy and Public Affairs*, 10.
—— (1981*b*). 'Equality of Resources', *Philosophy and Public Affairs*, 10.
Landesman, B. (1983). 'Egalitarianism', *Canadian Journal of Philosophy*, 13.
Nozick, Robert (1974). *Anarchy, State and Utopia*. New York: Basic Books.
Rawls, John (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
—— (1975). 'Fairness to Goodness', *Philosophical Review*, 84.
—— (1982). 'Social Unity and Primary Goods', in A. K. Sen and B. Williams, (eds.), *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.
—— (1985). 'Justice as Fairness, Political not Metaphysical', *Philosophy and Public Affairs*, 14.
—— (1986). 'Citizens' Needs and Primary Goods', unpublished.
Scanlon, T. M. (1975). 'Preference and Urgency', *Journal of Philosophy*, 72.
Sen, A. K. (1980). 'Equality of What?', in S. McMurrin (ed.), *Tanner Lectures on Human Values*, i. Cambridge: Cambridge University Press.
—— (1981). 'Ethical Issues in Income Distribution: National and International', in S. Grassman and E. Lundberg (eds.), *The World Economic Order: Past and Prospects*. London: Macmillan. Repr. in Sen (1984*a*).
—— (1982). *Choice, Welfare and Measurement*. Blackwell, Oxford.
—— (1983). 'Development: Which Way Now?', *Economic Journal*, 93.
—— (1983). 'Economics and the Family', *Asia Development Review*, 1.
—— (1984*a*). *Resources, Values and Development*. Oxford: Blackwell.
—— (1984*b*). 'The Living Standard', *Oxford Economic Papers*, 6.
—— (1985*a*). *Commodities and Capabilities*. Amsterdam: North-Holland.
—— (1985*b*). 'Rights and Capabilities', in T. Honderich (ed.), *Morality and Objectivity: A Tribute to J. L. Mackie*. London: Routledge and Kegan Paul. Repr. in Sen (1984*a*).
—— (1985*c*). 'Well-Being, Agency and Freedom: The Dewey Lectures 1984', *Journal of Philosophy*, 82.
—— (1987). 'Well-Being and Agency', unpublished.
—— *et al.* (1987). *The Standard of Living*. Cambridge: Cambridge University Press.
—— (1988). 'Capability and Well-being', unpublished draft of the paper that follows in this volume.

end p.29

# Capability and **Well□Being**

show chapter abstract and keywords
hide chapter abstract and keywords

**Amartya Sen**

## 1 Introduction

Capability is not an awfully attractive word. It has a technocratic sound, and to some it might even suggest the image of nuclear war strategists rubbing their hands in pleasure over some contingent plan of heroic barbarity. The term is not much redeemed by the historical Capability Brown praising particular pieces of *land*-not human beings-on the solid real-estate ground that they 'had capabilities'. Perhaps a nicer word could have been chosen when some years ago I tried to explore a particular approach to well-being and advantage in terms of a person's ability to do valuable acts or reach valuable states of being.[1]

[1] This was in a Tanner Lecture given at Stanford University in May 1979 ('Equality of What?'), later published as Sen (1980). The case for focusing on capability was introduced here in the specific context of evaluating inequality. I have tried to explore the possibility of using the capability perspective for analysing other social issues, such as well-being and poverty (Sen, 1982*a*, 1983*c*, 1985*b*), liberty and freedom (Sen, 1983*a*, 1988*a*, 1992), living standards and development (Sen, 1983*b*, 1984, 1987*b*, 1988*b*), gender bias and sexual divisions (Kynch and Sen, 1983; Sen, 1985*c*, 1990*b*), and justice and social ethics (Sen, 1982*b*, 1985*a*, 1990*a*).

The expression was picked to represent the alternative combinations of things a person is able to do or be-the various 'functionings' he or she can achieve.[2]

[2] Though at the time of proposing the approach, I did not manage to seize its Aristotelian connections, it is interesting to note that the Greek word *dunamin*, used by Aristotle to discuss an aspect of the human good, which is sometimes translated as 'potentiality', can be translated also as 'capability of existing or acting' (see Liddell and Scott, 1977: 452). The Aristotelian perspective and its connections with the recent attempts at constructing a capability-focused approach have been illuminatingly discussed by Martha Nussbaum (1988).

The capability approach to a person's advantage is concerned with evaluating it in terms of his or her actual ability to achieve various valuable functionings as a part of living. The corresponding approach to social advantage-for aggregative appraisal as well as for the choice of institutions and policy-takes the sets of individual capabilities as constituting an indispensable and central part of the relevant informational base of such evaluation. It differs from other approaches using other informational focuses, for example, personal utility (focusing on pleasures, happiness, or desire fulfilment), absolute or relative opulence (focusing on commodity bundles, real income, or real wealth), assessments of negative freedoms (focusing on procedural fulfilment of libertarian rights and rules of non-interference), comparisons of means of freedom (e.g. focusing on the holdings of 'primary goods', as in the Rawlsian theory of justice), and comparisons of resource holdings as a basis of just equality (e.g. as in Dworkin's criterion of 'equality of resources').

end p.30

Different aspects of the capability approach have been discussed, extended, used, or criticized by several authors, and as a result the advantages and difficulties of the approach have become more transparent.[3]

[3] See the contributions of Roemer (1982, 1986), Streeten (1984), Beitz (1986), Dasgupta (1986, 1988, 1989), Hamlin (1986), Helm (1986), Zamagni (1986), Basu (1987), Brannen and Wilson (1987), Hawthorn (1987), Kanbur (1987), Kumar (1987), Muellbauer (1987), Ringen (1987), B. Williams (1987), Wilson (1987), Nussbaum (1988, 1990), Griffin and Knight (1989*a*, 1989*b*), Riley (1988), Cohen (1990), and Steiner (1990). On related matters, including application, critique, and comparison, see also de Beus (1986), Kakwani (1986), Luker (1986), Sugden (1986), Asahi (1987), Delbono (1987), Koohi-Kamali (1987), A. Williams (1987), Broome (1988), Gaertner (1988), Stewart (1988), Suzumura (1988), de Vos and Hagennars (1988), Goodin (1985, 1988), Hamlin and Pettit (1989), Seabright (1989), Hossain (1990) and Schokkaert and van Ootegem (1990), among others.

There is, however, a need for a clearer and more connected account of the whole approach, particularly in view of some interpretational problems that have arisen in its assessment and use. This paper is an attempt at a clarificatory analysis at an elementary level. I shall also try to respond briefly to some interesting criticisms that have been made.

## 2 Functionings, Capability, and Values

Perhaps the most primitive notion in this approach concerns 'functionings'. *Functionings* represent parts of the state of a person-in particular the various things that he or she manages to do or be in leading a life. The *capability* of a person reflects the alternative combinations of functionings the person can achieve, and from which he or she can choose one collection.[4]

[4] If there are *n* relevant functionings, then a person's extent of achievement of all of them respectively can be represented by an *n*-tuple. There are several technical problems in the representation and analysis of functioning *n*-tuples and capability sets, on which see Sen (1985*b*: chs. 2, 4, and 7).

The approach is based on a view of living as a combination of various 'doings and beings', with quality of life to be assessed in terms of the capability to achieve valuable functionings.

Some functionings are very elementary, such as being adequately nourished, being in good health, etc., and these may be strongly valued by all, for obvious reasons. Others may be more complex, but still widely valued, such as achieving self-respect or being socially integrated. Individuals may, however, differ a good deal from each other in the weights they attach to these different functionings-valuable though they may all be-and the assessment of individual and social advantages must be alive to these variations.

In the context of some types of social analysis, for example, in dealing with extreme poverty in developing economies, we may be able to go a fairly long distance with a relatively small number of centrally important functionings and the corresponding basic capabilities (e.g. the ability to be well nourished and well sheltered, the capability of escaping avoidable morbidity and premature mortality, and so forth). In other contexts, including more general problems of economic development, the list may have to be much longer and much more diverse.

end p.31

Choices have to be faced in the delineation of the *relevant* functionings. The format always permits additional 'achievements' to be defined and included. Many functionings are of no great interest to the person (e.g. using a *particular* washing powder-much like other washing powders). [5]

[5] Bernard Williams (1987) raises this issue in his comments on my Tanner Lectures on the standard of living (pp. 98-101); on which see also Sen (1987*b*: 108-9). On the inescapable need for evaluation of different functioning and capabilities, see Sen (1985*b*: chs. 5-7). Just as the concentration on the commodity space in real-income analysis does not imply that every commodity must be taken to be equally valuable (or indeed valuable at all), similarly focusing on the space of functioning does not entail that each functioning must be taken to be equally valuable (or indeed valuable at all).

There is no escape from the problem of evaluation in selecting a class of functionings in the description and appraisal of capabilities. The focus has to be related to the underlying concerns and values, in terms of which some definable functionings may be important and others quite trivial and negligible. The need for selection and discrimination is neither an embarrassment, nor a unique difficulty, for the conceptualization of functioning and capability.

# 3 Value-Objects and Evaluative Spaces

In an evaluative exercise, we can distinguish between two different questions: (1) *What* are the objects of value? (2) *How valuable* are the respective objects? Even though *formally* the former question is an elementary aspect of the latter (in the sense that the objects of value are those that have positive weights), nevertheless the identification of the objects of value is *substantively* the primary exercise which makes it possible to pursue the second question.

Furthermore, the very identification of the set of value-objects, with positive weights, itself precipitates a 'dominance ranking' (*x* is at least as high as *y* if it yields at least as much of *each* of the valued objects). This dominance ranking, which can be shown to have standard regularity properties such as transitivity, can indeed take us some distance-often quite a long distance-in the evaluative exercise. [6]

[6] On this and on other formulations and uses of dominance ranking, see Sen (1970: chs. 1*, 7*, 9*).

The identification of the objects of value specifies what may be called an *evaluative space*. In standard utilitarian analysis, for example, the evaluative space consists of the individual utilities (defined in the usual terms of pleasures, happiness, or desire fulfilment). Indeed, a complete evaluative approach entails a class of 'informational constraints' in the form of ruling out *directly* evaluative use of various types of information, to wit, those that do not belong to the evaluative space.[7]

[7] On the crucial role of the informational basis, and on the formulation and use of informational constraints, see Sen (1970, 1977) and d'Aspremont and Gevers (1977).

The capability approach is concerned primarily with the identification of value-objects, and sees the evaluative space in terms of functionings and capabilities to function. This is, of course, itself a deeply evaluative exercise, but



answering question (1), on the identification of the objects of value, does not, on its own, yield a particular answer to question (2), regarding their relative values. The latter calls for a further evaluative exercise. Various substantive ways of evaluating functionings and capabilities can all belong to the general capability approach.

The selection of the evaluative space has a good deal of cutting power on its own, both because of what it *includes* as potentially valuable and because of what it *excludes*. For example, because of the nature of the evaluative space, the capability approach differs from utilitarian evaluation (more generally 'welfarist' evaluation[8]

[8] Welfarism requires that a state of affairs must be judged by the individual utilities in that state. It is one of the basic components of utilitarianism (the others being 'sum-ranking' and 'consequentialism'); on the factorization, see Sen (1982*a*) and Sen and Williams (1982).

) in making room for a variety of human acts and states as important in themselves (not just *because* they may produce utility, nor just to the *extent* that they yield utility).[9]

[9] Being happy and getting what one desires may be *inter alia* valued in the capability approach, but unlike in utilitarian traditions, they are not seen as the measure of all values.

It also makes room for valuing various freedoms-in the form of capabilities. On the other side, the approach does not attach direct-as opposed to derivative-importance to the *means* of living or *means* of freedom (e.g. real income, wealth, opulence, primary goods, or resources), as some other approaches do. These variables are not part of the evaluative space, though they can indirectly influence the evaluation through their effects on variables included in that space.

# 4 Capability and Freedom

The freedom to lead different types of life is reflected in the person's capability set. The capability of a person depends on a variety of factors, including personal characteristics and social arrangements. A full accounting of individual freedom must, of course, go beyond the capabilities of personal living and pay attention to the person's other objectives (e.g. social goals not directly related to one's own life), but human capabilities constitute an important part of individual freedom.

Freedom, of course, is not an unproblematic concept. For example, if we do not have the courage to choose to live in a particular way, even though we *could* live that way if we so chose, can it be said that we do have the freedom to live that way, i.e. the corresponding capability? It is not my purpose here to brush under the carpet difficult questions of this-and other-types. In so far as there are genuine ambiguities in the concept of freedom, that should be reflected in corresponding ambiguities in the characterization of capability. This relates to a methodological point, which I have tried to defend elsewhere, that if an underlying idea has an essential ambiguity, a precise formulation of that idea

must try to *capture* that ambiguity rather than hide or eliminate it. [10]

[10] On this, see Sen (1970, 1982*a*, 1987*a*). In many contexts, the mathematical representations should take the form of 'partial orderings' or 'fuzzy' relations. This is not, of course, a special problem with the capability approach, but applies generally to conceptual frameworks in social, economic, and political theory.

Comparisons of freedom raise interesting issues of evaluation. The claim is sometimes made that freedom must be valued independently of the values and preferences of the person whose freedom is being assessed, since it concerns the 'range' of choice a person has-*not* how she values the elements in that range or what she chooses from it. I do not believe for an instant that this claim is sustainable (despite some superficial plausibility), but had it been correct, it would have been a rather momentous conclusion, driving a wedge between the evaluation of *achievements* and that of *freedoms*. It would, in particular, be then possible to assess the freedom of a person independently of-or prior to-the assessment of the alternatives between which the person can choose. [11]

[11] The belief in this possibility seems to play a part in Robert Sugden's (1986) criticism of what he sees as my approach to capability evaluation, namely, a 'general strategy of trying to derive the value of a *set* of functioning vectors from prior ranking of the vectors themselves' (p. 821). He argues in favour of judging 'the value of being free to choose from a range of possible lives' *before* taking 'a view on what constitutes a valuable life'. This criticism is, in fact, based on a misunderstanding of the approach proposed, since it has been a part of my claim (on which more presently) that the judgement of the quality of life and the assessment of freedom have to be done *simultaneously* in an integrated way, and, in particular, that 'the quality of life a person enjoys is not merely a matter of what he or she achieves, but also of what options the person has had the opportunity to choose from' (Sen, 1985*b*: 69-70). But the point at issue in the present context is the possibility of judging a *range* of choice independently of the value characteristics of the *elements* in that range. It is this possibility that I am disputing.

.

How can we judge the goodness of a 'range' of choice independently of-or prior to-considering the nature of the alternatives that constitute that range? Some comparisons can, of course, be made in terms of set inclusion, for example, that reducing the 'menu' from which one can choose will *not* increase one's freedom. [12]

[12] Even this can be questioned when an expanded menu causes confusion, or the necessity to choose between a larger set of alternatives is a nuisance. But such problems can be dealt with through *appropriate* characterization of all the choices one has or does not have. This must include the consideration of the overall choice of having *or* not having to choose among a whole lot of relatively trivial alternatives (e.g. the choice of telling the telephone company to shut out mechanically dialled calls from sales agents offering a plethora of purchasing options). The issues involved in this kind of complex evaluation, incorporating choices over choices, are discussed in Sen (1992).

. But whenever neither set is entirely included in the other, we have to go beyond such 'subset reasoning'.

One alternative is simply to *count* the number of elements in the set as reflecting the value of the range of choice. [13]

[13] For an illuminating axiomatic derivation of the number-counting method of freedom evaluation, see Pattanaik and Xu (1990).

But this number-counting procedure leads to a rather peculiar accounting of freedom. It is odd to conclude that the freedom of a person is no less when she has to choose between three alternatives which she sees respectively as 'bad', 'awful', and 'gruesome' than when she has the choice between three alternatives which she assesses as 'good', 'excellent', and

end p.34

'superb'. [14]

[14] The unacceptability of this kind of number-counting evaluation of freedom is discussed in Sen (1985*b*). For an assessment of the axiomatic foundations of this and other methods of evaluation of freedom, see Sen (1991).

Further, it is always possible to add trivially to the number of options one has (e.g. tearing one's hair, cutting one's ears, slicing one's toes, or jumping through the window), and it would be amazing to see such additions as compensating for the loss of really valued options. [15]

[15] This type of case also shows why the set-inclusion ranking is best seen as a 'weak' relation of 'no worse than' or 'at least as good as', rather than as the 'strict' relation of 'better than'. Adding the option of 'slicing one's toes' to the set of valued options a person already has may not *reduce* her freedom (since one can reject toe-slicing), but it is hard to take it to be a strict *increase* in that person's freedom.

The assessment of the elements in a range of choice has to be linked to the evaluation of the freedom to choose among that range. [16]

[16] As was argued earlier, the relation is two-sided, and the evaluation of the freedom to lead a life and the assessment of the life led (including choosing freely) have to be done simultaneously, in a desegregated way.

# 5 Value-Purposes and Distinct Exercises

While the identification of value-objects and the specification of an evaluative space involve norms, the nature of the norms must depend on precisely what the purpose of the evaluation is. Assessing well-being may take us in one direction; judging achievement in terms of the person's *overall* goals may take us in a somewhat different direction, since a person can have objectives other than the pursuit of his or her own well-being. Judging achievement of either kind may also differ from the evaluation of the *freedom* to achieve, since a person can be advantaged in having more freedom and still end up achieving less.

We can make a fourfold classification of points of evaluative interest in assessing human advantage, based on two different distinctions. One distinction is between (1.1) the promotion of the person's *well-being*, and (1.2) the pursuit of the person's overall *agency goals*. The latter encompasses the goals that a person has reasons to adopt, which can *inter alia* include goals other than the advancement of his or her own well-being. It can thus generate orderings different from that of well-being. The second distinction is between (2.1) *achievement*, and (2.2) the *freedom to achieve*. This contrast can be applied both to the perspective of well-being and to that of agency. The two distinctions together yield four different concepts of advantage, related to a person: (1) 'well-being achievement', (2) 'agency achievement', (3) 'well-being freedom', and (4) 'agency freedom'. These different notions, which I have tried to discuss more extensively elsewhere, are not, of course, unrelated to each other, but nor are they necessarily identical. [17]

[17] Since a person's agency objectives will typically include, *inter alia*, his or her own well-being, the two will to some extent go together (e.g. an increase in well-being, other things being equal, will involve a higher agency achievement). In addition, a failure to achieve one's *non*-well-being objectives may also cause frustration, thereby reducing one's well-being. These and other connections exist between well-being and agency, but they do not make the two concepts congruent-nor isomorphic in the sense of generating the same orderings. Similarly, more freedom (either to have well-being or to achieve one's agency goals) may lead one to end up achieving more (respectively, of well-being or of agency success), but it is also possible for freedom to go up while achievement goes down, and vice versa. We have here four *interdependent* but *non-identical* concepts. These distinctions and their interrelations are discussed more fully in Sen (1985*a*, 1992).

end p.35

The assessment of each of these four types of benefit involves an evaluative exercise, but they are not the *same* evaluative exercise. They can also have very

disparate bearings on matters to which the evaluation and comparison of individual advantages are relevant. For example, in determining whether a person is deprived in a way that calls for assistance from others or from the state, a person's well-being may be, arguably, more relevant than his agency success (e.g. the state may have better grounds for offering support to a person for overcoming hunger or illness than for helping him to build a monument to his hero, even if he himself attaches more importance to the monument than to the removal of his hunger or illness). Furthermore, for adult citizens, *well-being freedom* may be more relevant to state policy, in this context, than *well-being achievement* (e.g. the state may have reason to offer a person adequate opportunities to overcome hunger, but not to insist that he must take up that offer and cease to be hungry). Interpersonal comparisons can be of many distinct types, with possibly dissimilar evaluative interests. Despite the interdependences between the different valuepurposes, they can generate quite distinct exercises with partly divergent concentration and relevance.

# 6 Well-Being, Agency, and Living Standards

The well-being achievement of a person can be seen as an evaluation of the 'wellness' of the person's state of being (rather than, say, the goodness of her contribution to the country, or her success in achieving her overall goals). The exercise, then, is that of assessing the constituent elements of the person's being seen from the perspective of her own personal welfare. The different functionings of the person will make up these constituent elements.

This does not, of course, imply that a person's well-being cannot be 'other-regarding'. Rather, the effect of 'other-regarding' concerns on one's well-being has to operate *through* some feature of the person's own being. Doing good may make a person contented or fulfilled, and these are functioning achievements of importance. In this approach, functionings are seen as central to the *nature* of well-being, even though the *sources* of well-being could easily be external to the person.

The functionings relevant for well-being vary from such elementary ones as escaping morbidity and mortality, being adequately nourished, having mobility, etc., to complex ones such as being happy, achieving self-respect, taking part in

end p.36

the life of the community, appearing in public without shame (the last a functioning that was illuminatingly discussed by Adam Smith [18]

[18] See Adam Smith (1776: Vol. ii, Bk V, ch. 2 (section on 'Taxes upon Consumable Commodities')), in Campbell and Skinner (1976), 469-71.

). The claim is that the functionings make up a person's being, and the evaluation of a person's wellbeing has to take the form of an assessment of these constituent elements.

If the value-purpose is changed from checking the 'well-ness' of the person's being to assessing the person's success in the pursuit of all the objectives that he has reason to promote, then the exercise becomes one of evaluation of 'agency achievement', rather than of well-being achievement. For this exercise, the space of functionings may be rather restrictive, since the person's goals may well include other types of objective (going well beyond the person's own state of being). Also, the difference between agency achievement and well-being achievement is not only a matter of *space* (the former taking us beyond the person's own life and functionings), but also one of differential *weighting* of the shared elements (i.e. for the functionings that are pertinent both to one's wellbeing and to one's other objectives, possibly different weights may be attached in agency evaluation *vis-à-vis* well-being appraisal).

The assessment of agency success is a broader exercise than the evaluation of well-being. It is also possible to consider 'narrower' exercises than the appraisal of well-being. A particularly important one is that of evaluating a person's *standard of living*. This, too, may take the form of focusing on the person's functionings, but in this case we may have to concentrate only on those influences on well-being that come from the nature of his *own* life, rather than from 'other-regarding' objectives or impersonal concerns. For example, the happiness generated by a purely other-regarding achievement (e.g. the freeing of political prisoners in distant countries) may enhance the person's well-being without, in any obvious sense, raising his living standard.

In the ethical context, the explicit recognition that one's well-being may often be affected by the nature of other people's lives is not, of course, new. Even Emperor Asoka, in the third century BC , noted the distinction clearly in one of his famous 'rock edicts' in the process of defining what should count as an injury to a person: 'And, if misfortune befalls the friends, acquaintances, companions and relations of persons who are full of affection [towards the former], even though they are themselves well provided for, [this misfortune] is also an injury to their own selves.' [19]

[19] Rock Edicts XIII at Erragudi, statement VII. For a translation and discussion, see Sircar (1979: 34).

The inability to be happy, which will be widely recognized as a failure of an important functioning (even though not the *only* important one, except in the hedonist version of utilitarianism), may arise either from sources within one's own life (e.g. being ill, or undernourished, or otherwise deprived), or from sources outside it (e.g. the pain that comes from sympathizing with others' misery). While both types of factor affect one's well-being, the case

end p.37

for excluding the latter from the assessment, specifically, of one's living standards would seem fairly reasonable, since the latter relates primarily to the lives of others, rather than one's own. [20]

[20] This view may be disputed by considering a different way of drawing the line between wellbeing and living standards. One common approach is to relate the assessment of living standards only to real incomes and to 'economic' or 'material' causes. On this see A. C. Pigou (1920); and on the conceptual differences see Bernard Williams (1987). But the Pigovian view has problems of its own. For example, if one has a disability that makes one get very little out of material income or wealth, or if one's life is shattered by an inconvenient and incurable illness (e.g. kidney problems requiring extensive dialysis), it is hard to claim that one's standard of living is high just because one is well heeled. I have discussed this question and related matters in Sen (1987*b*: 26-9, 109-10).

# 7 Why Capability, not Just Achievement?

The preceding discussion on the achievement of well-being and living standards has been related to functionings rather than to capabilities. This was done by design to introduce distinct problems in sequence, even though eventually an integrated view will have to be taken. In fact, the capability approach, as the terminology indicates, sees the capability set as the primary informational base. Why should we have to broaden our attention from functionings to capability?

We should first note that capabilities are defined derivatively from functionings. In the space of functionings any point, representing an *n*-tuple of functionings, reflects a combination of the person's doings and beings, relevant to the exercise. The capability is a *set* of such functioning *n*-tuples, representing the various alternative combinations of beings and doings any one (combination) of which the person can choose. [21]

[21] For formal characterizations, see Sen (1985*b*: chs. 2 and 7).

Capability is thus defined in the *space* of functionings. If a functioning achievement (in the form of an *n*-tuple of functionings) is a *point* in that space, capability is a *set* of

such points (representing the alternative functioning *n*-tuples from which one *n*-tuple can be chosen).

Note further that the capability set contains information about the actual functioning *n*-tuple chosen, since it too is obviously among the feasible *n*-tuples. The evaluation of a capability set may be based on the assessment of the particular *n*-tuple chosen from that set. Evaluation according to the achieved functioning combination is thus a 'special case' of evaluation on the basis of the capability set as a whole. In this sense, well-being achievement can be assessed on the basis of the capability set, even when no freedom-type notion influences that achievement. In this case, in evaluating the capability set for the value-purpose of assessing well-being achievement, we would simply have to identify the value of the capability set with the value of the achieved functioning *n*-tuple in it. The procedure of equating the value of the capability set to the value of *one* of the elements of that set has been called 'elementary evaluation'. [22]

[22] On this see Sen (1985*b*: 60-1). The distinguished element can be the *achieved* one (as in this case), or more specifically the *chosen* one (if there is a choice exercise in determining what happens), or the *maximal* one (in terms of some criterion of goodness). The three will coincide if what is achieved is achieved through choice, and what is chosen is chosen through maximization according to that criterion of goodness.

end p.38

Clearly, there is *at least* no informational loss in seeing well-being evaluation in terms of capabilities, rather than directly in terms of the achieved, or chosen, or maximal functioning *n*-tuple. While this indicates that the informational base of capability is at least as adequate as that of achieved functionings, the claim in favour of the capability perspective is, in fact, stronger. The advantages of the extension arise from two rather different types of consideration.

First, we may be interested not merely in examining 'well-being achievement', but also 'well-being freedom'. A person's actual freedom to live well and be well is of some interest in social as well as personal evaluation. [23]

[23] As was argued earlier in dealing with responsible adults, it may be appropriate to see the claims of individuals on society in terms of the *freedom* to achieve well-being (and thus in terms of real opportunities) rather than in terms of *actual achievements*. If the social arrangements are such that a responsible adult is given no less freedom (in terms of set comparisons) than others, but he still 'muffs' the opportunities and ends up worse off than others, it is possible to argue that no particular injustice is involved. On this and related matters, see Sen (1985*a*).

Even if we were to take the view, which will be disputed presently, that well-being achievement depends only on the achieved functionings, the 'well-being freedom' of a person will represent the freedom to enjoy the various possible well-beings associated with the different functioning *n*-tuples in the capability set. [24]

[24] The same capability set can than be used for the evaluation of both 'well-being achievement' (through *elementary evaluation*, concentrating on the achieved element) and 'well-being freedom' (through *non-elementary set evaluation*).

Second, freedom may have intrinsic importance for the person's well-being achievement. Acting freely and being able to choose may be directly conducive to well-being, not just because more freedom may make better alternatives available. This view is contrary to the one typically assumed in standard consumer theory, in which the contribution of a set of feasible choices is judged exclusively by the value of the best element available. [25]

[25] Thus, in standard consumer theory, set evaluation takes the form of elementary evaluation. For particular departures from that tradition, see Koopmans (1964) and Kreps (1979). In the Koopmans-Kreps approach, however, the motivation is not so much to see living freely as a thing of intrinsic importance, but to take note of uncertainty regarding one's own future preference by valuing—

instrumentally—the advantage of having more options in the future. On the motivational contrasts, see Sen (1985*a*, 1985*b*).

Even the removal of all the elements of a feasible set (e.g. of a 'budget set') other than the chosen best element is seen, in that theory, as no real loss, since the freedom to choose does not, in this view, matter in itself.

In contrast, if choosing is seen as a part of living (and 'doing *x*' is distinguished from 'choosing to do *x* and doing it'), then even 'well-being achievement' need not be independent of the freedom reflected in the capability set. [26]

[26] As was argued in an earlier paper, 'the "good life" is partly a life of genuine choice, and not one in which the person is forced into a particular life—however rich it might be in other respects' (Sen, 1985*b*): 69-70).

In that case, both 'well-being achievement' and 'well-being freedom' will have to be assessed in terms of capability sets. Both must then involve 'set evaluation' in a non-elementary way (i.e. without limiting the usable informational content of capability sets through elementary evaluation).

end p.39

There are many formal problems involved in the evaluation of freedom and the relationship between freedom and achievement. [27]

[27] See Sen (1985*b*, 1988*a*, 1991), Suppes (1987), Pattanaik and Xu (1990).

It is, in fact, possible to characterize functionings in a 'refined' way to take note of the 'counterfactual' opportunities, so that the characteristic of relating well-being achievement to functioning *n*-tuples could be retained without losing the substantive connection of well-being achievement to the freedom of choice enjoyed by the person. Corresponding to the functioning *x*, a 'refined' functioning (*x/S*) takes the form of 'having functioning *x* through choosing it from the set *S*'. [28]

[28] The characteristics and relevance of 'refined functioning' have been discussed in Sen (1985*a*, 1988*a*).

Sometimes even our ordinary language presents functionings in a refined way. For example, fasting is not just starving, but starving through rejecting the option of eating. The distinction is obviously important in many social contexts: we may, for example, try to eliminate involuntary hunger, but not wish to forbid fasting. The importance of seeing functionings in a refined way relates to the relevance of choice in our lives. The role of the choice involved in a capability set has been discussed above in the context of well-being only, but similar arguments apply to the assessment of agency achievement and the standard of living. [29]

[29] These issues are discussed in Sen (1985*a*, 1987*b*).

# 8 Basic Capability and Poverty

For some evaluative exercises, it may be useful to identify a subset of crucially important capabilities dealing with what have come to be known as 'basic needs'. [30]

[30] The 'basic needs' literature is extensive. For a helpful introduction, see Streeten *et al.* (1981). In a substantial part of the literature, there is a tendency to define basic needs in the form of needs for *commodities* (e.g. for food, shelter, clothing, health care), and this may distract attention from the fact that these commodities are no more than the *means* to real ends (inputs for valuable functionings and capabilities). On this question, see Streeten (1984). The distinction is particularly important since the relationship between commodities and capabilities may vary greatly between individuals even in the same society (and of course between different societies). For example, even for the elementary functioning of being well nourished, the relation between food intake and nutritional achievements varies greatly with metabolic rates, body size, gender, pregnancy, age, climatic conditions, epidemiological characteristics, and other factors (on these and related matters, see Drèze and Sen, 1989). The capability approach can accommodate the real issues underlying the concern for basic needs, avoiding the pitfall of 'commodity fetishism'.

There tends to be a fair amount of agreement on the extreme urgency of a class of needs. Particular moral and political importance may well be attached to fulfilling well-recognized, urgent claims.[31]

[31] The importance of socially recognized ideas of 'urgency' has been illuminatingly discussed by Thomas Scanlon (1975).

It is possible to argue that equality in the fulfilment of certain 'basic capabilities' provides an especially plausible approach to egalitarianism in the

end p.40

presence of elementary deprivation.[32]

[32] On this see Sen (1980). To avoid confusion, it should also be noted that the term 'basic capabilities' is sometimes used in quite a different sense from the one specified above, e.g. as a person's *potential* capabilities that *could* be developed, whether or not they are actually realized (this is the sense in which the term is used, for example, by Martha Nussbaum (1988).

The term 'basic capabilities', used in Sen (1980), was intended to separate out the ability to satisfy certain crucially important functionings up to certain minimally adequate levels. The identification of minimally acceptable levels of certain basic capabilities (below which people count as being scandalously 'deprived') can provide a possible approach to poverty, and I shall comment on the relation of this strategy to more traditional income-focused analyses of poverty. But it is also important to recognize that the use of the capability approach is not confined to basic capabilities only.[33]

[33] While the notion of basic capabilities was used in Sen (1980, 1983*c*), in later papers the capability approach has been used without identifying certain capabilities as 'basic' and others as not so (see e.g. Sen, 1984, 1985*a*, 1985*b*). This point is relevant to G.A. Cohen's distinction between focusing on what he calls 'midfare' and on functioning and capabilities. There are more important distinctions to explore (to be taken up in Section 9), but the contrasts look artificially sharper if the capability approach is seen as being confined *only* to the analysis of basic capabilities.

Turning to poverty analysis, identifying a minimal combination of basic capabilities can be a good way of setting up the problem of diagnosing and measuring poverty. It can lead to results quite different from those obtained by concentrating on inadequacy of income as the criterion of identifying the poor.[34]

[34] On this see Sen (1983*c*). See also Drèze and Sen (1989) and Hossain (1990).

The conversion of income into basic capabilities may vary greatly between individuals and also between different societies, so that the ability to reach minimally acceptable levels of basic capabilities can go with varying levels of minimally adequate incomes. The income-centred view of poverty, based on specifying an interpersonally invariant 'poverty line' income, may be very misleading in the identification and evaluation of poverty.

However, the point is sometimes made that poverty must, in some sense, be a matter of inadequacy of income, rather than a failure of capabilities, and this might suggest that the capability approach to poverty is 'essentially wrong-headed'. This objection overlooks both the motivational underpinning of poverty analysis and the close correspondence between capability failure and income inadequacy when the latter is defined taking note of *parametric variations* in income-capability relations. Since income is not desired for its own sake, any income-based notion of poverty must refer—directly or indirectly—to those basic ends which are promoted by income as means. Indeed, in poverty studies related to less developed countries, the 'poverty line' income is often derived explicitly with reference to nutritional norms. Once it is recognized that the relation between income and capabilities varies between communities and between people in the same community, the minimally adequate income level for reaching the same minimally acceptable capability levels will be

seen as variable—depending on personal and social characteristics. However, as long as minimal capabilities can be achieved

end p.41

by enhancing the income level (given the other personal and social characteristics on which capabilities depend), it will be possible (for the specified personal and social characteristics) to identify the minimally adequate income for reaching the minimally acceptable capability levels. Once this correspondence is established, it would not really matter whether poverty is defined in terms of a failure of basic capability or as a failure to have the *corresponding* minimally adequate income.[35]

[35] Technically, what is being used in this analysis is the 'inverse function', taking us back from specified capability levels to necessary incomes, given the other influences on capability. This procedure will not be usable, in this form, if there are people who are so handicapped in terms of personal characteristics (e.g. being a 'basket case') that no level of income will get them to reach minimally acceptable basic capabilities; such people would then be invariably identified as poor.

Thus, the motivationally more accurate characterization of poverty as a failure of basic capabilities can also be seen in the more traditional format of an income inadequacy. The difference in formulation is unimportant. What is really important is to take note of the interpersonal and intersocial variations in the relation between incomes and capabilities. That is where the distinctive contribution of the capability approach to poverty analysis lies.

## 9 Midfare, Functionings, and Capability

In this paper, I have so far been primarily concerned with clarifying and integrating the basic features of the capability approach, though I have taken the opportunity to address, in passing, some criticisms that have been made of this approach. In this section and in the next, I discuss two different lines of criticism — presented respectively by G.A. Cohen and Martha Nussbaum—arguing for different ways of analysing and assessing the problems of well-being and quality of life.

In his paper in this volume, and elsewhere (Cohen, 1989, 1990), G.A Cohen has provided a critical assessment of my writings on capability (and also of the theories of others—utilitarians, John Rawls, Ronald Dworkin, Thomas Scanlon, *et al.*), at the same time presenting his own answer to the question 'equality of what?'. Cohen is generous in giving credit where he reasonably can, and his assessment is positive in many ways, but the criticisms he makes, if sustained, would indicate a major motivational confusion as well as a conceptual inadequacy underlying the capability approach as I have tried to present it.

Cohen's main thesis is that in my paper 'Equality of What?' (Sen, 1980), I 'brought two distinct aspects of a person's condition under that single name [capability] and that this unnoticed duality has persisted in [my] subsequent writings'. 'Both aspects, or dimensions of assessment, should attract egalitarian interest, but one of them is not felicitously described as "capability" ' (p. 17). One aspect is concerned with 'a person being able to do certain basic things'. The other is what Cohen calls 'midfare', because 'it is in a certain sense midway between goods and utility'. 'Midfare is constituted of states of the person

end p.42

produced by goods, states in virtue of which utility levels take the values they do' (p. 18).

Cohen finds the dimension of midfare important for normative understanding and he notes, rightly, that I had put some emphasis on the *state* of the person, distinguishing it both from the *commodities* that help to generate that state, and from the *utilities* generated by the state. 'We must look, for example, at her nutrition level, and not just, as Rawlsians do, at her food supply, or, as welfarists do, at the utility she gets

out of eating food.' 'But', Cohen argues, 'this significant and illuminating reorientation is not equivalent to focusing on a person's capability.' 'Capability, and exercises of capability, form only one part of the intermediate midfare state' (p. 19); 'midfare, the product of goods which, in turn, generates utility, is not co-extensive with capability, and "capability" is therefore a bad name for midfare' (p. 20).

Is the distinction correct? I believe it is. The first thing to note is that Cohen's 'midfare' corresponds to what I have called a person's *functionings*, and not to *capability*. The two are related, but not meant to be the same. That distinction is, in fact, a basic part of the capability approach, and there is no embarrassment in acknowledging it. The real issue lies elsewhere, to wit, whether the capability set can have any relevance in analysing well-being, given the obvious connection between well-being and functionings (or midfare)—a connection that Cohen finds adequate for the analysis of well-being. This is an issue that was addressed in a less specific form earlier on in this paper (in Section 7). Cohen's preference for the perspective of midfare or functionings over that of capabilities relates to that substantive issue.

In Section 7, the relevance of the capability set for the analysis of well-being was defended on two different grounds, namely, (1) its connection with well-being *freedom* (even if well-being achievement depends only on the achieved functioning *n*-tuple), and (2) the possible importance of freedom (and thus of the capability set) for well-being *achievement* itself.

The second claim is the more controversial of the two. I believe it is correct, but I should also assert that even if it were incorrect, the capability approach would still be quite untarnished. As was discussed in Section 7, assessing well-being according to the achieved functioning *n*-tuple (or midfare) is a special case of the use of the capability perspective based on 'elementary evaluation' (focusing only on one distinguished element—the achieved functioning *n*-tuple—in the capability set). This point is obscured in Cohen's analysis by his conviction that 'the exercise of capability' must be a rather 'active' operation, and Cohen is misled by this diagnosis when he argues that he 'cannot accept . . . the associated athleticism, which comes when Sen adds that "the central feature of well-being is the ability to achieve valuable functionings" ' (p. 25). Cohen gives examples (e.g. small babies being well nourished and warm as a result of the activities of their parents) that clearly show that having midfare (or enjoying functionings) need not be a particularly athletic activity. I see no reason to object to this, since athleticism was never intended, despite the fact that Cohen has obviously been

end p.43

misled by my use of such words as 'capability' and 'achieving'. [36]

[36] Perhaps the word 'capability' *is* misleading, but I am not sure that this should be the case. The pieces of *land* to which Capability Brown attributed 'capability' could not have been much more active in looking after themselves than babies are. The crucial Greek word used in this context (by Aristotle among others), namely *dunamin*, can be translated as 'capability of existing or acting', and presumably 'existing' need not be the result of some vigorous 'exercise of capability'. Nor do I have any great difficulty in saying that the babies in question *did achieve* the state of being nourished and warm. Perhaps something else in my inept prose misled Cohen.

But let use move now from the minimalist defence to the claim that an active exercise of freedom might well be valuable for a person's quality of life and achieved well-being. Obviously, this consideration would be of no direct relevance in the case of babies (or the mentally disabled), who are not in a position to exercise reasoned freedom of choice (though babies *can* sometimes be amazingly cogent, choosy, and insistent). For people who are in a position to choose in a reasoned way and value that freedom to choose, it is hard to think that their well-being achievement would

never be affected if the freedom to choose were denied, even though the (unrefined) functioning vector (or midfare) were guaranteed by the actions of others. Even in Cohen's analysis of midfare, I should have thought that room would have to be found to see it in choice-inclusive terms, in much the same way that the functionings can be redefined in 'refined' terms (as discussed in Section 7). And if this is done, that would be isomorphic to including substantive consideration of the capability set, going beyond focusing exclusively on the achieved—unrefined—functioning vector (as was also discussed in Section 7).

Freedom has many aspects. Being free to live the way one would like may be enormously helped by the choice of others, and it would be a mistake to think of achievements only in terms of active choice *by oneself*. A person's ability to achieve various valuable functionings may be greatly enhanced by public action and policy,[37]

[37] On this see Drèze and Sen (1989).

and these expansions of capability are not unimportant for freedom for that reason. Indeed, I have argued elsewhere that 'freedom from hunger' or 'being free from malaria' need not be taken to be just rhetoric (as they are sometimes described); there is a very real sense in which the freedom to live the way one would like is enhanced by public policy that transforms epidemiological and social environments.[38]

[38] These issues are extensively discussed in Sen (1992). On related matters, see also Sen (1982*b*, 1983*a*, 1983*b*).

But the fact that freedom has that aspect does not negate the relevance of active choice by the person herself as an important component of living freely. It is because of the *presence* of this element (rather than the *absence* of others), that the act of choosing between the elements of a capability set has a clear relevance to the quality of life and well-being of a person.

But suppose we were to accept (wrongly, I believe) that this element of freedom really has no direct impact on the well-being of a person. In that case, the capability perspective could still be used to relate well-being achievement to

end p.44

achieved (unrefined) functionings (or midfare) through elementary evaluation. The need to relate well-being freedom to the capability set would also remain. That was indeed the first claim (in Section 7, pp. 38-9) in favour of the use of the capability set for analysing well-being (in this case, well-being freedom).

As was discussed earlier in this paper,[39]

[39] For a more extensive discussion, see Sen (1985*a*).

for many problems of individual behaviour and social policy, well-being freedom *is* a concept of relevance and importance. If achieved functionings (or midfare), defined in the 'unrefined' way, were all that mattered, we might be as worried about the rich person fasting as about the starving poor. If we are more concerned to eliminate the hunger of the latter, it is primarily because the former has the *capability* to be well nourished but chooses not to, whereas the latter lacks that capability and is forced into the state of starvation. Both may have the same midfare, but they differ in their capabilities. Capability does have importance in political and social analysis.

Motivationally, the focus on capability (in addition to achieved functionings) is, in fact, not altogether different from the concern that Cohen shows elsewhere for 'access to advantage'. Cohen notes that in his proposal

'advantage' is like Sen's 'functioning' in its wider construal, a heterogenous collection of desirable states of the person reducible neither to his resources bundle nor to his welfare level. And while 'access' includes what the term normally covers, I extend its

meaning under a proviso that anything which a person actually has counts as something to which he has access, no matter how he came to have it, and, hence, even if his coming to have it involved no exploitation of access in the ordinary sense (nor, therefore, any exercise of capability). If, for example, one enjoys freedom from malaria because others have destroyed the malaria-causing insects, then, in my special sense, such freedom from malaria is something to which one has access (Cohen, p. 28),

I do not see any great difficulty in 'extending' the meaning of 'access' in this way. An 'access' I enjoy may not have been created by me. But exactly the same applies to freedom and capability as well. The fact that a person has the freedom to enjoy a malaria-free life (or, to put it slightly differently, that his choice of a malaria-free life is feasible) may be entirely due to the actions of others (e.g., medical researchers, epidemiologists, public health workers), but that does not compromise the fact that he can indeed have a malaria-free life and has the capability (thanks largely to others) to achieve such a life. [40]

[40] On this see also Drèze and Sen (1989) and Sen (1992).

I don't even see that much 'extension' of ordinary usage is involved in such use of the terms freedom and capability (even though this is not the central issue in any case). [41]

[41] As was mentioned earlier (in footnote 2), in their well-known Greek-English lexicon, even Liddell and Scott (1977) had translated the Greek word *dunamin*, central to Aristotle's concept of human good, as 'capability of existing or acting' (p. 452).

Indeed, even the expression 'freedom from malaria', used also by Cohen, is a pointer to the fact that ordinary language takes a less narrow view of the use of the term freedom. Similarly, there is no underlying presumption that we have the capability to lead a malaria-free life only if we have

end p.45

ourselves gone around exterminating the malaria-causing insects.

Turning to a different issue also raised by Cohen, the really interesting question is not whether 'equality of access to advantage' coincides with 'capability' in general, since capability (as was discussed earlier) is a more versatile concept and its particular characterization has to be related to the 'evaluative purpose' of the exercise (e.g., whether 'agency' or 'well-being' is the focal concern in that exercise). But if advantage is seen specifically in terms of well-being (ignoring the agency aspect), then Cohen's 'equality of access to advantage' would indeed be very like equality of well-being freedom, defined in terms of evaluation of capability sets from that perspective. [42]

[42] Indeed, it is precisely as well-being freedom that 'advantage' was defined in Sen (1985*b*: 5-7, 59-71).

Cohen's analysis has brought out the distinctions between a number of different problems which are all addressed in the capability approach but which require separate treatment. While substantive differences may remain between his focus and mine (e.g. about the importance of *choosing* as a constitutive element in the quality of life), Cohen's analysis has greatly helped to pinpoint some focal issues and concerns, and the need to address them explicitly.

## 10 The Aristotelian Connections and Contrasts

In earlier writings I have commented on the connection of the capability approach with some arguments used by Adam Smith and Karl Marx. [43]

[43] See, particularly, Smith (1776) and Marx (1844). The connections are discussed in Sen (1984, 1985*a*, 1987*b*).

However, the most powerful conceptual connections would appear to be with the Aristotelian view of the human good. Martha Nussbaum (1988, 1990) has discussed illuminatingly the Aristotelian analysis of 'political distribution', and its relation to the capability approach. The Aristotelian account of the human good is explicitly linked with the necessity to 'first ascertain the function of man' and it then proceeds to explore 'life in the sense of activity'. [44]

[44] See particularly *The Nicomachean Ethics*, Bk I, s. 7; in the translation by Ross (1980: 12-14).

The basis of a fair distribution of capability to function is given a central place in the Aristotelian theory of political distribution. In interpreting Aristotle's extensive writings on ethics and politics, it is possible to note some ambiguity and indeed to find some tension between different propositions presented by him, but his recognition of the crucial importance of a person's functionings and capabilities seems to emerge clearly enough, especially in the political context of distributive arrangements.

While the Aristotelian link is undoubtedly important, it should also be noted that there are some substantial differences between the way functionings and capabilities are used in what I have been calling the capability approach and the way they are dealt with in Aristotle's own analysis. Aristotle believes, as Nussbaum (1988) notes, 'that there is just one list of functionings (at least at a certain level of generality) that do in fact constitute human good living' (p. 152).

end p.46

That view would not be inconsistent with the capability approach presented here, but *not*, by any means, *required* by it.

The capability approach has indeed been used (for example, in Sen, 1983*c*, 1984) to argue that while the *commodity* requirements of such capabilities as 'being able to take part in the life of the community' or 'being able to appear in public without shame' vary greatly from one community to another (thereby giving the 'poverty line' a relativist character in the space of commodities), there is much less variation in the *capabilities* that are aimed at through the use of these commodities. This argument, suggesting less variability at a more intrinsic level, has clear links with Aristotle's identification of 'non-relative virtues', but the Aristotelian claims of uniqueness go much further. [45]

[45] On this see Nussbaum (1990).

Martha Nussbaum, as an Aristotelian, notes this distinction, and also points to Aristotle's robust use of an objectivist framework based on a particular reading of human nature. She suggests the following:

It seems to me, then, that Sen needs to be more radical than he has been so far in his criticism of the utilitarian accounts of well-being, by introducing an objective normative account of human functioning and by describing a procedure of objective evaluation by which functionings can be assessed for their contribution to the good human life. [46]

[46] Nussbaum (1988: 176).

I accept that this would indeed be a systematic way of eliminating the incompleteness of the capability approach. I certainly have no great objection to anyone going on that route. My difficulty with accepting that as the *only* route on which to travel arises partly from the concern that this view of human nature (with a unique list of functionings for a good human life) may be tremendously overspecified, and also from my inclination to argue about the nature and importance of the type of objectivity involved in this approach. But mostly my intransigence arises, in fact, from the consideration that the use of the capability approach as such does not require taking that route, and the deliberate incompleteness of the capability approach

permits other routes to be taken which also have some plausibility. It is, in fact, the feasibility as well as the usefulness of a general approach (to be distinguished from a complete evaluative blueprint) that seems to me to provide good grounds for separating the general case for the capability approach (including, *inter alia*, the Aristotelian theory) from the special case for taking on *exclusively* this particular Aristotelian theory.

In fact, no matter whether we go the full Aristotelian way, which will also need a great deal of extension as a theory for practical evaluation, or take some other particular route, there is little doubt that the kind of *general* argument that Aristotle uses to motivate his approach does have a wider relevance than the defence of the particular form he gives to the nature of human good. This applies *inter alia* to Aristotle's rejection of opulence as a criterion of achievement (rejecting wealth and income as the standards), his analysis of *eudaimonia* in terms of valued activities (rather than relying on readings of mental states, as in some utilitarian procedures), and his assertion of the need to examine the

end p.47

processes through which human activities are chosen (thereby pointing towards the importance of freedom as a part of living).

# 11 Incompleteness and Substance

The Aristotelian critique points towards a more general issue, namely, that of the 'incompleteness' of the capability approach-both in generating substantive judgements and in providing a comprehensive theory of valuation. Quite different specific theories of value may be consistent with the capability approach, and share the common feature of selecting value-objects from functionings and capabilities. Further, the capability approach can be used with different methods of determining relative weights and different mechanisms for actual evaluation. The approach, if seen as a theory of algorithmic evaluation, would be clearly incomplete. [47]

[47] This relates to one part of the critique presented by Beitz (1986).

It may well be asked: why pause at outlining a general approach, with various bits to be filled in, rather than 'completing the task'? The motivation underlying the pause relates to the recognition that an agreement on the usability of the capability approach-an agreement on the nature of the 'space' of value-objects-need not *presuppose* an agreement on how the valuational exercise may be completed. It is possible to disagree both on the exact *grounds* underlying the determination of relative weights, and on the *actual* relative weights chosen, [48]

[48] On this see Sen (1985*b*: chs. 5-7).

even when there is reasoned agreement on the general nature of the value-objects (in this case, personal functionings and capabilities). If reasoned agreement is seen as an important foundational quality central to political and social ethics, [49]

[49] On this question, see Rawls (1971), Scanlon (1982), B. Williams (1985).

then the case for the pause is not so hard to understand. The fact that the capability approach is consistent and combinable with several different substantive theories need not be a source of embarrassment.

Interestingly enough, despite this incompleteness, the capability approach does have considerable 'cutting power'. In fact, the more challenging part of the claim in favour of the capability approach lies in what it denies. It differs from the standard utility-based approaches in not insisting that we must value *only happiness* (and sees, instead, the state of being happy as one among several objects of value), or *only*

*desire fulfilment* (and takes, instead, desire as useful but imperfect evidence-frequently distorted-of what the person herself values). [50]

[50] For comparisons and contrasts between the capability approach and utilitarian views, see Sen (1984, 1985*a*).

It differs also from other-non-utilitarian-approaches in not placing among value-objects *primary goods as such* (accepting these Rawlsian-focus variables only derivatively and instrumentally and only to the extent that these goods promote capabilities), or *resources as such* (valuing this Dworkinian perspective only in terms of the impact of resources on functionings and capabilities), and so forth. [51]

[51] See Rawls (1971, 1988*a*, 1988*b*), Dworkin (1971, 1988*a*, 1988*b*), Dworkin (1981), and Sen (1980, 1984, 1990*a*).

end p.48

A general acceptance of the intrinsic relevance and centrality of the various functionings and capabilities that make up our lives does have substantial cutting power, but it need not be based on a prior agreement on the relative values of the different functionings or capabilities, or on a specific procedure for deciding on those relative values.

Indeed, it can be argued that it may be a mistake to move on relentlessly until one gets to exactly one mechanism for determining relative weights, or — to turn to a different aspect of the 'incompleteness' — until one arrives at exactly one interpretation of the metaphysics of value. There are substantive differences between different ethical theories at different levels, from the meta-ethical (involving such issues as objectivity) to the motivational, and it is not obvious that for substantive political and social philosophy it is sensible to insist that all these general issues be resolved *before* an agreement is reached on the choice of an evaluative space. Just as the utilization of actual weights in practical exercises may be based on the acceptance of a certain *range* of variability of weights (as I have tried to discuss in the context of the *use* of the capability approach [52]

[52] See Sen (1985*b*); on the general strategy of using 'intersection partial orders', see Sen (1970, 1977).

), even the general rationale for using such an approach may be consistent with some ranges of answers to foundational questions.

# 12 A Concluding Remark

In this paper I have tried to discuss the main features of the capability approach to evaluation: its claims, its uses, its rationale, its problems. I have also addressed some criticisms that have been made of the approach. I shall not try to summarize the main contentions of the paper, but before concluding, I would like to emphasize the plurality of purposes for which the capability approach can have relevance.

There are different evaluative problems, related to disparate value-purposes. Among the distinctions that are important is that between well-being and agency, and that between achievement and freedom. The four categories of intrapersonal assessment and interpersonal comparison that follow from these two distinctions (namely, well-being achievement, well-being freedom, agency achievement, and agency freedom) are related to each other, but are not identical. The capability approach can be used for each of these different types of evaluation, though not with equal reach. It is particularly relevant for the assessment of well-being — in the form of both achievement and freedom — and for the related problem of judging living standards. As far as social judgements are concerned, the individual evaluations feed directly into social assessment. Even though the original motivation for using the capability

approach was provided by an examination of the question 'equality of what?' (Sen, 1980), the use of the approach, if successful for equality, need not

end p.49

be confined to equality only.[53]

[53] Corresponding to 'equality of what?', there is, in fact, also the question: 'efficiency of what?'

The usability of the approach in egalitarian calculus depends on the plausibility of seeing individual advantages in terms of capabilities, and if that plausibility is accepted, then the same general perspective can be seen to be relevant for other types of social evaluation and aggregation.

The potentially wide relevance of the capability perspective should not come as a surprise, since the capability approach is concerned with showing the cogency of a particular *space* for the evaluation of individual opportunities and successes. In any social calculus in which individual advantages are constitutively important, that space is of potential significance.

**Bibliography**

Aristotle (4th c. BC ). *The Nicomachean Ethics*: see Ross (1980).

Arneson, R. (1987). 'Equality and Equality of Opportunity for Welfare', *Philosophical Studies*, S4.

Asahi, J. (1987). 'On Professor Sen's Capability Theory', mimeographed, Tokyo.

Basu, K. (1987). 'Achievements, Capabilities and the Concept of Well-being', *Social Choice and Welfare*, 4.

Beitz, C. R. (1986). 'Amartya Sen's *Resources, Values and Development*', *Economics and Philosophy*, 2.

Brannen, J., and Wilson, G. (eds.) (1987). *Give and Take in Families*. London: Allen and Unwin.

Broome, J. (1988). Review of On Ethics and Economics and *The Standard of Living* in *London Review of Books*.

Campbell, R. H., and Skinner, A. S. (eds.) (1976). Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Clarendon Press.

Cohen, G. A. (1989). 'On the Currency of Egalitarian Justice', *Ethics*, 99.

—— (1990). 'Equality of What? On Welfare, Goods and Capabilities', *Recherches Economiques de Louvain*, 56.

Culyer, A. J. (1985). 'The Scope and Limits of Health Economics', *Okonomie des Gesundheitswesens*.

Dasgupta, P. (1986). 'Positive Freedom, Markets and the Welfare State', *Oxford Review of Economic Policy*, 2.

—— (1988). 'Lives and Well-being', *Social Choice and Welfare*, 5.

—— (1989). 'Power and Control in the Good Polity', in Hamlin and Pettit (1989).

d'Aspremont, C., and Gevers, L. (1977). 'Equity and the Informational Basis of Collective Choice', *Review of Economic Studies*, 46.

deBeus, Jos (1986). 'Sen's Theory of Liberty and Institutional Vacuum', University of Amsterdam.

deVos, K., and Hagenaars, Aldi J. M. (1988). 'A Comparison between the Poverty Concepts of Sen and Townsend', mimeographed, University of Leiden.

end p.50

Deaton, A., and Muellbauer, J. (1980). *Economics and Consumer Behaviour*. Cambridge: Cambridge University Press.

Delbono, F. (1987). Review article on *Commodities and Capabilities, Economic Notes*.

Drèze, J., and Sen, A. (1989). *Hunger and Public Action*. Oxford: Clarendon Press.

Dworkin, R. (1981). 'What is Equality? Part 2: Equality of Resources', *Philosophy and Public Affairs*, 10.
Elster, J., and Hylland, A. (eds.) (1986). *Foundations of Social Choice Theory*. Cambridge: Cambridge University Press.
Gaertner, W. (1988). Review of *Commodities and Capabilities* in *Zeitschriften fur National Okommnomia*, 48.
Goodin, R. E. (1985). *Political Theory and Public Policy*. Chicago and London: University of Chicago Press.
—— (1988). *Reasons for Welfare: Political Theory of the Welfare State*. Princeton: Princeton University Press.
Griffin, J. (1986). *Well-being*. Oxford: Clarendon Press.
Griffin, K., and Knight, J. (eds.) (1989*a*). 'Human Development in the 1980s and Beyond', special number, *Journal of Development Planning*, 19.
—— (1989*b*). 'Human Development: The Case for Renewed Emphasis', *Journal of Development Planning*, 19.
Hamlin, A. P. (1986). *Ethics, Economics and the State*. Brighton: Wheatsheaf Books.
Hamlin, A., and Pettit, P. (eds.) (1989). *The Good Polity: Normative Analysis of the State*. Oxford: Blackwell.
Hare, R. (1981). *Moral Thinking*. Oxford: Clarendon Press.
Harrison, R. (ed.) (1979). *Rational Action*. Cambridge: Cambridge University Press.
Hawthorn, G. (1987). 'Introduction' in Sen (1987*b*).
Helm, D. (1986). 'The Assessment: The Economic Border of the State', *Oxford Review of Economic Policy*, 2.
Hossain, I. (1990). *Poverty as Capability Failure*. Helsinki: Swedish School of Economics and Business Administration.
Kakwani, N. (1986). *Analysing Redistribution Policies*. Cambridge: Cambridge University Press.
Kanbur, R. (1987). 'The Standard of Living: Uncertainty, Inequality and Opportunity', in Sen (1987*b*).
Koohi-Kamali, F. (1988). 'The Pattern of Female Mortality in Iran and some of its Causes', Applied Economics Discussion Paper 62, Oxford Institute of Economics and Statistics.
Koopmans, T. C. (1964). 'On Flexibility of Future Preference', in M. W. Shelly and G. L. Bryan (eds.), *Human Judgments and Optimality*. New York: Wiley.
Kreps, D. (1979). 'A Representation Theorem for Preference for Flexibility', *Econometrica*, 47.
Kumar, B. G. (1987). Poverty and Public Policy: Government Intervention and Level in Kerala, India. D.Phil. dissertation, Oxford University.
Kynch, J., and Sen, A. (1983). 'Indian Women: Well-being and Survival', *Cambridge Journal of Economics*, 7.
Liddell, H. G., and Scott, R. (1977). *A Greek-English Lexicon*, extended by H. S. Jones and R. McKenzie. Oxford: Clarendon Press.
Luker, W. (1986). 'Welfare Economics, Positivist Idealism and Quasi-Experimental Methodology', mimeographed, University of Texas, Austin.
end p.51
Marx, K. (1844). *Economic and Philosophic Manuscript*. English translation. London: Lawrence and Wishart, 1977.
McMurrin, S. M. (ed.) (1980). *Tanner Lectures on Human Values*, i. Salt Lake City: University of Utah Press, and Cambridge: Cambridge University Press.
Muellbauer, J. (1987). 'Professor Sen on the Standard of Living', in Sen (1987*b*).

Nussbaum, M. (1988). 'Nature, Function, and Capability: Aristotle on Political Distribution', *Oxford Studies in Ancient Philosophy*, suppl. vol.
—— (1990). 'Non-Relative Virtues: An Aristotelian Approach', *Midwest Studies in Philosophy*, 13; revised version in this volume.
Pattanaik, P. K., and Xu, Yongsheng (1990). 'On Ranking Opportunity Sets in Terms of Freedom of Choice', *Recherches Economiques de Louvain*, S6.
Pigou, A. C. (1920). *The Economics of Welfare*. London: Macmillan.
Rawls, J. (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press, and Oxford: Clarendon Press.
—— (1988*a*). 'Priority of Right and Ideas of the Good', *Philosophy and Public Affairs*, 17.
—— (1988*b*). 'Reply to Sen', mimeographed, Harvard University.
—— *et al.* (1987). *Liberty, Equality, and Law: Selected Tanner Lectures on Moral Philosophy*, ed. S. McMurrin. Cambridge: Cambridge University Press, and Salt Lake City: University of Utah Press.
Riley, J. (1988). *Liberal Utilitarianism: Social Choice Theory and J. S. Mill's Philosophy*. Cambridge: Cambridge University Press.
Ringen, J. (1987). *The Possibility of Politics: A Study of the Economy of the Welfare State*. Oxford: Clarendon Press.
Roemer, J. (1982). *A General Theory of Exploitation and Class*. Cambridge, Mass.: Harvard University Press.
—— (1986). 'An Historical Materialist Alternative to Welfarism', in Elster and Hylland (1986).
Ross, D. (1980). Aristotle, *The Nicomachean Ethics*. The World's Classics. Oxford: Oxford University Press.
Scanlon, T. M. (1979). 'Preference and Urgency', *Journal of Philosophy*, 72.
—— (1982). 'Contractualism and Utilitarianism', in Sen and Williams (1982).
Schokkaert, E., and van Ootegem, L. (1990). 'Sen's Concept of the Living Standard Applied to the Belgian Unemployed', *Recherches Economiques de Louvain*, S6.
Seabright, P. (1989). 'Social Choice and Social Theories', *Philosophy and Public Affairs*, 18.
Sen, A. K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day. Republished Amsterdam: North-Holland, 1979.
—— (1977). 'On Weights and Measures: Informational Constraints in Social Welfare Analysis', *Econometrica*, 45.
—— (1980). 'Equality of What?' (1979 Tanner Lecture at Stanford), in McMurrin (1980); repr. in Sen (1982*a*) and Rawls *et al.* (1987).
—— (1982*a*). *Choice, Welfare and Measurement*. Oxford: Blackwell, and Cambridge, Mass.: MIT Press.
—— (1982*b*). 'Rights and Agency', *Philosophy and Public Affairs*, 11.
—— (1983*a*). 'Liberty and Social Choice', *Journal of Philosophy*, 80.
—— (1983*b*). 'Development: Which Way Now?', *Economic Journal*, 93; repr. in Sen (1984).
—— (1983*c*). 'Poor, Relatively Speaking', *Oxford Economic Papers*, 35; repr. in Sen (1984).

end p.52

—— (1984). *Resources, Values and Development*. Oxford: Blackwell, and Cambridge, Mass.: Harvard University Press.
—— (1985*a*). 'Well-being, Agency and Freedom: The Dewey Lectures 1984', *Journal of Philosophy*, 82.

—— (1985*b*). *Commodities and Capabilities*. Amsterdam: North-Holland.
—— (1985*c*). 'Women, Technology and Sexual Divisions', *Trade and Development*, 6.
—— (1987*a*). *On Ethics and Economics*. Oxford: Blackwell.
—— (1987*b*). *The Standard of Living* (1985 Tanner Lectures at Cambridge, with contributions by Keith Hart, Ravi Kanbur, John Muellbauer, and Bernard Williams, edited by G. Hawthorn). Cambridge: Cambridge University Press.
—— (1988*a*). 'Freedom of Choice: Concept and Content', *European Economic Review* 32.
—— (1988*b*). 'The Concept of Development', in H. Chenery and T.N. Srinivasan (eds.), *Handbook of Development Economics*. Amsterdam: North-Holland.
—— (1990*a*). 'Justice: Means versus Freedoms', *Philosophy and Public Affairs*, 19.
—— (1990*b*). 'Gender and Cooperative Conflicts', in Tinker (1990).
—— (1991). 'Preference, Freedom and Social Welfare', *Journal of Econometrics*, 50.
—— (1992). *Inequality Reexamined*. Oxford: Clarendon Press.
—— and Williams, B. (eds.) (1982). *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.
Sircar, D. C. (1979). *Asokan Studies*. Calcutta: Indian Museum.
Smith, Adam (1776). *An Inquiry into the Nature of Causes of the Wealth of Nations*: see Campbell and Skinner (1976).
Steiner, H. (1986). 'Putting Rights in their Place: An Appraisal of Amartya Sen's Work on Rights', mimeographed, University of Manchester.
Stewart, F. (1988). 'Basic Needs Strategies, Human Rights and the Right to Development', mimeographed, Queen Elizabeth House.
Streeten, P. (1984). 'Basic Needs: Some Unsettled Questions', *World Development*, 12.
Streeten, P., *et al.* (1981). *First Things First: Meeting Basic Needs in Developing Countries*. New York: Oxford University Press.
Sugden, R. (1986). Review of *Commodities and Capabilities*, *Economic Journal*, 96.
Suppes, P. (1987). 'Maximizing Freedom of Decision: An Axiomatic Analysis', in G. R. Feiwel (ed.), *Arrow and the Foundations of Economic Policy*. New York: New York University Press.
Suzumura, K. (1988). Introduction to the Japanese translation of *Commodities and Capabilities*. Tokyo: Iwanami.
Tinker, I. (ed.) (1990). *Persistent Inequalities*. New York: Oxford University Press.
Williams, A. (1985). 'Economics of Coronary Bypass Grafting', *British Medical Journal*, 291.
—— (1991). 'What is Health and Who Creates it?', in J. Hutton *et al.* (eds.), *Dependency to Enterprise*. London: Routledge.
Williams, B. (1985). *Ethics and the Limits of Philosophy*. London: Fontana, and Cambridge, MA: Harvard University Press.
—— (1987). 'The Standard of Living: Interests and Capabilities', in Sen (1987*b*).
Wilson, G. (1987). *Money in the Family*. Aldershot: Avebury.
Zamagni, S. (1986). 'Introduzione', in A. Sen, *Scelta, Benessere, Equita*. Bologna: Il Mulino.

end p.53

# G.A. Cohen: Equality of What? on Welfare, Goods and Capabilities Amartya Sen: Capability and Well-Being

show chapter abstract and keywords
hide chapter abstract and keywords

## Christine M. Korsgaard

These two papers propose ways to understand and measure certain important aspects of the quality of life. Gerald Cohen is concerned specifically with equality: what it is people must have an equal amount of, in order for them to be equal in the sense egalitarians ought to care about. His paper criticizes the views of Rawls and Sen, and offers his own answer, which is that people should be equal in their *access to advantage*.[1]

[1] Cohen defends his own view at greater length in his 'On the Currency of Egalitarian Justice', *Ethics*, 99 (July 1989), 906-44.

Both terms in this formula are meant to be eclectic. 'Advantage' includes both welfare and resources, and whatever else we might decide is a 'desirable state of the person'.[2]

[2] Cohen, 'Equality of What? On Welfare, Resources, and Capabilities', in this volume, 9-29.

You have 'access' to things you have or can get or are given to you. To say that people should be equal in their access to advantage, according to Cohen, is to say that any involuntary disadvantage-any disadvantage which either was not chosen or cannot be voluntarily overcome-ought to be eliminated or compensated.

Amartya Sen's view is that the quality of a person's life should be assessed in terms of the person's capabilities. A capability is the ability or potential to do or be something-more technically, to achieve a certain functioning. Functionings are divided into four overlapping categories, which Sen calls well-being freedom, well-being achievement, agency freedom, and agency achievement. Our capabilities are our potentials for all of these things. Sen's view, like Cohen's began as a thesis about the kind of value that egalitarians ought to be concerned about. People ought to be made equal in their capabilities, or at least in their *basic* capabilities. But Sen is now prepared to claim that his view provides a metric for other purposes as well.

We may assess any proposal about what constitutes the quality of life in various ways. First, we may assess it simply as a philosophical proposal about what a good life is. Second, we may assess it for its legitimacy as a *political* objective: whether it is the sort of thing we ought to bring about through political means. And third, we may assess it for its utility in determining actual political and economic policy-that is, whether it provides accurate enough measures to assess the effects of policy. Obviously, these three forms of assessment are not unrelated, and all are necessary. If capability and access to advantage were not at least important features of a good life, equalizing or maximizing them would not be politically desirable. If they were not legitimate political objectives, the fact that they enable us to make measurements, if it were one, would not carry any weight.

end p.54

Nevertheless, the three forms of assessment are separate. And the question I will focus on, in considering both access to advantage and capability, is their legitimacy as political objectives.

In assessing the legitimacy of a political objective, we must keep in mind that the state uses coercion to achieve its aims, and that the use of coercion is in general wrong. Only the pursuit of certain kinds of aims, in special circumstances, can justify the use of coercion; only these, therefore, can be legitimate as political objectives. There are two familiar and related devices to which political philosophers appeal in order to alleviate the moral problem created by the coercive nature of government. One is the idea of the consent of the governed, either actual or hypothetical, to the general institution of government. I take it for granted that since actual consent has not in fact been given, hypothetical consent is the best we can hope for. That is, the

best we can hope is to show that political institutions and their objectives are ones that it is reasonable for people to consent to. The other device is the mechanism of voting, as a way of approximating actual consent within society. Some philosophers believe that these two devices work together to dissipate the moral problem created by the use of state coercion. Each individual is supposed, hypothetically, to consent to submit to political decisions as long as she is allowed, actually, to contribute to those decisions by voting. But at least since Mill, few philosophers would say that voting is *sufficient* to justify the use of state coercion, for it leads to one version of what de Tocqueville called 'the tyranny of the majority'. [3]

[3] Mill, 1859: ch. 1.

Even in a democratic society, we must place limits on what may be achieved through political means. Political objectives must reflect the reasons people have for submitting to the coercive authority of the state in the first place.

The legitimacy of a political objective obviously depends on what sort of political philosophy one espouses, what one thinks the state is for. So I will begin by discussing the kinds of political objectives that have been thought legitimate, and sketching a view that I find plausible. I will then assess Sen's and Cohen's proposals in the light of that view. Following Rawls, I begin with the distinction between liberal and non-liberal political theories. [4]

[4] Rawls, 1982*b*: 159-85.

In a liberal theory, the purpose of the state is to allow each citizen to pursue his or her own conception of the good. In a non-liberal theory, some conception of the good is taken as philosophically established, and the goal of the state is to realize that conception. If a non-liberal theory is accepted, my first two forms of assessment are not after all separate: to show that something is a legitimate political objective, all we need do is show that it is indeed an established good. According to such theories, the state is in the business of bringing about the good. As Rawls points out, classical utilitarianism is strictly speaking a non-liberal theory, since it takes the goodness of the maximization of pleasure as both philosophically established and capable of justifying political policy. [5]

[5] Ibid., 160.

An Aristotelian theory that takes the purpose of the state to be to educate the citizens for a virtuous life, or a Marxist theory aimed

end p.55

at rendering us truly human, would also be non-liberal. These theories take a certain conception of the good life to be both established and capable of justifying the use of state coercion.

Rawls defines a liberal theory, by contrast, as one that 'allow[s] for a plurality of different and opposing, and even incommensurable, conceptions of the good'. [6]

[6] Ibid.

But the phrase 'allows for' is unfortunately ambiguous. Accordingly, there are two kinds of liberal. One kind of liberal agrees with the non-liberal that the purpose of the state is to enable the citizens to achieve a good life, but disagrees that there is just one established conception of the good life. It is important that each person choose, construct, and pursue her own conception of the good. There are a number of different reasons why one might hold this view. They correspond roughly to the various reasons for religious tolerance, and, like those, they are held by many liberals in a rather jumbled way. One reason is *scepticism*: there is no best life, or anyway we cannot prove there is, so we have no solid ground for forcing people to lead one kind of life rather than another. Another reason is *ethical individualism*. According to this

view, the goodness of a life essentially depends on its being chosen and constructed by the person who lives it. Your life, like your faith, must be your own spontaneous production if it is to be worth anything at all. A third reason, which lacks a theological analogue, is *epistemological individualism*, which consists of two propositions: first, there may be a best life for each person, but there is no one best life for everyone. Second, as a matter of fact, each person is best placed to *find out* for herself what the best life for her is. (It is by endorsing epistemological individualism that a utilitarian becomes a liberal.) What these views share is the idea that the direct realization of final goods, or best lives, is disqualified as a political goal. Either we do not know enough about final goods to use them in political justifications, or they are by their nature best left in the hands of individuals. The state can only be justified in controlling the distribution of instrumental or primary goods, the things that everyone pursuing a good life has reason to want. According to this view, the reason why people consent, or may be supposed hypothetically to consent, to political institutions, is because (i) they have a better life in society and (ii) the principles regulating society are such as they would have chosen. Because we must allow for a variety of conceptions of the good, however, the only legitimate way for the state to provide a better life is to increase and fairly distribute the stock of primary goods. For reasons that will become clear, I shall call this view 'the New Liberalism'.

There is another, older way to be a liberal which is a little different. According to Locke and Kant, the business of the state is to preserve and protect rights and freedom, *not* to facilitate the pursuit of a good life. These philosophers believed that it is of the nature of rights and freedom that their preservation justifies the use of coercion. Locke believed that you have natural rights which you are entitled to enforce: first, and innately, a right to your own labour, and

end p.56

second, by extension, a qualified right to what you mix your labour with.[7]

[7] Locke, 1690: ch. 5, s. 27.

Kant believed that everyone has an innate right to freedom, and, as a necessary extension of that right, to the acquisition of certain other rights, without which one cannot exercise one's freedom, and so which may be enforced.[8]

[8] Kant, 1797: 52; Prussian Academy page numbers (standard in most editions), 246. (Citations hereinafter given as e.g. 52/246.)

Kant argued that rights may be coercively enforced this way: rights are a necessary extension of freedom, so anyone who tries to interfere with rights tries to interfere with freedom. A rights violator is a hindrance to freedom. But anything that hinders a hindrance to freedom is consistent with freedom. And anything that is consistent with freedom is legitimate. The coercive enforcement of rights, being consistent with freedom, is therefore legitimate.[9]

[9] Kant, 1797: 35-6/231

These views also ground state coercive authority in a social contract, but the effect of the contract is different. In the New Liberalism, the hypothetical social contract seems to give rise to the coercive authority of the state. The justification of the use of coercion lies in the fact that people may plausibly be supposed to consent to arrangements that they would have chosen and that give them a better life. In the Old Liberalism, coercive authority is thought to attach to rights and freedom by their very nature — all that the contract does is *transfer* this already existing coercive authority to the state.[10]

[10] See e.g. Locke, 1690: ch. 7, s. 87. It follows that the goal of political society is the preservation of rights and property (ch. 9, s. 123). Locke, however, sometimes seems to give the government more

extensive powers in acting 'for the good' of the citizens, so long as their property is not violated. But what licenses this, I believe, is that Locke believes that consent is actual (express or tacit), not hypothetical, and also that there is a real possibility of withdrawing it. Kant portrays government explicitly as arising from a hypothetical contract investing coercive authority in a state (1797: 76-7/312; 80-1/315-16). In this case, 'transfer' is not exactly the right word for the effect of the contract on coercive authority. Kant thinks that in order to be morally legitimate, the coercive enforcement of rights must be 'reciprocal' and therefore must be accomplished through the state (36-7/232; 64-5/255-6). In other words, the *only* legitimate way to enforce your rights is to join in a political state with the person whom you claim has attacked your rights and so submit yourself as well as him to coercive enforcement (71-2/307-8).

Hypothetical consent is only allowed to determine *who* exercises coercive authority; it does not give rise to it.

A consequence is that, according to the Old Liberalism, state coercion can only be exercised in the protection and preservation of freedom. As in the New Liberalism, final goods are not legitimate political objectives, but there is a difference. In the New Liberalism, final goods are disqualified as political objectives because of their unknowability or variability. In the Old Liberalism, final goods are not so much disqualified as never qualified in the first place. I think the point is important because many people suppose that liberalism must be founded on philosophical scepticism about whether we can discover what the best life is. The Old Liberalism is consistent with the most thoroughgoing certainty about the best life. You can think, with Aristotle in book X of the *Nicomachean Ethics*, that you can prove that a contemplative life is best and still think that the state

end p.57

has no business getting people to contemplate. Only the protection of freedom, and not the achievement of the good, is grounds for the use of coercion.

There are two objections to the Old Liberalism, both of which the New Liberalism is designed partly to overcome. Both objections complain that the Old Liberalism is too conservative. First, according to the versions presented by Locke and Kant, extensive property rights are supposed to have existed in the state of nature, and the business of government is to protect those rights. This is objectionable both because it depends on a questionable premiss — that property rights exist independently of and prior to the state — and because it seems to saddle the government with an obligation to carry on and fortify inherited inequalities. I do not think this is essential to the view in its Kantian form. The only thing essential is that a right to freedom exist prior to the state, and that this be taken to be a natural human right. We can allow, with the New Liberalism, that we should determine what further rights will count as necessary extensions or realizations of freedom within the bounds of the state, by determining what people can reasonably agree to. But, the objection continues — and this is the second complaint — the Old Liberalism amounts to what some of us will regard as a dreary form of libertarianism. If the only thing that the state can guarantee is freedom, and not a good life, there will be no grounds for guaranteeing things that seem clearly to be part of the good and not of freedom — food, medical care, an economic minimum, and so forth. This kind of theory makes it hard to be a welfare liberal.

There is also a way to overcome this objection. It is to insist both on the necessity of employing a rich positive conception of freedom, and on the idea that certain welfare conditions must be met in order to achieve what Rawls calls the 'worth of liberty', the real possibility of taking advantage of one's rights and opportunities. [11]

[11] Rawls, 1971: s. 32, 204.

Rawls himself declines to treat the conditions of the worth of liberty as part of, or essential to, being free. But in fact it does not matter much whether we talk of the

worth of liberty or its reality. We cannot effectively guarantee liberty without guaranteeing its worth. The general idea behind this view, then, is that unless certain basic welfare conditions are met and resources and opportunities provided, we cannot seriously claim that society is preserving and protecting everyone's freedom. The poor, the jobless, the medically neglected, the unhoused, and the uneducated are not free no matter what rights they have been guaranteed by the constitution. There are two reasons for this. The first is their impaired capacity for formulating and pursuing a conception of the good. The second is just as important. A person who lacks these basic goods is subject to intimidation by the rich and powerful, especially if others depend on her. As unskilled woman labourer who puts up with a lower income, poor conditions, or even sexual harassment on the job because her only alternative is to let her children starve is not free. To fail to satisfy people's basic needs and provide

end p.58

essential skills and opportunities is to leave people without recourse, and people without recourse are not free. [12]

[12] My way of portraying the situation may suggest that I think of Rawls as the exemplary New Liberal. On the whole this is correct, but I also think that there are elements in Rawls's later writings that suggest a shift in the direction of the welfarist version of Old Liberalism. A particular issue illustrates the point. In *A Theory of Justice*, Rawls treats liberties as items on the list of primary goods. In the general conception of justice, they may be traded off for other sorts of primary goods (s. 11, 62-3). Rawls then has to provide an explanation of the priority of liberty in the special conception of justice. That is, he has to explain why, in favourable conditions, the liberties come to acquire the special status of goods that citizens will not trade off for anything else (s. 82, 541-8). In his later writings, Rawls instead places more emphasis on the conception of the citizen as a moral person, one of whose highest-order interests is in the exercise of autonomy itself: not just achieving her conception of the good, but freely choosing and revising it. The citizen so understood views the state not only as a *locus* for pursuit of her determinate conception of the good, but also as a *locus* for the exercise of autonomy. See Rawls, 1980 and 1982*a*. This is a shift in the direction of the Old Liberalism, and among other things it makes the priority of liberty easier to explain. See esp. 1982*a*: 27 ff. This does not mean that the general conception of justice has disappeared from Rawls's account; rather, he has made it clearer why this conception is only acceptable when the worth of liberty cannot be established.

I think that there are some important philosophical advantages in a welfarist version of the Old Liberalism. For if we must justify state coercion by appeal to the notion of hypothetical consent, we need some way to limit what sort of thing is a candidate for such consent. We cannot suppose that people hypothetically consent to coercion in the name of just anything society decides is good. The older, freedom-based theory allows the state to coerce people only for the sake of something for which they may legitimately be coerced anyway — the achievement of freedom for every person. It explains the priority of liberty over other goods, and yet at the same time its emphasis on positive freedom and the worth of liberty explains why the guarantee of basic welfare and opportunities is essential. The result of employing these notions is a very large coincidence in practice between the New Liberalism and the Old. But the justification is still different. In the Old Liberalism, primary goods are justified as essential to the worth, or the reality, of freedom in a positive sense, rather than as general means to various conceptions of the good.

Now I return to the proposals of Sen and Cohen. First, the view of legitimate political objectives that I have sketched leads immediately to a strong although qualified endorsement of Sen's proposal that we should distribute with an eye to capabilities. For Sen argues that the idea of capabilities gives us a way of understanding the idea of positive freedom, and I think that this is correct: to make people capable of effectively realizing their goals and pursuing their well-being is to make them positively free. The qualification is this. In his paper Sen considers whether and to

what extent his view can be justified by the idea that human well-functioning, suitably defined, corresponds to some philosophical ideal of the final good: for instance that of Aristotle or Marx. [13]

[13] Sen, 'Capability and Well-Being', in this volume, 30-53.

On the view I have sketched, this correspondence, if it were one, would play no role in justifying the distribution of capabilities as a political objective.

What I have to say about Cohen's paper is more complicated. First, I want to



address Cohen's claim that 'midfare' is a legitimate political objective in its own right and his criticism of Sen's view on this point. [14]

[14] See Cohen, 'Equality of What?', 18 ff. 'Midfare' is Cohen's term for the effect of goods on the person.

Cohen accuses Sen of 'athleticism', noticing places where Sen seems to say that the point of providing someone with, say, food is that they may *feed themselves* rather than just that they may *get fed*, whoever is the agent in the case. Cohen argues that it is important that people get fed, not just that they feed themselves, and supposes that Sen misses this point because he is focused too much on freedom and activity. But a person's getting fed may be justified politically by the contribution to his freedom that being well nourished makes. We do not have to choose between giving nourishment a political weight which is unrelated to freedom and the 'athleticism' of which Cohen accuses Sen. Feeding yourself is not the only free activity for which the provision of an adequate diet is essential.

Cohen's own proposal gives a different role to freedom from those of Sen and Rawls. Cohen believes that it is unjust that people should be disadvantaged in ways they do not freely choose. He thinks that this idea is what gives intuitive force to the 'expensive tastes' criticism of welfarism propounded by Rawls and others. Cohen accuses Rawls of switching back and forth between a deterministic and a libertarian view of human nature. When he attacks the political use of the notion of desert, Rawls uses deterministic arguments, saying that if someone is more diligent or ambitious than others, these virtues are most likely the product of a favourable upbringing. On the other hand, Rawls claims that people should be held responsible for their tastes, and should be regarded as the autonomous authors of their own conceptions of the good. What Rawls ought to say, according to Cohen, is just that it is difficult to *tell* to what extent people are responsible for either their efforts or their preferences. [15]

[15] Ibid.

But the fact that it is difficult to tell is no reason not to reward people to the extent that they *are* responsible for their efforts, and, more importantly to Cohen, no reason not to compensate them for disadvantageous preferences that they cannot help. Cohen acknowledges that his proposal requires assessment of the extent to which disadvantages are voluntary in individual cases, and that such judgements are not easy to make. But he says that 'there is no antecedent reason whatsoever for supposing that judgements of justice, at a fine grain degree of resolution, are easy'. [16]

[16] Ibid.

There are several things wrong here. First, Cohen assumes that there is a single answer, in any given case, to the metaphysical question of the extent to which one's choice or effort was free. But this ignores a certain complication. Freedom of the will may *itself* be the result of a favourable upbringing and social conditions. We may believe that a human being is free, if ever, when she not only has a range of options

but an education that enables her to recognize those options as such and the self-respect that makes her choice among them a real one. Ignorance, lack of imagination, and lack of self-respect are not just external constraints

end p.60

on the range of your options: they can cripple the power of choice itself. The possession of freedom of the will may itself be lucky.

If this is right, there is something better for the state to do than the acts of individual compensation Cohen proposes. We can set society up so that people's choices will be autonomous and free. And here it becomes important to insist, along with Rawls, that the subject of justice is the basic structure of society. Cohen envisions a world in which government officials make judgements about the extent to which people are responsible for their preferences. But part of the reason why Rawls thinks that we should focus on the basic structure is that this way we can avoid having to make moralizing judgements about individual cases. By setting up the basic structure of society as what Rawls calls a system of 'pure procedural justice', we avoid having to ask intractable questions about whether particular individuals deserve the positions in which they have landed. [17]

[17] Rawls, 1971: ss 14, 48.

In a similar way, we should avoid having to make particular metaphysical judgements about whether persons have formed their preferences autonomously or not. Society should be set up so that we can assume, as far as possible, that they have done so. This is not just for the pragmatic reason that metaphysical judgements about free will are hard to make in particular cases. In the conceptions of Rawls and Sen, freedom is regarded as something society should bring about, not just as the occasion for judgements about what people deserve. And this leads me to a final point, which is that there is also a moral reason for working through the basic structure, and so avoiding particular judgements of the sort Cohen has in mind. Judgements about whether others have freely chosen their conceptions of the good are not only ones we cannot very easily make, they are ones we ought not to make. Such judgements are disrespectful. If one of our goals is to make it possible for the members of society to have decent *moral* relations with one another, this is an additional reason for making freedom appear instead as a *consequence* of justice; something that results, that is, from a just basic structure of society.

**Bibliography**

Kant, Immanuel (1797). *The Metaphysical Elements of Justice*, trans. John Ladd (1965). Indianapolis: Bobbs-Merrill Library of Liberal Arts.

Locke, John (1690). *Second Treatise of Government*. Indianapolis: Hackett Publishing.

Mill, John Stuart (1859). *On Liberty*. Indianapolis: Hackett Publishing.

Rawls, John (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.

—— (1980). 'Kantian Constructivism in Moral Theory: The Dewey Lectures 1980', *Journal of Philosophy*, 77 (Sept. 1980).

—— (1982*a*). 'The Basic Liberties and Their Priority', in *Tanner Lectures on Human Values*, iii. Salt Lake City: University of Utah Press.

—— (1982*b*). 'Social Unity and Primary Goods', in *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams. Cambridge: Cambridge University Press.

end p.61

# Quality of Life Measures in Health Care and Medical Ethics

show chapter abstract and keywords
hide chapter abstract and keywords

**Dan Brock**

# 1 Introduction

There has been considerable philosophical work during the last two decades, especially in the United States but not limited to there, in a relatively new field called medical ethics. My aim in this paper is to explore what illumination that body of work might offer to our understanding of the quality of life. If one looks only to the medical ethics literature explicitly addressing the notion of the quality of life, there are few sustained analyses of it and of its role in various medical and health care contexts. Consequently, it is necessary to look more broadly to issues and areas of research that often do not explicitly address the quality of life, but that nevertheless have an important bearing on it. I believe there are two main areas of work in medical ethics that fit this criterion. The first is work on ethical frameworks for medical treatment decision-making in a clinical context, including accounts of informed consent and life-sustaining treatment decisions. The second is the development of valuational measures of outcomes of health care treatments and programmes; these outcome measures are designed to guide health policy and so must be able to be applied to substantial numbers of people, including across or even between whole societies. The two main parts of this paper will address these two main bodies of work. Before doing so, however, several preliminary issues need to be briefly addressed.

I have mentioned that the literature that I will be summarizing and drawing on often does not explicitly address the concept of the 'quality of life', but instead uses other notions that are either closely related or roughly equivalent in the context. Sometimes a notion of 'health' is employed, particularly in its broader interpretations, as exemplified in the World Health Organization definition of 'health' as a state of complete physical, mental, and social well-being.[1]

[1] Breslow, 1972: 347-55.

The notion of patient 'well-being', independent of its use within a definition of health, is also often employed for evaluation of outcomes in health care. Another conceptual framework commonly employed for evaluating health care outcomes is the assessment of the benefits and burdens of that care for the patient (and sometimes for others as well). Still another common conceptual framework often employed looks to the effects of health care on patients' interests, with a best interests standard particularly prominent for patients whose preferences cannot be determined. These and other conceptual schemes are not fully interchangeable in health care, much less in broader contexts. Nevertheless, they all have in common their use in evaluating health care outcomes for patients and their

end p.95

employment as at least part of a comprehensive account of a good life for persons. I shall freely draw here on each of these conceptual frameworks, and others, though indicating where differences between them become important.

The 'quality of life' can be given a number of more or less broad interpretations, depending on the scope of the evaluative factors concerning a person's life that it is taken to include. Medicine and health care often affect a person's life in only some limited areas or respects. Nevertheless, my concern will be with the broadest conception of, in Derek Parfit's words, 'what makes a life go best', and I shall try to show that medicine and health care may affect and illuminate more aspects of that question than might at first be thought.[2]

[2] Parfit, 1984, cf. esp. app. I.

No concept is entirely apt or widely accepted in either philosophical or common usage for this broad role, but I shall use the concept of a 'good life' to refer to the quality of life of persons in its broadest interpretation.

It is common in much philosophical work on theories of the good for persons or of a good life to distinguish three broad kinds of theory. While this classification misses some distinctions important for my purposes here, it provides a natural starting point. These three alternative theories I will call the hedonist, preference satisfaction, and ideal theories of a good life. [3]

[3] See T.M. Scanlon's discussion in this volume of these alternative theories. What I call preference satisfaction and ideal theories he calls desire and substantive good theories.

Much of the philosophical work on these theories has been in the service of developing an account of 'utility', broadly construed for employment in consequentialist moral theories. [4]

[4] Some time ago I discussed these as alternative interpretations of utility (Brock, 1973). The most subtle and detailed recent discussion of these alternative theories is Griffin, 1986, chs. 1-4.

What is common to hedonist theories, as I will understand them here, is that they take the ultimate good for persons to be the undergoing of certain kinds of conscious experience. The particular kinds of conscious experience are variously characterized as pleasure, happiness, or the satisfaction or enjoyment that typically accompanies the successful pursuit of our desires. Particular states of the person that do not make reference to conscious experience, such as having diseased or healthy lungs, and particular activities, such as studying philosophy or playing tennis, are part of a good life on this view only to the extent that they produce the valuable conscious experience.

Preference satisfaction theories take a good life to consist in the satisfaction of people's desires or preferences. I here understand desires or preferences as taking states of affairs as their objects: for example, my desire to be in Boston on Tuesday is satisfied just when the state of affairs of my being in Boston on Tuesday obtains. This is to be distinguished from any feelings of satisfaction, understood as a conscious experience of mine, that I may experience if I am in Boston on Tuesday. The difference is clearest in cases in which my desire is satisfied, but I either do not or could not know that it is and so receive no satisfaction from getting what I desire: for example, my desire that my children should have long and fulfilling lives, a state of affairs that will only fully obtain

end p.96

after my death. For preference satisfaction theories of a good life to be at all plausible, they must allow for some correcting or 'laundering' of a person's actual preferences. [5]

[5] Virtually all discussions of desire or preference satisfaction theories of the good contain some provision for correcting preferences. One of the better treatments, with extensive references to the literature, is Goodin, 1986: 75-101.

The most obvious example is the need to correct for misinformed preferences: for example, my desire to eat the sandwich before me not knowing that its ingredients are bad and will make me ill. Other corrections of preferences have also been supported by proponents of the preference satisfaction theory that are compatible with its underlying idea that ultimately what is good for persons is that they should get what they most want or prefer.

The third kind of theory holds that at least part of a good life consists neither of any conscious experience of a broadly hedonist sort nor of the satisfaction of the person's corrected preferences or desires, but instead consists of the realization of specific, explicitly normative ideals. [6]

[6] What I call ideal theories are what Parfit (1984) calls 'objective list' theories. I prefer the label 'ideal theories', because what is usually distinctive about this kind of theory is its proposal of specific, normative ideals of the person.

For example, many have held that one component of a good life consists in being a self-determined or autonomous agent, and that this is part of a good life for a person even if he or she is no happier as a result and has no desire to be autonomous. Ideal theories will differ both in the specific ideals the theories endorse and in the place they give to happiness and preference satisfaction in their full account of the good for persons. There is a strong tendency in much of the philosophical literature to seek a simple, comprehensive theory, such as the hedonist or preference satisfaction theories: proponents of ideal theories commonly acknowledge a plurality of component ideals that place constraints on and/or supplement the extent to which happiness and/or preference satisfaction serves a person's good. The account I will develop of quality of life judgements in health care strongly suggests that it is a mistake to let the attractions of a simple, unified theory of a good life force a choice between the hedonist and preference satisfaction theories. Instead, these quality-of-life judgements suggest the importance of giving independent place to the considerations singled out by each of the three main alternative theories, as ideal theories do, in any adequate overall account of the quality of life or of a good life for persons. The quality of life judgements made in medicine and health care also help some to fill in the content of a theory of a good life.

A major issue concerning ethical judgements generally, and judgements concerning a good life in particular, is the sense and extent to which such judgements are objective or subjective. A number of different senses have been given to the notions of 'objectivity' and 'subjectivity' in these contexts, and other essays in this volume take up some of these general theoretical issues and will develop some of these alternative senses.[7]

[7] See the papers by H. Putnam, R.A. Putnam, and M. Walzer.

I will not attempt an extended analysis of



this general theoretical issue here. Nevertheless, one sense in which what constitutes a good life for a particular person is believed to be subjective or objective mirrors the distinction between hedonist and preference satisfaction theories on the one hand, and ideal theories on the other. Hedonist and preference theories are both subjective in the sense that both hold that what is good for a particular person depends on what in fact makes that person happy or what that person in fact (with appropriate corrections) desires. (This is compatible, of course, with acknowledging that what will make a particular person happy or satisfy his or her preferences is an 'objective matter of fact', even if often an extremely difficult one to determine.)

Ideal theories are objective, or at least contain objective components, in the sense that they hold a good life for a person is, at least in part, objectively determined by the correct or justified ideals of the good life, and does not in those respects depend either on what makes that person happy or on what that person's (even corrected) preferences happen to be. The question of whether accounts of a good life are objective or subjective is, then, an explicitly normative issue about what is the correct or most justified substantive theory of a good life. This sense of the objective—subjective dispute has been a central concern in the debates in medical ethics and health care about the quality of life. Interestingly, I believe that medicine and health care provide some of the most persuasive instances for both the objective *and* the subjective components of a good life, and so point the way towards a theory that incorporates hedonist, preference, and ideal components.

Haavi Morreim has distinguished a different sense in which quality of life judgements in medicine are either objective or subjective.[8]

[8] Morreim, 1986: 45-69.

In her account, objective quality of life judgements are made on the basis of intersubjectively observable, material facts about a person (facts concerning his or her body, mind, functional capabilities, and environment), together with a socially shared evaluation of those facts, specifically of how those facts determine the person's quality of life. Subjective quality of life judgements also appeal to material facts about a person and his or her condition (though these may also include facts about the person's private psychological states), together with *that person's* value judgements about how those facts affect his or her quality of life. According to this account, the essential issue that determines whether a quality of life judgement is objective or subjective is whether the evaluative judgements concerning a particular individual's quality of life are and must be shared by some wider group or are, instead, only the individual's own. Since there are many possible wider social groups, one respect in which one could make sense of degrees of this kind of objectivity is in terms of the size, breadth, or nature of the wider social group; important variants include an individual's community or larger society, and a maximally wide group might be all humans or rational agents. It should be obvious that my and Morreim's senses of the objective—subjective

end p.98

distinction are independent: the individual whose quality of life is in question might hold any of the three substantive theories of the good for persons distinguished above, as might any wider social group.

A full conception of a good life for a person that does not reduce to a single property like happiness or preference satisfaction must assign a weight to the various components that contribute to that life's being good, though there may not be full comparability between different components and so in turn only partial comparability between different possible life courses for a person. Amartya Sen has suggested in several places the formal device of understanding these different components as independent vectors, each of which contributes to an overall assessment of the degree to which a person has a good life.[9]

[9] Sen's main discussion of the 'vector view' applied to the notion of utility is in Sen, 1980. I am much indebted to Sen's subtle discussions in a number of places of distinctions of importance to conceptions of the quality of life and of a good life. Besides his essay in this volume, see esp. Sen, 1985*a*, 1985*b*, 1987.

There are several benefits to an analysis of what constitutes a good life into a number of independent vectors. First, it allows us to accept part of what proponents of each of the three traditional theories of a good life have wanted to insist on, namely the theoretical independence of those components. The three components of happiness, preference satisfaction, and ideals of a good life can each be represented by their individual vectors, or subdivided further into distinct vectors within each component, having independent weight within an overall account of a good life. Second, the vector approach quite naturally yields the possibility of two senses of partial comparability of the quality of different lives. For a single individual, alternative possible lives may be only partially commensurable if one alternative life provides a greater value on one vector, but a lesser value on another vector, than another possible life. But for two different persons it is important that at least partial comparability between their lives may be possible, contrary to the dogma about the impossibility of interpersonal comparisons of utility, by comparing common vectors or by comparing different changes in common vectors making up a good life for each.

Medicine and health care provide strong grounds for insisting on these independent vectors and, perhaps more important, also suggest a content and structure to the ideals along the lines proposed by Sen in his work on agency and capabilities, which drew on setting largely outside of health care.

We also need to distinguish between the relative importance of a particular feature or condition, say as represented by a specific vector, in its contribution to a person having a good life, compared with what I shall call its broader moral importance. A simple example will suffice. One condition that may plausibly contribute to a person's quality of life or good life is his or her physical mobility. It may be possible to specify roughly a normal level of physical mobility for persons of a similar age at a particular historical stage and in a particular society, and then to specify roughly levels of mobility say 25 per cent below and 25 per cent above the norm, such that the effect on a person's quality of life in moving

end p.99

from 25 per cent below the norm up to the norm is quantitatively roughly the same as moving from the norm to 25 per cent above it. While the degree or importance of the two changes in a person's quality of life or good life may be roughly the same, it can none the less be consistently held that these two comparable effects on the person's quality of life have different *moral* importance or priority. It might be held, for example, that on grounds of equality of opportunity bringing a person's mobility from 25 per cent below the norm up to the norm has greater moral priority than increasing his mobility from the norm to 25 per cent above it. The general point is that aspects of a person's quality of life may play a role not only in judgements about his quality of life or about how good a life he has, but also in other distinct moral and political judgements, or in the application of independent moral principles such as a principle of equal opportunity. This is, of course, a thoroughly familiar point in moral and political philosophy generally, and concerning consequentialist moralities in particular, against which it is often objected that they ignore the moral importance of whether the good is fairly or justly distributed. In the present context its importance is in reminding us to distinguish judgements concerning the improvement or reduction of people's quality of life from other independent moral evaluations of those same changes so as not to confuse needlessly the nature of quality of life judgements in health care.

## 2 Ethical Frameworks for Health Care Treatment Decision-Making

The first broad area of work within medical ethics bearing on the concept of the quality of life concerns the aims of medicine and the account of medical treatment decision-making appropriate to those aims. It may be helpful to begin with a natural objection to thinking that these issues in medical ethics will illuminate any broad notion of the good life. On the contrary, as Leon Kass has argued, medicine's proper end is the much narrower one of health, or the healthy human being, and other goals such as happiness and gratifying patient desires are false goals for medicine.[10]

[10] Kass, 1985.

Kass understands health to be a naturalistically defined property of individual biological organisms, organisms, which must be understood as organic wholes, and whose parts

have specific functions that define their nature as parts: the bone marrow for making red blood cells; the lungs for exchange of oxygen and carbon dioxide; the heart for pumping the blood. Even at a biochemical level, every molecule can be characterized

in terms of its function. The parts, both macroscopic and microscopic, contribute to the maintenance and functioning of the other parts, and make possible the maintenance and functioning of the whole. [11]

[11] Ibid., 171.

What constitutes well-functioning varies with the particular biological species in question, but Kass is at pains to argue that 'health is a natural standard or

end p.100

norm—not a moral norm, not a "value" as opposed to a "fact", not an obligation—a state of being that reveals itself in activity as a standard of bodily excellence or fitness'. [12]

[12] Ibid., 173. For a more philosophically sophisticated analysis of the concept of health that also construes it in functional terms as a natural, biological norm not involving value judgements, see Boorse, 1975, 1977. One of the most useful collections of papers on concepts of health is Caplan, Engelhardt, and McCartney, 1981.

Kass's work constitutes one of the more ambitious attempts to justify two common-sense beliefs about the 'objectivity' of medicine: that the aim of medicine is and should be the patient's health, and that health is a biologically determined, objective matter of fact. If so, then physicians, with their impressive body of scientific knowledge concerning human biological functioning and the impact of therapeutic interventions on diseases and their natural courses, would seem to be the proper judges of whether we are healthy and, if we are not, what therapeutic interventions will be likely to make us more so. This hardly begins to do justice to the subtlety of Kass's view—though it is a view that I believe to be fundamentally mistaken—but it does bring out why one might think medicine, properly aimed only at human health defined in terms of biological functioning, has little to teach us regarding broader social issues about the quality of life. I believe it is fair to say that the main body of work in medical ethics within the last two decades has rejected Kass's view that the sole proper aim of medicine is health, defined in naturalistic, biological terms, and the ethical framework for medical treatment decision-making that it would seem to imply. We need to see how the alternative, broader view of the aims of medicine that should guide medical treatment decision-making bears on an understanding of the quality of life.

It has become a commonplace, at least in the developed countries, that medicine has achieved the capacity commonly to offer to patients suffering from particular diseases a number of alternative treatments, and to extend patients' lives in circumstances in which the benefit to the patient of doing so is increasingly problematic. In the United States this has led to patients pursuing various means of gaining control over decisions about their treatment. In the case of competent patients, a broad consensus has developed that such patients have the right to decide about their care in a process of shared decision-making with their physicians and to reject any proferred treatment. In the case of incompetent patients, an analogous consensus has been developing that an incompetent patient's surrogate, seeking to decide as the patient would have decided in the circumstances if competent, is likewise entitled to decide about the patient's care with the patient's physician and to reject any care the patient would not have wanted—though the consensus concerning incompetent patients is less broad and more ringed with qualifications. Each consensus is reflected in a large medical ethics literature, a growing body of legal decisions, legal mechanisms such as Living Wills and Durable Powers of Attorney for Health Care, whose purpose is to ensure patients' control over their care, pronouncements and

end p.101

studies of authoritative bodies and commissions, policies of health care institutions, and the practice of health care professionals.[13]

[13] I make no attempt here to provide any more than a few representative references to this very large literature. Probably the single best source for the medical ethics literature in this area is the *Hastings Center Report*. In the medical literature, see Wanzer, 1984, and Ruark, 1988. For a good review of most of the principal legal decisions in the United States concerning life-sustaining treatment, see Annas and Glantz 1986. The most influential treatment of these issues by a governmental body in the United States is the report of the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1983). See also the recent report by the Hastings Center (1987). For discussions of Living Wills and Durable Powers of Attorney see Steinbrook and Lo, 1984, and Schneiderman and Arras, 1985. An application to clinical practice of the consensus that patients should have rights to decide about their care is Jonsen, Siegler, and Winslade, 1982.

The common view is now that health care decision-making should be a process of shared decision-making between patient (or the patient's surrogate in the case of an incompetent patient) and physician.[14]

[14] An influential statement of the shared decision-making view is another report of the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1982). A sensitive discussion of the difficulties of achieving shared decision-making in clinical practice is Katz, 1984.

Each is seen as indispensable to sound decision-making. The physician brings his or her training, knowledge, and expertise to bear for the diagnosis of the patient's condition, the estimation of the patient's prognosis with different alternative treatments, including the alternative of no treatment, and a recommendation regarding treatment. The patient brings the knowledge of his or her aims, ends, and values that are likely to be affected by different courses of treatment, and this enables a comparative evaluation of different possible outcomes to be made. As alternative treatments have multiplied and become possible in circumstances promising increasingly marginal or questionable benefits, both physicians and patients are called upon to make increasingly difficult judgements about the effects of treatment on patients' quality of life. It is worth noting that proponents of shared decision-making need not reject the functional account of health as a biological norm defended by Kass and others. What they can reject is the claim that the only proper goal of medicine is health. Instead, medicine's goal should be to provide treatment that best enables patients to pursue successfully their overall aims and ends, or life plans. It is the relative value of health, and of different aspects of health, as compared with other ends, that varies for different persons and circumstances.

Most patients' decisions about life-sustaining treatment will be based on their judgement of the benefits and burdens of the proposed treatment and the life it sustains, though in some instances patients may give significant weight to other factors such as religious obligations, the emotional burdens and financial costs for their families, and so forth. Except for patients who hold a form of vitalism according to which human life should or must be sustained at all costs and whatever its quality, these decisions by competent patients must inevitably involve an assessment of their expected quality of life if life-sustaining treatment



is employed, though, as I shall note shortly, of only a very restricted sort.

Some have rejected the acceptability of quality of life judgements in the case of incompetent patients unable to decide for themselves, for whom others must therefore make treatment decisions.[15]

[15] E.g. Ramsey, 1978: 206-7.

One version of the objection is that no one should decide for another whether that other's quality of life is such that it is not worth continuing it. More specifically, the

objection is that it is unacceptable to judge that the quality of another person's life is so poor that it is not worth the cost and effort to others to sustain that person's life. This objection, however, is not to making quality of life judgements in this context generally, but only to concluding that a person's life is not worth sustaining because its poor quality makes it not of value, but instead a burden, to *others*. The sound point that this objection confusingly makes is that quality of life judgements concerning a particular person should address how the conditions of a person's life affect its quality or value to *that person*, and not its value to others. Moreover, persons might judge their quality of life to be low and nevertheless value their lives as precious. In economic and policy analysis one version of the so-called human capital method of valuing human life, which values a person's life at a given point in time by his or her expected future earnings minus personal consumption, in effect values a person's life in terms of its economic value to others. [16]

[16] I have discussed some of the ethical implications of different measures for valuing lives found in the economic and policy literatures (Brock, 1986).

But there is no reason to reject the soundness of any evaluation by one person of another's quality of life simply because some might draw a further unjustified conclusion that if its quality is sufficiently low to make it on balance a burden to others, it ought not to be sustained.

The quality of life judgement appropriate to life-sustaining treatment decisions, whether made by a competent patient or an incompetent patient's surrogate, should thus assess how the conditions of the patient's life affect the value of that life to that patient. Nevertheless, even properly focused in this way, the role of quality of life judgements in decisions about whether to withhold or withdraw life-sustaining treatment is extremely limited. This quality of life judgement focuses only on which side of a *single threshold* a person's quality of life lies. The threshold question is: 'Is the quality of the patient's life so poor that for that person continued life is worse than no further life at all?' Or, in the language of benefits and burdens commonly employed in this context: 'Is the patient's quality of life so poor that the use of life-sustaining treatment is unduly burdensome, that is, such that the burdens to the patient of the treatment and/or the life that it sustains are sufficiently great and the benefits to the patient of the life that is sustained sufficiently limited, to make continued life on balance no longer a benefit or good to the patient?' [17]

[17] It has been argued that this is the proper understanding of the distinction between 'ordinary' and 'extraordinary' treatment. That is, extraordinary treatment is treatment that for the patient in question and in the circumstances that obtain is unduly burdensome. Cf. President's Commission, 1983: 82-9.

The only discrimination in quality



of life required here is whether the quality of the life is on balance sufficiently poor to make it worse than non-existence to the person whose life it is.

Some have objected that this judgement is incoherent since, though it is possible to compare the quality of two lives, or of a single person's life under different conditions, it is not possible to make the quality of life comparison needed here because one of the alternatives to be compared is non-existence. If a person no longer exists, there is no life that could possibly have any quality so as to enter into a comparison with the quality of the life sustained by treatment. This objection does correctly point out that the judgement in question cannot involve a comparison of the quality of two alternative periods of life, though it could compare two possible lives, one that ends at that time and another that continues longer. However, it does not follow that there is no sense to the question of whether the best life possible for a person with some form of life-sustaining treatment is of sufficiently poor quality, or sufficiently

burdensome, to be worse for that person than no further life at all. Perhaps the most plausible example is the case of a patient suffering from an advanced stage of invariably fatal cancer, who is virtually certain to die within a matter of days whatever is done, and whose life will be filled in those remaining days with great and unrelievable pain and suffering. (With the appropriate use of presently available measures of pain relief, it is in fact only very rarely the case that great pain and suffering in such cases cannot be substantially relieved.) The burden of those remaining days may then be found by the patient to be virtually unbearable, while the life sustained provides nothing of value or benefit to the patient. This judgement addresses the quality of the life sustained and appears to be a sensible judgement. It is just the judgement that patients or their surrogates commonly understand themselves to be making when they decide whether to employ or continue life-sustaining treatment.

Alternatively, the objection to someone ever making such a judgement for another may not be based on any putative incoherence of such judgements, but may express instead a concern about the difficulty of ever reliably deciding how *another* would in such circumstances decide, due perhaps to the diversity and unpredictability of people's actual decisions for themselves. Moreover, if the difficulty of reliably making such judgements for others is in fact this great, then we might well have a related practical concern that the interests of others, which may be in conflict with those of the patient, may consciously or unconsciously infect judgements about what is best for the patient.

Despite these difficulties, there have been attempts to formulate some general substantive standards to determine when an incompetent patient's quality of life is so poor that withholding or withdrawing further life-sustaining treatment is justified. Nicholas Rango, for example, has proposed standards for nursing home patients with dementia. [18]

[18] Rango, 1985.

He emphasizes the importance of being clear about the purposes for which care is provided and distinguishes three forms

end p.104

of care: (1) palliative care aimed at relieving physical pain and psychological distress; (2) rehabilitative care aimed at identifying and treating 'excess disabilities, the gap between actual level of physical, psychological or social functioning and potential functioning capacity', [19]

[19] Ibid., 836, quoting E.M. Brody.

and (3) medical care aimed at reducing the risk of mortality or morbidity. He emphasizes the importance of therapeutic caution because a seriously demented patient will not be able to understand the purposes of painful or invasive interventions, and so presumably cannot choose to undergo and bear burdensome treatment for the sake of promised benefits. [20]

[20] The importance of this factor was stressed in a widely publicized legal decision concerning the use of painful chemotherapy for a man suffering from cancer who had been severely retarded from birth (*Superintendent of Belchertown State School* v. *Saikewicz*, 1977).

Rango propose two conditions, either of which is sufficient to justify forgoing further treatment of a chronic medical condition or a superimposed acute illness: (*a*) when the patient is burdened by great suffering despite palliative and rehabilitative efforts; (*b*) when the dementia progresses 'to a stuporous state of consciousness in which the person lives with a negligible awareness of self, other, and the world'. [21]

[21] Rango, 1985: 838.

Even within the relatively narrow focus of life-sustaining treatment decisions for demented patients, Rango's proposal can be seen to include three different kinds of components of quality of life assessments. The first, covered by treatment aim (1) and patient conditions (*a*) and (*b*), concerns the quality of the patient's conscious experience. The second, covered especially by treatment aim (2) and patient condition (*b*), concerns the patient's broad functional capacities. The third, covered especially by patient condition (*b*) and by the patient's ability to understand the purpose of treatment and in turn to choose to undergo it, concerns the centrality to quality of life of the capacity to exercise choice in forming and pursuing an integrated and coherent life plan. I shall argue below that each of these three kinds of condition is an essential component of an adequate account of the quality of life.

At the other end of life, the debate in the United States about treatment for critically ill new-born infants has also focused on the role of quality of life considerations in determining when life-sustaining treatment is a benefit for the infant. One influential attempt to bring quality of life considerations into these decisions is the proposal of the moral theologian, Richard McCormick, that a new-born infant's life is a value that must be preserved only if the infant has the potential for a 'meaningful life'. [22]

[22] McCormack, 1974.

A meaningful life is one that contains some potential for human relationship: anencephalic new-borns, for example, wholly lack this potential, while those with Down's syndrome or spina bifida (to cite two of the most discussed kinds of case) normally do not. Nancy Rhoden has developed a more detailed proposal along similar lines regarding life-sustaining treatment for new-born infants: aggressive treatment is not mandatory if an infant: (1) is in the process of dying; (2) will

end p.105

never be conscious; (3) will suffer unremitting pain; (4) can only live with major, highly restrictive technology which is not intended to be temporary (e.g., artificial ventilation); (5) cannot live past infancy (i.e., a few years); or (6) lacks potential for human interaction as a result of profound retardation. [23]

[23] Rhoden, 1985. Another sensitive discussion of the need for quality of life judgements in treatment decisions for imperilled new-born babies is Arras, 1984.

Rhoden's proposal is typical of those accepting the use of quality of life considerations in decisions about life-sustaining treatment for new-born infants in its focus on the infant's at least minimal capacity for positive conscious experience (conditions 2 and 3): the infant's capacities for physical, mental, and social functioning (conditions 2, 3, 4, and 6), and the infant's capacities to live far enough into childhood to begin a life that can be viewed and experienced 'from the inside' by the child as a life lived in a biographical, not biological, sense (conditions 1 and 4; I will say more about this special feature of infant 'life years' below in discussing the relevance of mortality data to good lives). However, the very limited, single-threshold character of the quality of life assessments required in decisions as to whether to forgo or to employ life-sustaining treatment, whether for adults or new-born infants, takes us only a little way in understanding quality of life assessments in health care treatment decision-making.

It is necessary, consequently, to broaden the focus from life-sustaining treatment decisions to medical treatment decisions generally. Here, as noted earlier, there is a widespread consensus that competent patients are entitled, in a process of shared decision-making with their physicians, to decide about their treatment and to refuse any proferred or recommended treatment. In the United States, the doctrine of

informed consent, both in medical ethics and in the law, requires that treatment should not be given to a competent patient without that patient's informed and voluntary consent.[24]

[24] The most comprehensive treatment of the informed consent doctrine is Faden and Bauchamp, 1986. See also President's Commission, 1982. The exceptions to the legal requirement of informed consent are discussed in Meisel, 1979.

What does this doctrine, which lodges decision-making authority with the patient, imply about the nature of judgements concerning the patient's quality of life? An argument that it presupposes the normatively subjective, preference satisfaction account of a good life might, in rough outline, go as follows.

Each of the requirements of the informed consent doctrine can be understood as designed to provide reasonable assurance that the patient has chosen the treatment alternative most in accord with his or her own settled preferences and values. If the patient's decision is not informed—specifically, if the patient is not provided in understandable form with information regarding his or her diagnosis, the prognosis when different treatment alternatives (including the alternative of no treatment) are pursued, including the expected risks and benefits, with their attendant probabilities, of treatment alternatives—then the patient will lack the information needed to select the alternative most in accord with his or her settled preferences. If the patient's decision is not voluntary,



but is instead forced, coerced, or manipulated by another, then it is likely not to be in accordance with the patient's settled preferences, but instead will forward another's interests, or another's view of what is best for the patient. If the patient is not competent to make the choice in question, then he or she will lack the ability to use the information provided to deliberate about the alternatives and to select the one most in accord with his or her settled preferences. When all three requirements are satisfied—the decision is informed and voluntary, and the patient competent—others can be reasonably assured that the patient's choice fits the patient's own conception of a good life, as reflected in his or her settled preferences.[25]

[25] When patients' settled preferences are not in accord with their values, then their informed and voluntary choices may not reflect their conception of their good. One of the clearest examples is the patient who is addicted to morphine, hates his addiction and tries unsuccessfully to resist it, but in the end is overpowered by his desire for the morphine and takes it. This in essence is Frankfurt's example in his classic paper (Frankfurt, 1971). Frankfurt's analysis is in terms of first- and second-order desires, but can also be put in terms of the desires a person in fact has as opposed to the desires the person values and wants to have. When these are in conflict his informed and voluntary choice may not reflect the values that define his own conception of his good. There is a sense in which his choice in these conditions is involuntary, so it would be possible to extend the informed consent doctrine's requirement of voluntariness to include this sense. Alternatively, it might be possible to interpret the requirement of competence in a way that makes the morphine addict incompetent to decide whether to continue using morphine.

Viewed in this way, the informed consent doctrine may appear to be grounded in a preference satisfaction account of the good or the quality of life, and so not to require any more complex vector account of the sort suggested earlier. Even in this very crude form, however, this argument can be seen to be unsound if one asks what values the informed consent doctrine and the account of shared decision-making in medicine are usually thought to promote, and what values support their acceptance. The most natural and obvious first answer has already implicitly been given: the informed consent doctrine in health-care treatment decision-making is designed, when its three requirements are satisfied, to serve and promote the patient's well-being, as defined by the patient's settled, uncoerced, and informed preferences. If

this were the only value at stake, or at least clearly the dominant value, then it would be plausible to argue that the informed consent doctrine rests, at least implicitly, on a preference satisfaction account of the good life. However, it is not the only value at stake. Usually regarded as at least of roughly commensurate importance is respecting the patient's self-determination or autonomy.[26]

[26] This account of the principal values underlying the informed consent doctrine as patient well-being and self-determination is common to many analyses of that doctrine; cf. President's Commission, 1982. I have employed it in Brock, 1987.

The interest in self-determination I understand to be the interest of persons, broadly stated, in forming, revising over time, and pursuing in their choices and actions their own conception of a good life; more narrowly stated for my specific purposes here, it is people's interest in making significant decisions affecting their lives, such as decisions about their medical care, for themselves and according to their own values. Sometimes this is formulated as the right to self-determination.

end p.107

Whether interest or right, however, the greater the moral weight accorded to individual self-determination as one of the values underlying and supporting the informed consent doctrine, the weaker the basis for inferring that the doctrine presupposes the normatively subjective, preference satisfaction account of the good life. This is because the greater the moral weight accorded to individual self-determination, the more self-determination can explain the requirement of informed consent, even assuming the patient chooses in a manner sharply contrary to his or her own well-being. Moreover, there is substantial reason to suppose that the doctrine does in fact largely rest on respect for patients' self-determination. In the celebrated 1914 legal case of *Schloendorff* v. *Society of New York Hospitals*, usually cited as the first important enunciation of the legal requirement of consent for medical care, Justice Cardozo held that: 'Every human being of adult years and sound mind has a right to determine what shall be done to his own body; and a surgeon who performs an operation without his patient's consent commits an assault, for which he is liable in damages.'[27]

[27] *Schloendorf* v. *Society of New York Hospital*, 1914.

I shall make no attempt here to trace the development of the legal doctrine of informed consent since *Schloendorff*, but it is probably fair to say that no other subsequent case has been as influential or as often cited in that development. And the later cases, in one form or another, repeatedly appeal to a right to self-determination to support that developing doctrine.[28]

[28] Cf. Faden and Beauchamp, 1986: ch. 4.

Nearly half a century later, for example, in the important 1960 case of *Natanson* v. *Kline* the Kansas Supreme Court made an equally ringing appeal to self-determination:

Anglo-American law stars with the premise of thorough-going self-determination. It follows that each man is considered to be master of his own body, and he may, if he be of sound mind, expressly prohibit the performance of life-saving surgery, or other medical treatment. A doctor might well believe that an operation or form of treatment is desirable or necessary but the law does not permit him to substitute his own judgment for that of the patient by any form of artifice or deception.[29]

[29] *Natanson* v. *Kline*, 1960.

The philosophical tradition regarding the problem of paternalism is equally bound up in a commitment to the importance of individual self-determination or autonomy.

Here, the *locus classicus* is John Stuart Mill's renowned assertion of the 'one very simple principle' that

the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise or even right. [30]

[30] Mill, 1859: 13.

The voluminous subsequent philosophical literature on paternalism certainly suggests that this principle is not as simple as Mill supposed, but it also makes

end p.108

clear that, even for one like Mill who in other contexts was an avowed utilitarian, the case for non-interference with individual self-determination and liberty of action does not rest on any claim that doing so cannot be for a person's 'good', understood in a normatively subjective interpretation. Quite to the contrary, Mill and the many who have followed him have been at pains to insist that such interference is not justified even when it would truly be for the good of the one interfered with. [31]

[31] The most detailed recent account of justified paternalism in the spirit of Mill's position is VanDeVeer, 1986. The most sophisticated development of a Millian position on paternalism in criminal law is Feinberg, 1986. I have explored some of the issues between rights-based and consequentialist accounts of paternalism in Brock, 1983 and 1988.

The result is that it is not possible to draw any firm conclusion from the doctrine of informed consent in medicine that patients' well-being or quality of life is understood according to a normatively subjective interpretation. Individual self-determination can serve as the foundation for the informed consent doctrine and can make that doctrine compatible with any of the three main alternative accounts of the good or quality of life that I have distinguished.

What is the relation between these two values of patient self-determination and well-being, commonly taken as underlying the informed consent doctrine, and the broad concept of a good life? The conventional view, I believe, is that the patient's well-being is roughly equivalent to the patient's good and that individual self-determination is a value independent of the patient's well-being or good. Respecting the patient's right to self-determination, then, at least sometimes justifies respecting treatment choices that are contrary to the patient's well-being or good. [32]

[32] This conventional view is reflected in the independent ethical principles of beneficence and autonomy in Beauchamp and Childress, 1979. This book has probably been the account of moral principles most influential with people in medicine and health care without philosophical training in ethics.

Respecting self-determination is commonly held to be what is required by recognizing the individual as a person, capable of forming a conception of the good life for him or herself. If personal self-determination is a fundamental value—fundamental in that it is what is involved in respecting persons—however, then I suggest that our broadest conception of a good life should be capable of encompassing it rather than setting it off as separate from and in potential conflict with a person's well-being or good, as in the conventional account of informed consent. What we need is a distinction between a good life for a person in the broadest sense and a person's personal well-being, such that only personal well-being is independent of and potentially in conflict with individual self-determination.

We should think of being self-determined as central to—a central part of—having a good life in the broadest sense. It is in the exercise of self-determination that we maintain some control over, and take responsibility for, our lives and for what we will become. This is not to deny, of course, that there are always substantial limits and constraints within which we must exercise this judgement and choice. But it is to say that showing respect for people through respecting their self-determination acknowledges the fundamental place of

end p.109

self-determination in a good life. I take this to be essentially what Rawls intends by his claim that people have a highest-order interest in autonomy and what Sen means by his notion of agency freedom.[33]

[33] Cf. Rawls, 1980, and Sen, 1985, esp. pp. 203-4.

We do not want this broad conception of a good life, however, to prevent our making sense of persons freely and knowingly choosing to sacrifice their own personal well-being for the sake of other persons. To cite an extreme case, a parent might knowingly and freely choose not to pursue expensive life-sustaining treatment such as a heart transplant in order to preserve financial resources for his children's education. Given his love and sense of responsibility for his children, he would judge his life to be worse if he had the transplant at the expense of his children's education, though certainly his health and his personal well-being would be improved. We might say here that he values his personal well-being in these circumstances less than the well-being of his children. In its requirement that such a choice be respected, the informed consent doctrine implicitly accepts that the best life for a person is a life of self-determination or choice, even if the exercise of that self-determination or choice results in a lessened state of personal well-being. Precisely how to make this distinction between the good life, as opposed to the personal well-being, of an individual raises difficulties that I cannot pursue here. The rough idea is that personal well-being makes essential reference to the states of consciousness, activities, and capacities for functioning of the person in question (I will pursue this further in Section 2 of this paper), and it is these that are worsened when the parent pursues his conception of a good life in sacrificing his personal well-being for that of his child.[34]

[34] If personal well-being is understood in this way, it suggests that satisfaction of a person's non-personal desires that make no such reference to him do not increase the person's well-being. Consider a loyal fan of the Boston Red Sox baseball team, who wants the Red Sox to win the pennant. On the last day of the season, tied for first place with the New York Yankees, the Red Sox beat the Yankees and win the pennant. Suppose the fan is travelling in a remote area of Alaska on the day of the big game and a week later, before getting out of the wilderness area and learning of the Red Sox victory, he is killed by a rock slide. Was his personal well-being increased at all simply because the state of affairs he desired—the Red Sox victory—obtained? I believe the answer should be no.

The harder question is whether, in our broader sense of a good life, he had a better life even unbeknownst to him. And was the quality of his life any better? Certainly his life *as experienced by him* was no better and not of higher quality. Even in our broad sense of a good life, his life may seem not to have gone better, but perhaps that is only because this is a relatively unimportant desire. Suppose instead, to adapt an example of Parfit's (1984), a person devotes fifty years of his life to saving Venice and then, confident that it is safe, goes on vacation to the Alaskan wilderness. While he is there a flood destroys Venice, but, like the Red Sox fan, he never learns of it because a week after the flood and before getting out of the wilderness he is killed by a rock slide. Parfit notes that it is plausible to say of the destruction of Venice both that it has made the person's life go less well because he had invested his life in this goal and his life's work is now in vain, but also that it cannot lower the quality of his life if it does not affect the quality of his experience. This suggests a point where a broad notion of a good life may diverge from the notion of the quality of life. Since I believe that medicine and medical

ethics have little illumination to offer on this point, I set it aside here and shall in the body of the paper continue to use the broad notion of a good life largely interchangeably with the quality of life.

Medical treatment decisions must often be made for patients who are not

end p.110

themselves competent to make them. There has been considerable discussion both in medical ethics and in the law regarding appropriate ethical standards for such decisions.[35]

[35] Allen Buchanan and I have discussed ethical issues in decision-making for incompetent persons in our paper (Buchanan and Brock, 1986: 17-94) and in our book (Buchanan and Brock, 1989).

Since quality of life considerations are virtually always relevant to these decisions, the ethical frameworks developed for these decisions must employ, either explicitly or implicitly, a conception of the quality of life of patients. There is considerable consensus that if the patient in question, while still competent, formulated and left an explicit advance directive clearly and unambiguously specifying his or her wishes regarding treatment in the circumstances now obtaining, then those wishes should be followed, at least within very broad limits, by those treating the patient. At the present time in the United States, most states have adopted legislation giving the force of law to one or another form of so-called Living Will, which allows people to give binding instructions about their treatment should they become incompetent and unable to decide for themselves. Several other states have more recently enacted legislation permitting people to draw up a Durable Power of Attorney for Health Care, which combines the giving of instructions about the person's wishes regarding treatment with the designation of who is to act as surrogate decision-maker, and so to interpret those instructions, should one become incompetent to make the decisions oneself.

In the usual case in which an incompetent patient has left no formal advance directive, two principles for the guidance of those who must decide about treatment for the patient have been supported—the *substituted judgement* principle and the *best interests* principle. The substituted judgement principle requires the surrogate to decide as the patient would have decided if competent and in the circumstances that currently obtain. The best interests principle requires the surrogate to make the treatment decision that best serves the patient's interests. This has the appearance of a dispute between what I earlier called normatively subjective and normatively objective accounts of a good life, since the only point of the best interests principle as an alternative to substituted judgement might seem to be that it employs a normatively objective standard of the person's good that does not depend on his or her particular subjective preferences and values. However, this appearance is misleading. These two principles of surrogate decision-making are properly understood, in my view, not as competing alternative principles to be used for the same cases, but instead as an ordered pair of principles to cover all cases of surrogate decision-making for incompetent patients in which an advance directive does not exist, with each of the two principles to apply in a different subset of these cases. (This is not to say that these two principles are always in fact understood in medical ethics, the law, or health care practice as applying to distinct groups of cases; the treatment of these two principles is rife with confusion.)

The two groups of cases are differentiated with regard to the information

end p.111

available or obtainable concerning the patient's general preferences and values that has some bearing on the treatment choice at hand. The two principles are an ordered pair in the sense that when sufficient information is available about the relevant preferences and values of the patient to permit a reasonably well-grounded application of the substituted judgement principle, then surrogate decision-makers for

the patient are to use that information and that principle to infer what the patient's decision would have been in the circumstances if he or she were competent. [36]

[36] Rebecca Dresser has developed a Parfitian challenge to the substituted judgement principle as well as to the authority of advance directives in cases in which the conditions creating the patient's incompetence also reduce or eliminate the psychological continuity and connectedness necessary for personal identity to be maintained. Cf. Dresser, 1986. Cf. also Buchanan, 1988.

In the absence of such information, and only then, surrogate decision-makers for the patient are to select the alternative that is in the best interests of the patient, which is usually interpreted to mean the alternative that most reasonable and informed persons would select in the circumstances. Thus, these two principles are not competing principles for application in the same cases, but alternative principles to be applied in different cases.

Nevertheless, it might seem that the best interests standard remains a normatively objective account when it is employed. However, this need not be so. If the best interests standard is understood as appealing to what most informed and reasonable persons would choose in the circumstances, it employs the normatively subjective preference standard. And it applies the choice of most persons to the patient in question because in the absence of any information to establish that the patient's relevant preferences and values are different than most people's, the most reasonable presumption is that the patient is like most others in the relevant respects and would choose like those others. Thus, the best interests standard, like the two other standards of choice for incompetent patients—the advance directive and substituted judgement standards—can be understood as requiring the selection of the alternative the patient would most likely have selected, with the variations in the standards suited to the different levels of information about the patient that is available. Just as with the informed consent doctrine that applies to competent patients, so these three principles—advance directives, substituted judgement, and best interests—guiding surrogate choice for incompetent patients can all be understood as supported and justified by the values of patient well-being *and* self-determination. Thus, each of these three principles implicitly employs an account of a good life that is a life of choice and self-determination concerning one's aims, values, and life plan.

Before leaving the ethical frameworks that have been developed in the medical ethics, legal, and medical literatures for treatment decision-making for competent and incompetent patients, I want to make explicit an indeterminacy in these frameworks concerning the nature of the ethical theory they presuppose. I have noted that it is common to base these ethical frameworks on two central values in a good life: patient well-being and self-determination. What is commonly left unclear, however, is the foundational status of the ethical value of



individual self-determination in the underlying ethical theory supporting these decision-making frameworks. Self-determination might be held to have only derivative or instrumental value within a broadly consequentialist moral theory. Specifically, it might be held to be instrumentally valuable for the fundamental value of happiness or preference satisfaction within normatively subjective theories of a good life of a hedonist or preference satisfaction sort. If, as seems true at least for many social conditions and historical periods, most people have a relatively strong desire to make significant decisions about their lives for themselves, then it is at least a plausible presumption that their doing so will generally promote their happiness or the satisfaction of their desires. If self-determination is only valuable at bottom in so far as it leads to happiness or preference satisfaction, then it will not be part of an

ideal of the person that is objective in the sense of its value not being entirely dependent on happiness or satisfaction. Since on most plausible theories of the good, happiness and desire satisfaction are significant *part* of a good life, and since self-determination does commonly make a significant contribution to people's happiness or desire satisfaction, self-determination will commonly have significant instrumental value on any plausible theory of a good life.[37]

[37] Scanlon argues in this volume that desire satisfaction is not itself a basic part of individual well-being, but is dependent on hedonistic or ideal (what he calls substantive good) reasons for its support.

As a result, to single out self-determination as one of the two principal values underlying the informed consent doctrine in medical ethics as it applies both to competent and incompetent patients is not to make clear at a foundational level of ethical theory whether self-determination is held to have only instrumental value, or also significant non-instrumental value as an important component of an objective ideal of the person. The vast majority of ethical discussions of informed consent and health care treatment decision-making simply do not either explicity address this foundational question of ethical theory or even implicitly presuppose a particular position on it. From a practical perspective, this foundational indeterminacy has the value of allowing proponents of incompatible ethical theories, for example consequentialists and rights-based theorists, to agree on the fundamental importance of self-determination and choice in a good life.

There is one final difficulty to be noted in attempting to infer the account of a good life at the level of basic ethical theory from the ethical frameworks for treatment decision-making advocated for and employed in medical practice. There is a general difficulty in inferring the underlying values or ethical principles that support social practices. A social practice like that of informed consent and shared decision-making in medicine must guide over time a very great number of treatment decisions carried out in a wide variety of circumstances by many and diverse patients, family members, and health care professionals. A well-structured practice must take account of and appropriately minimize the



potential in all involved parties for well-intentioned misuse or ill-intentioned abuse of their roles. Institutional constraints may thus be justified, though on some particular occasions they will produce undesirable results, because in the long run their overall results are better than those of any feasible alternatives. For example, even if we appeal only to the value of patient well-being and leave aside any independent value of self-determination, a strong right of patients to refuse any treatment might be justified if most of the time people are themselves the best judges of what health care treatment will best promote their happiness or satisfy their enduring preferences and values. Alternatively, that same strong right to refuse treatment might be justified within a normatively objective ideal theory of a good life for persons because, though individuals can be mistaken about their own good when they pursue their happiness or seek to satisfy their desires, no reliable alternative social and legal practice is feasible that will produce better results in the long run, even judged by that ideal theory of a good life. The imperfections and limitations of people and institutions may lead supporters of quite different accounts of a good life to support roughly the same institutions in practice. In a more general form, this is a thoroughly familiar point in moral philosophy where defenders of fundamentally different moral theories, such as consequentialists and rights theorists, may converge on the institutions justified by their quite different theories. Without explicitly uncovering the justificatory rationales for specific social institutions accepted by particular persons, we cannot confidently

infer the ethical principles or judgements, and specifically the conception of a good life, that they presuppose.

# 3 Health Policy Measures of the Quality of Life

I want now to shift attention from the account of the quality of life presupposed by ethical frameworks for medical treatment decision-making to more explicit measures used to assess health levels and the quality of life as it is affected by health and disease within larger population groups. Early measurement attempts focused on morbidity and mortality rates in different populations and societies. These yield only extremely crude comparisons, since they often employ only such statistics as life expectancy, infant mortality, and reported rates of specific diseases in a population. Nevertheless, they will show gross differences between countries, especially between economically developed and underdeveloped countries, and between different historical periods, in both length and quality of life as it is affected by disease. Major changes in these measures during this century, as is well known, have been due principally to public health measures such as improved water supplies, sewage treatment, and other sanitation programmes and to the effects of economic development in improving nutrition, housing, and education; improvements in the quality of and access to medical care have been less important. In recent decades, health policy researchers have developed a variety of measures that go substantially beyond crude morbidity

end p.114

and mortality measures. Before shifting our attention to them, however, it is worth underlining the importance of mortality measures to the broad concept of a good life. When quality and quantity of life are distinguished, both are relevant to the degree to which a person has a good life. People whose lives are of high quality, by whatever measure of quality, but whose lives are cut short well before reaching the normal life span in their society, have had lives that have gone substantially less well, because of their premature death, than reasonably might have been expected. People typically develop, at least by adolescence, more or less articulated and detailed plans for their lives; commonly, the further into the future those plans stretch, the less detailed, more general, and more open-ended they are. Our life plans undergo continuous revision, both minor and substantial, over the course of our lives, but at any point in time within a life people's plans for their lives will be based in part on assumptions about what they can reasonably expect in the way of a normal life span.[38]

[38] Two of the more important discussions of life plans and of how they can give structure and coherence to life are Fried, 1970: ch. 10, and Rawls, 1971: ch. 7.

When their lives are cut short prematurely by illness and disease they lose not just the experiences, happiness, and satisfactions that they would otherwise have had in those lost years, but they often lose as well the opportunity to complete long-term projects and to achieve and live out the full shape, coherence, and conclusion that they had planned for their lives. It is this rounding out and completion of a life plan and a life that helps enable many elderly people, when near death, to feel that they have lived a full and complete life and so to accept their approaching death with equanimity and dignity. The loss from premature death is thus not simply the loss of a unit of a good thing, so many desired and expected-to-be-happy life-years, but the cutting short of the as yet incompletely realized life plan that gave meaning and coherence to the person's life.

The importance of life plans for a good life suggests at least two other ways in which different mortality rates within societies affect the opportunities of their members to

attain good lives. Citizens of economically underdeveloped countries typically have shorter life expectancies than do citizens of the developed countries. Thus, even those who reach a normal life span will have less time to develop and enjoy a richly complex and satisfying life than will those who reach a normal, but significantly longer, life span in the developed countries. Mortality data indicate that citizens of underdeveloped countries typically have less good lives as a result of inadequate health care *both* because of their shorter life expectancies *and* because of their increased risk of not living to even the normal life span in their own society.

One final relation between mortality data and the importance of life plans in a good life concerns infant and extremely early childhood mortality. It is common to view such early mortality as particularly tragic, both because of

end p.115

the greater amount of expected life-years lost to the individual and because the life was cut short just as it was getting started. The loss is often deep for the parents, in part because of the hopes and plans for the infant's future that *they* had for the infant. But the death of an infant or extremely young child before he or she has developed the capacity to form desires, hopes, and plans for the future cuts that life short—'life' understood as a connected plan or unfolding biography with a beginning, middle, and end—before it has begun; the infant is alive, but does not yet have a life in this biographical sense.[39]

[39] The notion of a biographical life is employed in the medical ethics literature by Rachels, 1986: 5-6, by Singer and Kuhse, 1985: 129-39, and by Callahan, 1987. It is also implicit in Tooley's (1983) account of the right to life.

From the perspective of this biographical sense of having a life as lived from the 'inside', premature death in later childhood, adolescence, or early adulthood commonly makes a life that has got started go badly, whereas infant death does not make a life go badly, but instead prevents it from getting started.

There is a voluminous medical and health policy literature focused on the evaluation of people's quality of life as it is affected by various disease states and/or treatments to ameliorate or cure those diseases.[40]

[40] Among the useful papers consulted for this third section on health policy measures of quality of life, and not cited in other notes, are: Anderson 1986; Berg, 1986; Bergner, 1976; Calman, 1984; Cohen, 1982; Cribb, 1985; Editorial, 1986; Edlund and Tancredi, 1985; Flanagan, 1982; Gehrmann, 1978; Gillingham and Reece, 1980; Grogono, 1971; Guyatt, 1986; Hunt and EcEwen, 1980; Katz, 1963; Kornfeld, 1982; Klotkem, 1982; Liang, 1982; Najman and Levine, 1981; Pearlman and Speer, 1983; Presant, 1984; Report, 1984; Starr, 1986; Sullivan, 1966; Thomasma, 1984, 1986; Torrance, 1972, 1976*a*, 1976*b*.

The dominant conception of the appropriate aims of medicine focuses on medicine as an intervention aimed at preventing, ameliorating, or curing disease and its associated effects of suffering and disability, and thereby restoring, or preventing the loss of, normal function or of life. Whether the norm be that of the particular individual, or that typical in the particular society or species, the aim of raising people's function to *above* the norm is not commonly accepted as an aim of medicine of equal importance to restoring function *up to* the norm. Problematic though the distinction may be, quality of life measures in medicine and health care consequently tend to focus on individuals' or patients' *dysfunction* and its relation to some such norm. At a deep level, medicine views bodily parts and organs, individual human bodies, and people from a functional perspective. Both health policy analysts and other social scientists have done considerable work constructing and employing measures of health and quality of life for use with large and relatively diverse populations. Sometimes these measures explicitly address only part of an overall evaluation of people's quality of life, while in other instances they address something

like overall quality of life as it is affected by disease. A closely related body of work focuses somewhat more narrowly on an evaluation of the effect on quality of life of specific modes of treatment for specific disease states. This research is more clinically oriented, though the breadth of impact on quality of life researchers seek to measure does vary to

end p.116

some extent, depending often on the usual breadth of impact on the person of the disease being treated. Generally speaking, the population-wide measures tend to be less sensitive to individual differences as regards both the manner and the degree to which a particular factor affects people's quality of life. It will be helpful to have before us a few representative examples of the evaluative frameworks employed.[41]

[41] An example of a broad quality of life measure not focused on health care and disease can be found in the Swedish Level of Living Surveys discussed in the essay by Erikson in this volume.

The Sickness Impact Profile (SIP) was developed by Marilyn Bergner and colleagues to measure the impact of a wide variety of forms of ill health on the quality of people's lives.[42]

[42] Bergner, 1981.

Table 1 enumerates the items measured.

A second example of an evaluative framework is the Quality of Life Index (QLI) developed by Walter O. Spitzer and colleagues to measure the quality of life of cancer patients (see Table 2).[43]

[43] Spitzer, 1981.

A third prominent measure developed by Milton Chen, S. Fanshel and others is the Health Status Index (HSI), which measures levels of function along certain dimensions (see Table 3).[44]

[44] Chen, Bush, and Patrick, 1975.

It would be a mistake, of course, to attempt to infer precise and comprehensive philosophical theories of the quality of life or of a good life from measures such as these. The people who develop them are commonly social scientists and health care researchers who are often not philosophically sophisticated or concerned with the issues that divide competing philosophical accounts of a good life. The practical and theoretical difficulties in constructing valid measures that are feasible for large and varied populations require compromises with and simplifications of—or simply passing over—issues of philosophical importance. Nevertheless, several features of these measures are significant in showing the complexity of the quality of life measures employed in health care and, I believe, of any adequate account of the quality of life or of a good life.

First, the principal emphasis in each of the three measures of quality of life is on function, and functions of the 'whole person' as opposed to body parts and organ systems. In each case the functions are broadly characterized so as to be relevant not simply to a relatively limited and narrow class of life plans, but to virtually any life plan common in modern societies. Following the lead of Rawls's notion of 'primary goods', I shall call these 'primary functions'.[45]

[45] Rawls, 1971: 62, 90-5.

In the SIP, the categories of sleep and rest, and eating are necessary for biological function. The categories of work, home management, and recreation and pastimes are central activities common in virtually all lives, though the relative importance they have in a particular life can be adjusted for by making the measure relative to what had been the individual's normal level of activity in each of these areas prior to

sickness. The two broad groups of functions, physical and psychosocial, are each broken down into several distinct components. For

end p.117

**Table 1 Sickness Impact Profile Categories and Selected Items**

| Dimension | Category | Items describing behaviour related to: | Selected items |
|---|---|---|---|
| Independent categories | SR | Sleep and rest | I sit during much of the day |
| | | | I sleep or nap during the day |
| | E | Eating | I am eating no food at all, nutrition is taken through tubes or intravenous fluids |
| | | | I am eating special or different food |
| | W | Work | I am not working at all |
| | | | I often act irritable towards my work associates |
| | HM | Home management | I am not doing any of the maintenance or repair work around the house that I usually do |
| | | | I am not doing heavy work around the house |
| | RP | Recreation and pastimes | I am going out for entertainment less |
| | | | I am not doing any of my usual physical recreation or activities |
| I. Physical | A | Ambulation | I walk shorter distances or stop to rest often |
| | | | I do not walk at all |
| | M | Mobility | I stay within one room |
| | | | I stay away from home only for brief periods of time |
| | BCM | Body care and movement | I do not bath myself at all, but am bathed by someone else |
| | | | I am very clumsy in my |

| | | | |
|---|---|---|---|
| | | | body movements |
| II. Psychosocial | SI | Social interaction | I am doing fewer social activities with groups of people |
| | | | I isolate myself as much as I can from the rest of the family |
| | AB | Alertness behaviour | I have difficulty reasoning and solving problems, e.g. making plans, making decisions, learning new things |
| | | | I sometimes behave as if I were confused or disoriented in place or time, e.g. where I am, who is around, directions, what day it is |
| | EB | Emotional behaviour | I laugh or cry suddenly |
| | | | I act irritable and impatient with myself, e.g. talk badly about myself, swear at myself, blame myself for things that happen |
| | C | Communication | I am having trouble writing or typing |
| | | | I do not speak clearly when I am under stress |

end p.118

each primary function, the SIP measures the impact of sickness by eliciting information concerning whether activities typical in the exercise of that function continue to be performed, or have become limited. Even for primary functions, about which it is plausible to claim that they have a place in virtually any life, the different functions can have a different *relative* value or importance within different lives, and the SIP makes no attempt to measure those differences. The QLI likewise addresses a person's levels of activity in daily living, specifically measuring the presence of related behaviours in the relevant areas. In measuring health and outlook, the primary concern is with subjective feeling states of the person, though here too there is concern with relevant behaviour. The category of support addresses both the social behaviour of the individual and the availability of people in the individual's environment to provide such relationships. This category illustrates the important point that most primary functional capacities require both behavioural capacities in the individual and relevant resources in the individual's external environment. The

HSI addresses three broad categories of primary function—mobility, physical activity, and social activity—with evidence of current functional capacity found in current levels of activity. It is noteworthy that even this index, which focuses explicitly on the health status of individuals, does not measure the presence or absence of disease, as one might expect given common understandings of 'health' as the absence of disease; like the SIP and QLI, it too measures levels of very broad primary functions. A second important feature of these measures shows up explicitly only in the first two—SIP and QLI—and is best displayed in the 'outlook' category of the QLI, though it is also at least partly captured in the 'emotional behaviour' category of the SIP. Both these categories can be understood as attempts to capture people's subjective response to their objective physical condition and level of function, or, in short, their level of happiness or satisfaction with their lives, though the actual measures are far too crude to measure happiness with much sensitivity. The important point is that the use of these categories represents a recognition that *part* of what makes a good life is that the person in question is happy or pleased with how it is going; that is, subjectively experiences it as going well, as fulfilling his or her major aims, and as satisfying. This subjective happiness component is not unrelated, of course, to how well the person's life is going as measured by the level of the other objective primary functions. How happy we are with our lives is significantly determined by how well our lives are in fact going in other objective respects. Nevertheless, medicine provides many examples that show it is a mistake to assume that the subjective happiness component correlates closely and invariably with other objective functional measures. In one study, for example, researchers found a substantial relation between different objective function variables and also between different subjective response or outlook variables, but only a very limited relation between objective and subjective variables. [46]

[46] Evans, 1985.

These data reinforce the importance of

end p.119

**Table 2 Quality of Life Index: Formal of the Final Version Adopted**

Study No _____

Age _____ ☐☐
See $M_1$ $F_2$ (Ring appropriate letter) _____ ☐
Primary Problem or Diagnosis _____
_____ ☐☐☐
Secondary Problem or Diagnosis, or complication (if appropriate) _____ ☐☐☐
Scorer's Speciality _____ ☐☐☐

Scoring Form

*Score each Reading 2, 1 or 0 according to your most recent assessment of the patient.*

**ACTIVITY** *During the last week, the patient*

- has been working or studying full time, or nearly so, in usual occupation; or managing own household, or participating in unpaid or voluntary activities, whether retired or not ..... 2
- has been working or studying in usual occupation or managing own household or participating in unpaid or voluntary activities, but requiring major assistance or a significant reduction in hours worked or a sheltered situation or was on sick leave ..... 1 ☐
- has not been working or studying in any capacity and not managing own household ..... 0

**DAILY LIVING** *During the last week, the patient*

- has been self-reliant in eating, washing, toiletting, and dressing: using public transport or driving own car ..... 2
- has been requiring assistance (another person or special equipment) for daily activities and transport but performing light tasks ..... 1 ☐
- has not been managing personal care or light tasks and/or not leaving own home or institution at all ..... 0

**HEALTH** *During the last week, the patient*

- has been appearing to feel well or reporting feeling 'great' most of the time ..... 2
- has been lacking energy or not feeling entirely 'up to par' more than just occasionally ..... 1 ☐
- has been feeling very ill or 'lousy', seeming weak and washed out most of the time, or was unconscious ..... 0

**SUPPORT** *During the last week*

- the patient has been having good relationships with others and receiving strong support from at least one family member and/or friend ..... 2
- support received or perceived has been limited from family and friends and/or by the patient's condition ..... 1 ☐
- support from family and friends occurred infrequently or only when absolutely necessary or patient was unconscious ..... 0

**OUTLOOK** *During the past week the patient*

- has usually been appearing calm and positive in outlook, accepting and in control of personal circumstances, including surroundings ..... 2
- has sometimes been troubled because not fully in control of personal circumstances or has been having periods of obvious anxiety or depression ..... 1 ☐
- has been seriously confused or very frightened or consistently anxious and depressed or unconscious ..... 0

QL INDEX TOTAL ☐☐

How confident are you that your scoring of the preceding dimensions is accurate? Please ring the appropriate category.

| Absolutely Confident | Very Confident | Quite Confident | Not Very Confident | Very Doubtful | Not at all Confident |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 ☐ |

end p.120
end p.121

## Table 3 Scales and Definitions for the Classification of Function Levels

| ScaleStep | | Definition |
|---|---|---|
| | | *Mobility scale* |
| 5 | Travelled freely | Used public transportation or drove alone. For below 6 age group, travelled as usual for age. |
| 4 | Travelled with difficulty | (a) Went outside alone, but had trouble getting around community freely, or (b) |

| | | |
|---|---|---|
| | | required assistance to use public transportation or automobile. |
| 3 | In house | (a) All day, because of illness or condition, or (b) needed human assistance to go outside. |
| 2 | In hospital | Not only general hospital, but also nursing home, extended care facility, sanatorium, or similar institution. |
| 1 | In special unit | For some part of the day in a restricted area of the hospital such as intensive care, operating room, recovery room, isolation ward, or similar unit. |
| | Death | |
| | | *Physical activity scale* |
| 4 | Walked freely | With no limitations of any kind. |
| 3 | Walked with limitations | (a) With cane, crutches, or mechanical aid, or (b) limited in lifting, stooping, or using stairs or inclines, or (c) limited in speed or distance by general physical condition. |
| 2 | Moved independently in wheelchair | Propelled self alone in wheelchair. |
| 1 | In bed or chair | For most or all of the day. |
| | Death | |
| | | *Social activity scale* |
| 5 | Performed major and other activities | *Major* means specifically: play for below 6, school for 6-17, and work or maintain household for adults. *Other* means all activities not classified as major, such as athletics, clubs, shopping, church, hobbies, civic projects, or games as appropriate for age. |
| 4 | Performed major but limited in other activities | Played, went to school, worked, or kept house but limited in other activities as defined above. |
| 3 | Performed major activity with limitation | Limited in the amount or kind of major activity performed, for instance, needed special rest periods, special school, or special working aids. |
| 2 | Did not perform major activity but | Did not play, go to school, work or keep house, but dressed, bathed, and fed |

| | | |
|---|---|---|
| | performed self-care activities | self. |
| 1 | Required assistance with self-care activities | Required human help with one or more of the following — dressing, bathing, or eating — and did not perform major or other activities. For below 6 age group, means assistance not usually required for age. |
| | Death | |

end p.122

including both objective function and subjective response categories in a full conception of the quality of life, since neither is a reliable surrogate for the other. Given this at least partial independence between happiness and function variables, what is their relative weight in an overall assessment of a good life? Here, too, medicine brings out forcefully that there can be no uniform answer to this question. In the face of seriously debilitating injuries, one patient will adjust her aspirations and expectations to her newly limited functional capacities and place great value on achieving happiness despite these limitations. Faced with similar debilitating injuries, another patient will assign little value to adjusting to the disabilities in order to achieve happiness in spite of them, stating that she 'does not want to become the kind of person who is happy in that debilitated and dependent state'. [47]

[47] The main character in the popular play and subsequent film *Whose Life is it Anyway*?, having become paralysed from the neck down, displayed this attitude of not wanting to become a person who had adjusted to his condition.

There are other important qualifications of the generally positive relation between this happiness component of a good life and both the other primary function components and the overall assessment of how good a life it is. These qualifications are not all special to health care and the quality of life, but some are perhaps more evident and important in the area of health care than elsewhere. The first qualification concerns people's adjustments to limitations of the other primary functions. Sometimes the limitation in function, or potential limitation, is due to congenital abnormalities or other handicaps present from birth. For example, an American television programme recently reported on a follow-up of some of the children, now young adults, born to pregnant women who had taken the drug Thalidomide in the late 1950s. [48]

[48] *60 Minutes*, CBS television network programme, 21 February 1988.

The people reported on had suffered no brain damage but had been born with serious physical deformities, including lacking some or any arms and legs. While this placed many impairments in the way of carrying out primary functions such as eating, working, home management, physical mobility, and ambulation in the manner of normal adults, these people had made remarkable adjustments to compensate for their physical limitations: one was able to perform all the normal functions of eating using his foot in place of missing arms and hands; another made his living as an artist painting with a brush held between his teeth; another without legs was able to drive in a specially equipped car; and a mother of three without legs had adapted so as to be able to perform virtually all the normal tasks of managing a family and home. These were cases where physical limitations that commonly restrict and impair people's primary functional capacities and overall quality of life had been so well compensated for as to enable them to perform the *same* primary functions, though in different ways, as well as normal, unimpaired persons do. While a few life plans

possible for others remained impossible for them because of their limitations (for example, being professional athletes), their essentially

end p.123

unimpaired level of primary functions as a result of the compensations they had made left them with choice from among a sufficiently wide array of life plans that it is probably a mistake to believe that their quality of life had been lowered much or at all by their impairments. These cases illustrate that even serious physical limitations do not always lower quality of life if the disabled persons have been able or helped sufficiently to compensate for their disabilities so that their level of primary functional capacity remains essentially unimpaired; in such cases it becomes problematic even to characterize those affected as disabled.

In other cases, compensating for functional disabilities, particularly when they arise later in life, may require adjustments involving substantial changes in the kind of work performed, social and recreational activities pursued, and so forth. When these disabilities significantly restrict the activities that had been and would otherwise have been available to and pursued by the person, they will, all other things being equal, constitute reductions in the person's quality of life. If they do so, however, it will be because they significantly restrict the choices, or what Norman Daniels has called the normal opportunity range, available to the persons, and not because the compensating paths chosen need be, once entered on, any less desirable or satisfying.[49]

[49] Daniels, 1985: chs. 2 and 3.

The *opportunity for choice* from among a reasonable array of life plans is an important and independent component of quality of life: it is insufficient to measure only the quality of the life plan the disabled person now pursues and his or her satisfaction with it. Adjustments to impairments that leave primary functions undiminished or that redirect one's life plan into areas where function will be better — both central aims of rehabilitative medicine — can, however, enhance quality of life even in the face of a diminished opportunity range.

In his theory of just health care Daniels uses the notion of an *age-adjusted* normal opportunity range, which is important for the relation between opportunity and quality of life or a good life. Some impairments in primary functions occur as common features of even the normal aging process, for example, limitations in previous levels of physical activity. Choosing to adjust the nature and level of our planned activities to such impairments in function is usually considered a healthy adjustment to the aging process. This adjustment can substantially diminish the reduction in the person's quality of life from the limitations of normal aging. Nevertheless, even under the best of circumstances, the normal aging process (especially, say, beyond the age of 80), does produce limitations in primary functions that will reduce quality of life. Thus, while quality of life must always be measured against normal, primary functional capacities for humans, it can be diminished by reductions both in individual function below the age-adjusted norm and by reductions in normal function for humans as they age.

I have suggested above that adjustments in chosen pursuits as a result of

end p.124

impairments in primary, or previously pursued individual, functions can compensate substantially (fully, in the effects on happiness) for impaired function, but will often not compensate fully for significant reductions in the range of *opportunities* available for choice, and so will not leave quality of life undiminished. In some cases, however, a patient's response and adjustment to the limitations of illness or injury may be so complete, as regards his commitment to and happiness from the new chosen life

path, that there is reason to hold that his quality of life is as high as before, particularly as he gets further away in time from the onset of the limiting illness or injury and as the new life becomes more securely and authentically the person's own. An undiminished or even increased level of happiness and satisfaction, together with an increased commitment to the new life, often seem the primary relevant factors when they are present. But we must also distinguish different *reasons* why the affective or subjective component of quality of life, which I have lumped under the notion of happiness, may remain undiminished, since this is important for an evaluation of the effect on quality of life.

A person's happiness is to some significant extent a function of the *degree* to which his or her major aims are being at least reasonably successfully pursued. Serious illness or injury resulting in serious functional impairment often requires a major revaluation of one's plan of life and its major aims and expectations. Over time, such revaluations can result in undiminished or even increased levels of happiness, despite decreased function, because the person's aspirations and expectations have likewise been revised and reduced. The common cases in medicine in which, following serious illness, people come to be satisfied with much less in the way of hopes and accomplishments illustrate clearly the incompleteness of happiness as a full account of the quality of life. To be satisfied or happy with getting much less from life, because one has come to expect much less, is still to get *less* from life or to have a less good life. (The converse of this effect is when rising levels of affluence and of other objective primary functions in periods of economic development lead to even more rapidly rising aspirations and expectations, and in turn to an *increasing gap* between accomplishments and expectations.) Moreover, whether the relation of the person's choices to his aspirations and expectations reflects his exercise of self-determination in response to changed circumstances is important in an overall assessment of his quality of life and shows another aspect of the importance of self-determination to quality of life.

Illness and injury resulting in serious limitations of primary functions often strike individuals without warning and seemingly at random, and are then seen by them and others as a piece of bad luck or misfortune. Every life is ended by death, and few people reach death after a normal life span without some serious illness and attendant decline in function. This is simply an inevitable part of the human condition. Individual character strengths and social support services enable people unfortunately impaired by disease or injury to adjust their aims and expectations realistically to their adversity, and then to get on

end p.125

with their lives, instead of responding to their misfortune with despair and self-pity. Circumstances beyond individuals' control may have dealt them a cruel blow, but they can retain dignity as self-determining agents capable of responsible choice in directing and retaining control over their lives within the limits that their new circumstances permit. We generally admire people who make the best of their lot in this way, and achieve happiness and accomplishment despite what seems a cruel fate. This reduction in aims and expectations, with its resultant reduction in the gap between accomplishments and aims, and the in turn resultant increase in happiness, is an outcome of the continued exercise of self-determination. It constitutes an increase in the happiness and self-determination components of quality of life though, of course, only in response to an earlier decrease in the person's level of primary functions.

Other ways of reducing this gap between accomplishments and expectations bypass the person's self-determination and are more problematic as regards their desirability and their effect on a person's quality of life. Jon Elster has written, for example, outside of the medical context, of different kinds of non-autonomous preferences and preference change.[50]

[50] Elster, 1982.

Precisely characterizing the difference between what Elster calls non-autonomous preferences and what I have called the exercise of self-determination in adjusting to the impact of illness and injury raises deep and difficult issues that I cannot pursue here. Nevertheless, I believe that response to illness through the exercise of self-determined choice, in the service of protecting or restoring quality of life, is one of the most important practical examples of the significance for overall assessments of the quality of life of *how* to achieve the reasonable accord between aims and accomplishments that happiness requires.

# 4 Conclusion

Let us tie together some of the main themes in accounts of the quality of life or of a good life suggested by the literature in medical ethics and health policy. While that literature provides little in the way of well-developed, philosophically sophisticated accounts of the quality of life or of a good life, it is a rich body of analysis, data, and experience on which philosophical accounts of a good life can draw. I have presented here at least the main outlines of a general account of a good life suggested by that work. The account will be a complex one which, among the main philosophical theories distinguished earlier, probably most comfortably fits within ideal theories. I have suggested that we can employ Sen's construction of a plurality of independent vectors, each of which is an independent component of a full assessment of the degree to which a person has a good life.

The ethical frameworks for medical treatment decision-making bring out the centrality of a person's capacity as a valuing agent, or what I have called self-determination,



in a good life. The capacity for and exercise of self-determination can be taken to be a — or I believe *the* — fundamental ideal of the person within medical ethics. The exercise of self-determination in constructing a relatively full human life will require in an individual four broad types of primary functions: biological, including, for example, well-functioning organs; physical, including, for example, ambulation; social, including, for example, capacities to communicate; mental, including, for example, a variety of reasoning and emotional capabilities. There are no sharp boundaries between these broad types of primary function, and for different purposes they can be specified in more or less detail and in a variety of different bundles. The idea is to pick out human functions that are necessary for, or at least valuable in, the pursuit of nearly all relatively full and complete human life plans. These different functions can be represented on different vectors and they will be normatively objective components of a good life, though their relative weight within any particular life may be subjectively determined.

There are in turn what we can call agent-specific functions, again specifiable at varying levels of generality or detail, which are necessary for a person to pursue successfully the particular purposes and life plan he or she has chosen: examples are functional capacities to do highly abstract reasoning of the sort required in mathematics or philosophy and the physical dexterity needed for success as a musician, surgeon, or athlete. Once again, these functions can be represented on independent vectors, though their place in the good life for a particular person is

determined on more normatively subjective grounds depending on the particular life plan chosen. The relative weight assigned to agent-specific functions and, to a substantially lesser degree, to primary functions, will ultimately be determined by the valuations of the self-determining agent, together with factual determinations of what functions are necessary in the pursuit of different specific life plans. The centrality of the valuing and choosing agent in this account of a good life gives both primary and, to a lesser extent, agent-specific functional capacities a central place in the good life because of their necessary role in making possible a significant range of opportunities and alternatives for choice.

At a more agent-specific level still are the particular desires pursued by people on particular occasions in the course of pursuing their valued aims and activities. Different desires and the degree to which they can be successfully satisfied can also be represented using the vector approach. It bears repeating that the level of a person's primary functional capacities, agent-specific functional capacities, and satisfaction of specific desires will all depend both on properties of the agent and on features of his or her environment that affect those functional capacities and desire pursuits. The inclusion of primary functions, agent-specific functions, and the satisfaction of specific desires all within an account of the good life allows us to recognize both its normatively objective and normatively subjective components. Analogously, these various components show why we can expect partial, but only partial, interpersonal comparability of the quality of life or of good lives — comparability will require interpersonal

end p.127

overlapping of similarly weighted primary functions, agent-specific functions, and specific desires. The importance of functional capacities at these different levels of generality reflects the centrality of personal choice in a good life and the necessity for a choice of alternatives and opportunities.

Finally, there will be the hedonic or happiness component of a good life, that aspect which represents a person's subjective, conscious response in terms of enjoyments and satisfactions to the life he or she has chosen and the activities and achievements it contains. These may be representable on a single vector or on a number of distinct vectors if the person has distinct and incommensurable satisfactions and enjoyments. Happiness will usually be only partially dependent on the person's relative success in satisfying his or her desires and broader aims and projects. Once again, it is the valuations of the specific person in question that will determine the relative weight the happiness vector receives in the overall account of a good life for that person.

Needless to say, in drawing together these features of an account of the quality of life or of a good life from the medical ethics and health policy literatures, I have done no more than sketch a few of the barest bones of a full account of a good life. However, even these few bones suggest the need for more complex accounts of the quality of life than are often employed in programmes designed to improve the quality of life of real people.

**Bibliography**

Anderson, John P., *et al.* (1986). 'Classifying Function for Health Outcome and Quality-of-Life Evaluation', *Medical Care*, 24, 454-71.

Annas, George, and Glantz, Leonard (1986). 'The Right of Elderly Patients to Refuse Life-Sustaining Treatment', *Milbank Quarterly*, vol. 64, suppl. 2 95-162.

Arras, John (1984). 'Toward an Ethic of Ambiguity', *Hastings Center Report*, 14(Apr.), 25-33.

Beauchamp, Tom L., and Childress, James F. (1979). *Principles of Biomedical Ethics*. New York: Oxford University Press.
Berg, Robert L. (1986). 'Neglected Aspects of the Quality of Life', *Health Services Research*, *21*, 391-5.
Bergner, Marilyn, *et al.* (1976). 'The Sickness Impact Profile: Conceptual Formulation and Methodology for the Development of a Health Status Measure', *International Journal of Health Services*, 6, 393-415.
—— (1981). 'The Sickness Impact Profile: Development and Final Revision of a Health Status Measure', *Medical Care*, 19, 787-805.
Boorse, Christopher (1975). 'On the Distinction between Disease and Illness', *Philosophy and Public Affairs*, 5, 49-68.
—— (1977). 'Health as a Theoretical Concept', *Philosophy of Science*, 44.
Breslow, Lester (1972). 'A Quantitative Approach to the World Health Organization Definition of Health: Physical, Mental and Social Well-being', *International Journal of Epidemiology*, 1, 347-55.
end p.128
Brock, Dan W. (1973). 'Recent Work in Utilitarianism', *American Philosophical Quarterly*, 10, 241-76.
—— (1983). 'Paternalism and Promoting the Good', in Rolf Sartorius (ed.), *Paternalism*. Minneapolis: University of Minnesota Press.
—— (1986). 'The Value of Prolonging Human Life', *Philosophical Studies*, 50, 401-28.
—— (1987). 'Informed Consent', in Tom Regan and Donald VanDeVeer (eds.), *Health Care Ethics*. Philadelphia: Temple University Press.
—— (1988). 'Paternalism and Autonomy', *Ethics*, 98, 550-65.
Buchanan, Allen (1988). 'Advance Directives and the Personal Identity Problem', *Philosophy and Public Affairs* 17, 277-302.
—— and Brock, Dan W. (1986). 'Deciding for Others', *Milbank Quarterly*, vol. 64, suppl. 2 17-94.
—— —— (1989). *Deciding For Others*. Cambridge: Cambridge University Press.
Callahan, Daniel (1987). *Setting Limits*. New York: Simon and Schuster.
Calman, K. C. (1984). 'Quality of Life in Cancer Patients: An Hypothesis', *Journal of Medical Ethics*, 10, 124-7.
Caplan, Arthur L., *et al.* (1981). *Concepts of Health and Disease: Interdisciplinary Perspectives*. Reading, Mass.: Addison-Wesley.
Chen, Milton M., Bush, J. W., and Patrick, Donald L. (1975). 'Social Indicators for Health Planning and Policy Analysis', *Policy Sciences*, 6, 71-89.
Cohen, Carl (1982). 'On the Quality of Life: Some Philosophical Reflections', *Circulation*, vol. 66 suppl. 3 29-33.
Cribb, Alan (1985). 'Quality of Life: A Response to K. C. Calman', *Journal of Medical Ethics*, 11, 142-5.
Daniels, Norman (1985). *Just Health Care*. Cambridge: Cambridge University Press.
Dresser, Rebecca (1986). 'Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law', *Arizona Law Review*, 28, 373-405.
Editorial (1986). 'Assessment of Quality of Life in Clinical Trials', *Acta Medica Scandia*, 220, 1-3.
Edlund, Mathew, and Tancredi, Lawrence (1985). 'Quality of Life: An Ideological Critique', *Perspectives in Biology and Medicine*, 28, 591-607.

Elster, Jon (1982). 'Sour Grapes: Utilitarianism and the Genesis of Wants', in Amartya Sen and Bernard, Williams (eds.), *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.
Evans, Roger W. (1985). 'The Quality of Life of Patients with End Stage Renal Disease', *New England Journal of Medicine*, 312, 553-9.
Faden, Ruth R., and Beauchamp, Tom L. (1986). *A History and Theory of Informed Consent*. New York: Oxford University Press.
Feinberg, Joel (1986). *Harm to Self*. New York: Oxford University Press.
Flanagan, John C. (1982). 'Measurement of Quality of Life: Current State of the Art', *Archives of Physical Rehabilitation Medicine*, 63, 56-9.
Frankfurt, Harry (1971). 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy*, 68, 5-20.
Fried, Charles (1970). *An Anatomy of Values*. Cambridge, Mass.: Harvard University Press.

end p.129

Gehrmann, Friedhelm (1978). ' "Valid" Empirical Measurement of Quality of Life', *Social Indicators Research*, 5, 73-109.
Gillingham, Robert, and Reece, William S. (1980). 'Analytical Problems in the Measurement of the Quality of Life', *Social Indicators Research*, 7, 91-101.
Goodin, Robert (1986). 'Laundering Preferences', in Jon Elster and Aanund Hylland (eds.), *Foundations of Social Choice Theory*. Cambridge: Cambridge University Press.
Griffin, James (1986). *Well-Being*. Oxford: Oxford University Press.
Grogono, A. W. (1971). 'Index for Measuring Health', *Lancet*, vol. 2 for 1971, 1024-6.
Guyatt, Gordon H., *et al.* (1986). 'Measuring Disease-Specific Quality of Life in Clinical Trials', *Canadian Medical Association Journal*, 134, 889-95.
Hastings Center (1987). *Guidelines on the Termination of Treatment and the Care of the Dying*. Briarcliff Manor, NY: The Hastings Center.
Hunt, Sonya, and McEwen, James (1980). 'The Development of a Subjective Health Indicator', *Sociology of Health and Illness*, 2, 203-31.
Jonsen, Albert R., Siegler, Mark, and Winslade, William J. (1982). *Clinical Ethics*. New York: Macmillan.
Kass, Leon (1985). *Toward a More Natural Science*. New York: Free Press.
Katz, Jay (1984). *The Silent World of Doctor and Patient*. New York: Free Press.
Katz, Sidney, *et al.* (1963). 'Studies of Illness in the Aged. The Index of ADL: A Standardized Measure of Biological and Psychosocial Function', *JAMA*, 185, 914-19.
Klotkem, Frederic J. (1982). 'Philosophic Considerations of Quality of Life for the Disabled', *Archives of Physical Rehabilitation Medicine*, 63, 59-63.
Kornfeld, Donald S., *et al.* (1982). 'Psychological and Behavioral Responses After Coronary Artery Bypass Surgery', *Circulation*, vol. 66, suppl. 3 24-8.
Liang, Mathew *et al.* (1982). 'In Search of a More Perfect Mousetrap (Health Status or Quality of Life Instrument)', *Journal of Rheumatology*, 9, 775-9.
McCormack, S. J., Richard J. (1974). 'To Save or Let Die: The Dilemma of Modern Medicine', *JAMA*, 229, 172-6.
Meisel, Alan (1979). 'The "Exceptions" to the Informed Consent Doctrine: Striking a Balance Between Competing Values in Medical Decision-making', *Wisconsin Law Review*, 413-88.
Mill, J. S. (1859). *On Liberty*. Indiana/New York: Bobbs-Merrill, 1956.
Morreim, E. Haavi (1986). 'Computing the Quality of Life', in G. J. Agich and C. E. Begley (eds.), *The Price of Health*. Dordrecht: D. Reidel.

Najman, Jackob, and Levine, Sol (1981). 'Evaluating the Impact of Medical Care and Technologies on the Quality of Life: A Review and Critique', *Social Science and Medicine*, 15F, 107-15.
*Natanson* v. *Kline* (1960) 186 Kan. 393, 350 p. 2d 1093.
Parfit, Derek (1984). *Reason and Persons*. Oxford: Oxford University Press.
Pearlman, Robert, and Speer, James (1983). 'Quality of Life Considerations in Geriatric Care', *Journal of the American Geriatrics Society*, 31, 113-20.
Presant, Cary A. (1984). 'Quality of Life in Cancer Patients', *American Journal of Clinical Oncology*, 7, 571-3.
President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1982). *Making Health Care Decisions*. Washington: US Government Printing Office.
—— (1983). *Deciding to Forgo Life-Sustaining Treatment*. Washington: US Government Printing Office.

end p.130

Rachels, James (1986). *The End of Life*. Oxford: Oxford University Press.
Ramsey, Paul (1978). *Ethics at the Edge of Life*. New Haven, Conn.: Yale University Press.
Rango, Nicholas (1985). 'The Nursing Home Resident with Dementia', *Annals of Internal Medicine*, 102,m 835-41.
Rawls, John (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
—— (1980). 'Kantian Constructivism in Moral Theory', *Journal of Philosophy*, 77, 515-72.
Report (1984). 'The 1984 Report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure', *Archives of Internal Medicine*, 144, 1045-57.
Rhoden, Nancy K. (1985). 'Treatment Dilemmas for Imperiled Newborns: Why Quality of Life Counts', *Southern California Law Review*, 58, 1283-347.
Ruark, John E., *et al.* (1988). 'Initiating and Withdrawing Life Support', *New England Journal of Medicine* 318 (Jan.), 25-30.
*Schloendorf* v. *Society of New York Hospital* (1914). 211 N.Y. 125, 105 N.E. 92, 95.
Schneiderman, Lawrence, and Arras, John (1985). 'Counseling Patients to Counsel Physicians on Future Care in the Event of Patients Incompetence', *Annals of Internal Medicine*, 102, 693-8.
Sen, Amartya (1980). 'Plural Utility', *Proceedings of the Aristotelian Society*, 81, 193-218.
—— (1985*a*). *Commodities and Capabilities*. Amsterdam: Elsevier Science Publishers.
—— (1985*b*). 'Well-being, Agency and Freedom: The Dewey Lectures', *Journal of Philosophy*, 82, 169-221.
—— (1987). *The Standard of Living*. Cambridge: Cambridge University Press.
Singer, Peter, and Kuhse, Helga (1985). *Should This Baby Live?* Oxford: Oxford University Press.
Spitzer, Walter O., *et al.* (1981). 'Measuring the Quality of Life of Cancer Patients: A Concise QL-Index for Use by Physicians', *Journal of Chronic Disease*, 34, 585-97.
Starr, T. Jolene, *et al.* (1986). 'Quality of Life and Resuscitation Decisions in Elderly Patients', *Journal of General Internal Medicine*, 1, 373-9.

Steinbrook, Robert, and Lo, Bernard (1984). 'Decision-making for Incompetent Patients by Designated Proxy: California's New Law', *New England Journal of Medicine*, 310, 1598-601.
Sullivan, Daniel F. (1966). 'Conceptual Problems in Developing an Index of Health', *Vital and Health Statistics*, 2, 1-18.
*Superintendant of Belchertown State School* v. *Saikewicz* (1977). 370 N.E. 2d 417.
Thomasma, David C. (1984). 'Ethical Judgment of Quality of Life in the Care of the Aged', *Journal of the American Geriatrics Society*, 32, 525-7.
—— (1986). 'Quality of Life, Treatment Decisions, and Medical Ethics', *Clinics in Geriatric Medicine*, 2, 17-27.
Tooley, Michael (1983). *Abortion and Infanticide*. Oxford: Oxford University Press.
Torrance, George W. (1972). 'Social Preferences for Health States: An Empirical Evaluation of Three Measurement Techniques', *Socio-Economic Planning Sciences*, 10, 129-36.
—— (1976) 'Toward A Utility Theory Foundation for Health Status Index Models', *Health Services Research*, 10, 129-36.

end p.131

—— *et al.* (1972). 'A Utility Maximization Model for Evaluation of Health Care Programs', *Health Services Research*, 6, 118-33.
VanDeVeer, Donald (1986). *Paternalistic Intervention*. Princeton, NJ: Princeton University Press.
Wanzer, Sidney H., *et al.* (1984). 'The Physicians' Responsibility Toward Hopelessly Ill Patients', *New England Journal of Medicine*, 310 (Apr.), 955-9.

end p.132

# Dan Brock: Quality of Life Measures in Health Care and Medical Ethics

show chapter abstract and keywords
hide chapter abstract and keywords

**James Griffin**

Professor Brock asks, What does the literature of medical ethics have to tell us about the quality of life? He answers, Not much. That is not surprising. What makes up the quality of a life? It is a tough question. Doctors' dilemmas make it urgent, but not easier. And medical practice cannot do without settling at least a bit of theory. Medical administrators need to know to what extent the values that go to make up the quality of life are commensurable, and how commensuration in their case works, and how interpersonal comparisons work too.

Can these administrators get by with only a fairly narrow conception of well-being? Well, the aim of medicine is health, and it would be wrong, I think, to let the notion of 'health' swell to include everything that bears on the quality of life. Which patients should get scarce treatment? Should we do fewer expensive heart transplants and more inexpensive hip replacements? When should we pull the plug? Medical decisions — at any rate, decisions about when, and how far, and in what way, and for whom, to pursue health — need to use much wider normative notions than that of health alone. Brock says: 'my concern will be with the broadest conception of . . . "what makes a life go best" '. To my mind, he is right to be concerned with the broadest conception. But perhaps we should also be concerned with narrower conceptions. Perhaps the notion of the quality of life fragments into several notions appropriate to different sorts of social decision, even decisions about health. It is wrong of us, I think, to expect one notion of the quality of life to be up to answering all the questions that we try to answer with it. I shall come back to this shortly.

Brock thinks that we can extract from the literature on medical ethics 'the main outlines of a general account of a good life'. I think that the literature suggests his account no more than it suggests any one of many others. Still, his account is an interesting possibility. What I want to do is to express doubts about it.

Brock divides theories of a good life into three: (1) hedonist, (2) preference satisfaction or desire fulfilment, and (3) ideal. He does not think that we should force a choice between them but, rather, give independent place to items from each, as (he says) ideal theories can do. So Brock's is a kind of ideal theory. It can also be seen as a 'vector view' of the sort that Amartya Sen anatomized in his article 'Plural Utility'. [1]

[1] Sen, 1980-1.

Brock goes on to identify four 'components' of a good life:

end p.133

1. primary functions (such as mobility and communication, which are used in carrying out just about any life plan);
2. agent-specific functions (ones that are used in carrying out one agent's life plan in particular);
3. desire fulfilment, and
4. happiness.

Then, in a particularly fundamental place, there will be autonomy.

Now these items are very different in category. Happiness and autonomy *are* substantive prudential values. Functions, as Brock defines them, however, are not; they are characteristic means to realizing such values. And desire fulfilment, I should say, is not a substantive prudential value at all but rather a meta-ethical feature of those values. So, Brock's scheme does not touch rock bottom (that is, intrinsic prudential values), and I think, for reasons that I shall come to, that we ought to touch ground before taking off again.

Take desire fulfilment. Brock seems to me to treat it as a prudential value. At least, he speaks at one point of autonomy's being perhaps only instrumentally valuable, valuable because it leads either to happiness or to desire fulfilment. But I think that desire fulfilment is at some remove from prudential values. The desire fulfilment account of value, as Brock points out, has to be concerned with desires that are corrected in some way, because the fulfilment of actual desires comes nowhere near being a plausible candidate as a prudential value. But once we start talking about *corrected* desires, once we ask how stiff a demand 'corrected' represents, we move into territory where old maps do not help a lot. Now, how stiff a requirement should it represent? The rough idea of a corrected desire is one formed by an appropriate appreciation of the nature of its object, so it has to include whatever is necessary for that. 'Corrected' might be taken (as, for example, Richard Brandt takes it [2]

[2] Brandt, 1979: 10, but also chs. 2-7 *passim*.

) to require only that desires survive criticism by facts and logic, where it is meant to be a question of fact — largely of psychological fact — whether a desire is corrected. But a particularly irrational desire — say, one planted deep when one was young — might well survive criticism by facts and logic, and its mere persistence does not seem to guarantee that its fulfilment will make one better off. For instance, I might wish to hog the limelight on all occasions. I might have learned from long experience that succeeding does not work to my advantage, but still want to. I might not react appropriately, or strongly enough, to this important piece of self-knowledge. So

'corrected' cannot mean just 'formed while possessed of factual knowledge and correct logic'; it must mean something stronger such as 'formed in proper appreciation of the nature of the object'.
There are appropriate responses to certain pieces of knowledge, and 'appropriate' cannot here be just 'most common' because most of us sometimes go on wanting certain things — self-assertion, say — too much even when our desires



are formed while possessed of factual knowledge and correct logic. 'Appropriate' has to mean something close to 'correct'. But this shifts importance away from the mere presence of a desire to the proper appreciation of its object. It gives us, I should say, some sort of ideal theory of prudential value. In that I agree with Brock. Where I disagree is with his regarding desire fulfilment as itself one of the values. Desire fulfilment seems to me best seen not as a substantive value, but as a formal notion of what it is for something to be prudentially valuable: it is for it to be (not entirely trivially) the object of a sufficiently well informed desire. A person can move from a desire's being unfulfilled to its being fulfilled and become a lot worse off as a result. He can make the same move, even with a fully enlightened desire, and be no better off. It is not the state of a desire's being fulfilled that is valuable. What are valuable are certain objects of desire. True, one can see various things that might make it *seem* that desire fulfilment is itself a prudential value — the frustration, for instance, that sometimes comes from non-fulfilment, the boost that can, although need not, come with fulfilment, and the autonomy that we respect in respecting the fulfilment of a person's *actual* (not *corrected*) desires. And frustrations, boosts, and autonomy all have to do with substantive values. Desire fulfilment counts towards the quality of life, not as such, but because when suitably corrected and restricted, desires are linked to objects subsumable under some desirability characteristic (autonomy, enjoyment, deep personal relations, accomplishing something with one's life, etc.). This, of course, changes the face of desire accounts a lot. But this is what they have to become for there to be a plausible link between fulfilment of desire and prudential value. I must apologize for whizzing through, and dogmatizing about, immensely difficult matters. I have, however, discussed much of this elsewhere.[3]

[3] In Griffin, 1986: Part I; also, more recently, in Griffin, forthcoming *a*; Griffin, forthcoming *b*.

Take another item from Brock's list of the 'components' of a good life: functions. What Brock refers to as 'functions' are not, I take it, the same as what Sen calls 'functionings' or, as Sen also puts it, achievement of 'beings' and 'doings'.[4]

[4] See e.g. Sen, 1985, esp. chs. 2 and 4; Sen, 1987*a*, esp. ch. 2; Sen, 1987*b*, esp. lect. II.

Brock, I think, is interested in the *possession* of certain pretty *basic* abilities, while Sen, though his list includes those, wants to include much more besides — for example, the *exercise* of decidedly *non-basic* abilities. So Sen's category is so wide that it includes both means to substantive goods and the substantive goods themselves.
Now, what one includes in an account of a good life will, I think, depend upon what the notion of a good life is used for. We need a full notion in taking decisions about how we want to live our own lives, and a doctor probably needs the same full notion in taking decisions about what the best trade-offs are for a particular patient. It is, I think, a different, perhaps narrower, notion of a good life that comes into play in many of a government's decisions



about how it ought to allocate resources. And it is not a notion of a good life at all, but some accessible, tolerably reliable indicator of it, that we often have to fall back on. What Brock calls primary functions, for instance, are so central to most people's lives

that it is entirely reasonable for them to feature in 'Sickness Impact Profiles' (Table 1). Even though they are just means to prudential values, they are accessible and measurable, and the link is close. So it makes sense to give them prominence in measures of quality of life that we must carry out in everyday settings. Still, they are at some remove from actual prudential values. The same handicap — say, the loss of a finger — can devastate some lives but have only minor cosmetic disadvantages in others. Medical decisions have often to be tailored to a particular patient; often it is only having seen the values at stake in *that* life that we can decide what to sacrifice for what. We have often got to get behind functions to the real prudential values that they are means to.

Suppose, then, we were to take not 'functions', as Brock uses the term, but the much more capacious class of 'functionings', as Sen uses the term. A large part of the interest in both functions and functionings stems from the belief that the value space cannot be identified with either material or social goods (e.g. Rawls's primary goods) or subjective responses to them (e.g. the classical utilitarians' utility), but is to be found somewhere between them. That seems to me right, but it is nothing new. Though classical utilitarians often wrote as if they saw utility in terms of mental states, at other times they suggested something quite different (recall Mill on qualitative differences in utility: the higher one is that to which '*all* or *almost all* who have experience of both give a decided *preference*';[5]

[5] Mill, 1863: ch. 2.

and recall the example of the informed preference for Socrates dissatisfied — short on pleasures, in the ordinary sense of the term, and short on desire fulfilment — over the Fool satisfied; the values referred to here are far from subjective responses and close to, perhaps identical to, beings and doings). And certain preference utilitarians long ago moved utility into the realm of beings and doings.[6]

[6] This is obscured in Sen's writings by his generally construing 'utility' in quite a narrow way; he sees 'utility' as committed to a 'metric of happiness or desire-fulfilment', which (given his narrow interpretation of these two notions) he rightly says has obvious limitations as an explanation of human well-being (Sen, 1987*a*: 45). It is not that Sen is unaware of the great variety of possible, or actual past, uses of 'utility' — that 'versatile name', as he remarks (Sen, 1985: 17); far from it. But he thinks that happiness and desire fulfilment are 'the traditional meanings of utility' (Sen, 1985: 3), and he generally uses the term in that traditional sense. Still, his objections to utility as an interpretation of well-being show just how narrowly he uses 'happiness' and 'desire fulfilment'. People often become reconciled to deprivation, he rightly observes; they stop minding or hoping. So using either happiness or desire fulfilment as a metric may distort how well off a person actually is. That seems to me right and important, but the point I want to make here is that this works as an objection only if one uses especially narrow interpretations. Sen even seems to confine desires to *actual* desires, ignoring the tradition of corrected or rational or informed desires, and certainly ignoring any especially strong standard for a desire's being informed. See Sen, 1985: 20-2, 52-3; Sen, 1987*a*: 45-6; Sen 1987*b*: 7-12.

With their notion of a 'rational' desire, they took the desire fulfilment account of utility well down the road towards, if



not quite all the way to, the list of prudential goods that seems to me the inevitable destination of a fully worked out desire account.[7]

[7] See e.g. Brandt, 1979, esp. ch. 6.

Then there was G.E. Moore's ideal version of utilitarianism as early as 1903. Prudential values can indeed be seen as beings and doings. Even 'happiness', in the sense most relevant to the quality of life, is probably best understood, for reasons that Aristotle gives, as falling into the category of *activity*, and so is caught by 'doings'. The talk of beings and doings leaves us well short of a delineation of value

space, because those terms include vastly more than just values. We need a *differentia*. Without it, the job of locating the value space is virtually undone.
For Sen, nourishment and health are paradigmatic functionings.[8]
[8] See the list of functions in Sen, 1985: 46.
But neither is itself a prudential value. It would make perfectly good sense to choose to be, at least for a while, undernourished through fasting, if one's spiritual life were thereby enhanced. And I do not think that it is just the presence of choice here that gives undernourishment through fasting a different moral status from undernourishment from starvation; to think so would be to overlook what makes sense of the choice. What makes sense of it is that fasting (and the discomforts of hunger that go with it) is to be seen as part of a process of reaching a larger good — say, spiritual growth. That is why the case of a child in a family that fasts who does not himself choose to fast is morally different from starvation too. Discomfort is small stuff compared to a spiritual life. We understand these cases best, I think, in terms of the various substantive values in play. It might also make good sense to accept some avoidable forms of ill health. It would make good sense not to spend long periods exercising one's legs to combat a wasting disease, if ease of walking had little relevance to the quality of one's particular life, while using the exercise time instead for one's work did have. Nourishment and health are both at some remove from prudential values. Unless we acknowledge that, we cannot explain the rationality of much prudential deliberation.
Why does this remoteness from real prudential values matter? Why does it matter whether we deliberate in terms of functions, desire fulfilment, and subjective reactions or in terms of what seem to me to be prudential values: enjoyment, accomplishment, deep personal relations, the elements of human dignity (autonomy, liberty), and so on? It seems to me to matter because we have to get a sense of the whole deliberative project that we are engaged in. We have to see how to compare one value with another, and one person's values with another's, and how far commensurability and comparability go. And we shall not find out — we shall not get even our crudest bearings — until we see what is prudentially valuable and what makes it valuable. Brock speaks of the vector view. But choosing vectors that are at some remove from actual prudential values does not allow us to get at how the vectoring actually works. For example, how important is *this* primary function (say, legs that work) to



*this* person (say, Itzhak Perlman)? What is the significance of *this* person's contentment with a disability? Contentment with one's lot, as Brock notes, can variously be both a desirable constituent of the good life and an undesirable obstacle to a better. We cannot proceed, I think, until we have got the right materials to work with, and that means going deeper to the prudential values at stake.
This is not an objection to the vector view. On the contrary, as a view about the broad notion of the quality of life (as opposed to a moral view, which is an entirely different matter), some form of it must be right. There may be several irreducibly different prudential values. They may have different weights in different cases; they may vector into a single direction. But all of this just mentions possibilities. We need to see how it actually works. Well, suppose we use the vectors 'desire fulfilment' and 'subjective reaction'. Let me take an example of vectoring from Sen's paper 'Plural Utility'.[9]
[9] Sen, 1980-1: 203-4.

Suppose someone prefers bitter truth to comforting delusion but is palmed off with the latter. We feel sorry for him. Then later he learns that he was deceived. We feel sorrier for him. Sen says that bare desire fulfilment and experienced desire fulfilment (that is, subjective reaction) are both relevant, and that to insist on a choice between them seems 'arbitrary and uncalled for'. [10]

[10] Ibid., 203.

But the trouble with leaving it at that is that at least sometimes we have to supply some weighting for various vectors, and the weighting ought to be neither arbitrary nor left to haphazard intuition. I doubt that the categories 'bare desire fulfilment' and 'experienced desire fulfilment' will help us get at the weightings that we are after. We are after a comparison of *values*, and neither 'bare desire fulfilment' not 'experienced desire fulfilment' even names a value.

Suppose, then, that we use instead actual prudential values as our vectors. One range of prudential values, I think (and shall have just to go on dogmatizing), comprises the elements of human dignity — that is, such things as our being able autonomously to choose a course through life and being at liberty to follow it. We also value pleasure and the avoidance of its opposites. With these in focus, with some understanding of their bounds and grounds, we can hope to make progress towards weightings. We value our liberty to carry out our plans, greatly in the case of our most central plans and less when the plans are more peripheral. Following these lines, we might hope eventually to explain how much weight to attach to the desire to live out the most central features of one's life plan, and how this weight compares with the upset and distress it might cause others. We can hope also to get an explanation of why liberty has different weights on different occasions — though the desire to carry out the central parts of our life plan generally outweigh our neighbours' huffiness over our doing so, the desire to swim nude, being not all that central, might not outweight the upset *it* might cause.

In arguing for the need to identify actual prudential values, I am not claiming

end p.138

that their identification is exclusively, or even supremely, important. We need the broad conception of the quality of life (that is, the list of prudential values) for our reasoning about how to make our own lives go best. Doctors need it for decisions about certain patients. But there are many reasons, both moral and practical, to work with a narrower conception of the quality of life in taking certain social decisions. And it may be (I strongly suspect it is) that we need several different conceptions of the quality of life for different sorts of social decision. That is, we need both the broad conception and also an understanding of the various considerations at work in generating narrower conceptions.

**Bibliography**

Brandt, R. B. (1979). *A Theory of the Good and the Right*. Oxford: Clarendon Press.

Griffin, James (1986). *Well-Being*. Oxford: Clarendon Press.

——(forthcoming *a*). 'Against the Taste Model', in J. Elster and J. Roemer (eds.), *Interpersonal Comparisons of Well-Being*. Cambridge: Cambridge University Press.

——(forthcoming *b*). 'Value: Reduction, Supervenience, and Explanation by Ascent', in K. Lennon and D. Charles (eds.), *Reduction, Explanation, and Realism*. Oxford: Clarendon Press.

Mill, J. S. (1863). *Utilitarianism*. London.

Sen, Amartya (1980-1). 'Plural Utility', *Proceedings of the Aristotelian Society*, 81.

——(1985). *Commodities and Capabilities*. Amsterdam: North-Holland.

——(1987*a*). *On Ethics and Economics*. Oxford: Blackwell.

——(1987*b*). *The Standard of Living*. Cambridge: Cambridge University Press.

end p.139

**Part II Traditions, Relativism, and Objectivity**

end p.141
end p.142

# Objectivity and the Science—Ethics Distinction

show chapter abstract and keywords
hide chapter abstract and keywords

**Hilary Putnam**

## 1 The Fact-Value Dichotomy: Background

The Logical Positivists argued for a sharp fact-value dichotomy in a very simple way: scientific statements (outside of logic and pure mathematics), they said, are 'empirically verifiable' and value judgements are 'unverifiable'. This argument continues to have wide appeal to economists (not to say laymen), even though it has for some years been looked upon as naïve by philosophers. One reason that the argument is naïve is that it assumes that there is such a thing as 'the method of verification' of each isolated *scientifically meaningful* sentence. But this is very far from being the case. Newton's entire theory of gravity, for example, does not *in and of itself* (i.e. in the absence of suitable 'auxiliary hypotheses') imply any testable predictions whatsoever (cf. Putnam, 1974). As Quine has emphasized (Quine, 1951), reviving arguments earlier used by Duhem, scientific statements 'meet the test of experience as a corporate body'; the idea that each scientific sentence has its own range of confirming observations and its own range of disconfirming observations, independent of what other sentences it is conjoined to, is wrong. If a sentence that does not, in and of itself, by its very meaning, have a 'method of verification' is meaningless, then most of theoretical science turns out to be meaningless!

A second feature of the view that 'ethical sentences are cognitively meaningless because they have no method of verification' is that even if it had been correct, what it would have drawn would not have been a *fact-value* dichotomy. For, according to the positivists themselves, metaphysical sentences are cognitively meaningless for the same reason as ethical sentences: they are 'unverifiable in principle'. (So are poetic sentences, among others.) The positivist position is well summarized by Vivian Walsh (Walsh, 1987):

Consider the 'putative' proposition 'murder is wrong'. What empirical findings, the positivists would ask, tend to confirm or disconfirm this? If saying that murder is wrong is merely a misleading way of reporting what a given society believes, this is a perfectly good sociological fact, and the proposition is a respectable empirical one. But the person making a moral judgement will not accept this analysis. Positivists then wielded their absolute analytic/synthetic distinction: if 'murder is wrong' is not a synthetic (empirically testable) proposition it must be an analytic proposition, like (they believed) those of logic and mathematics — in effect, a tautology. The person who wished to make the moral judgement would not accept this, and was told that the disputed utterance was a 'pseudo-proposition' like those of poets, theologians and metaphysicians.

As Walsh goes on to explain, by the end of the fifties 'most of the theses

end p.143

necessary for this remarkable claim' had been abandoned. The positivist theory of 'cognitive significance' had fallen. The absolute analytic-synthetic distinction was seen to fail as an account of how scientific theories are actually put together. Writing in a volume honouring Carnap (Quine, 1963), Quine summed up its demise, writing, 'the lore of our fathers is black with fact and white with convention, but there are no

*completely* white threads and no quite black ones.' Explaining the impact of all this, Walsh writes:

Another retreat, forced upon logical empiricism by the *needs* of pure science, opened the way for a further rehabilitation of moral philosophy. The old positivist attack on the status of moral judgements had required the claim that each *single* proposition must, at least in principle, be open to test. It became evident that many of the propositions of which the higher theory of pure science are composed could not survive this demand. Theoretical propositions, the logical empiricists decided, became 'indirectly' meaningful if part of a theory which possessed (supposed) observation statements which had empirical confirmation to some degree (never mind that the theoretical statement/observation statement dichotomy itself broke down!); but the clear fact/value distinction of the early positivists depended upon being able to see if *each single* proposition passed muster. To borrow and adapt Quine's vivid image, if a theory may be black with fact and white with convention, it might well (as far as logical empiricism could tell) be red with values. Since for them confirmation *or* falsification had to be a property of a theory *as a whole*, they had no way of unraveling this whole cloth. Yet even today economists whose philosophical ancestry is logical empiricism still write as if the old positivist fact/value dichotomy were beyond challenge.

The collapse of the grounds on which the dichotomy was defended during the period Walsh is describing has not, however, led to a demise of the dichotomy, even among professional philosophers. What it has led to is a change in the nature of the *arguments* offered for the dichotomy. Today, it is defended more and more on metaphysical grounds. At the same time, even the defenders of the dichotomy concede that the old arguments for the dichotomy were bad arguments. For example, when I was a graduate student, a paradigmatic explanation and defence of the dichotomy would have been Charles Stevenson's. I attacked Stevenson's position at length in a book published some years ago (Putnam, 1981). When Bernard Williams's last book (Williams, 1985) appeared, I found that Williams gave virtually the same arguments against this position. Yet Williams still defends a sharp 'science-ethics' dichotomy; and he regards his science-ethics dichotomy as capturing something that was essentially right about the old 'fact-value' dichotomy.

Something else has accompanied this change in the way the dichotomy is defended. The old position, in its several versions — emotivism, voluntarism, prescriptivism — was usually referred to as 'non-cognitivism'. 'Non-cognitivism' was, so to speak, the generic name of the position, and the more specific labels were the proprietary names given the position by the various distributors. And the generic name was appropriate, because all the various slightly different formulations of the generic product had this essential ingredient in common:

end p.144

ethical sentences were 'non-cognitive', that is to say, they were neither true nor false. Today, philosophers like Williams [1]

[1] The philosopher whose views are closest to Williams is, perhaps, David Wiggins (cf. Wiggins, 1987).

do not deny that ethical sentences can be true or false; what they deny is that they can be true or false *non-perspectivally*. Thus, the position has been (appropriately) renamed: while the proprietary versions of the new improved drug still have various differences one from the other, they all accept the name Relativism. *Non-cognitivism has been rebaptized as Relativism*.

## 2 The Entanglement of Fact and Value

Just why and how non-cognitivism has given way to relativism is a complicated question, and it is not the purpose of this paper to explore it in detail. But one reason is surely an increased appreciation of what might be called the *entanglement* of fact and value. That entanglement was a constant theme in John Dewey's writing. But this aspect of pragmatism was neglected in Anglo-American philosophy after Dewey's death, in spite of Morton White's valiant effort to keep it alive (White, 1956), and it was, perhaps, Iris Murdoch who reopened the theme in a very different way. Murdoch's three essays, published together in Murdoch, 1971, contain a large number of valuable insights and remarks: two have proved especially influential. Murdoch was the first to emphasize that languages have two very different sorts of ethical concepts: abstract ethical concepts (Williams calls them 'thin' ethical concepts), such as 'good', and 'right', and more descriptive, less abstract concepts (Williams calls them 'thick' ethical concepts) such as, for example, *cruel, pert, inconsiderate, chaste*. Murdoch (and later, and in a more spelled-out way, McDowell (1978 and 1979)) argued that there is no way of saying what the 'descriptive component' of the meaning of a word like 'cruel' or 'inconsiderate' is without using a word of the same kind; as McDowell put the argument, a word has to be connected to a certain set of 'evaluative interests' in order to function the way such a thick ethical word functions; and the speaker has to be aware of those interests and be able to identify imaginatively with them if he is to apply the word to novel cases or circumstances in the way a sophisticated speaker of the language would. The attempt of non-cognitivists to split such words into a 'descriptive meaning component' and a 'prescriptive meaning component' founders on the impossibility of saying what the 'descriptive meaning' of, say, 'cruel' is without using the word 'cruel' itself, or a synonym. Secondly, Murdoch emphasized that when we are actually confronted with situations requiring ethical evaluation, whether or not they also require some action on our part, the sorts of descriptions that we need—descriptions of the motives and character of human beings, above all—are

end p.145

descriptions in the language of a 'sensitive novelist', not in scientistic or bureaucratic jargon. When a situation or a person or a motive is appropriately described, the decision as to whether something is 'good' or 'bad' or 'right' or 'wrong' frequently follows automatically. For example, our evaluation of a person's moral stature may critically depend on whether we describe her as 'impertinent' or 'unstuffy'. Our Life-world, Murdoch is telling us, does not factor neatly into 'facts' and 'values'; we live in a messy human world in which seeing reality with all its nuances, seeing it as George Eliot, or Flaubert, or Henry James, or Murdoch herself can, to some extent, teach us to see it, and making appropriate 'value judgments' are simply not separable abilities. I confess that when I read *The Sovereignty of 'Good'* I thought that Murdoch gave a perceptive description of the sphere of private morality (which is, of course, the sphere with which a novelist has to deal), but that she too much ignored the public sphere, the sphere in which issues of social justice arise and must be worked out. But more recently I have come to think that a similar entanglement of the factual and the ethical applies to this sphere as well. It is all well and good to describe hypothetical cases in which two people 'agree on the facts and disagree about values', but in the world in which I grew up such cases are unreal. When and where did a Nazi and an anti-Nazi, a communist and a social democrat, a fundamentalist and a liberal, or even a Republican and a Democrat, agree on the facts? Even when it comes to one specific policy question, say, what to do about the decline of American education, or about unemployment, or about drugs, every argument I have

ever heard has exemplified the entanglement of the ethical and the factual. There is a weird discrepancy between the way philosophers who subscribe to a sharp fact-value distinction *make* ethical arguments sound and the way ethical arguments *actually* sound. (Stanley Cavell once remarked (Cavell, 1979) that Stevenson writes like someone who has *forgotten* what ethical discussion is like.)

# 3 Relativism and the Fact-Value Dichotomy

According to Bernard Williams, a properly worked-out relativism can do justice to the way in which fact and value can be inseparable; do justice to the way in which some statements which are both descriptive and true ('Caligula was a mad tyrant') can also be value judgements. The idea is to replace the fact-value distinction by a very different distinction, the distinction between *truth* and *absoluteness*.

Although Williams does not explain what he understands truth to be very clearly, he seems to thinks truth is something like right assertability in the local language game; that is, if the practices and shared values of a culture determine an established use for a word like 'chaste'—a use which is sufficiently definite to permit speakers to come to agreement on someone's chastity or lack of chastity (or whatever the example of a 'thick ethical concept' may be)—then

end p.146

it can be simply true that a person in the culture is 'chaste' (or 'cruel', or 'pious', or whatever). Of course, if I do not belong to the culture in question and do not share the relevant evaluative interests, then I will not describe the person in question as 'chaste', even if I know that that is a correct thing to say in that culture; I will be 'disbarred' from using the word, as Williams puts it. As he also puts it (with deliberate paradox), that so-and-so is chaste is possible *knowledge* for someone in the culture, but not possible knowledge for *me*.

If truth were the only dimension with respect to which we could evaluate the cognitive credentials of statements, then Williams would be committed to ethical realism or at least to the rejection of ethical *anti*-realism. For, on his view, 'Mary is chaste', 'Peter is cruel', 'George is a perfect knight', can be true in the very same sense in which 'Snow is white' is true, while still being ethical utterances. But there *is* an insight in non-cognitivism, these philosophers claim, even if non-cognitivism was mistaken in what it took to be its most essential thesis, the thesis that ethical sentences are not capable of truth (or, alternatively, the thesis that an ethical sentence has a distinct 'value component', and this 'value component' is not capable of truth). That thesis (or those theses) are rejected by Williams. As I said, he accepts the arguments of Murdoch and McDowell against the 'two components' theory; he recognizes the way in which fact and value are entangled in our concepts; and he agrees that ethical sentences can be true. How then can he maintain that there was an insight contained in non-cognitivism? What *was* the insight that the fact-value distinction tried to capture?

According to Williams, there are truths and truths. If I say that grass is green, for example, I certainly speak the truth; but I do not speak what he calls the *absolute* truth. I do not describe the world as it is '*anyway*', independently of any and every '*perspective*'. The concept 'green' (and possibly the concept 'grass' as well) are not concepts that finished science would use to describe the properties that things have apart from any 'local perspective'. Martians or Alpha Centaurians, for example, might not have the sorts of eyes we have. They would not recognize any such property as 'green' (except as a 'secondary quality' of interest to human beings, a disposition to affect the sense organs of *homo sapiens* in a certain way), and 'grass' may be too unscientific a classification to appear in their finished science. Only concepts that

would appear in the (final) description of the world that any species of determined natural researchers is destined to converge to can be regarded as telling us how the world is 'anyway' ('to the maximum degree independent of perspective'). Only such concepts can appear in statements which are 'absolute'. And the philosophically important point—or one of them, for there is something to be added—is that while value judgements containing thick ethical concepts can be true, they cannot be absolute. The world, as it is in itself, is *cold*. Values (like colours) are *projected* on to the world, not discovered in it.

What has to be added is that, on Williams's view, values are even *worse off*

end p.147

than colours in this respect. For the discovery that green is a secondary quality has not undermined our ability to use the word. We no longer think that colours are non-dispositional properties of external things, but this in no way affects the utility of colour classification. But the realization that value attributes, even 'thick' ones ('chaste', 'cruel', 'holy'), are projections has a tendency to cause us to lose our ability to use those terms. If we become reflective to too great a degree, if we identify ourselves too much with the point of view of the universe, we will no longer be able to employ our ethical concepts. The realization that ethical concepts are projections places us in a ticklish position: we cannot stop being reflective, but we cannot afford to be (very much of the time) *too* reflective. We are in an unstable equilibrium.

The reason for this difference between ordinary secondary qualities like green and thick ethical attributes like chastity, according to Williams, is that the interests which colour classification subserves are universal among human beings, whereas the interests that thick ethical concepts subserve are the interests of one human community (one 'social world') or another. Even if different cultures have somewhat different colour classifications, there is no *opposition* between one culture's colour classifications and those of another culture. But the interests which define one social world may be in conflict with the interests which define a different social world. And realizing that my ethical descriptions are in this way parochial (however 'true' they may also be) is decentring.

Williams believes that coming to realize by just how far ethical description misses describing the world as it is 'absolutely' not only does but *should* affect our first-order ethical judgements. There are *moral* consequences to the 'truth in relativism' (speaking, of course, from within *our* social world). The moral consequence (and perhaps also the metaphysical consequence), according to Williams, is that moral praise or condemnation of another way of life loses all point when that other way of life is too distant from ours. (Too distant in the sense that neither way of life is a live option for the other.) It makes no sense to try to evaluate the way of life of the ancient Aztecs, for example, or of the Samurai, or of a bronze age society. To ask whether their ways of life were *right*, or their judgements *true*, is (or should be) impossible for us; the question should lapse, once we understand the non-absoluteness of ethical discourse. And the fact that the question lapses *constitutes* 'the truth in relativism'.

## 4 Absoluteness

This dichotomy between what the world is like independent of any local perspective and what is projected by us seems to me utterly indefensible. I shall begin by examining the picture of science which guides Williams.

The picture of science is that science converges to a single true theory, a

end p.148

single explanatory picture of the universe. But one is hard put to know why one should believe this.

If we start at the level of common-sense objects, say stones, it suffices to notice that, in rational reconstruction, we can take a stone to be an aggregation—or as logicians say a 'mereological sum' of time-slices of particles (or, alternatively, of field-points—notice that these are incompatible but equally good choices!)—or we can take a stone to be an individual which consists of different particles in different possible worlds (and also occupies different locations in space in different possible worlds) while remaining self-identical. If a stone consists of *different* time-slices of particles in different possible worlds, then it cannot (as a matter of modal logic) be *identical* with an aggregation (mereological sum) of time-slices of particles [2]

[2] This argument is due to Kripke (in unpublished lectures).

, and obviously it makes no sense to say that a collection of space-time points could have occupied a different location than it did. So, if it is simply a matter of how we formalize our language whether we say (with Saul Kripke) that stones and animals and persons, etc., are *not* identical with mereological sums at all, or say (as suggested in Lewis, 1973) that they *are* mereological sums (and take care of Kripke's difficulty by claiming that when we say that 'the' stone consists of different particle slices in different possible worlds, then what that means is that the various modal 'counterparts' of the stone in different possible worlds consist of different particle slices, and not that the self-identical stone consists of different particle slices in different possible worlds—and to me this certainly looks like a mere choice of a formalism, and not a question of fact) we will be forced to admit that it is partly a matter of our conceptual choice which scientific object a given common-sense object—a stone or a person—is identified with.

Nor is the situation any better in theoretical physics. At the level of space-time geometry, there is the well-known fact that we can take points to be individuals *or* we can take them to be mere limits. States of a system can be taken to be quantum mechanical superpositions of particle interactions (*à la* Feynman) or quantum mechanical superpositions of field states. (This is the contemporary form of the wave-particle duality.) And there are many other examples.

Not only do single theories have a bewildering variety of alternative rational reconstructions (with quite different ontologies); but there is no evidence at all for the claim (which is essential to Williams's belief in an 'absolute conception of the world') that science converges to a *single* theory. I do not doubt that there is some convergence in scientific knowledge, and not just at the observational level. We know, for example, that certain *equations* are approximately correct descriptions of certain phenomena. Under certain conditions, the Poisson equation of Newtonian gravitational theory gives an approximately correct description of the gravitational field of a body. But the theoretical

end p.149

picture of Newtonian mechanics has been utterly changed by General Relativity; and the theoretical picture of General Relativity may in turn be utterly replaced by Supergravitation theory, or by some theory not yet imagined. We simply do not have the evidence to justify speculation as to whether or not science is 'destined' to converge to some one definite theoretical picture. It could be, for example, that although we will discover more and more approximately correct and increasingly accurate equations, the *theoretical picture* which we use to explain those equations will continue to be upset by scientific revolutions. As long as our ability to predict, and to mathematize our predictions in attractive ways, continues to advance science will 'progress' quite satisfactorily; to say, as Williams sometimes does, that convergence to one big picture is required by the very concept of knowledge, is sheer dogmatism.

Yet, without the postulate that science 'converges' to a single definite theoretical picture with a unique ontology and a unique set of theoretical predicates, the whole notion of 'absoluteness' collapses. It is, indeed, the case that ethical knowledge cannot claim 'absoluteness'; but that is because the notion of 'absoluteness' is incoherent. Mathematics and physics, as well as ethics and history and politics, show our conceptual choices; the world is not going to impose a single language upon us, no matter what we choose to talk about.

# 5 More About Absoluteness

The notion of absoluteness has further properties that we should be clear about. According to Williams, what makes the truth of a statement 'absolute' is not the fact that scientists are destined to 'converge' on the truth of that statement, that is to say, admit it to the corpus of accepted scientific belief in the long run, but the *explanation* of the fact of convergence. We converge upon the statement that *S* is true, where *S* is a statement which figures in 'the absolute conception of the world', because 'that is the way things are' (independently of perspective). But what sort of an explanation is *this*?

The idea that some statements force themselves upon us because 'that is how things are' is taken with immense seriousness by Williams: indeed, it is the centre of his entire metaphysical picture. Sometimes when I don't want to give a reason for something I may shrug my shoulders and say, 'Well, that's just how things are'; but that is not what Williams is doing here. 'That is how things are' (independently of perspective) is supposed to be a *reason* (Williams calls it an 'explanation') not a refusal to give a reason.

The idea that some statements get recognized as true (if we investigate long enough and carefully enough) because they simply describe the world in a way which is independent of 'perspective' is just a new version of the old 'correspondence theory of truth'. As we have already seen, Williams does not claim that truth is correspondence—for him, truth is rather right assertability in the language game. But *some* truths—the 'absolute' ones—are rightly assertable

end p.150

in the language game *because* they correspond to the way things (mind-independently) are. Even if correspondence is not the definition of truth, it is the *explanation* of absolute truth. And I repeat my question: What sort of an explanation is this?

The idea of a statement's corresponding to the way things are, the idea of a term's having a correspondence to a language-independent class of things, and the idea of a predicate's having a correspondence to a language-independent attribute are ideas which have no metaphysical force at all unless the correspondence in question is thought of as a genuine relation between items independent of us and items in language, a correspondence which is imposed by the world, as it were, and not just a tautological feature of the way we talk about talk. What I have in mind by this perhaps puzzling-sounding remark is this: if you think it is just a *tautology* that 'snow' corresponds to snow, or that 'Snow is white' is true if and only if snow is white, then you regard the 'correspondence' between the word 'snow' and snow as a correspondence *within* language. Within our language we can talk about snow and we can talk about the word 'snow' and we can *say* they correspond. To this even a philosopher who rejects the very idea of a substantive notion of 'truth' or a substantive notion of reference can agree. But if, as Williams believes, the fact that we are 'fated' to accept the sentence 'Snow is white' is *explained* by something 'out there' and by the fact that the sentence corresponds to that something 'out there',

then the correspondence too must be 'out there'. A *verbal* correspondence cannot play this kind of explanatory role. Williams's picture is that there is a *fixed* set of objects 'out there', the 'mind independent objects', and a fixed relation—a relation between words and sentences in *any* language in which 'absolute' truths can be expressed, any language in which science can be done, and those fixed mind-independent Reals—and that this relation *explains* the (alleged) fact that science converges. If this picture is unintelligible, then the notion of an 'absolute conception of the world' must also be rejected as unintelligible.

Now, I have argued for a number of years that this picture *is* unintelligible. First, I contend that there is not *one* notion of an 'object' but an open class of possible uses of the word 'object'—even of the technical logical notion of an object (value of a variable of quantification). The idea that reality itself fixes the use of the word 'object' (or the use of the word 'correspondence') is a hangover from pre-scientific metaphysics (Putnam, 1987). Secondly, the idea of the world 'singling out' a correspondence between objects and our words is incoherent. As a matter of model-theoretic fact, we know that even if we somehow fix the intended truth-values of our sentences, not just in the actual world but in all possible worlds, this does *not* determine a unique correspondence between words and items in the universe of discourse (Putnam, 1981). Thirdly, even if we require that words not merely 'correspond' to items in the universe of discourse but be causally connected to them in some way, the required notion of 'causal connection' is deeply *intentional*. When we say

end p.151

that a word and its referent must stand in a 'causal connection of the appropriate kind', then, even in cases where this is true, the notion of 'causal connection' being appealed to is fundamentally the notion of *explanation*. And explanation is a notion which lies in the same circle as reference and truth (Putnam, 1989). [3]

[3] Cf. my writings from 1978 to 1989, listed in the bibliography.

But why should this be a problem? Why should Williams and other meta-physical realists not just say, 'Very well, then. The ultimate description of the world—the world as it is in itself—requires intentional notions.' (In fact, Williams does not say this; Williams ends his book on Descartes (Williams, 1978) with an endorsement of Quine's criticism of intentional notions!) The answer, of course, is that a science of the intentional is a we-know-not-what. According to Williams, what gives the notion of an absolute conception of the world clout, what saves this notion from being a 'we-know-not-what', is that we have a good idea of what an absolute conception of the world would look like in *present-day physics*. But Williams does not expect present-day physics, or anything that looks like present-day physics, to yield an account of the intentional. He is thus caught in the following predicament: a correspondence theory of truth requires a substantive theory of reference. (And, I have argued, a belief in such a theory is hidden in Williams's talk of 'the way things are' *explaining* why we will come to believe 'the absolute conception of the world'.) If we say, 'Well, who knows, perhaps future science—we know not how—will come up with such a theory',—then we abandon the claim that we know the *form* of the 'absolute conception of the world' *now*. The absolute conception of the world becomes a 'we-know-not-what'. If we say, on the other hand, 'Reference can be reduced to physical parameters', then we commit ourselves to refuting the arguments (e.g. Putnam, 1988) against the possibility of a physicalist reduction of semantic notions. But Williams clearly does not wish to undertake any such commitment.

Instead, Williams's suggestion is that the intentional (or the 'semantic') is itself perspectival, and the absolute conception will some day explain why this kind of talk is useful (as it explains why talk of 'grass' and 'green' is useful, even though 'grass' and 'green' are not notions that figure in the absolute conception of the world). But here Williams shows a wobbly grasp of the logical structure of his own position. For the absolute conception of the world was *defined* in terms of the idea that some statements describe the world with a minimum of 'distortion', that they describe it 'as it is', that they describe it 'independently of perspective'—and what does any of this talk mean, unless something like a correspondence theory of truth is in place? Williams tacitly assumes a correspondence theory of truth when he defines the absolute conception, and then forgets that he did this when he suggests that we do not need to assume that such semantic notions as the 'content' of a sentence will turn out to figure in the absolute conception itself.

end p.152

# 6 Metaphysics and Entanglement

What led Williams to defend this complicated metaphysical theory was the desire to assert a 'truth in relativism' while resisting relativism in science. But in the process of building up this intricate construction with its two kinds of truth (ordinary and 'absolute'), its perspectivalism about secondary qualities and ethics (and, oddly, also about the intentional) and its anti-perspectivalism about physics, he often ignores the entanglement of the factual and the ethical—although he himself stresses that entanglement at other points in his discussion. Consider, for example, the question as to whether we can condemn the Aztec way of life, or, more specifically, the human sacrifice that the Aztecs engaged in. On Williams's view, the Aztec belief that there were supernatural beings who would be angry with them if they did not perform the sacrifices was, as a matter of scientific fact, wrong. This belief we *can* evaluate. It is simply false; and the absolute conception of the world, to the extent we can now approximate it, tells us that it is false. But we cannot say that 'the Aztec way of life' was wrong. Yet, the feature of the Aztec way of life that troubles us (the massive human sacrifice) and the belief about the world that conflicts with science were interdependent. If we can say that the Aztec belief about the Gods was false, why can we not say that the practice to which it led was wrong (although, to be sure, understandable, given the false factual belief)? If we are not allowed to call the practice wrong, why are we allowed to call the belief false? The so-called 'absolute' and the ethical are just as entangled as the 'factual' and the ethical.

For a very different sort of example, consider the admiration we sometimes feel for the Amish (traditional Mennonite) way of life. Even atheists sometimes admire the community solidarity, the helpfulness, and the simplicity of the Amish way. If a sophisticated atheist who felt this way were asked why he or she admires the Amish, they might say something like this: 'I am not necessarily saying we should give up our individualism altogether. But the kind of individualism and competitiveness which has brought so much scientific and economic progress, also brings with it egotism, arrogance, selfishness, and downright cruelty. Even if the Amish way of life rests on what I regard as false beliefs, it does show some of the possibilities of a less competitive, less individualistic form of life; and perhaps we can learn about these possibilities from the Amish without adopting their religion.' Now, Williams does not deny that we can say things like this; that we can learn from cultures to which we stand in the relation he calls 'the relativity of distance', cultures which are not 'real options' for us. But how does this differ from saying, 'Some of the Amish beliefs are false, but other of their beliefs may be true?' Many of Williams's examples load the

dice in favour of relativism by taking science to consist of individual judgements which may be called true or false, while taking 'cultures' to offer only 'take it as a whole or reject it as a whole' options.

The problem with the whole enterprise lies right here: Williams wants to

end p.153

acknowledge the entanglement of fact and value and hold on to the 'absolute' character of (ideal) scientific knowledge at the same time. But there is no way to do this. It cannot be the case that scientific knowledge (future fundamental physics) is absolute and nothing else is; for fundamental physics cannot explain the possibility of *referring to* or *stating* anything, including fundamental physics itself. So, if everything that is *not* physics is 'perspectival', then the notion of the 'absolute' is itself hopelessly perspectival! For that notion, as I have already pointed out, is explained (albeit in a disguised way) in terms of notions which belong to the theory of reference and truth, and not to physics. And the idea of a 'relativism of distance' which applies to ethics but not to science also fails, because ethics and science are as entangled as ethics and 'fact'. What we have in *Ethics and the Limits of Philosophy* is, in fact, *not* a serious argument for ethical relativism, but rather an expression of a mood. Reading *Ethics and the Limits of Philosophy*, one gets the feeling that one is being told that ethical relativism is the 'sophisticated' point of view, the 'modern' point of view, and that what is being offered is a *sophisticated reflection on the consequences of this presupposition*. But the presupposition itself does not stand up to any kind of examination—or at least, the way Williams defends the presupposition crumbles the moment one tries to subject it to any sort of careful examination.

## 7 Entanglement and Positivism

Relativism appeals to sophisticated people for different reasons. It appeals to Williams because the idea of ethical objectivity is metaphysically unacceptable. He does not see how we could *know* objective ethical truths if there were any. This metaphysical (or is it epistemological?) appeal is one I do not myself feel. It is not that I *do* know how I know that, for example, human dignity and freedom of speech are better than the alternatives, except in the sense of being able to offer the sorts of arguments that ordinary non-metaphysical people with liberal convictions can and do offer. If I am asked to explain how ethical knowledge is possible at all in 'absolute' terms, I have no answer. But there are all sorts of cases in which I have to say, 'I know this, but I don't know how I know it.' Certainly physics doesn't tell me how I know anything.

Another, very different, appeal is to those who fear that the alternative to cultural relativism is cultural imperialism. But recognizing that my judgements claim objective validity and recognizing that I am shaped by a particular culture and that I speak in a particular historical context are not incompatible. I agree with Williams that it would be silly to ask if the way of life of an eighteenth-century Samurai is 'right' or 'wrong'; but the reason this is a silly question isn't that we are too 'distant', or that becoming eighteenth-century Samurai isn't a 'real option' for us. In my view, it would be a silly question if we *were* eighteenth-century Samurai. Indeed, 'Is our own way of life right

end p.154

or wrong?' is a silly question, although it isn't silly to ask if this or that particular feature of our way of life is right or wrong, and 'Is our view of the world right or wrong?' is a silly question, although it isn't silly to ask if this or that particular belief is right or wrong. As Dewey and Peirce taught us, real questions require a context and a point. But this is as true of scientific questions as it is of ethical ones. Instead of trying once again to discover some deep truth contained in positivism—in the fact-

value dichotomy, or in 'non-cognitivism', or in the verifiability theory of meaning—we should break the grip of positivism on our thinking once and for all.

The failure of the latest attempt to find some deep truths in positivism is no accident. Although Williams tries to do justice to the entanglement of fact and value, he fails do so, because positivism was fundamentally a denial of entanglement, an insistence on sharp dichotomies: science-ethics, science-metaphysics, analytic-synthetic. The science-ethics dichotomy that Williams wants to preserve presupposed the science-metaphysics and analytic-synthetic distinctions he rejects. This is why Williams's book-length attempt to spell out his position is either self-contradictory or hopelessly ambiguous at every crucial point.

Recognizing that the entanglement of fact and value, as well as of science and ethics, science and metaphysics, analytic and synthetic, is here to stay may also help us to see our way past another contemporary shibboleth: the supposed incompatibility of universalist (or 'enlightenment') and parochial values. Recently I was struck by something Israel Scheffler has written (Scheffler, 1987):

I have always supposed that the universal and the particular are compatible, that grounding in a particular historical and cultural matrix is inevitable and could not conceivably be in conflict with universal principles. I have thus belonged to both sides of a divide which separated most Jewish academics and intellectuals of my generation.

When we argue about the universal applicability of principles like freedom of speech or distributive justice we are not claiming to stand outside of our own tradition, let alone outside of space and time, as some fear; we are standing within a tradition, and trying simultaneously to learn what in that tradition we are prepared to recommend to other traditions *and* to see what in that tradition may be inferior—inferior either to what other traditions have to offer, or to the best we may be capable of. Williams is right when he says that this kind of reflection may destroy what we have taken to be ethical knowledge; it may certainly lead us to re-evaluate our beliefs, and to abandon *some* of them; but he is wrong when he fears that the most ultimate kind of reflective distance, the kind which is associated with the 'absolute conception of the world', will destroy *all* ethical knowledge. Here he is worrying about a distance which is wholly illusory. No conception of the world is 'absolute'.

Williams describes the 'absolute conception of the world' as something required by the very concept of knowledge. What this transcendental moment

end p.155

in Williams's argument shows is that, for him, there is no conceivable alternative to the idea of an absolute conception of the world—or no alternative save, perhaps, a scepticism as total as that of the ancient Greeks. But we are not forced to choose between scientism and scepticism in the way Williams thinks. The third possibility is to accept the position we are fated to occupy in any case, the position of beings who cannot have a view of the world that does not reflect our interests and values, but who are, for all that, committed to regarding some views of the world—and, for that matter, some interests and values—as better than others. This may be giving up a certain metaphysical picture of objectivity, but it is not giving up the idea that there are what Dewey called 'objective resolutions of problematical situations'—objective resolutions to problems which are *situated*, that is, in a place, at a time, as opposed to an 'absolute' answer to 'perspective-independent' questions. And that is objectivity enough.

**Bibliography**

Cavell, Stanley (1979). *The Claim of Reason*. Oxford: Oxford University Press.

Lewis, David (1976). *Counterfactuals*. Cambridge, Mass.: Harvard University Press.
McDowell, John (1978). 'Are Moral Requirements Hypothetical Imperatives?', *Proceedings of the Aristotelian Society*, suppl. vol. 52.
—— (1979). 'Virtue and Reason', *Monist*, 62.
Murdoch, Iris (1971). *The Sovereignty of 'Good'*. London: Routledge and Kegan Paul.
Putnam, Hilary (1974). 'The Corroboration of Theories', in P.A. Schilpp (ed.), *The Philosophy of Karl Popper*. LaSalle, Ill.: Open Court.
—— (1978). *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
—— (1980). 'Models and Reality', *Journal of Symbolic Logic*, 45 (repr. in Putnam, 1983).
—— (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
—— (1983). *Realism and Reason* (vol. iii of my *Philosophical Papers*). Cambridge: Cambridge University Press.
—— (1986). 'Information and the Mental', in E. Lepore (ed.), *Truth and Interpretation*. Oxford: Oxford University Press.
—— (1987). *The Many Faces of Realism*. LaSalle, Ill.: Open Court.
—— (1988). *Representation and Reality*. Cambridge, Mass.: Bradford Books.
—— (1989). 'Model Theory and the Factuality of Semantics', in Alex George (ed.), *Reflections on Chomsky*. Oxford: Blackwell.
Quine, Willard V. (1951). 'Two Dogmas of Empiricism', first published in *Philosophical Review*; repr. in *From a Logical Point of View* (revised edn., 1961). Cambridge, Mass.: Harvard University Press.
—— (1963). 'Carnap on Logical Truth', in P.A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. LaSalle, Ill.: Open Court.
Scheffler, Israel (1987). 'Teachers of My Youth' (unpublished). Copyright, I. Scheffler.)

end p.156

Walsh, Vivian (1987). 'Philosophy and Economics', in J. Eatwell, M. Milgate, and P. Newman (eds.), *The New Palgrave: A Dictionary of Economics*, iii. London: Macmillan, and New York: Stockton Press.
White, Morton (1956). *Towards Reunion in Philosophy*. Cambridge, Mass.: Harvard University Press.
Wiggins, David (1987). *Needs, Values, Truth*. Oxford: Oxford University Press.
Williams, Bernard (1978). *Descartes: The Project of Pure Enquiry*. Harmondsworth: Penguin Books.
—— (1985). *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.

end p.157

# Hilary Putnam: Objectivity and the Science-Ethics Distinction

show chapter abstract and keywords
hide chapter abstract and keywords

**Lorenz Krüger**

What I admire and appreciate very much about Putnam's paper on the science-ethics distinction—indeed about his more recent philosophical work (1981, 1987) in general—is the fresh and original attack on a key problem of the Western scientific-philosophical tradition: the problem of establishing a unified and coherent cognitive enterprise which comprises values and facts, decisions and insights, ethics and science alike. If he should succeed, we would—*vis-à-vis* the problem of quality of

life—be in a very different situation, because we could settle disputes on values by co-operative *investigations* rather than controversial *negotiations*.

The commentator, it seems to me, must inevitably ask the following critical question: has Putnam succeeded in showing that unified concepts of objectivity and rationality can be developed for science *and* ethics? Indeed, is it even desirable or advisable to insist on the ideal of the unified scientific-ethical enterprise?

In pursuing these questions, I shall not adopt a strategy that might seem to be suggested, if not demanded, by the arrangement of Putnam's argument; that is, I shall *not* try to examine the possible defences of the opponents he attacks in his paper, notably Bernard Williams. I cannot do this, simply because I happen to accept Putnam's main argument against Williams. Like him, I think that the distinction between relative and absolute truth (in the sense he attributes to Williams) is untenable. But, unlike him, I see different reasons for this untenability; hence I am led to adopt different conclusions—conclusions that are at variance with the idea of a unified sphere of objectivity for science and ethics. Leaving other aspects of Putnam's paper aside, I turn immediately to the disputed distinction between the relative and the absolute.

Relative truths are such as will be accepted by some community or other of rational believers but remain dependent on certain limited cultural contexts. (What the limits of these contexts and of these cognitive communities are remains an open question to be explored for each case in its own right.) Absolute truths, however, are context-independent; and this they are said to be by virtue of their correspondence to the things 'out there', or, more precisely, by their corresponding to such features of reality as could invariantly and equivalently be expressed in any other sufficiently elaborate language.

Favoured candidates for the status of such features are, of course, suitably chosen objects as described by physics. Candidates for relative truths, on the other hand, are ordinary statements of everyday life such as 'the grass is green', 'our neighbour is a malicious person', or 'the tax laws are just'. What counts as 'grass' or 'green', let alone as a 'malicious character' or a 'just tax law', is

end p.158

hardly culturally invariant, though such statements may be recognized as true or false, once a cultural context is taken for granted.

Now Putnam opposes this view for at least two important reasons. First, he thinks that the conception of absolute truth is untenable, indeed unintelligible; second, he thinks that relative truth is less than we need and less than we can attain. I fully agree with both points.

My misgivings with Putnam's approach concern his views on the connection between his two critical points and the conclusions he derives from this connection. In my reading of his paper the connection is the following: If you grant the first point, that is, if you drop absolute truth and the correspondence between beliefs and things out there altogether, you are left with nothing but relative truths. Yet these — as far as their epistemic status is concerned — have already been granted to form a homogeneous field. That the neighbour is a malicious person is a truth, or a falsity, of the same kind and epistemological status as the assertion that the grass is green. Putnam, his opponents, and further authors (some cited in his paper, e.g. Iris Murdoch), have accumulated an impressive sample of statements that do not admit of an analysis into a 'factual' and an 'evaluative' component. The examples range from recognizing a person as pert or polite to condemning human sacrifice as utterly evil. All these cases display an inextricable 'entanglement of fact and value'.

Nevertheless, they are all plausibly presented as acceptable candidates for the assignment of a truth-value, though Putnam does not deny that the actual ascription of either truth or falsity to such assertions about value-facts may depend on comprehensive ideologies and forms of life, for example, the Aztec religion, which demanded human sacrifice.

We can now see how the rejection of an admittedly ill-conceived and untenable partition of the set of truths combines with a positive characterization of those that fall on the relativistic side. As far as Putnam's present paper goes, it is this combination that leads on to further conclusions, conclusions which I believe to be unacceptable. The most important of these is the desired unity of our cognitive field. Another important conclusion (which, if true, would support the first) is the following: given that there is no absolute truth because there are no fixed objects 'out there', the notion of correspondence between statements or beliefs and transcendent 'facts' becomes an empty or even an unintelligible concept. It is replaced by rational acceptability.

Here are my objections. The first conclusion — the existence of a unified domain of rationally assertable truths — is an all-too-sweeping extrapolation of a correct observation: that of inextricable entanglements of fact and value in many life issues. This extrapolation in its turn is encouraged by the second conclusion, which is that truth can *never* consist in correspondence to external facts. To this I object that the inference from 'no correspondence to something fixed out there' to 'no correspondence at all' is a *non sequitur*. Now, if there were *some* kind of correspondence in some cases and no correspondence of *any* kind in some other cases, a new form of the fact-value distinction and

end p.159

of the ethics-science distinction might be established — a form that is freed from the metaphysical burden of absoluteness.

I take up the second point first, because it lies close to the centre of Putnam's theoretical philosophy, that is, his internal or pragmatic realism. To avoid misunderstandings, I want to stress that I second Putnam's criticism of external, absolute, or metaphysical realism, and that I am all for internal, knower-related, or pragmatic realism. Moreover, I share Putnam's hope of nevertheless resisting a sweeping cultural (let alone subjective or personal) relativism. I part company with him, however, on at least two issues. First, I do not believe that we can secure a non-relativistic realism without some notion of 'correspondence' to external reality. Second, I do not believe that all rationally acceptable propositions have the same epistemological status, rather that some are acceptable because of their correspondence, others for entirely different reasons. The first kind can be called 'true' (or 'false') in the standard meaning of the terms, the second kind cannot. If truth talk is applied to the second kind, this has to be done by tactful extension of, or on analogy with, the first type. But there are limits to these extensions or analogies, limits that are to be assessed in a rational analysis. In other words: I think that the field of rationality is much wider than that of truth, and that this circumstance is of vital importance for the survival of rationality.

What is correspondence if it is not to be absolute or metaphysical? It is a relationship between our statements or beliefs, on the one hand, and our perceptions and actions, on the other. Both relata are experiential and free from metaphysical suspicion. Take a simple example, for instance, the statement that there is now an even number of people in this room. If it should happen to be true, that would mean, among other things, that you could pair off all people present, so that no one is left

alone. Here is an obvious correspondence between our verbal or mental representation and our material, that is, non-verbal and non-representational, contacts with a piece of reality.
Now, it may conceivably be the case that we could have a system of representations that lacks the conceptual distinction between 'odd' and 'even'. One may, therefore, argue — and this is quite in line with internal realism — that even my banal example illustrates the fact that our conceptual system (presumably together with our imaginative possibilities) irrevocably prejudges what can ever become an object for us. But it is equally clear from the example that it would be absurd to collapse the doubleness of representation and material contact.
So far everything seems trivial. But perhaps the next step is already less innocuous. We all believe that people come in countable integer units and that they cannot penetrate walls, so that it lies in the nature of things — that is, of our material contacts with the world — that there is actually no choice between alternative representational systems when it comes to the question of how many people have gathered in a closed room. Statements concerning such matters are simply true or false; and they are that, first, because of a certain

end p.160

correspondence with reality and, second, regardless of the cultural context, provided it is developed enough to discuss such matters at all.
Again, no more than a trivial point? Yes, as far as normal people's concerns go. But no, as soon as subtle philosophy is brought into play. What I am implying is no more, but also no less, than an assumption to the effect that the niceties of a logically refined ontology are completely irrelevant to our problem of correspondence. It is part of our material world contact that we do not encounter, and cannot count, time-slices of people, or experience them as portions of some stuff, of a kind of human mush. Nor do we have a choice between an individualistic ontology *à la* Carnap and a mereological ontology according to the Polish logicians (Putnam, 1987: 17-21) in this case. Brevity forces me to make an unduly perfunctory statement here: correspondence, and with it the reference of terms, does not suffer in the least from ontological relativity, at least not in our humdrum examples.
But how far do they carry us into more complex and more abstract matters? The lack of serious alternative descriptions smacks of absoluteness again. But here we must guard against the philosopher's professional disease, of turning universal experiences into conceptual necessities. I take the lack of serious alternative representations of simple matters to be no more but also no less than a contingent, if enduring, fact about those matters and about humans who take cognizance of them. Keeping this point in mind, let us now turn to a different example, for instance, the proposition that all the people in this room are independent and honest thinkers. There is — we will probably concede — some fact of the matter concerning this statement, that is, circumstances that justify its assertion or denial, as the case may be. But it is also clear that this statement is much less stable than the one about the even number, especially across time and across cultural boundaries. Now, I submit that the difference is not just one of precision or accessibility to examination or the like, but a difference in subject-matter and, by implication, in our relation to that subject-matter. (Here I find myself in partial, though not complete, agreement with Williams, as rendered by Putnam on pp. 146-56.)
What lies behind this apparent difference? It is not of our doing that humans — or shall I artificially say: instantiations of humanity — come in discrete physical units, whereas codes of intellectual independence, still more of honesty, are products of

intentional and, within limits, free human life. True, they are far from arbitrary, and, as with everything historical, any single person will usually feel almost powerless to change them. But (*pace* Plato) numbers, whether in 'pure' mathematics or in experience, and honesty are worlds apart in kind. One may also express the contrast in a Kantian language, using a key term of his philosophy: with respect to some matters we are (or could and should be) *autonomous*, with respect to others we are not. This is what sharpens the point of the famous confrontation of the moral law in us with the starry sky above us. To be sure, we conceive of both *within the limits of our intellect and reason*, but we do not impose on ourselves, in Kant's terms, the laws of the

end p.161

intellect, whereas we do impose on ourselves the laws of our actions. In other words, what I want to say is that the distinction between facts and values can and should be reconstructed *within* the framework of internal realism.

In the spirit of the rejection of pre-scientific metaphysics — which I share with Putnam — I feel driven towards and into history for further instruction and examination of this claim. What I believe I see there is the following pattern (I hope the reader will forgive me for its crudeness, which I cannot avoid owing to lack of space and competence): to build the world of values and forms of life according to the order of the cosmos is a fundamental characteristic of our philosophical-scientific tradition — born with Pythagorean speculation, nourished by Plato (though not, or so it seems, by Socrates), pronounced the guideline of social order by the Stoics, surviving in the Christian idea of the participation of the rational human soul in the divine order of things, finally entering into the optimistic research programmes of the modern philosopher-scientists like Galileo, Descartes, Hobbes, and Hume. But precisely in its modern enlightened form the great programme has come to grief. We can only register, or so it seems to me, a deep disappointment with its outcome. The disillusionment is, first, that no scientific ethics, as projected by the philosopher-scientists of the New Age, has been established; second, that the same holds for a scientific insight into 'the laws' of history, if there are such at all. On the contrary, current and recurrent attempts at providing a specifically scientific basis for the responsible and rational investigation of history and ethics — sociobiology or certain orthodox versions of Marxism are examples — are at best helpless *vis-à-vis* our problems and at worst dangerously ideological.

My statement of the historical lesson is, of course, a sketch, no more. It lacks the elaboration and the defences it would need to stand up to scrutiny. But it may help to see what I am up to: in our age of science we cannot but connect even the most elementary statements with their actual or potential theoretical context. But if we try to do this with my two examples — the even number on the one hand and the intellectual independence and honesty on the other — we find ourselves confronted with two deeply different sorts of context: on the one side there is that growing web of scientific theories about that part of our experience that we cannot produce, change, or destroy. Let us call it 'non-human nature' (though it comprises, of course, a good part of our physical and mental apparatus). True, we can exploit this nature, but only, as Francis Bacon said, by obeying it. On the other side there is that part of our experience that we can change, one is tempted to say, that part of it that *is* us. And here obedience does not seem to be the right, perhaps not even a possible, attitude. We want to have, and to an impressive degree do have, disciplined research traditions and institutions that explore this domain of our experienced reality; but they would be ill advised, should they suppose as their object a human nature in any

sense comparable to the non-human nature which is the supposed object of the natural sciences.

end p.162

For easiness of reference, I may perhaps call the view I have just outlined 'historical realism'. I think of it as an elaboration, but also as a correction, of internal realism. The name is to remind us that we learn what is real in the course of our history. One lesson — the one I have mentioned — appears to be that self-conscious and potentially autonomous human beings are a peculiar reality of their own — a reality to which we do not react in the same way as that in which we react or respond to non-human nature, so that our beliefs or statements about this reality cannot correspond to their contents in a way in which our beliefs or statements about non-human nature can.

The significance of this lesson, if it is one, hardly needs to be pointed out. If we have a claim to autonomy, narrow as its bounds may be, it is important to recognize that, if for beliefs and statements belonging under its rule there is truth or falsity, then it is not truth or falsity that relates to 'correspondence' in the sense outlined above. I do not want to dispute anyone's use of 'true' or 'false'. We may find that all the people in this room *are* independent thinkers, and we may wish to say that a statement to this effect is *true*. But here 'true' will presumably mean something like 'rationally acceptable' or even 'not rationally rejectable'. Any attempt at ascertaining this kind of truth, however, will involve us in explications of what it is to be independent. These explications, in turn, will involve questions about what balance we want, or ought, to strike between individual originality and mutual social adjustment. In short, the answer to the question as to whether or not we are independent thinkers is not independent from the question as to what kind of people we want to be or what type of society we want to have.

For reasons like these, there can be no objectivity in the domain of autonomy, since there is no object. Instead, there is a world of subjects. Needless to say, I am not here denying the possibility of, indeed the urgent need for, intersubjectivity and rationality. What I claim is only that Putnam's renewed attempt at unifying human reason has not really advanced beyond Kant's great divide of theoretical and practical reason. But how could it have? And would that have been desirable? Besides the by now global unification of our substantial science plus its material technology, we will want to preserve cultural pluralism, an option I share with Putnam (see e.g. 1981: 147-9). It is true, he does not seem to fear that the continuity between science and ethics will work against pluralism. But I do fear that it would work against cultural pluralism, if only it were effectively realizable. And that would not be desirable.

Thus, I conclude that to reach reasonable agreement — or reasonable disagreement, wherever the circumstances leave room for it — is one thing; to claim objectivity or truth is another.

end p.163

**Bibliography**

Putnam, Hilary (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.

—— (1987). *The Many Faces of Realism*. La Salle, Ill.: Open Court.

end p.164

# Objectivity and Social Meaning

show chapter abstract and keywords
hide chapter abstract and keywords

**Michael Walzer**

# [1]

I probably do not have an objective view of objectivity. Having been accused so often of disdaining it, I come to it now with some trepidation; I want to make a cautious approach, repressing for a while the uneasy sense that the conjunction in my title misrepresents the likely outcome of my argument. Let me begin with a strong, simplistic, and usefully wrong definition of objectivity: a given perception, recognition, or understanding can be called 'objective' if its content is wholly or largely determined by its object—so that a range of human subjects, differently placed, with different personalities and different, even conflicting, interests, would agree on the same content so long as they attended to the same object. The table determines the objective perception of the table. What makes for objectivity is simply this: the object imposes itself. The subject is passive and undiscriminating, a promiscuous consumer of available 'data'.

For reasons philosophers have long understood, that cannot be right. Human beings are active subjects. Our faculties of perception and cognition help to determine whatever it is that we finally see or recognize or understand. But we are still inclined to call the perception 'objective' so long as these faculties are so widely shared as to constitute what we might call a normal subject. Then perception is objective when it is jointly determined by the object and the normal subject. If someone without depth perception reports on the existence of a table different from the one the rest of us see, his is the subjective report. The table and the normal person looking at the table (who represents 'the rest of us') together determine what the table objectively is (looks like). The object still imposes itself, but perception is conditioned by the character of the receptive organism, and the idea of 'objectivity' incorporates the results of this conditioning.

But that cannot be right either, and this time for reasons that have generated a long series of complex, difficult, and sometimes high-flying philosophical arguments. We do not come to the object with faculties alone but with interests and ideas too. And what we see, recognize, and understand depends (with a strong but not absolute dependency) on what we are looking for, our cognitive *concerns*, and the ways we have of describing what we find, our conceptual *schemes*. Given our concerns and our schemes, what opportunity is there now for the object to impose itself? We seem armoured against imposition, shaping the world to our own purposes.

But I do not mean to surrender objectivity so quickly—indeed, it is the scientific perception of the world, driven by a strong purposefulness and

end p.165

structured by elaborate and highly speculative schemes, that makes the most insistent claim these days to be called an objective perception. The claim takes many different forms, but in all its forms it must hold that if the object does not impose itself it is still recalcitrant to conceptual and purposive impositions. Scientific concepts must accommodate the object—not as the object appears, perhaps, but as it *really is*. I am not going to comment on that last assertion, except to say that for most of us, at least, appearance is an important aspect of reality. But I want to accept the claim that objectivity hangs (somehow) on the accommodation of the object by a knowing, inquiring subject. The knowing subject shapes the object, but he cannot shape it however he likes; he cannot just decide that a table, say, has a circular or a square shape without reference to the table. Similarly, someone self-confidently applying a conceptual scheme that divided the world into friends, enemies, reading matter, and edible plants would get the table wrong (objectively wrong), or he would miss the

table entirely, and deny its reality, and that would be a merely idiosyncratic (subjective) denial.

This is still a very simple account of objectivity, a rough, commonsensical approach to philosophical difficulties that lie beyond my immediate congnitive concerns and possibly also beyond the conceptual schemes at my disposal. But the account works, more or less well, for simple objects-in-the-world. The question that I want to ask now is whether it works at all for objects to which we assign use and value, objects that carry 'social meanings'. [1]

[1] Are there any objects without social meaning? Perhaps the phrase 'simple objects-in-the-world' names a null set. But I am going to assume that there are such things, which we accommodate and shape directly, without any necessary reference to their sociological significance. Stones are, for my purposes, simple objects-in-the-world—until they are made into cornerstones, tombstones, grindstones, milestones, stepping stones; or doorsteps (or, more dramatically, until they are used for coronations or set up as markers for a sacred history). As for tables, see the argument below.

This term, borrowed from anthropology, seems to cast a cloud over all claims to objective knowledge. Social meanings are constructions of objects by sets of subjects, and once such constructions are, so to speak, in place, the understanding of the object has been and will continue to be determined by the subjects. New sets of subjects learn the construction and then respect or revise it with only a minimal accommodation of the object. The object may or may not limit the constructive work in which they are engaged. Obviously, the table cannot be constructed as an intercontinental ballistics missile. But it can become a desk, a workbench, a butcher's block, or an altar, and each of these can take on meanings to which the 'mere' table gives us no positive clue. Can perceptions of objects like these, objects-with-meanings, ever be called objective? It is easy enough to imagine situations in which one person's altar is another person's butcher's block. But we do accept reports on social constructions. Now objectivity (in the reports) hangs on an acknowledgement of the construction. Our shared understanding of what an altar is, what we have made it for, determines our perception of the table-that-is-an-altar. The holiness of the altar is similarly

end p.166

objective, since it is part of the same construction. All normal persons living within the system of social meanings would deliver similar reports on the objective reality of tables-that-are-altars-that-are-holy.

But this may go too far. Suppose that there are dissenting voices within the society where some tables are holy altars, people who deny the construction, who announce, 'There's nothing there but an old table.' That is also an objective report of a kind. Can we say that it is an incomplete report, that it misses something of real importance? Imagine a fuller report: 'Some people claim that it's an altar and treat it as if it were holy, but there's nothing there but an old table.' No incompleteness now, and now only the disagreement can be objectively reported: '*I think* there's nothing there but an old table.' Nothing in the nature of the table will lead us to say that it is or is not a holy altar. The altar is objectively there only for those who understand it to be objectively there. It is holy only for those who acknowledge its holiness. And what they will have to say, if they are to report objectively, is that it is holy *for them*. Believers will want to say more than that. They will want to say that God has sanctified the table and made it into a holy altar; hence everyone who knows how things really are in the world will acknowledge its holiness. But I shall take it as given that altars and holiness are alike human creations. The believers are wrong, then to take the holiness of their altar as a universally recognizable (objective) fact. The altar is holy only because and only in so far as they have made it so. With regard to such

creations, the rest of us are not bound by majority rule; only the voice of the people as a whole resembles the voice of God. Social constructions must reflect a general agreement—or, better, since no vote is ever taken, there must be a consensus—if there is ever to be an unqualified objectivity, an objectivity without pronouns, in our reports about them. (Reports from outside observers will always need pronouns: 'their altars' or, in more extended form, 'these tables, which they use as altars'.) The more complex and specific the construction the more surprising it is when a consensus is actually reached. The social processes that make this possible are mixed processes, involving force and fraud, debate and consent, long periods of habituation; overall, they remain mysterious.

Compared to 'altar', 'table' is both uncomplicated and indeterminate; hence its meanings will rarely provoke significant or stirring dissent. Someone who says, 'That's not a table', while pointing to a flat piece of wood with a supporting structure of appropriate height, will probably lead us to talk about mistakes, not disagreements. (I will not stop here to imagine bizarre cases of table-like objects that are rarely not tables.) We would suspect some failure of normal understanding. A table is indeed a social construction, as well as a physical construction, but the socially constructive work is so rudimentary that we are unlikely to recognize much more in it than the assignment of a general name to the object. And then we expect people of normal understanding to remember the name. Nothing much follows from remembering it; the construction of the flat piece of wood, etc., as a table does not require

end p.167

us to use or to value the table in a certain way. More specific constructions, by contrast, have normative consequences.

Tables-that-are-altars-that-are-holy must be treated in accordance with certain principles and rules. I cannot use the altar, for example, as a desk on which to write profane essays on social meaning—not because the altar will resist the use or God strike me dead, but because it would be wrong to do that given what an altar is in my society (for me and my fellow members). Nor can I chop it up for firewood, even if the church is very cold; or trade it for personal profit—a new suit, say, or a season ticket to the opera, or a place on the Stock Exchange: it would be the wrong thing to do. But would it be objectively wrong? It seems to me that it is possible to ask questions like that too soon. Clearly the rules of use and value are not determined by the 'mere' table; nor are they jointly determined by the table and a normal person looking at the table; nor do they represent an accommodation of the table by a knowing subject or a scientific observer. The rules follow from the social construction of the table as a holy altar and they would seem to be objective rules only for those men and women who join in the construction or acknowledge its results. The other might be bound by some notion of 'decent respect' for the opinions of their fellows, but not by the idea of holiness.

But perhaps we can go a little further than this. If we think of the holiness of the altar not as an isolated construction but as one feature of a more complex whole, a cultural system or a way of life, then the force of the rules is considerably enhanced. Imagine the table-that-is-an-altar-that-is-holy within a set of connected constructions: socially meaningful occasions (holy days), spaces (churches), officials (priests and bishops), performances (religious services), texts (scriptures, prayers, homilies, catechisms), and beliefs (theologies or cosmologies), and the result is something from which individuals cannot so easily opt out. Some day there will be alternative occasions, spaces, officials, performances, texts, and beliefs, arising out of a long process of social change ('secularization', say); and then people will be able to

explain to their fellows why the altar is not (really) holy. But now the refusal of some dissident or rebel to treat the altar in accordance with its rules of use and value is probably not a straightforward denial of its holiness but a specific act of desecration—literally, an effort to reverse the process through which this particular altar has been consecrated. And the religious rebel committed to desecration is likely to appeal, in much the same way as the early Protestants appealed, to other features of the existing cultural system or way of life, features that give him reasons, so he says, for what he does. The system as a whole still has objective value for him; he lives within the set of social constructions. Where else can he live?

## [2]

We can still ask whether this is objectively the best place to live. But I want, again, to postpone that question in order to explore more fully the crucial

end p.168

implication of my argument thus far: that social construction is also moral legislation. The meanings with which we invest objects have normative consequences. I have been calling these norms 'rules of use and value'; they are also rules of distribution, that is, they regulate our relations not only with things but also with other people. Any number of philosophers have argued that morality is a human invention, writing mostly as if what we invent are the rules that govern a moral life. We leap to principles like equality before God or personal autonomy or the greatest happiness; and then we make lists, like the *Decalogue*. Maybe we do that, sometimes; but the thickness of the moral world and the density of our relationships suggest a radically different kind of invention. One of the ways we reach that thickness and density is through the social construction of objects (of all kinds). Social construction makes for a complex and rich world, many features of which will seem so obvious to us that we will not be prompted to ask whether they are, of all possible features of all possible worlds, objectively best. They will have a more immediate objectivity. So we will use and value objects in accordance with the meaning they have in our world, and we will exchange, share, and distribute them in accordance with their use and value. We will know what objects we owe to other people as soon as we understand what those objects (really) are and what they are for. And a great part of our conduct towards other people will be governed by these distributive entailments of social meanings.

At this point, it will be useful to take up another example, even though the difficulties of the table-that-is-an-altar-that-is-holy have by no means been exhausted. I want to consider the construction of a human life—not a biological but a social life, not a life span but a life course in a particular society, namely, our own. What we have constructed is a life-that-is-a-career-that-is-open-to-talents. Obviously, there is nothing in the nature of a human life that determines its construction as a career. Any given version of the life course is conditioned by the life span, so that youth, maturity, and age give rise to a pattern like training, professional practice, and retirement; but these latter three do not by themselves constitute a career. A career is an individual achievement; it is constituted by choice and qualification. Though career patterns may be collectively established and repetitively enacted, a career is none the less a projection of the self into a chosen and uncertain future. What makes this projection possible is the opening up of certain sorts of places and positions (professional or bureaucratic), which I will call 'offices'. Offices are the objects of careers. The social construction of the two goes hand in hand, like altars and offerings. If careers are open to talents, then offices must be distributed on meritocratic principles to qualified persons. If we imagine individual men and women planning careers-open-to-talents, we must also imagine competitions for office. If there are competitions, there must be

rules protecting the competitors, not only against violence but also against discrimination, that is, against any refusal to attend honestly to their qualifications. Once careers and offices are in place, nepotism becomes a wrongful practice. It would be wrong for me, the member of some search committee, say, to

end p.169

favour my brother over a more qualified candidate. It does not matter that I have a very strong, and to my mind overriding, belief in family loyalty; I am caught up in a complex set of social constructions that has normative entailments. Someone who fails to respect the table-that-is-an-altar-that-is-holy does no injury to the table, and in so far as the failure is private he does no injury (causes no offence) to other men and women either. His is one of the minor sins. But once constructions determine distributions, private refusals make for a more serious wrong. And when behaviour is in question, general agreement is no longer a necessary condition of rightness or wrongness; the rule against nepotism, for example, is binding even on individuals who argue that offices are family holdings and not objects of careers. There will not be many people, however, who will actually *argue* that offices are family holdings—except in the unlikely case that a strong familial idealism is part of the same set of social constructions as the career-open-to-talents, and then we might well recognize the public nepotist as a conscientious objector. Acts of refusal and opposition commonly have a basis of this sort, in the coexistence of contradictory constructions. Then people have to choose, guided only by their best understanding of the complex social world they inhabit.

I want to stress that it would not be objectively wrong to adopt the argument for offices-as-family-holdings. Majority rule does not govern arguments about social meaning; it only governs behaviour. The rules of behaviour, then, are objectively right relative to the prevailing meanings, but the prevailing meanings are not objectively right (or wrong). They are only objectively *there*, the objects, that is, of more or less accurate reports. The life-that-is-a-career could, over time, be constructed in an entirely different way, and offices could be reconstructed to match the difference, and no wrong would have been done. It would not be the case that lives or offices had somehow been misunderstood; nor would the men and women leading lives and holding (or not holding) offices under the new dispensation have been treated unjustly.

It is not my claim that the whole of morality is objectively relative (relatively objective?) in this way, only whatever part of it is entailed by the social construction of objects. Even here we might plausibly ask whether there are cases where construction is jointly determined by its objects and its human agents in such a way that the same normative entailments appear again and again, in all or almost all human societies. Then the same behaviour would be wrongful for the same reasons in all human societies; morality would lose its particularist character without ceasing to be relative to social construction. The easiest case has to do with the things we call 'food': given the human body, the construction of edible objects is not an entirely free construction—though people in different cultures do choose different things to eat and not eat, edibility itself is (in part) socially determined. In any case, the experience or expectation of hunger and the possibility of eating certain things work together to turn some of those things into human provisions, and it would seem to follow from this that provisions should be provided for those in need;

end p.170

food belongs to the hungry. (Who should do the providing and at whose expense are questions not so easily answered.) More complex and specific constructions will still be culturally relative: we save certain food for festive occasions or we burn it before

the gods or we waste it at extravagant banquets. But the original construction of things-that-are-food-for-the-hungry entails certain distributive rules that have, I suspect (this could be checked), always been recognized. Hoarders in time of famine act wrongly, for example, given what food is for.

I shall assume that reiterated social construction rather than diffusion from an authoritative centre is the preferred explanation for the appearance of identical or similar used and valued objects in different societies. There is no authoritative centre, no Jerusalem from which meanings go forth. The list of similarly constructed uses and values, then, constitutes what we might think of as a universal and objective morality—relative to social construction where construction repetitively takes the same form, relative to the prevailing argument where the same argument always prevails. We could go on to further explanations: if certain things-in-the-world are constructed in the same way again and again, presumably there is something in the nature of the things and/or something in the nature of the human agents that accounts for the construction. As the example of food suggests, the account is likely to be a naturalistic one. But I doubt that the list of similar constructions would be very long; nor would it include the complex and specific constructions that make for the thickness of moral life: food for eating would get on the list, not food for offerings. This is what it means to say that complexity is free: the more complex the construction the more room there is for cultural difference. Complex constructions do not turn up again and again, and they do not have plausible or satisfying naturalistic explanations.

There is no universal model for social construction, and the range of difference among actual outcomes is very wide. It might be argued, however, that this is so only because the constructive work takes place under a great variety of adverse and advantageous (mostly adverse) conditions. Only a common necessity, like the need for nourishment, makes for sameness. But if we imagine social construction in ideal conditions (and if ideal conditions are a single set of conditions), then we will get a model outcome, that is, a free construction that is at the same time the best construction. I am afraid that this is an impossible dream. For we can replace actual with hypothetical social construction only if we know, and not hypothetically, what conditions are ideal. And if we know that, then we already know the model outcome. We simply pull into our account of the imagined (ideal, original, natural) conditions all those materials, and only those materials, out of which we want society constructed. We might as well draw a blueprint of the good society and give up the idea that construction is free.

end p.171

## [3]

But if we do not have a model outcome, how can we ever criticize actual outcomes? This question is motivated, I think, by a misunderstanding of what an actual outcome is. Social construction is first of all conceptual in character. Holy altars and careers-open-to-talents are ideas, and the distributive norms that follow from them are also ideas. These ideas are never more than partially instantiated in the world; holiness and openness are more often than not honoured in the breach. What social critics commonly do is to hold the idea, or some more or less elaborated interpretation of the idea, over against the instance of the idea. Or, just as commonly, they hold some other idea or complex set of ideas, also the product of social construction, over against this idea and its instances. They say, if careers are open to talents, then why are they not open to the talents of Jews, or blacks, or women? Or, if our society is a

union of families or a democratic and co-operative community of citizens, how can we tolerate the intense competitiveness generated by careers-open-to-talents? Criticism of this sort depends on objective values, where objectivity is a true report on social meaning. The criticism itself, however, is not objectively true or false, for it also depends on an interpretation of social meaning, and interpretations are (except at the margins) only more or less persuasive and illuminating. But surely there are times when we want to say something stronger than this. We want to say that though the report is objectively true, the meaning is wrong (and not just *wrong for us*). Or we want to say that that is not the way we ought to think about altars or careers or whatever. Or even, that is not what an altar or a career *really is*.

Is it possible for a whole society to get things wrong in this fundamental way? This is the question that I have been postponing, and it is time now to try to deal with it. I want to make sure, though, that we understand exactly what the question is. Clearly, it is possible for individuals within a society to get things wrong, even fundamentally wrong, and it is also possible for groups of individuals to do the same thing. We should think of the Nazi case in these terms. It would strain the imagination to describe a fully elaborated world of complex meanings of a Nazi sort; in any case, no such world has ever existed. Within German or European or Western culture, the Nazis were an aberration, and in so far as we can make out their distributive principles—air for Aryans, gas for Jews—we can readily say that these are objectively wrong, immoral, monstrous. All the resources necessary for a judgement of this sort are already available, the products of a long history of social construction. It is a great mistake to make of the Nazis a hard case. The hard case comes when we begin to think that a long history of social construction has somehow gone awry. Consider, then, those societies where women (all women) seem to have been socially constructed as objects of exchange and where rules of exchange follow from the construction. I will not attempt an internal account of the exchanges

end p.172

that actually take place or of the meanings attached to them. Perhaps our understanding of an 'object' and even of an 'exchange' is not available to the participants. All I will say about the social construction at this point is that women are transferred among households, from one patriarchal jurisdiction to another, as if they were objects of exchange. What should we think of that? Are exchanges conducted in accordance with the rules objectively just?

There are a number of possibilities here: either women have or have not played a part in the constructive work; either they agree or they do not agree to its outcomes. Or, in a language less marked by our own conceptions of moral agency and lives-as-careers, they acquiesce or do not acquiesce, go along or do not go along with the outcomes. If they have played no part and do not go along, then the exchanges cannot be described as just. We can only report on the disagreement. Or perhaps we can say, as I would be inclined to say, that the exchanges are unjust, because in this case the objects are also human subjects, capable (as tables and lives are not) of going along or not, and the resistance of the constructed object nullifies the construction. It does not matter if the resistance is inarticulate, passive, hidden, or private. So long as we can in one way or another discover it, so long as we have probable cause to believe in its reality, the social construction fails.

The women involved may or may not be able to describe themselves as persons-engaged-in-social-construction; the vocabulary that I have deployed here is presumably not their own vocabulary. But we can see how their resistance 'works' in the world and why the construction fails. The unanimity or consensus principle plays

a part in explaining this failure, but something more is involved. Constructions of persons are not free—and not only in the obvious sense that we cannot make women or men into intercontinental ballistics missiles. The theory of social construction implies (some sort of) human agency and requires the recognition of women and men as agents (of some sort). We might say, looking at the idea itself as something we have made, that the construction of social-construction-with-human-agents has certain moral entailments. Among these is the right of subjective nullification, the right of the agents to refuse any given object status—as commodities, 'hands', slaves, or whatever.[2]

[2] A friend writes in criticism of this 'right' that some participants in *any* social system 'will resist or resent or reject the position or identity they are given . . . Not all resistance nullifies.' So we have to decide in each case whether the resistance is legitimate or not, and this requires standards that are extrinsic to the business of social construction. Yes, we will have to decide whether the resistance is just a way of avoiding particular obligations incurred by particular men or women. Agency itself cannot be denied, and promises made by agents are not subject to unilateral repudiation. But the denial of agency can always be repudiated, with the consequences described further on in the text. The right of nullification is simply the agent's right to claim her agency against any social process of objectification — and it does follow, I think, from the view of objectification as the work of human agents.

But what if women, for whatever reasons, actually agree that they are objects of exchange and live willingly by the rules of exchange? The phrase 'for

end p.173

whatever reasons' conceals a problem here, which philosophers who are quick with hypothetical examples are prone to ignore. What reasons could these women possibly have? We can easily see the reasons they might have for concealing disagreement, for stifling anger, for expressing resentment only in private or only in the company of other women. But if the experience of being treated as an object of exchange is the sort of experience we think it is, and if the women being exchanged are beings like us, what reasons could they have for agreeing? If, on the other hand, the experience does not match our understanding of it, and if these hypothetical women are beings of a different sort, then what is the philosophical issue here? What can we say, why should we want to say anything at all, about experiences and beings of which we are entirely ignorant?

Still, let us accept the hypothesis in its strongest form: here is a society in which women really do agree to the construction of themselves as objects of exchange. They do not agree because they have been brainwashed, because some chemical process or some hitherto unknown social process has turned them into moral robots or made servitude a reflex—for then, whatever they did or said, it would not constitute *agreeing*. Nor do they agree because they have no choice or because they are physically coerced or because they find themselves in desperate difficulties from which agreement is the only escape—like the women who sells herself into slavery in order to feed her children. For their agreement in these circumstances would not count as the construction of themselves as objects of exchange; it would represent only a reluctant and resentful acceptance of a pretence, a role that they could not refuse or escape. We must imagine reasons of a different sort: that the exchange of women brings some benefits to at least some women (even if the benefits are much greater for men); that it is only one part of a larger pattern of relationship, fitted to a system of beliefs, symbolically represented, ritually enacted and confirmed, handed down from mothers to daughters over many generations. So women accept the construction, even participate in it. What normative consequences follow?

One possible response is that no consequences follow, for agency is inalienable. This is Rousseau's argument, not applied by him to the self-subordination of women but obviously applicable: 'To renounce liberty is to renounce being a man, to

surrender the rights of humanity and even its duties . . . Such a renunciation is incompatible with man's nature; to remove all liberty from his will is to remove all morality from his acts.'[3]

[3] Jean Jacques Rousseau, *The Social Contract*, trans. G. D. H. Cole (London: Dent), 9.

Since human beings are agents by nature, and necessarily responsible for the worlds they make, the surrender of agency simply does not count; it is a gesture without effect. The argument from social construction is harder than this since it cannot refer to a universal and unconditioned *moral* agency. Now agents are socially produced, themselves involved in the production. It is still true that we (with our perceptions,

end p.174

understandings, theories) can recognize the-woman-who-is-an-object-of-exchange as a social construction with moral entailments only if we also recognize the same woman as a moral agent capable of agreeing (or not) to the construction. She can only be (morally) an object if she is simultaneously a subject confirming her object status. She is constituted by a contradiction—in so far as her subordinate status depends (morally) on her own agreement or acquiescence and is therefore inconsistent with subordination itself—and therein lies her freedom. She can never become just an object of exchange; the proof of this is that if she ever repudiates her object status, she is immediately and wholly a subject; the rules of exchange instantly lose their force. But so long as she confirms them (and even if her confirmation takes, as it commonly will, some other form than explicit agreement), they retain their force: she is partly an object.[4]

[4] I do not mean to say more than this. I do not mean that the construction is right because the woman confirms it, only that it is effective and consequential in the moral world. Her agreement (or acquiescence) had evidentiary, not legitimizing, force. Agreement makes for rightness only within moral systems where it is understood to do that, and in such systems it is commonly hedged with qualifications as to the freedom of subjects, the knowledge available to them, and so on. Hence women-who-are-objects-of-exchange can be exchanged justly or unjustly; but the objectification itself is not justified by their agreement.

There is nothing in the nature of a woman, or a man, that rules out contradictions of this sort. (The case is somewhat similar, I think, to Kant's means-ends polarity. We do not have to treat every person we meet, on every occasion, as an end-in-himself, for persons can agree to be means, like good civil servants who make themselves into the instruments of their fellow citizens, even surrendering some of their civil rights. But they can always resign from instrumentality.)

So long as the woman-who-is-an-object-of-exchange confirms her object status, the contradiction in her being is an objective contradiction. We can give a true account of it. Someone who claims that she is wholly an object is wrong. But so is someone who claims that she is wholly a subject (this would be roughly analogous to insisting that we must always, on every occasion, treat the civil servant as an end-in-himself). This last turn of the argument may well seem to many readers too relativistic, a surrender to what Marxists call 'false consciousness'. But once we have ruled out brainwashing and coercion, I see no morally acceptable way of denying the woman-who-is-an-object-of-exchange her own reasons and her own place in a valued way of life. That does not mean that we cannot argue with her, offering what we take to be better reasons for the repudiation of (what we take to be) object status. It does mean that, once the argument begins, she has to choose what *she* thinks are the better reasons, without any certainty as to which ones are objectively best. But we can say, and this seems to me all that we should want to say, that the choice is truly hers.

Is this not a plausible account of social construction seemingly gone awry? If nature provided a blueprint for construction, the process would not go awry as often as it does (seems to do). If something like gender equality were

end p.175

a simple entailment of the constructive process and every internal contradiction were ruled out a priori, then arguments for equality would be much easier than they are. When we encounter a complex set of social meanings, we enter a moral world, and it is no tribute to the creators of that world to deny its reality. Social meanings are constructed, accepted, and revised for reasons, and we have to engage those reasons. When we engage them from the outside, as in the case of women-who-are-objects-of-exchange, we are like missionaries preaching a new way of life to the natives, and we would do best, morally and politically, to try to work out what they find valuable or satisfying in their old way of life. More often, and more importantly, criticism of the old ways comes from within, as the result of long processes of social change. For the construction of objects-with-meanings and thus of moral worlds goes on and on; it is a continuous process in which we are all engaged. Conservatives try to freeze the process, but that effort is only one more instance of constructive activity (it has its reasons), one more expression of human agency. Criticism is no different in form.

Consider, for example, the construction of lives-that-are-careers-that-are-open-to-talents in a society where women are still objects of exchange. Over a period of time, institutions and practices take shape that make it possible (or necessary) for some members of the society, mostly men, to plan their lives-and, over the same period, lives of this sort are discussed, argued about, rendered meaningful. In the course of this process women will find that they have a new reason to repudiate their object status, for only by doing so can they undertake careers of their own. Some of them will seize upon this reason, and then more and more of them; at some point women-who-are-objects-of-exchange will be relics, sad memories, their agreement to subordination hard to understand. If a few people try to act out the rules of exchange, they will appear quixotic, not so much defenders of old ways as fools of time. In similar fashion, an archaeological guide might say to us: 'These were the holy altars of Xanadu, in the days when holiness reigned in the city.'

## [4]

I come back at last to my initial reservations. The kind of objectivity that I have attached to social meanings is probably not the kind that philosophers seeking objectivity are interested in. They are in search of things as they really are or as they must be. But I know very little about things as they really are, social construction apart. It is true (we can give objective reports) that particular constructions are reiterated in one social setting after another. The extent of the reiteration and the reasons for it-these are empirical matters. Exactly what evidence would lead us to say that such and such a construction could not or should never be otherwise, I do not know; in any case, that sort of evidence will not often be available. Interesting objects, all the more complex

end p.176

constructions, can always be otherwise. Tables need not be altars; lives need not be careers.

But is it not objectively true that meanings are always constructed? Men and women who claimed to have discovered meaning in nature, say, would surely be misreporting (misconstructing) their own activity-as if they were telling us that it was not Adam but God who named the animals. Even if no particular meaning were

objectively true or right or necessary, it would still be the case that the construction of meaning is a real process. Men and women really have made tables into holy altars and lives into careers. This has been the presupposition of my argument, and I have even pointed to its possible moral entailments; I do not want to run away from it now. But it is a strange 'objectivity' that leaves us adrift in a world we can only make and remake and never finish making or make correctly.

*Note*: In this paper I have tried to sketch an account of 'social meaning' that might underpin and uphold the theory of distributive justice presented a few years ago in my book *Spheres of Justice* (New York: Basic Books, 1983). My views on objectivity have been guided, stimulated, and provoked by recent philosophical and anthropological work that I can only acknowledge in a general way by listing a few crucial books: Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981) and *The Many Faces of Realism* (La Salle, IU.: Open Court, 1988); Nelson Goodman, *Ways of Worldmaking* (Indianapolis: Hackett, 1978); Thomas Nagel, *The View from Nowhere* (Oxford: Oxford University Press, 1986); Clifford Geertz, *Local Knowledge* (New York: Basic Books, 1986); and the essays collected in *Rationality and Relativism*, ed. Martin Hollis and Steven Lukes (Cambridge, Mass.: MIT Press, 1982) and in *Objectivity and Cultural Divergence*, ed. S. C. Brown (Cambridge: Cambridge University Press, 1984). I am grateful to Ruth Anna Putnam, Alan Wertheimer, John Goldberg, and Thomas Nagel, who read the paper in an earlier version and told me what was wrong with it. Martha Nussbaum suggested that I write the paper and provided its title, but she is responsible only for its existence, not for its argument.

end p.177

# Michael Walzer: Objectivity and Social Meaning

show chapter abstract and keywords
hide chapter abstract and keywords

**Ruth Anna Putnam**

That social meanings lack objectivity, that we are 'adrift' in the world, matters to Michael Walzer because it threatens to deprive us of any foundation from which to criticize injustice in cultures other than our own. So Walzer uses the theory of the social construction of meanings (values) itself to provide a new foundation for intercultural social criticism. I do not think that that attempt is successful; nor do I think it is needed. I shall modify Walzer's initial account of objectivity, I shall argue that the modified notion is the only notion of objectivity we need, and I shall conclude that we are not adrift in the world although we do keep making and remaking it.

Walzer's first sense of 'objective' is this: 'objectivity hangs (somehow) on the accommodation of the object by a knowing, inquiring subject.' What is involved here is a double accommodation. First, the conceptual scheme itself must accommodate the world: someone using the scheme 'friends, enemies, reading matter, and edible plants' is said to 'get the table wrong (objectivity wrong), or he would miss the table entirely, and deny its reality, and that would be a merely idiosyncratic (subjective) denial'. In contrast, 'the scientific perception of the world . . . makes the most insistent claim these days to be called an objective perception'. Second, given the framework, we cannot just decide how things are, our beliefs must accommodate the object.

The second point is, of course, correct. Given a conceptual scheme which allows for the distinction, we can distinguish between objective beliefs and all sorts of other things: imaginative constructions, illusions and hallucinations, neurotic fantasies, etc. In so far as these can be communicated at all, they depend on the shared conceptual scheme relative to which they fail to be objectively true.

But what is objectively true relative to one scheme may be objectively false relative to another-relative to the common-sense scheme this table is solid, relative to some scientific scheme it consists mainly of empty space. We, not individually but as communities, make and modify conceptual schemes; we do so in response to concerns which we have. So we have to specify the concerns relative to which a particular scheme accommodates the world. And we should remember that we never stand outside all conceptual schemes.

Standing within the common-sense scheme, we see that objective truth *relative to it* is most likely to be achieved by disinterested inquiry. That does not mean that the scientific schemes with which we replace the common-sense scheme as a result of these inquiries are more objective *sans phrase*. We can say only that given certain cognitive concerns, one or another scientific scheme

end p.178

is best; but none of these schemes can be called 'the scientific perception of the world.'

So I shall rephrase Walzer's first account of objectivity as follows: A belief (perception, recognition, understanding) is objective relative to a conceptual framework, if its truth or falsehood relative to that framework depends on how the world is rather than on what the knower thinks it is.

Walzer holds that his first concept of objectivity 'works, more or less well, for simple objects-in-the-world' but wonders whether 'it works at all for objects to which we assign use and value, objects that carry "social meanings' ". In contrast, I shall argue that this notion (at least as I have reformulated it) is the only notion of objectivity he needs. I am inclined to reject any dichotomy between objects-that-carry-social-meaning and simple-objects-in-the-world, or between objects that entail moral legislation and objects than do not. There is a difference between trees and tables: human beings make tables out of trees, human beings do not, in that sense of 'make', make trees; there is no comparable difference between tables and altars. We recognize trees as trees-we have the concept 'tree'-because they are important to us, important as resources and, sometimes, as obstacles. Trees answer many of our needs, including the needs which we meet by making tables. We make and recognize tables because we need objects with flat surfaces to put things on, to work on, to consecrate as altars. Finally, trees, tables, and altars entail moral legislation: trees are to be protected from various kinds of blight, tables are not to be chopped up for firewood, altars are not to be used as desks, etc.

Walzer points out that some social constructions (e.g. food) are reiterated in culture after culture and encourages us to think of the list of reiterated constructions 'as a universal and objective morality', not, to be sure, objective absolutely but 'relative to social construction where construction repetitively takes the same form'. Such universal constructions can be, he thinks, explained naturalistically; they respond to deep and pervasive human needs and give rise to a (to be sure, rebuttable) presumption that they best meet these needs. Conversely, objects/meanings which exist in only one, or a few, cultures may not be necessary for human life. So universal values, although socially constructed, appear to Walzer to be more like simple objects-in-the-world, for they accommodate the world that includes our needs, while the social constructions that are peculiar to a single culture are 'free'. And so universality of construction appears to him as a new sort of objectivity.

From my perspective, no new sense of objectivity is required. To say that interculturally shared values-food, for example-are objective while those peculiar to one culture-altar-to-the-Virgin or freedom-of-the-press-are not is to say that most

reports about something being food are objectively true or false relative to the intersection of most communities' conceptual schemes while reports about altars-to-the-Virgin or freedom-of-the-press are not simply because these concepts are not found in that intersection. But then objectivity is still the same old thing; why should the extent of agreement, or the extent

end p.179

of the possibility of agreement, change the ontological or epistemological status of a thing?

Interculturality is a form of intersubjectivity. Intersubjectivity is a familiar criterion of objectivity. A perception or judgement can be more or less widely shared; the more widely it is shared, the more objective it is said to be. We suspect perceptions and judgements that are tainted by interests, fears, and other subjective influences, we distrust what is seen from one point of view only. Beliefs so formed, we think, accommodate the knower more than the object. But in opting for intersubjectivity we are the losers. Intersubjectivity is not achieved by superposition of subjective views: that results in cubist paintings but is not a way in which we can manage to see the world, so why should it be an objective picture? Intersubjectivity can also be achieved by eliminating whatever fails to agree with all the subjective perceptions; that results in a thin and transparent rather than a thick and dense moral world; it results in the cold, clean world of science rather than the warm, messy world of everyday life. Does that really accommodate the object (the world and us, us-in-the-world) best?

The thickness and complexity of our moral world are given us by constructions that are peculiar to a given culture. Walzer's paper is essentially a search for an objectivity suitable to these. The question he asks is, I think, this: take a set of concepts that are peculiar to a given culture-a particular notion of the divine, of altars holy to a divine being, of worship required by that divinity, of behaviour that consecrates tables into altars, and of behaviour that would reverse that transformation. When we now say that a certain table is objectively an altar, are we 'accommodating the object'? Since there is nothing about tables that requires some of them to become altars, how could 'altar' accommodate or fail to accommodate a particular table? I would like to suggest that that is the wrong question; 'altar' accommodates the object table-which-has-undergone-the-prescribed-consecration-rites. That is why for both the members of the culture in question, and the anthropologist who studies them, the thing is objectively an altar. For the members of the culture, the complex I have described is part of their common-sense world; for the anthropologist that same complex is part of the culture under investigation.

Still, there are such things as religious strife. Because of that possibility, Walzer says that 'Social constructions must reflect a . . . consensus if there is ever to be an unqualified objectivity in our reports about them'. Why so? Dissension in the culture is no obstacle to objectivity in the reports of 'outsiders', of anthropologists; they can describe divided opinions as easily, as objectively, as unanimity. Dissension casts a shadow over the reports of 'insiders' only. Recognizing a particular table as an altar, treating it as altars in that culture are to be treated, will be a matter of course only if there exists some kind of religious consensus in that culture, a consensus which may range from everyone worshipping the same deity or deities in the same way to everyone respecting each other's different worship of different deities. But what happens when

end p.180

conflicting faith communities exist within one society in which there is no mutual respect? The members of each faith community will say of their altars that they are objectively altars, and of those of other communities perhaps that they are mere

tables, perhaps (therein lie the roots of religious strife) that they are abominations. The members of each group maintain that their perspective somehow accommodates the world while that of their opponents does not. Walzer maintains, in effect, that the claims of both groups are subjective. I want to take issue with that claim.

To be sure, there is nothing in the nature of a table that necessitates its becoming an altar; but there is something in our nature that gives rise to a sense of awe and that in turn frequently gives rise to the construction of altars. So one can ask whether altars (of religions in general, or of a particular religion) and the ways of life in which they are embedded accommodate the sense of awe and the need to give it expression. More radically, one can ask whether the sense of awe and all it brings in its train is a good thing in human lives or whether our children would be better off were we to raise them so as to be immune to it. Of course, those living within and affirming a religious tradition will not ask these questions; but my point here is simply that just as some conceptual schemes are worse than others because they frustrate our cognitive needs-they misclassify tables or deny their existence-so some sets of social meanings are worse than others because they frustrate other human needs-they misclassify not only objects but human beings, their needs, their emotions, etc., or they deny the existence or legitimacy of these. Social meanings are deeply embedded in our conceptual schemes-even scientific schemes involve the notions of relevance, warrant, truth, etc.-and value judgements are not neatly separable from descriptions. One has to know facts to know whether something is an altar, whether a certain type of behaviour is nepotism; but saying that something is an altar or nepotism is at the same time to evaluate and to prescribe. Walzer is exactly right when he says that these objects carry social meaning; he is wrong when he fears that that deprives them of a sort of objectivity that other things have.

Let me then return briefly to the contending groups, Protestants and Catholics, or, more dramatically, Christians and Aztecs. How could the Christians say (how can we say) that it was wrong for the Aztecs to practise human sacrifice unless they (we) recognize their practice as sacrifice? For if they say, simply, that the Aztecs committed murder, they have obviously failed to understand the Aztec practice. (Notice that even a pacifist does not say, simply, that soldiers are murderers, though he holds that taking a human life while fighting as a soldier is just as heinous, just as sinful, as ordinary murder, perhaps even more so.) But if they recognize the practice as sacrifice, on what basis may they criticize it? This raises the issue of the possibility of social criticism.

There is no puzzle concerning the possibility of social criticism from within, though I am puzzled by Walzer's claim that it depends on objective values but is not itself objective. I shall let that pass. There is a puzzle concerning

end p.181

criticism from without. Since there is no Archimedean point from which to raise such a critique, we can only offer it from within our own culture. But we are wary of cultural imperialism, of attempting to impose our way of life on others who are doing very well by their own lights; and, in some cases and in some sense, even by ours-I am thinking of the Amish that Hilary Putnam mentioned. There is also the sheer fact that often we do not understand one another; William James talks about that in his wonderful little essay 'On a Certain Blindness in Human Beings' (James, 1983). There is, finally, the fact that some things done in another culture strike us as so wrong that we would be doing wrong were we not to criticize it, where criticism is the least we can do and a precondition of doing anything more.

It is this last that prompts Walzer to develop his arguments against the practices of the patriarchal society, the society which passes women from one male jurisdiction to another and does not recognize women as agents. There are two possibilities: either the women accept the construction or they do not. If they do not, then the resources for internal criticism are at hand; but in both cases Walzer is prepared to offer critical arguments from 'outside', from his, and that means our, perspective. I want to look at these arguments.

In the case where the women do not accept their construction as objects, Walzer says, 'The theory of social construction implies (some sort of) human agency and requires the recognition of women and men as agents (of some sort) . . . the construction of social-construction-with-human-agents has certain moral entailments. Among these is . . . the right of the agents to refuse any given object status.' I find this argument puzzling. Of course, if one construes morality as a social construction, then one construes human beings as agents, although I do not see how it follows that one construes *all* human beings as agents; but I shall put that concern aside. Instead I wish to inquire after the point of the argument. How is this argument supposed to enable us to take sides *inter*culturally with the women in their *intra*cultural struggle? How is it supposed to enable us to rebut charges of cultural imperialism and to claim a transcultural objectivity for our view?

Walzer would reply, I think, that the claim that meanings are socially constructed is objectively true, that is, part of the scientific world view. If it follows from this that human beings are agents, then that claim too is objectively true in this strong sense of objective. And if it follows from this that they are not to be construed as mere objects, then that too will be objective, as objective as anything can be. But if, as I have argued, the scientific perspective does not carry with it a particularly robust objectivity-and here we deal not even with the hard sciences but with quite speculative explanations of very complex social phenomena-then Walzer's argument is simply one more liberal argument against patriarchy and the charge of cultural imperialism remains undefeated if the argument is to be exported to other cultures. But, of course, better a little cultural imperialism than a lot of oppression.

When Walzer considers a society in which women accept themselves as

end p.182

objects of exchange, his argument becomes more complicated and I become more puzzled. In this situation the woman is said to be constituted by a contradiction: on the one hand, she is (objectively) an object of exchange, on the other hand, the construction succeeds only if she is (objectively) a moral agent capable of consenting (or not) to the construction. But there is no contradiction here: she is an object of exchange relative to the world view of her society, which she shares. She is a moral agent relative to Michael Walzer's world view, which she does not share. There is no more contradiction here than in saying that this table is solid (in the common-sense view) and that it is mostly empty space (from the viewpoint of quantum physics). Perhaps I am obtuse. Perhaps the 'objective contradiction' in this woman's being, which exists as long as she affirms her object status, consists in this: the woman cannot simultaneously think of herself as an object and as someone who agrees to/disagrees with anything. But if she cannot do that, neither can Walzer. And if he cannot do it, how can he say: 'I see no morally acceptable way of denying the-woman-who-is-an-object-of-exchange her own reasons and her own place in a valued way of life'? How can he say that 'the choice is (truly) hers'? I am genuinely puzzled. This seems a far too complicated way of saying that before we preach a new way of life to the natives 'we would do best, morally and politically, to try to work

out what they find valuable or satisfying in their old way of life'. If we regard human beings as free and equal moral agents, then for us it follows that we are to respect them, that we are not to impose our morality on them but rather to reason with them. In every society some human beings are regarded as more than mere objects; that seems to me to be a basis from which one might begin to argue 'from inside' against patriarchy and other forms of oppression and enslavement. This is the sense in which intercultural meanings are important: they enable us to engage in intercultural criticism, to structure it as internal criticism. Or so it would be if ways of life were always based on consensus. When they are not, when there is oppression, when the dictator, or the oppressing group, is not amenable to reason, we face a different problem: the problem of how to deal with evil. Nothing is gained by calling it 'objective evil', and there is no single answer.

Recognizing that there is only one kind of objectivity, that all criticism is internal but that the resources for such criticism are greater than they would be if there were a dichotomy of simple-objects-in-the-world and objects-which-carry-social-meanings, should alleviate the feeling of being 'adrift in a world we can only make and remake and never finish making or make correctly'. We are not adrift, we are anchored by the world, by our needs, by how we have understood and made the world until now.

I want to conclude by saying briefly how all this connects with the concerns of this volume, that is, with the quality of life. Certain needs are universal, and one can establish 'scientifically' what will satisfy these needs — for example, the quantity and kind of food that will keep a person not just alive but healthy.

end p.183

As soon as we go beyond this, disagreements arise. What is seen as a necessity in one society—universal education up to a fairly advanced level, for example—is regarded as a luxury in another; what is valued highly here is perceived as a threat to a valued way of life there—pluralism, for instance. Measures that seem absolutely essential if the quality of life is to improve, such as limiting family size, may run counter to fundamental values. I take it that Michael Walzer and I agree that these are matters about which we should reason together; but because I believe that one cannot separate facts and values, cognitive and other concerns, because I believe that our descriptions and our evaluations of features of the world are inextricably intertwined, I am more optimistic than he concerning the possibilities of cross-cultural agreement. We live in one world not only because of the extensive interdependence of our economies, not only because all of us together may become victims of ecological or nuclear disasters; we live in one world because to a large extent we share our understanding of that world, and on the basis of that shared understanding we can come to further agreements.

**Bibliography**

James, William (1983). *Talks to Teachers on Psychology and to Students on Some of Life's Ideals*. Cambridge, Mass.: Harvard University Press.

end p.184

# Explanation and Practical Reason

show chapter abstract and keywords
hide chapter abstract and keywords

**Charles Taylor**

## [1]

Our modern conceptions of practical reason are shaped—I might say distorted—by the weight of moral scepticism. Even conceptions which intend to give no ground to scepticism have frequently taken form in order best to resist it, or to offer the least

possible purchase to it. In this practical reason falls into line with a pervasive feature of modern intellectual culture, which one could call the primacy of the epistemological: the tendency to think out the question of what something *is* in terms of the question of how it is *known*.

The place of what I call scepticism in our culture is evident. By this I do not mean just a disbelief in morality, or a global challenge to its claims—though the seriousness with which a thinker like Nietzsche is regarded shows that this is no marginal position. I am also thinking of the widespread belief that moral positions cannot be argued, that moral differences cannot be arbitrated by reason, that when it comes to moral values, we all just ultimately have to plump for the ones which feel/seem best to us. This is the climate of thought which Alasdair MacIntyre calls (perhaps a bit harshly) 'emotivist', [1]

[1] Alasdair MacIntyre, *After Virtue* (University of Notre Dame Press, 1981) ch. 2.

which at least ought to be called in some sense 'subjectivist'. Ask any undergraduate class of beginners in philosophy, and the majority will claim to adhere to some form of subjectivism. This may not correspond to deeply felt convictions. It does seem to reflect, however, what these students regard as the intellectually respectable option.

What underpins this climate? Some fairly deep metaphysical assumptions, when one gets down to it. But certainly, on the immediate level, it is fostered by the actual experience of moral diversity. On an issue like abortion, for instance, it does not seem to be possible for either side to convince the other. Protagonists of each side tend to think that their position is grounded in something self-evident. For some it just seems clear that the foetus is not a person, and that it is absurd to ruin the life of some being who undeniably has this status in order to preserve it. For others it is absolutely clear that the foetus is both a life and human, and so terminating it cannot be right unless murder is. Neither side can be budged from these initial intuitions, and once one accepts either one the corresponding moral injunctions seem to follow.

If the seeming helplessness of reason tells us something about its real limits, then a worrying thought arises: what if some people came along who just failed to share our most basic and crucial moral intuitions? Suppose some people thought that innocent human beings could be killed in order to achieve some advantage for the others, to make the world more aesthetically pleasing, or something

end p.208

of the sort? And have we not actually experienced people who stepped way outside the bounds of our core morality: the Nazis, for instance? Is reason as powerless before such people as it seems to be to arbitrate the dispute about abortion? Is there no way to show them they are wrong?

This is where our implicit model of practical reason begins to play an important role. If 'showing them' means presenting facts or principles which they cannot but accept, and which are sufficient to disprove their position, then we are indeed incapable of doing this. But one could argue that this is a totally wrong view of practical reason. Faced with an opponent who is *unconfusedly* and *undividedly* convinced of his position, one can indeed only hope to move him rationally by arguing from the ground up, digging down to the basic premisses we differ on, and showing him to be wrong there. But is this really our predicament? Do we really face people who quite lucidly reject the very principle of the inviolability of human life?

In fact, this does not seem to be the case. Intellectual positions put forward to justify behaviour like that of the Nazis — to the extent that any of their ravings justify this appellation at all — never attack the ban on murder of conspecifics frontally. They are always full of special pleadings: for example, that their targets are not really of

the same species, or that they have committed truly terrible crimes which call for retaliation, or that they represent a mortal danger to others, and so on. This kind of stuff is usually so absurd and irrational that it comes closer to raving than to reason. And, of course, with people like this reason is in fact ineffective as a defence. But this is not to say that reason is powerless to show them they are wrong. Quite the contrary. The fact that these terrible negations of civilized morality depend so much on special pleading, and of a particularly mad and irrational sort, rather suggests that there are limits beyond which *rational* challenges to morality have great trouble going.

This might indicate a quite different predicament of, and hence task for, practical reasoning. Its predicament would be defined by the fact that there are limits to what people can unconfusedly and undividedly espouse; so that, for instance, in order to embrace large-scale murder and mayhem, they have to talk themselves into some special plea of the sort mentioned above, which purports to square their policies with some recognized version of the prohibition against killing. But these pleas are vulnerable to reason, and in fact barely stand up to the cold light of untroubled thought.

The task of reasoning, then, is not to disprove some radically opposed first premiss (e.g. killing people is no problem), but rather to show how the policy is unconscionable on premisses which both sides accept, and cannot but accept. In this case, its job is to show up the special pleas.

On this model — to offer here at any rate a first approximation — practical argument starts off on the basis that my opponent already shares at least some of the fundamental dispositions towards good and right which guide me. The error comes from confusion, unclarity, or an unwillingness to face some of what he cannot lucidly repudiate; and reasoning aims to show up this error. Changing



someone's moral view by reasoning is always at the same time increasing his self-clarity and self-understanding.

There are two quite different models of practical reason, let us call them the apodeictic and the *ad hominem*, respectively. I think that John Stuart Mill was making use of a distinction of this kind, and opting for the second, in his famous (perhaps notorious) remarks in *Utilitarianism*. 'Questions of ultimate ends are not amenable to direct proof', he avers, and yet 'considerations may be presented capable of determining the intellect either to give or to withhold its assent to the doctrine [namely, of utility]; and this is the equivalent to proof'.[2] This may sound like someone trying to squirm his way out of a contradiction, but the distinction is quite clear and sound. You cannot argue someone into accepting an ultimate end, utility or any other, if he really rejects it. But in fact, the whole case of utilitarians is that people do *not* reject it, that they all really operate by it, albeit in a confused and therefore self-defeating fashion. And this is why there may be 'considerations . . . capable of determining the intellect'. In fact, Mill shows us what he thinks these are in chapter 4, where he goes on to argue that what people in fact desire is happiness.[3]

[3] *Ibid.*, 34.

The appeal is to what the opponent already seeks, a clear view of which has to be rescued from the confusions of intuitionism.

But, it might be thought, this invocation of Mill is enough to discredit the *ad hominem* model irremediably. Is this not exactly where Mill commits the notorious naturalistic fallacy, arguing from the fact that men desire happiness to its desirability, on a

glaringly false analogy with the inference from the fact that men see an object to its visibility? [4]

[4] Ibid.

Derisive hoots have echoed through philosophy classes since G.E. Moore, as first-year students cut their teeth on this textbook example of a primitive logical error. There is no doubt that this argument is not convincing as it stands. But the mistake is not quite so simple as Moore claimed. The central point that the Moorean objection indicates is the special nature of moral goals. This is a phenomenon which I have tried to describe with the term 'strong evaluation'. [5]

[5] Cf. e.g. 'What is Human Agency?' in *Human Agency and Language* (Cambridge University Press, 1985).

Something is a moral goal of ours not just in virtue of the fact that we are *de facto* committed to it. It must have this stronger status, that we see it as demanding, requiring, or calling for this commitment. While some goals would have no more claim on us if we ceased desiring them, for example, my present aim of having a strawberry ice cream after lunch, a strongly evaluated goal is one such that, were we to cease desiring it, *we* would be shown up as insensitive or brutish or morally perverse.

That is the root of our dissatisfaction with Mill's argument here. We feel that just showing that we always desire something, even that we cannot help

end p.210

desiring it, by itself does nothing to show that we *ought* to desire it, that it is a moral goal. Supposing I were irremediably addicted to smoking. Would that prove that I ought to smoke? Clearly not. We understand smoking from the beginning as a weakly evaluated end. We have to distinguish between showing of some end that we cannot help desiring it, and showing that all our strong evaluations presuppose it, or involve it, once we overcome our confusions about them. In the second case, we would have demonstrated that we cannot be lucid about ourselves without acknowledging that we value this end. This is the sense in which it is inescapable, not after the fashion of some *de facto* addiction. Whereas addictions are rightly declared irrelevant to moral argument, except perhaps negatively, the proof of inescapable commitment is of the very essence of the second, *ad hominem* mode of practical reasoning, and is central to the whole enterprise of moral clarification.

Mill plainly had some intuition to this effect when he deployed the argument in *Utilitarianism*. One of the things he was trying to show is that everyone else's commitment collapses into his. But the argument is botched because of a crucial weakness of the doctrine of utility itself, which is based on the muddled and self-defeating attempt to do away with the whole distinction between strong and weak evaluation. The incoherence of Mill's defence of the 'higher' pleasures, on the grounds of mere *de facto* preference by the 'only competent judges', [6]

[6] *Utilitarianism*, 8-11.

is also a testimony to the muddles and contradictions which this basically confused theory gives rise to.

But this does point to one of the most important roots of modern scepticism. We can already see that people will tend to despair of practical reason to the extent that they identify its mode of argument as apodeictic. This clearly sets an impossible task for it. But this will be accepted to the degree that the alternative, *ad hominen* mode of argument appears inadequate or irrelevant. And this it is bound to do, as long as the distinction between strong and weak evaluation is muddled over or lost from sight.

This confusion can only breed bad arguments *à la* Bentham and Mill, and these, once denounced, discredit the whole enterprise.

But utilitarianism does not come from nowhere. The whole naturalist bent of modern intellectual culture tends to discredit the idea of strong evaluation. The model for all explanation and understanding is the natural science which emerges out of the seventeenth-century revolution. But this offers us a neutral universe: it has no place for intrinsic worth, or goals which make a claim on us. Utilitarianism was partly motivated by the aspiration to build an ethic which was compatible with this scientific vision. But to the extent that this outlook has a hold on the modern imagination, even beyond the ranks of utilitarianism, it militates in favour of accepting the apodeictic model, and hence of a quasi-despairing acquiescence in subjectivism.

The link between naturalism and subjectivism is even clearer when looked

end p.211

at from another angle. The seventeenth-century scientific revolution destroyed the Platonic-Aristotelian conception of the universe as the instantiation of Forms, which defined the standards by which things were to be judged. The only plausible alternative construal of such standards in naturalist thought was as projections of subjects. They were not part of the fabric of things, but rather reflected the way subjects react to things, the pro- or con-attitudes they adopt. Now perhaps it is a fact that people's attitudes tend to coincide — a happy fact, if true; but this does nothing to show that this point of coincidence is more right than any other possible one. [7]

[7] See J.L. Mackie, *Ethics* (Penguin Books, 1977), for an excellent example of the consequences of uncompromisingly naturalist thinking.

The opposition to this naturalist reduction has come from a philosophical stance which might in a broad sense be called 'phenomenological'. By that I mean a focus on our actual practices of moral deliberation, debate, and understanding. The attempt is to show, in one way or another, that the vocabularies we need to explain human thought, action, and feeling, or to explicate, analyse, and justify ourselves or each other, or to deliberate on what to do, all inescapably rely on strong evaluation. Or, to put it negatively, that the attempt to separate out a language of neutral description, which combined with commitments or pro/con-attitudes might recapture and make sense of our actual explanations, analysis, deliberations, etc., leads to failure and will always lead to failure. It seems to me that this case has been convincingly made out, in a host of places. [8]

[8] Cf. e.g. John McDowell, 'Virtue and Reason', *The Monist*, 62 (1979), 331-50; and also MacIntyre, *After Virtue*. Bernard Williams makes the case very persuasively in his *Ethics and the Limits of Philosophy* (Harvard University Press, 1985), ch. 8 See also my 'Neutrality in Political Science', in *Philosophy and the Human Sciences* (Cambridge University Press, 1985).

This kind of argument is, of course, not only a justification of the very foundation of the *ad hominem* mode of reasoning, but an example of it. It tries to show us that in all lucidity we cannot understand ourselves, or each other, cannot make sense of our lives or determine what to do, without accepting a richer ontology than naturalism allows, without thinking in terms of strong evaluation. This might be thought to beg the question, establishing the validity of a mode of argument through a use of it. But the presumption behind this objection ought to be challenged: what in fact ought to trump the ontology implicit in our best attempts to understand or explain ourselves? Should the epistemology derived from natural science be allowed to do so, so that its metaphysical bias in favour of a neutral universe overrules our most lucid self-understandings in strongly evaluative terms? But does this not rather beg the crucial question, namely, whether and to what extent human life is to be explained in terms

modelled on natural science? And what better way to answer this question than by seeing what explanations actually wash?

end p.212

## [2]

Enough has been said in the above, I hope, to show that one of the strongest roots of modern scepticism and subjectivism in regard to ethics is the naturalist temper of modern thought. This tends to discredit in advance the *ad hominem* mode of argument, which actually might hold out the hope of settling certain moral issues by reason, and leaves only the apodeictic model in the field, which clearly sets an impossible standard. Within a human situation inescapably characterized in strongly evaluative terms, we can see how argument aimed at self-clarification might in principle at least bring agreement. In a neutral universe, what agreement there is between attitudes seems merely a brute fact, irrelevant to morals, and disagreement seems utterly inarbitrable by reason, bridgeable only by propaganda, arm-twisting, or emotional manipulation.

But this analysis brings to mind another source of modern scepticism, constituted by the independent attractions of the apodeictic model itself. This is where we really measure the tremendous hold of epistemology over modern culture.

This model emerges *pari passu* with and in response to the rise of modern physical science. As we see it coming to be in Descartes and then Locke, it is a foundationalist model. Our knowledge claims are to be checked, to be assessed as fully and responsibly as they can be, by breaking them down and identifying their ultimate foundations, as distinct from the chain of inferences which build from these towards our original unreflecting beliefs. This foundationalist model can easily come to be identified with reason itself. Modern reason tends to be understood no longer substantively but procedurally, and the procedures of foundationalism can easily be portrayed as central to it. But from the foundationalist perspective, only the apodeictic mode of reasoning is really satisfactory; the appeal to shared fundamental commitment seems simply a recourse to common prejudices. The very Enlightenment notion of prejudice encapsulates this negative judgement.

This brings us to another aspect. Foundationalist reasoning is meant to shake us loose from our parochial perspective. In the context of seventeenth-century natural science this involved in particular detaching us from the peculiarly human perspective on things. The condemnation of secondary qualities is the most striking example of this move to describe reality no longer in anthropocentric but in 'absolute' terms.[9]

[9] I borrow the term from the interesting discussion in Bernard Williams's *Descartes* (Penguin Books, 1978), 66-7. See also Thomas Nagel, *The View from Nowhere* (Oxford University Press, 1986). I have discussed this issue in my 'Self-interpreting Animals', in *Human Agency and Language*.

But if the canonical model of reasoning involves breaking us free from our perspective as much as possible, then the *ad hominem* mode cannot but appear inferior, since by definition it starts from what the interlocutor is already committed to. And here a particularly important consideration comes into play. Starting from where your interlocutor is not only seems an inferior mode of

end p.213

reasoning in general, but it can be presented as a peculiarly bad, and indeed vicious, form of practical reason. For all those whose instinct tells them that the true demands of morality require radical change in the way things are, and the way people have been trained to react to them, starting from the interlocutor's standpoint seems a formula for conservatism, for stifling at the start all radical criticism, and foreclosing all the really important ethical issues.

This has always been one of the strongest appeals of utilitarianism, and one of the greatest sources of self-congratulation on the part of partisans of utility. It is not only that their theory seems to them the only one consonant with science and reason, but also that they alone permit of reform. J.S. Mill argues against views based on mere 'intuition' that they freeze our *axiomata media* for ever, as it were, and make it impossible to revise them, as mankind progresses and our enlightenment increases. 'The corrolaries from the principle of utility, . . . admit of indefinite improvement, and, in a progressive state of the human mind, their improvement is perpetually going on.'[10]

[10] *Utilitarianism*, 24.

Here, then, is a source of modern scepticism and subjectivism which is as powerful as naturalism, and tends to operate closely in tandem with it, namely, the belief that a critical morality, by its very nature, rules out the *ad hominem* mode of practical reasoning. Naturalism and the critical temper together tend to force us to recognize the apodeictic mode as the only game in town. The obvious severe limitations of this mode in the face of ethical disagreement then push us towards a half-despairing, half-complacent embracing of an equivocal ethical subjectivism.

I have tried to show elsewhere that this identification of the demands of critical morality with a procedural understanding of reason and the apodeictic mode is deeply mistaken.[11]

[11] 'Justice after Virtue', unpublished MS.

But erroneous or not, it has clearly been immensely influential in our intellectual culture. One can see this in the way people unreflectingly argue in terms of this model.

Discuss the question of arbitrating moral disputes with any class, graduate or undergraduate, and very soon someone will ask for 'criteria'. What is aimed at by this term is a set of considerations such that, for two explicitly defined, rival positions *X* and *Y*, (1) people who unconfusedly and undividedly espouse both rival positions *X* and *Y* have to acknowledge them, and (2) they are sufficient to show that *Y* is right and *X* is wrong, or vice versa. It is then driven home, against those who take an upbeat view of practical reason, that for any important moral dispute, no considerations have both *a* and *b*. If the rift is deep enough, things which are *b* must fail of *a*, and vice versa.

The problem lies with the whole unreflecting assumption that criteria in this sense are what the argument needs. We shall see, as we explore this further, that this assumption, as it is usually understood in the context of foundationalism, amounts to ruling out the most important and fruitful forms of the *ad hominem* mode.

end p.214

But this whole assumption that rational arbitration of differences needs criteria has become very problematic, not only for practical reasons. It is a notorious source of puzzlement and sceptical challenges in the history of science as well. It is some underlying assumption of this kind that has driven so many people to draw sceptical conclusions from the brilliant work of Thomas Kuhn (conclusions to which Kuhn himself has sometimes been drawn, without ever succumbing to them). For what Kuhn persuasively argued was the 'incommensurability' of different scientific outlooks which have succeeded each other in history. That is, their concepts are non-intertranslatable, and—what is even more unsettling—they differ as to what features or considerations provide the test of their truth. The considerations each recognizes as having *b* are diverse. There are no criteria. And so the radical inference of a Feyerabend has seemed widely plausible: 'anything goes'.

But, as Alasdair MacIntyre has argued in another work, [12]

[12] 'Epistemological Crises, Dramatic Narrative, and the Philosophy of Science', *The Monist*, 60 (1977), 453-72.

it is clear that what needs revision here is our metatheory of scientific reasoning, rather than, for example, our firmly established conviction that Galileo made an important step forward relative to Aristotelian physics. The blind acceptance of a foundationalist, apodeictic model of reasoning is perhaps just as damaging here as in ethics. Calling to mind how inadequate the model is here can both help to weaken its hold on us in general, and allow us to see more exactly what is truly peculiar to practical reason.

MacIntyre argues very convincingly that the superiority of one scientific conception over another can be rationally demonstrated, even in the absence of what are normally understood as criteria. These are usually seen as providing some externally defined standard, against which each theory is to be weighted independently. But what may be decisive is that we are able to show that the *passage* from one to the other represents a gain in understanding. In other words, we can give a convincing narrative account of the passage from the first to the second as an advance in knowledge, a step from a less good to a better understanding of the phenomena in question. This establishes an asymmetrical relation between them: a similarly plausible narrative of a possible transition from the second to the first could not be constructed. Or, to put it in terms of a real historical transition, portraying it as a *loss* in understanding is not on. [13]

[13] For a parallel notion of the asymmetrical possibilities of transition, this time applied to practical reason, see Ernst Tugendhat's notion of a possible *Erfahrungsweg* from one position to another, in *Selbstbewusstsein und Selbstbestimmung* (Frankfurt, 1979), 275.

What I want to take from this is the notion that one can sometimes arbitrate between positions by portraying *transitions* as gains or losses, even where what we normally understand as decision through criteria—*qua* externally defined standards—is impossible. I should like to sketch here three argument forms, in ascending order of radical departure from the canonical, foundationalist mode.

1. The first takes advantage of the fact that we are concerned with transitions,



that the issue here is a comparative judgement. On the standard, unreflecting assumptions of foundationalism, comparative judgements are usually secondary to absolute ones. Rival positions *X* and *Y* are checked against the facts, and one is shown to be superior to the other because it predicts or explains certain facts which the other does not. The comparative judgement between the two is based on absolute judgements concerning their respective performance in the face of reality, just as in a football game, the comparative verdict: team *X* won, is founded on two absolute assessments: team *X* scored three goals, and team *Y* scored two. The role of criteria here is taken by facts, observations, protocols, or perhaps by standards to be applied to explanations of facts—such as elegance or simplicity. The most popular theory of scientific reasoning with this traditional structure, Popper's, does indeed resemble the eliminative rounds in a championship. Each theory plays the facts, until it suffers defeated, and is then relegated.

But, as MacIntyre shows, comparative reasoning can draw on more resources than this. What may convince us that a given transition from *X* to *Y* is a gain is not only or even so much how *X* and *Y* deal with the facts, but how they deal with each other. It may be that from the standpoint of *Y*, not just the phenomena in dispute, but also the history of *X*, and its particular pattern of anomalies, difficulties, makeshifts, and breakdowns, can be greatly illuminated. In adopting *Y*, we make better sense not just

of the world, but of our history of trying to explain the world, part of which has been played out in terms of *X*.

The striking example, which MacIntyre alludes to, is the move from Renaissance sub-Aristotelian to Galilean theories of motion. The Aristotelian conception of motion, which entrenched the principle: no motion without a mover, ran into tremendous difficulty in accounting for 'violent' motion, for example, the motion of a projectile after it leaves the hand, or the cannon mouth. The Paduan philosophers and others looked in vain for factors which could play the continuing role of movers in pushing the projectile forward. What we now see as the solution did not come until theories based on inertia altered the entire presumption of what needs explaining: continued rectilinear (or for Galileo circular) motion is not an explanandum.

What convinces us still today that Galileo was right can perhaps be put in terms of the higher 'score' of inertial theories over Aristotelian ones in dealing with the phenomena of motion. After all this time, the successes of the former are only too evident. But what was and is also an important factor—and which obviously bulked relatively larger at the time—is the ability of inertial theories to make sense of the whole pattern of difficulties which had beset the Aristotelians. The superiority is registered here not simply in terms of their respective 'scores' in playing 'the facts', but also by the ability of each to make sense of itself and the other explaining these facts. Something more emerges in their stories about each other than is evident in a mere comparison of their several performances. This show an asymmetrical relation between them: you

end p.216

can move from Aristotle to Galileo realizing a gain in understanding, but not vice versa.

2. This is still not a radical departure from the foundational model. True, decisive criteria are not drawn from the realm of facts, or of universally accepted principles of explanation. But the crucial considerations are still accessible to both sides. Thus the pre-Galileans were not unaware of the fact that they had a puzzling problem with violent motion. To speak Kuhnian language, this was an 'anomaly' for them, as their intellectual perplexity and the desperate expedients they resorted to testify. The decisive arguments are transitional—they concern what each theory has to say about the other, and about the passage from its rival to itself—and this takes us beyond the traditional way of conceiving validation, both positivist and Popperian. But in the strict sense of our definition above, there are still criteria here, for the decisive considerations are such that both sides must recognize their validity.

But, it can be argued, if we look at the seventeenth-century revolution from a broader perspective, this ceases to be so. Thus if we stand back and compare the dominant models of science before and after the break, we can see that different demands were made on explanation. The notion of a science of nature, as it came down from Plato, and especially from Aristotle, made explanation in terms of Forms (*eidē* or species) central, and beyond that posited an order of Forms, whose structure could be understood teleologically, in terms of some notion of the good, or of what ought to be. Principles like that of plenitude, which Lovejoy identifies and traces, make sense on that understanding: we can know beforehand, as it were, that the universe will be so ordered as to realize the maximum richness.[14]

[14] Arthur Lovejoy, *The Great Chain of Being* (Harper Torchbook Edition, 1960).

Similarly, explanations in terms of correspondences are possible, since it follows from the basic conception that the same constellation of ideas will be manifested in all the different domains.

Now if science consists of a grasp of order of this kind, then the activity of explaining why things are as they are (what we think of as science) is intrinsically linked to the activity of determining what the good is, and in particular how human beings should live through attuning themselves to this order. The notion that explanation can be distinct from practical reason, that the attempt to grasp what the world is like can be made independent of the determination of how we should stand in it, that the goal of understanding the cosmos can be uncoupled from our attunement to it, made no sense to the pre-modern understanding.

But notoriously the seventeenth-century revolution brought about an uncoupling of just this kind. The turn to mechanism offered a view of the universe as neutral; within it cause-effect relations could be exploited to serve more than one purpose. Galileo and his successors, we might say, turned towards an utterly different paradigm of explanation. If scientific explanation can always be roughly understood as in some sense rendering the puzzling comprehensible by showing how the phenomenon to be explained flows from mechanisms or

end p.217

modes of operation which we understand, then the seventeenth century sees a massive shift in the kind of understanding which serves as the basic reference point.

There is certainly one readily available mode of human understanding which the Platonic-Aristotelian tradition drew on. We are all capable of understanding things in terms of their place in a meaningful order. These are the terms in which we explain the at first puzzling behaviour of others, or social practices which at first seem strange, or some of the at first odd-seeming details of a new work of art, or the like.

In another quite different sense of 'understanding', we understand an environment when we can make our way about in it, get things done in it, effect our purposes in it. This is the kind of understanding a garage mechanic has, and I unfortunately lack, of the environment under the hood of my car.

One of the ways of describing the seventeenth-century revolution in science is to say that one of these paradigms of understanding came to take the place of the other as the basic reference point for scientific explanation of nature.[15]

[15] I am borrowing here from Max Scheler's analysis in his essays 'Soziologie des Wissens' and 'Erkenntnis und Arbeit'.

But this has as an ineluctable consequence the distinction of explanation from practical reason I mentioned above. Only the first type of understanding lends itself to a marriage of the two.

But once we describe it in this way, the scientific revolution can be made to seem not fully rationally motivated. Of course, we all accept today that Galileo was right. But can we *justify* that preference in reason? Was the earlier outlook shown to be inferior, or did its protagonists just die off? If you ask the ordinary person today for a quick statement of why he holds modern science to be superior to the pre-modern, he will probably point to the truly spectacular technological pay-off that has accrued to Galilean science. But this is where the sceptic can enter. Technological pay-off, or the greater ability to predict and manipulate things, is certainly a good criterion of scientific success on the post-Galilean paradigm of understanding. If understanding is knowing your way about, then modern technological success is a sure sign of progress in knowledge. But how is this meant to convince a pre-Galilean? He will be operating with a quite different paradigm of understanding, to which manipulative capacity is irrelevant, which rather proves itself through a different ability, that of discovering our proper place in the cosmos, and finding attunement with it. And, it

could be argued, modern technological civilization is a spectacular failure at *this*, as ecological critics and green parties never tire of reminding us.
Is the argument then to be considered a stand-off between the two, judged at the bar of reason? Here the sceptical spin-off from Kuhn's work makes itself felt. Once one overcomes anachronism and comes to appreciate how different earlier theories were, how great the breaks are in the history of knowledge—and this has been one of the great contributions of Kuhn's work—then it can appear that no

end p.218

*rational* justification of the transitions is possible. For the considerations that each side takes into account diverge. Each theory carries with it its own built-in criteria of success—moral vision and attunement in one case, manipulative power in the other—and is therefore invulnerable to the other's attack. In the end, we all seem to have gone for manipulative power, but this has to be for some extra-epistemic consideration, that is, not because this mode of science has been shown superior as *knowledge*. Presumably, we just like that pay-off better. In terms of my earlier discussion, what we lack here are 'criteria': there are no decisive considerations which *both* sides must accept.
Some people are driven by their epistemological position to accept some account of this kind. [16]

[16] For instance, this seems to be implicit in Mary Hesse's view: see her 'Theory and Value in the Social Sciences', in C. Hookway and P. Pettit (eds.), *Action and Interpretation* (Cambridge University Press, 1979). She speaks there of prediction and control as 'pragmatic' criteria of scientific success (p. 2).

But this seems to me preposterous. Once more, the account can appear plausibly only because it fails to consider the transition between the two views. It sees each as assessing a theory's performance in face of reality by its own canons. It does not go further and demand of each that it give an account of the existence of the other: that is, not just explain the world, but explain also how this other, rival (and presumably erroneous) way of explaining the world could arise.
Once one makes this demand, one can appreciate the weakness of pre-Galilean science. There is a mode of understanding which consists of knowing one's way about. This is universally recognized. In making another mode a paradigm for scientific explanation, pre-Galilean science drew on a set of assumptions which entailed that this manipulative understanding would never have a very big place in human life. It always allowed for a lower form of inquiry, the domain of 'empirics', who scramble around to discover how to achieve certain effects. But the very nature of the material embodiment of Forms, as varying, approximate, never integral, ensured that no important discoveries could be made here, and certainly not an exact and universal body of findings. Consequently, the very existence of such a body of truths, and the consequent spectacular manipulative success, represents a critical challenge for pre-modern science. Indeed, it is difficult to see how it could meet this challenge. On its basic assumptions, modern science should not have got off the empiric's bench, emerged from the dark and smelly alchemist's study to the steel-and-glass research institutes that design our lives.
So the problem is not some explanatory failure on its own terms, not some nagging, continuing anomaly, as in the narrower issue above of theories of motion; it is not that pre-Galilean science did not perform well enough by its own standards, or that it did not have grounds within itself to downgrade the standards of its rivals. If we imagine the debate between the two theories being carried on timelessly on Olympus, before any actual results are obtained by one or the other, then it is indeed a stand-off. But what the earlier science cannot

end p.219

explain is the very success of the later *on the later's own terms*. Beyond a certain point, you just cannot pretend any longer that manipulation and control are not relevant criteria of scientific success. Pre-Galilean science died of its inability to explain or assimilate the actual success of post-Galilean science, where there was no corresponding symmetrical problem. And this death was quite rationally motivated. On Olympus the grounds would have been insufficient; but, faced with the actual transition, you are ultimately forced to read it as a gain. Once again, what looks like a stand-off when two independent, closed theories are confronted with the facts, turns out to be conclusively arbitrable in reason when you consider the transition. [17]

[17] I have discussed this point at somewhat greater length in my 'Rationality', in M. Hollis and S. Lukes (eds.), *Rationality and Relativism* (Cambridge, Mass., 1982), reprinted in my *Philosophy and the Human Sciences* (Cambridge University Press, 1985).

I have been arguing in the above that the canonical, foundationalist notion of arbitrating disputes through criteria generates scepticism about reason, which disappears once we see that we are often arguing about transitions. And we have seen that this scepticism affects some of the more important transitions of science just as much as it does the disputes of morality, and for the same reason, namely, the seeming lack of common criteria. In particular, it tends to make the history of science seem less rational than it has in fact been.

The second case is in a sense a more radical departure from the canonical model than the first. For the defeat does not come from any self-recognized anomaly in the vanquished theory. Nevertheless, there was *something* which the losing theory had to recognize outside the scope of its original standards, namely, that the very success of mechanistic science posed a problem. If we ask why this is so, we are led to recognize a human constant, namely, a mode of understanding of a given domain, *D*, which consists in our ability to make our way about and effect our purposes in *D*. We might borrow a term from Heidegger, and call this understanding as we originally have it prior to explicitation or scientific discovery 'pre-understanding'. One of the directions of increasing knowledge of which we are capable consists in making this pre-understanding explicit, and then in extending our grasp of the connections which underlie our ability to deal with the world as we do. Knowledge of this kind is intrinsically linked with increased ability to effect our purposes, with the acquisition of potential recipes for more effective practice. In some cases, it is virtually impossible to extend such knowledge without making new recipes available; and an extension of our practical capacities is therefore a reliable criterion of increasing knowledge. Because of these links between understanding and practical ability, we cannot deny whatever increases our capacities its title as a gain in knowledge in some sense. We can seek to belittle its significance, or deem it to be by nature limited, disjointed, and lacunary, as Plato did. But then we have to sit up and take notice when it manages to burst the bounds we set for it—and this is what has rendered the transition to Galilean science a rational one.

end p.220

The mediating element is something deeply embedded in the human life form, of which we are all implicitly aware, and which we have to recognize when made explicit: the link between understanding (of a certain kind) and practical capacity. But then is not the predicament of reason here coming to look analogous to the description I offered above of moral disputes? The task is not to convince someone who is undividedly and unconfusedly attached to one first principle that he ought to shift to an entirely different one. So described, it is impossible; rather, we are always

trying to show that, granted what our interlocutor already accepts, he cannot but attribute to the acts or policies in dispute the significance we are urging on him. Now here it has been a question of altering the first principles of science — the paradigms of understanding underlying it and the standards of success. And we can see a rational path from one to the other, but only because in virtue of what the pre-Galilean already accepts he cannot but recognize the significance of Galilean science's massive leap forward. No more in one case than in the other is it a question of radical conversion from one ultimate premiss to the other. That would indeed be irrational. Rather we show that the pre-Galilean could not undividedly and unconfusedly repudiate the deliverance of post-Galilean science as irrelevant to the issue that divides them.

Perhaps, then, those ultimate breakpoints we speak of as 'scientific revolutions' share some logical features with moral disputes. They are both rendered irrational and seemingly inarbitrable by an influential but erroneous model of foundationalist reasoning. To understand what reason can do in both contexts, we have to see the argument as being about transitions. And, as the second case makes plain, we have to see it as appealing to our implicit understanding of our form of life.

This brings to the fore one of the preconceptions which has bedevilled our understanding here and fostered scepticism. On the standard foundationalist view, the protagonists are seen as closed explicit systems. Once one has articulated their major premisses, it is assumed that all possible routes of appeal to them have been defined. So the pre-Galilean model, with its fixed standards of success, is seen as impervious to the new standards of prediction and control. But the real positions held in history do not correspond to these water-tight deductive systems, and that is why rational transitions are in fact possible.

We could argue that there are also moral transitions which could be defended in a way very analogous to the scientific one just described. When one reads the opening pages of Foucault's *Surveiller et Punir*, [18]

[18] Paris, 1975.

with its riveting description of the torture and execution of a parricide in the mid-eighteenth century, one is struck by the cultural change we have gone through in post-Enlightenment Western civilization. We are much more concerned about pain and suffering than

end p.221

our forebears, and we shrink from the infliction of gratuitous suffering. It would be hard to imagine people taking their children to such a spectacle in a modern Western society, at least openly and without some sense of unease and shame.

What has changed? It is not that we have embraced an entirely new principle, that our ancestors would have thought the level of pain irrelevant, providing no reason at all to desist from some course of action involving torture or wounds. It is rather that this negative significance of pain was subordinated to other weightier considerations. If it is important that punishment in a sense undo the evil of the crime, restore the balance — what is implicit in the whole notion of the criminal making '*amende honorable*' — then the very horror of parricide calls for a particularly gruesome punishment. It calls for a kind of theatre of the horrible as the medium in which the undoing can take place. In this context, pain takes on a different significance: there has to be lots of it to do the trick. The principle of minimizing pain is trumped.

But then it is possible to see how the transition might be assessed rationally. If the whole outlook which justifies trumping the principle of minimizing suffering — which involves seeing the cosmos as a meaningful order in which human society is

embedded as a microcosm or mirror — comes to be set aside, then it is rational to be concerned above all to reduce suffering. Of course, our ultimate judgement will depend on whether we see the change in cosmology as rational; and that is, of course, the issue I have just been arguing in connection with the scientific revolution. If I am right there, here too the transition can perhaps be justified.

Of course, I am not claiming that all that has been involved in this important change has been the decline of the earlier cosmology. There are other, independent grounds in modern culture which have made us more reluctant to inflict pain. Some of them may have sinister aspects, if we believe Foucault himself. I have not got space to go into all this here. [19]

[19] I have discussed this at greater length in my 'Foucault on Freedom and Truth', in *Philosophy and the Human Sciences*.

But surely we must recognize the decline of the older notion of cosmic/social order as *one* consideration which lends a rational grounding to modern humanitarianism. This change would not only be linked to that in scientific theory, it would also be analogous to it in rational structure: to something which has always been recognized, although formerly in a subordinate place (the link between understanding and practice, the good of reducing pain), we are now constrained to give a more central significance because of changes which have taken place.

But the analogy I have been trying to draw between the justification of some scientific and moral revolutions cannot hide the fact that a great many moral disputes are much more difficult to arbitrate. To the extent that one can call on human constants, these are much more difficult to establish. And the suspicion dawns that in many cases such constants are of no avail. The differences between some cultures may be too great to make any *ad hominem* form of argument valid

end p.222

between them. Disputes of this kind would be inarbitrable.

3. But this form of argument, from the constants implicitly accepted by the interlocutor, does not exhaust the repertoire of practical reason. There is one more form, which is also an argument about transitions, but an even more striking departure from the canonical model. In both the above two forms the winner has appealed to some consideration which the loser had to acknowledge: his own anomalies, or some implicit constant. In the light of this consideration it was possible to show that the transition from *X* to *Y* could be seen as a gain, but not the reverse. So there is still something like a criterion operating here.

But we can imagine a form or argument in which no such consideration is invoked. The transition from *X* to *Y* is not shown to be a gain because this is the only way to make sense of the key consideration: rather it is shown to be a gain directly, because it can plausibly be described as mediated by some error-reducing move. This third mode of argument can be said to reverse the direction of argument. The canonical, foundationalist form can only show that the transition from *X* to *Y* is a gain in knowledge by showing that, say, *X* is false and *Y* true, or *X* has probability *n*, and *Y* has 2*n*. The two forms we have been considering focus on the transition, but they too only show that the move from *X* to *Y* is a gain because we can make sense of this transition from *Y*'s perspective but not of the reverse move from *X*'s perspective. We still ground our ultimate judgement in the differential performance of *X* and *Y*.

But consider the possibility that we might identify the transition directly as the overcoming of an error. Say we know that it consisted in the removing of a contradiction, or the overcoming of a confusion, or the recognition of a hitherto ignored relevant factor. In this case, the order of justifying argument would be

reversed. Instead of concluding that *Y* is a gain over *X* because of the superior performance of *Y*, we would be confident of the superior performance of *Y* because we know that *Y* is a gain over *X*.

But are we ever in a position to argue in this direction? In fact, examples abound in everyday life. First take a simple case of perception. I walk into a room, and see, or seem to see, something very surprising. I pause, shake my head, rub my eyes, and observe carefully. Yes, there really is a pink elephant with yellow polka dots in the class. I guess someone must be playing a practical joke.

What has gone on here? In fact, I am confident that my second perception is more trustworthy, but not because it scores better than the first on some measure of likelihood. On the contrary, if what I got from the first look was something like: 'maybe a pink elephant, maybe not', and from the second: 'definitely a pink elephant with yellow polka dots', there is no doubt that the first must be given greater antecedent probability. It is after all a disjunction, one of whose arms is overwhelmingly likely in these circumstances. But in fact I trust my second perception, because I have gone through an ameliorating transition. This is something I know how to bring off, it is part of my know-how as a perceiver. And that is what I in fact bring off by shaking my head (to clear the dreams), rubbing my eyes (to get the rheum out of them), and setting myself to observe with

end p.223

attention. It is my direct sense of the transition as an error-reducing one which grounds my confidence that my perceptual performance will improve.

Something similar exists in more serious biographical transitions. Joe was previously uncertain whether he loved Anne or not, because he also resents her, and in a confused way he was assuming that love is incompatible with resentment. But now he sees that these two are distinct and compatible emotions, and the latter is no longer getting in the way of his recognizing the strength of the former. Joe is confident that his present self-reading (I certainly love Anne) is superior to his former self-reading (I'm not sure whether I love Anne), because he knows that he passed from one to the other via the clarification of a confusion, that is, by a move which is in its very nature error-reducing.

Some of our gains in moral insight prove themselves to us in just this way. Pete was behaving impossibly at home, screaming at his parents, acting arrogantly with his younger siblings; he felt constantly resentful and very unhappy. He felt a constant sense of being cheated of his rights, or at least that is how it was formulated by his parents to the social worker. Now things are much better. Pete himself now applies this description to his former feelings. In a confused way, he felt that something more was owed to him as the eldest, and he resented not getting it. But he never would have subscribed to any such principle, and he clearly wants to repudiate it now. He thinks his previous behaviour was unjustified, and that one should not behave that way towards people. In other words, he has gone through a moral change: his views of what people owe each other in the family have altered. He is confident that this change represents moral growth, because it came about through dissipating a confused, largely unconsciously held belief, one which could not survive his recognizing its real nature.

These three cases are all examples of my third form of argument. They are, of course, all biographical. They deal with transitions of a single subject, whereas the standard disputes I have been discussing occur between people. And they are often (in the first case, always) cases of inarticulate, intuitive confidence, and hence

arguably have nothing to do with practical *reason* at all, if this is understood as a matter of forms of *argument*.

These two points are well taken. I have chosen the biographical context, because this is where this order of justification occurs at its clearest. But the same form can be and is adapted to the situation of interpersonal argument. Imagine I am a parent, or the social worker, reasoning with Pete before the change. Or say I am a friend of Joe's talking out his confused and painful feelings about Anne. In either case, I shall be trying to offer them as an interpretation of themselves which identifies these confused feelings as confused, and which thus, if accepted, will bring about the self-justifying transition.

This is, I believe, the commonest form of practical reasoning in our lives, where we propose to our interlocutors transitions mediated by such error-reducing moves — the identification of contradiction, the dissipation of confusion, or rescuing from (usually motivated) neglect a consideration whose significance

end p.224

they cannot contest. But this is a form of argument where the appeal to criteria, or even to the differential performance of the rival views in relation to some decisive consideration, is quite beside the point. The transition is justified by the very nature of the move which effects it. Here the *ad hominem* mode of argument is at its most intense, and its most fruitful.

## [3]

I would like in conclusion to try to draw together the threads of this perhaps too rambling discussion. I argued at the outset that practical arguments are in an important sense *ad hominem*. As a first approximation, I described these as arguments which appeal to what the opponent is already committed to, or at the least cannot lucidly repudiate. The notion that we might have to convince someone of an ultimate value premiss which he undividedly and unconfusedly rejects is indeed a ground for despair. Such radical gaps may exist, particularly between people from very different cultures; and in this case, practical reason is certainly powerless.

The discussion in the second part allows us to extend our notion of this kind of argument. It is not just cases where we can explicitly identify the common premiss from the outset that allow of rational debate. This was in fact the case with my opening example. Both the Nazi and I accept some version of the principle 'Thou shalt not kill', together with a different set of exclusions. Rational argument can turn on why he can permit himself the exclusions he does, and in fact, this historic position does not stand up long to rational scrutiny. It was really mob hysteria masquerading as thought.

But our discussion of transitions shows how debate can be rationally conducted even where there is no such explicit common ground at the outset. Now these arguments, to the effect that some transition from *X* to *Y* is a gain, are also *ad hominem*, in two related ways. First, they are specifically directed to the holders of *X*, in a way that apodeictic arguments never are. A foundational argument to the effect that *Y* is the correct thesis shows its superiority over the incompatible thesis *X* only incidentally. The proof also shows *Y*'s superiority over all rivals. It establishes an absolute, not just a comparative claim. If I establish that the correct value for the law of attraction is the inverse square and not the inverse cube of the distance, this also rules out the simple inverse, the inverse of the fourth power, etc.

It is crucial to transition arguments that they make a more modest claim. They are inherently comparative. The claim is not that *Y* is correct *simpliciter*, but simply that, whatever is 'ultimately' true, *Y* is better than *X*. It is, one might say, less false. The

argument is thus specifically addressed to the holders of *X*. Its message is: whatever else turns out to be true, you can improve your epistemic position by moving from *X* to *Y*; this step is a gain. But nothing need follow from this for the holders of a third, independent position. Above all, there is no claim

end p.225

to the effect that *Y* is the ultimate resting point of inquiry. The transition claim here is perfectly compatible with a further one which in turn supersedes *Y*. As MacIntyre puts it:

we are never in a position to claim that now we possess the truth or now we are fully rational. The most that we can claim is that this is the best account which anyone has been able to give so far, and that our beliefs about what the marks of 'a best account so far' are will themselves change in what are at present unpredictable ways.[20]

[20] Cf. my 'Epistemological Crisis', 455.

Second, these arguments all make their case by bringing to light something the interlocutor cannot repudiate. Either they make better sense of his inner difficulties than he can himself (case 1); or they present him with a development which he cannot explain on his own terms (case 2); or they show that the transition to *Y* comes about through a move which is intrinsically error-reducing (case 3). In relation to the original example of arguing with a Nazi, these greatly extend the range of rational debate. For what they appeal to in the interlocutor's own commitments is not explicit at the outset, but has to be brought to light. The pattern of anomalies and contradictions only becomes clear, and stands out as such, from the new position (case 1); the full significance of a hitherto marginalized form of understanding only becomes evident when the new position develops it (case 2); the fact that my present stance reposes on contradiction, confusion, and screening out the relevant only emerges as I make the transition — indeed, in this case, making the transition is just coming to recognize this error (case 3).

The range of rational argument is greatly extended, in other words, once we see that not all disputes are between fully explicated positions. Here the canonical, foundationalist model is likely to lead us astray. As we saw above with the second case, pre-Galilean science is indeed impregnable if we think only of its explicit standards of success: it has no cause to give any heed to technological pay-off. But in fact this pay-off constitutes a devastating argument, which one can only do justice to by articulating implicit understandings which have hitherto been given only marginal importance. Now I would argue that a great deal of moral argument involves the articulation of the implicit, and this extends the range of the *ad hominem* argument far beyond the easy cases where the opponent offers us purchase in one of his explicit premisses.

Naturally none of the above shows that all practical disputes are arbitrable in reason. Above all, it does not show that the most worrying cases, those which divide people of very different cultures, can be so arbitrated. Relativism still has something going for it, in the very diversity and mutual incomprehensibility of human moralities. Except in a dim way, which does more to disturb than enlighten us, we have almost no understanding at all of the place of human sacrifice, for instance, in the life of the Aztecs. Cortés simply thought that these people worshipped the devil, and only our commitment to a sophisticated pluralism

end p.226

stops us making a similar lapidary judgement.

And yet, I want to argue that the considerations above show that we should not give up on reason too early. We do not need to be so intimidated by distance and incomprehensibility that we take them as sufficient grounds to adopt relativism. There

are resources in argument. These have to be tried in each case, because nothing assures us a priori that relativism is false either. We have to try and see.

Two such resources are relevant to this kind of difference. First, there is the effect of working out and developing an insight which is marginally present in all cultures. In its developed form, this will make new demands, ones which upset the moral codes of previous cultures. And yet the insight in its developed form may carry conviction; that is, once articulated, it may be hard to gainsay. This is analogous to case 2 above, where the spectacular development of technology makes post-Galilean science hard to reject.

Second, the practices of previous cultures which are so challenged will often make sense against the background of a certain cosmology, or of semi-articulate beliefs about the way things have to be. These can be successfully challenged, and shown to be inadequate. Something of the kind was at stake in the discussion above of our changed attitude to suffering. Indeed, that case seems to show both these factors at work: we have developed new intuitions about the value and importance of ordinary life,[21] and, at the same time, we have fatally wounded the cosmology which made sense of the earlier gruesome punishments. These two together work to feed our convictions about the evil of unnecessary suffering.

Perhaps something similar can make sense of and justify our rejection of human sacrifice, or—to take a less exotic example—of certain practices involving the subordination of women. In this latter case, the positive factor—the developed moral insight—is that of the worth of each human being, the injunction that humans must be treated as ends, which we often formulate in a doctrine of universal rights. There is something very powerful in this insight, just because it builds on a basic human reaction, which seems to be present in some form everywhere: that humans are especially important, and demand special treatment. (I apologize for the vagueness in this formulation, but I am gesturing at something which occurs in a vast variety of different cultural forms.)

In many cultures, this sense of the special importance of the human is encapsulated in religious and cosmological outlooks, and connected views of human social life, which turn it in directions antithetical to modern rights doctrines. Part of what is special about humans can be that they are proper food for the gods, or that they embody cosmic principles differentially between men and women, which in fact imposes certain roles on each sex.

The rights doctrine presents human importance in a radical form, one that is hard to gainsay. This latter affirmation can be taken on several levels. Just



empirically there seems to be something to it, although establishing this is not just a matter of counting heads, but of making a plausible interpretation of human history. One that seems plausible to me goes something like this: recurrently in history new doctrines have been propounded which called on their adherents to move towards a relatively greater respect for human beings, one by one, at the expense of previously recognized forms of social encapsulation. This has been generally true of what people refer to as the 'higher' religions. And, of course, it has been the case with modern secular ideologies like liberalism and socialism. Where these have appeared, they have exercised a powerful attraction for human beings. Sometimes their spread can be explained by conquest, for example, Islam in the Middle East, liberalism in the colonial world, but frequently this is not the case, for example, Buddhism in India, Christianity in the Mediterranean world, Islam in Indonesia. Disencapsulated respect for the human seems to say something to us humans.

But, of course, this is a remark from the 'external' perspective, and does not by itself say anything about the place of reason. Can we perspicuously reconstruct these transitions in terms of *arguments*? This is hazardous, of course, and what follows could only at best be a crude approximation. But I think it might be seen this way. Disencapsulated respect draws us, because it articulates in a striking and far-reaching form what we already acknowledge in the sense I vaguely indicate with the term 'human importance'. Once you can grasp this possibility, it cannot help but seem prima-facie right. A demand is 'prima-facie right', when it is such as to command our moral allegiance, if only some other more weighty considerations do not stand in our way. Probably most of us feel like this about the ideal anarchic communist society; we'd certainly go for it, if only . . .

But, of course, the condition I mentioned: 'if you can grasp the possibility', is no pro forma one. For many societies and cultures, a disencapsulated view is literally unimaginable. The prescriptions of general respect just seem like perverse violations of the order of things.

Once one is over this hump, however, and can imagine disencapsulation, a field of potential argument is established. Universal respect now seems a conceivable goal, and one that is prima-facie right, if only. . . . The argument now turns on whatever fits into this latter clause. Yes, women are human beings, and there is a case therefore for giving them the same status as men, but unfortunately the order of things requires that they adopt roles incompatible with this equality, or they are crucially weaker or less endowed, and so cannot hack it at men's level, etc.

Here reason can get a purchase. These special pleadings can be addressed, and many of them found wanting, by rational argument. Considerations about the order of things can be undermined by the advance of our cosmological understanding. Arguments from unequal endowment are proved wrong by trying it out. Inequalities in capacity which seem utterly solid in one cultural setting just dissolve when one leaves this context. No one would claim that argument alone

end p.228

has produced the revolution in the status of women over the last centuries and decades in the West. But it all had something to do with the fact that the opponents of these changes were thrown on to a kind of strategic defensive; that they had to argue about the 'if onlys' and 'but unfortunatelys'. They had a position which was harder and harder to defend in reason.

But, one might argue, this is exactly where one is in danger of falling into ethnocentrism. The plight of, say, nineteenth-century opponents of women's franchise is utterly different from that of, say, certain Berber tribesmen today. On one account, the Berbers see the chastity of their womenfolk as central to the family honour, to the point where there can be a recognized obligation even to kill a kinswoman who has 'lost' her honour. Try telling them about Kant's *Critique of Practical Reason* or the works of John Stuart Mill, and you will get a different reaction from that of mainstream politicians in nineteenth-century Europe.

The gap can seem unbridgeable; there is this claim about honour, and what can you say to that? Honour has to do with avoiding shame, and can you argue with people about what they find shameful? Well, yes and no. If honour and shame are taken as ultimates, and if the fact that they are differently defined in different societies is ignored or discounted as just showing the depravity of the foreigner, then no argument is possible. But if one takes seriously the variety of definitions, and at the same time acknowledges that there are other moral or religious demands with which honour must be squared, then questions can arise about what really should be a

matter of honour, what is true honour, and the like. The thought can arise: maybe some other people have a better conception of honour, because theirs can be squared with the demands of God, say, or those of greater military efficacy, or control over fate.

The watershed between these two attitudes is more or less the one I mentioned above, whereby one becomes capable of conceiving of disencapsulated conditions, or at least of seeing one's society as one among many possible ones. This is undoubtedly among the most difficult and painful intellectual transitions for human beings. In fact, it may be virtually impossible, and certainly hazardous, to try to *argue* with people over it. But what does this say about the limitations of reason? Nothing, I would argue. The fact that a stance is hard to get to does not in any way show that it is not a more rational stance. In fact, each of our cultures is one possibility among many. People can and do live human lives in all of them. To be able to understand this sympathetically—or at least to understand some small subset of the range of cultures, and realize that one ought ideally to understand more—is to have a truer grasp of the human condition than that of people for whom alternative ways are utterly inconceivable. Getting people over this hump may require more than argument, but there is no doubt that the step is an epistemic gain. People may be unhappier as a result, and may lose something valuable that only unreflecting encapsulation gives you, but none of that would make this encapsulation any the less blind.

Even the most exotic differences do not therefore put paid to a role for reason.

end p.229

Of course, no one can show in advance that the 'if onlys' or 'but unfortunatelys' which stand in the way of universal rights can be rationally answered. It is just conceivable that some will arise which will themselves prove superior, more likely that there will be some where reason cannot arbitrate; and almost certain that we pay a price for our universalism in the loss of some goods which were bound up with earlier, more encapsulated forms of life. But none of this gives us cause a priori to take refuge in an agnostic relativism.

Unless, that is, we have already bought the faulty meta-ethic I have been attacking here. I want to end with the basic claim with which I started, and which underlies this whole exploration: and that is that modern philosophy, and to some extent modern culture, has lost its grip on the proper patterns of practical reason. Moral argument is understood according to inappropriate models, and this naturally leads to scepticism and despair, which in turn has an effect on our conception of morality, gives it a new shape (or misshapes it). We are now in a better position to see some of the motivations behind this misunderstanding.

I believe that we can identify in the above discussion three orders of motivation which combine to blind us. First, the naturalist outlook, with its hostility to the very notion of strong evaluation, tends to make the *ad hominem* argument seem irrelevant to ethical dispute. To show that your interlocutor is really committed to some good proves nothing about what he ought to do. To think it does is to commit the 'naturalistic fallacy'.

Second, naturalism together with the critical outlook have tended to brand *ad hominem* arguments as illegitimate. Reason should be as disengaged as possible from our implicit commitments and understandings, as it is in natural science, and as it must be if we are not to be victims of the status quo with all its imperfections and injustices. But once we neutralize our implicit understandings, by far the most

important field of moral argument becomes closed and opaque to us. We lose sight altogether of the articulating function of reason.
This distorts our picture not only of practical reason, but also of much scientific argument. And this brings us to the third motive: the ascendancy of the foundationalist model of reasoning, which comes to us from the epistemological tradition. This understands rational justification as (*a*) effected on the basis of criteria, (*b*) judging between fully explicit positions, and (*c*) yielding in the first instance absolute judgements of adequacy or inadequacy, and comparative assessments only mediately from these. But we have just seen what an important role in our reasoning is played by irreducibly comparative judgements—judgements about transitions—in articulating the implicit, and in the direct characterization of transitional moves which make no appeal to criteria at all. To block all this from view through an apodeictic model of reasoning is to make most moral discussion incomprehensible. Nor does it leave unimpaired our understanding of science and its history, as we have amply seen. The connections between scientific explanation and practical reason are in fact close: to lose sight of the one is to fall into confusion about the other.

end p.230

**Bibliography**

Foucault, Michel (1975). *Surveillir et punir: naissance de la prison*. Paris: Gallimard.
Hesse, Mary (1978). 'Theory and Value in the Social Sciences', in C. Hookway and P. Pettit (eds.), *Action and Interpretation. Studies in the Philosophy of the Social Sciences*. Cambridge: Cambridge University Press, 1979.
Lovejoy, Arthur O. (1942). *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, Mass.: Harvard University Press.
MacIntyre, Alasdair C. (1981). *After Virtue: A Study in Moral Theory*. Notre Dame, Ind.: University of Notre Dame Press.
Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
McDowell, John (1979). 'Virtue and Reason'. *The Monist*, 62, 331-50.
Mill, John Stuart (1863). *Utilitarianism*. Edited with introduction by G. Sher. Indianapolis: Hackett.
Scheler, Max (1926*a*). 'Erkenntnis und Arbeit', in *Die Wissensformen und die Gesellschaft*. Leipzig: Der Neue Geist Verlag, 233-486.
——(1926*b*). 'Die Soziologie des Wissens', in ibid., 47-229.
Taylor, Charles (1977). 'Epistemological Crisis, Dramatic Narrative, and the Philosophy of Science'. *The Monist*, 60, 453-72.
——(1985*a*). *Human Agency and Language. Philosophical Papers I*. Cambridge: Cambridge University Press.
——(1985*b*). *Philosophy and the Human Sciences. Philosophical Papers II*. Cambridge: Cambridge University Press.
——(1989). *Sources of the Self: The Making of Modern Identity*. Cambridge, Mass., and London: Harvard University Press.
—— 'Justice after Virtue' (unpublished MS).
Tugendhat, Ernst (1979). *Selbstbewusstein und Selbstbestimmung: Sprachanalytische Interpretatinen*. Frankfurt am Main: Suhrkamp Verlag.
Williams, Bernard A. O. (1978). *Descartes: The Project of Pure Inquiry*. Atlantic Highlands, NJ: Humanities Press.

end p.231



show chapter abstract and keywords

hide chapter abstract and keywords

## Martha Nussbaum

[1]

Charles Taylor's argument is, I believe, both convincing and important. And it is especially important for this project, since it shows an illuminating way of looking at the difficult evaluative disputes we encounter in thinking about the development of societies that are both different from one another and also (in most cases) internally heterogeneous. Development is itself an evaluative concept: it implies a progression from one situation to another that is (allegedly) in some ways better or more complete. Sometimes this issue of evaluation is ignored. Sometimes policy-makers and social scientists proceed as if it were perfectly clear to everyone what the values involved in development are-or even as if there were no evaluative question involved at all, but only certain facts which are alleged to be measurable independently of evaluation. Several of the papers in this volume, notably Hilary Putnam's, but also, in a very different way, Robert Erikson's, convincingly show the incoherence and barrenness of that way of thinking. [1]

[1] See also the papers in this volume by Brock, Scanlon, and Sen, and the comments by Griffin (on Brock) and Sen (on Bliss). For related philosophical discussion, see Sen and Williams (1982), Sen (1980, 1985, 1987), Williams (1973, 1985), and Wiggins (1987).

But if we do accept the fact that we are grappling with a difficult evaluative question here, we then need to know how to reflect about such questions. We need to have a clear conception of what it is to ask and answer them, how to argue about them where there is disagreement, and what sort and degree of success to expect from reason and rational argument. Providing a powerful account of rational ethical argument seems to me to be one of the central challenges for a practical philosophy, a philosophy that will really help people to make progress on troublesome human problems. The need for such an account in connection with development problems was certainly one of the primary motivations for bringing philosophers together with economists and other social scientists in this project. And Taylor's account seems to me among the most powerful of those currently before us in philosophy, so his participation in the project is especially valuable. I do not find much to criticize in the paper: I am in sympathy with most of what Taylor says. So in commenting I want simply, first, to say a little more about why I think Taylor's work in general, and this paper in particular, is significant for the concerns and projects of development studies. Then I shall briefly add two points to Taylor's analysis: one about the history of science, one about moral psychology.

end p.232

[2]

A person who studies the history of philosophical approaches to the social sciences over the past several decades in search of a model for the analysis of development would be likely to conclude that she is faced with just two alternatives, each in its own way unpalatable. On the one hand, she would find an approach that conceives of social science as a kind of natural science, [2]

[2] Notice, however, that the conception of natural science employed here is not above reproach: see below, and Putnam's paper in this volume; also Putnam (1981).

and of the reasoned understanding of human beings that is the goal of social science as an understanding detached from the commitments and self-understandings that are characteristic of human beings in their daily lives. Such approaches usually involve some sort of reduction of qualitative distinctions to quantitative distinctions; and they attach a great deal of importance to the simplified mathematical

representation of complex human matters. This approach, which has had enormous influence in shaping economic approaches to development, seems to have the advantage of promising truly rational solutions to difficult problems of choice. But it may, in the end, seem unacceptable because of the way in which it obscures or denies the richness and plurality of human values and commitments (both across societies and within each single society), and because of its reductive understanding of what human beings and communities are.

On the other hand, she would find a reaction against this approach, a reaction by now itself well entrenched in the social sciences. The alternative approach insists on restoring human self-interpretations to the sphere of social analysis in all their richness and variety. But its proponents frequently give up on practical reason, holding that there is no way in which reason can really resolve evaluative disputes. It is held that once we understand that the points of view of the participants in the dispute, to be correctly represented, must be represented from within the participants' own perspectives on the world, and once we understand, in addition, that cultural value schemes are highly various and largely incommensurable with one another, we will realize that practical reason has no effective part to play in such disputes. If it tries to take up a position of neutrality, detaching itself from all the competing conceptions, it will be unable to do so coherently, since no such external standpoint is available. If, on the other hand, it remains within the perspective of one of the parties, it seems that it must prove unfair and insensitive to the concerns of the other party, and be, really, nothing more than an attempt to dominate the other party. At the bottom of all so-called reasoning, then, is nothing but power. The work of Michel Foucault is frequently invoked in defence of this pattern of reasoning-although I believe that this is in some ways a misleading oversimplification of Foucault's contribution. [3]

[3] See, for example, Foucault (1984), discussed in my paper in this volume.

This alternative approach has been highly influential, especially in anthropology,

end p.233

literary theory, and discussions of legal interpretation. It has an obvious appeal, where development studies are concerned, since it restores to the field of analysis so much of human life that the other approach omits. [4]

[4] For a variant of this approach in literary theory, see Derrida (1976, 1979), Fish (1980, 1985); in legal studies, see Fish (1982), and many others. In the development context, a related approach is powerfully presented in Marglin and Marglin (1990). A number of these writers have been influenced by R. Rorty (1979, 1982).

And yet it may well seem in the end as unpalatable as the other, since it tells us that we cannot succeed in establishing by practical reasoning any conclusions critical of things we might like to criticize in societies whose traditions we are examining. To use Taylor's example, it tells us that any attempt we might make to criticize another society's treatment of women, or to hold that real development for that society must include some changes in women's position, is and can be nothing more than a kind of cultural imperialism.

For the past twenty-five years, Charles Taylor has been developing a distinctive position in this dispute, one that combines, I believe, the best features of the two approaches, and also reinterprets their disagreement with one another in an illuminating way. [5]

[5] Views with which Taylor's might be fruitfully compared include those of Putnam (1981, 1987), Davidson (1984), and, in literary theory, Wayne Booth (1988). For a related discussion of the law, which makes a powerful case for a richer descriptive language, see White (1989); the implications of Booth's work for legal studies are discussed in Nussbaum (1988), repr. in Nussbaum (1990).

Beginning with *The Explanation of Behaviour* in 1964, [6]

[6] Taylor (1964).

Taylor has consistently offered arguments of very high quality against a reductive natural science approach to the study of human beings and human action, insisting that any human science worthy of being taken seriously must include, and indeed base itself upon, the sense of value and the commitments that human beings actually display as they live and try to understand themselves. On the other hand, he has also consistently argued, as he does here, that this does not leave practical reason nowhere to go. In a series of papers on anthropological understanding of different cultures, he has shown that it is possible to take very seriously the data of cultural anthropology and of history, and the differences between conceptual schemes, and yet to hold that, in certain ways and under certain circumstances, practical reason can legitimately criticize traditions. [7]

[7] Taylor (1985*a*, 1985*b*).

The combination of anthropological sensitivity with philosophical precision in these papers gives them a more or less unique place in the debate, and they are certainly an invaluable guide to anyone working on development issues. [8]

[8] See also Taylor (1989). a major philosophical account of the development of modern conceptions of the self.

The present paper is, I think, of special interest in Taylor's defence of practical reason. According to Taylor's own account, in order to argue successfully against an opponent he needs to be able, among other things, to give a plausible account of the opponent's error. And in this paper he does just that. First, he argues effectively against the opponent of practical reason, presenting forms of unmistakably

end p.234

rational argument that do not depend on the starting points that this opponent has held, plausibly, to be unavailable. Then he goes on to tell us where the opponent went wrong, and in a very interesting way. For the error of the sceptical opponent of practical reason consists in remaining too much in the grip of the very picture of rational argument that is allegedly being criticized. While objecting to the hegemony of the natural sciences over the human sciences, and while seeking to restore to the human sciences their own rich humanistic character, the opponent has, presumably without full awareness, imported into her analysis one very central part of the natural science model, namely, its understanding of what constitutes a rational argument. For she seems to assume that rational argument requires neutrality, and deduction from premisses that are external to all historical perspectives. If this is not available, then we can say goodbye to reason itself. It is only because of this residual commitment to the rejected model, and her consequent neglect of other forms of rational arguing, that the opponent has been able so quickly to conclude that, the ethical domain being what it is, there are no good rational arguments to be found in it. This, I think, is a profound diagnosis. In various forms this problem is present in many contemporary positions that end up embracing some form of subjectivism or scepticism about practical reason: in Foucault's work, at least sometimes; in the work of Jacques Derrida, who seems to argue that without unmediated access to the world as it is in itself we have no arguments, nothing but the free play of interpretations; in other ways, in the work of numerous others who have influenced recent thinking in the human sciences. [9]

[9] See, for example, my criticism of Fish in Nussbaum (1985), repr. in Nussbaum (1990).

What Taylor does is, first of all, to point out the opponent's reliance on a model of arguing drawn from natural science; second, to argue from the history of science that

this is not even a very persuasive picture of the way scientific argument goes, especially in times of scientific change; and, third, to offer several models of reasoning that both explain the recalcitrant scientific cases and offer examples of ethical rationality. Like Hilary Putnam, Taylor shows that a really good account of science and scientific progress will not yield as hard a distinction between the scientific and the ethical as has been defended by some philosophers. And he shows that in both areas progress can be achieved by a complex and patient type of self-clarification, by patterns of argument in which implicit commitments are brought to light. Such arguments might well be said to be as old as Socratic cross-examination, and to have as their goal something like the self-understanding sought by Socrates.[10]

[10] See the illuminating account of Socratic procedure in Vlastos (1983).

Several morals for development studies might be drawn from this analysis. First, that we ought to reject disengaged pseudoscientific understandings of the human being, in favour of conceptions that give a larger role to people's own commitments and self-understandings. Second that when we do so we need not and should not give up on rational argument. But we must not expect the rational



arguments we use to be like those we associate with a certain view of natural science. They are likely to be piecemeal rather than global; they will be very much rooted in the particularities of people's historical situations; and frequently, as Taylor has said, they will be biographical rather than abstract. They will describe progress in a way that may seem to lack neatness and simplicity. For they will be, as he says, inherently comparative in their understandings of development, rather than absolute. And since they are attempts to bring to light what is deep and incompletely perceived in the thought of the person or group in question, they will be, frequently, both highly concrete and somewhat indirect. They may, for example, tell stories, appeal to the imagination and the emotions-tapping, through a very non-scientific use of language, people's intuitions about what matters most.[11]

[11] The case for such language in the law is powerfully made in White (1989), a critical review of Posner (1988); White contrasts the language of literature with the language characteristic of much of economics. The case for the use of literary language in moral reflection is made in Nussbaum (1990); see also Nussbaum and Sen (1989).

All of this I find exemplified in a striking way in Robert Erikson's account of a sociological approach that is both humanistic and committed to practical reason; and I think that Erikson should perhaps be less defensive than he is about the apparent messiness and complexity of his descriptions. There is precision in the lucid depiction of a highly complex and indeterminate situation; there is evasion and vagueness in the simple schematic description of a multi-faceted concrete case.

Finally, Taylor points out that the process of rational argument is frequently associated with raising the level of discontent and unhappiness in the people who are doing the arguing. People who, as a result of arguments such as those he describes, become aware of the variety of human societies and lose the isolation of what Taylor calls their 'encapsulated' condition frequently feel pain: both the pain of a new dissatisfaction with current arrangements and the pain of reflection itself. It seems to me that this is an important observation, and one that arises in development in significant ways. If one is committed to measuring development in terms of utility-construed either as happiness (pleasure) or as the satisfaction of current desires and preferences-one will be bound to judge that self-understanding is inimical to development, in such cases. Taylor seems to me to be right that self-understanding has a value of its own, apart from any utility (so construed) it brings. And his example

of the changing position of women shows us one case where the pain of discontent has had a definite positive link with development. This connection is also supported by Amartya Sen's data on women's changing perceptions of their health situation. [12]

[12] See Sen (1985: app. B).

## [3]

My one quarrel with the picture of scientific change presented by Taylor is that the picture he opposes is actually much *less* plausible, even initially, than his

end p.236

account makes it seem. Taylor's final answer to the alleged conceptual discontinuity between ancient Greek science and Galilean science is, I think, the right one: that there were all along interests in practical control and being able to lead a flourishing life that motivated all parties to the dispute, interests in terms of which partisans of the new science could show its superiority, in a way that the old-time Aristotelian could not deny. My objection is only that the fact that ancient science did have an interest in practical control is unmistakably clear; the opponent's account of the alleged conceptual discontinuity has simply obscured it by describing the ancient picture in too narrow and monolithic a way.

In fact, if one looks not only at Plato, but also at the people who were actually doing science in the ancient world-above all at the development of Greek medicine-one finds ubiquitous reference to practical manipulation and control, and one finds that this is one of the primary hallmarks of the scientific. Doctors regularly defend their procedures on the grounds that they work, and oppose rival medical procedures on the grounds that they are too abstract and schematic to be useful. [13]

[13] See especially the treatises 'On the Art' and 'On Ancient Medicine'; some of the relevant issues are discussed in Nussbaum (1986; ch. 4).

Even mathematics is repeatedly presented as a discipline whose primary point and motivation is its practical usefulness, in measuring and navigation and so forth. Prometheus, in Aeschylus' *Prometheus Bound*, calls it 'chief of all the clever stratagems'. [14]

[14] Aeschylus, *Prometheus Bound*, line 459.

Even if the later Aristotelians of Taylor's story played down this aspect of science, they could hardly have been ignorant of it, since Aristotle himself gives it prominence. [15]

[15] Cf. Aristotle, *Metaphysics*, I. 1.

So it seems to me that the opponent's story does not really get off the ground; the alleged rupture is just bad history.

Why is this worth mentioning here? It is worth mentioning because so much of the historiography on which contemporary debates about conceptual discontinuity are based is bad in exactly this way. So many alleged discontinuities look like discontinuities because the historian in question has made a highly selective use of texts, or has stuck to theoretical work, neglecting the history of popular thought. Much of Foucault's work is flawed in this way, [16]

[16] The *History of Sexuality*, for example, relies above all on Plato and Xenophon-both philosophers and both from wealthy oligarchic backgrounds. Both are extremely unrepresentative of popular thought.

although I still believe it to be important work. My experience is that if one studies any single society in sufficient depth one finds in it a rich plurality of views and conceptions, frequently in active debate with one another; and frequently these debates themselves are very important in explaining such large-scale conceptual shifts as do take place. This is essential to any good history of the later reception of

ancient Greek ideas. Amartya Sen and I have argued that it is also essential to good work on India, where the existence of debate and internal criticism needs emphasis.[17]

[17] See Nussbaum and Sen (1989).

If I have any slight criticism of Taylor's historiography in several of his papers,



it is that things sometimes look simpler and more monolithic than in fact they were. Appreciating their complexity gives us new ways of understanding how practical reason can function in social argument, and reinforces the general picture Taylor presents here.[18]

[18] Taylor (1989) is therefore especially to be welcomed, since its dense and complex historical account lays out the issues with a fullness impossible in the shorter papers.

## [4]

Now I want to make a point about psychology. Despair about the efficacy of practical reason, in the recent philosophical literature, frequently takes a form slightly different from the form that Taylor criticizes here (the form of despair that seeks scientific neutrality, even while it holds it to be unavailable). Frequently the problem of practical reason is put in terms of an alleged gap between reason and motivation. The argument imagines some bad or immoral person, and then says, 'All right. Even supposing that we can show to *our* satisfaction that this person is engaging in bad reasoning, even if we can convince ourselves *about* her by an argument that moves *us*, isn't there something troublesome about the fact that the argument doesn't do anything *for that person*? Have we really given that person reasons for action, if they are not reasons that have some force with that person, in the sense that they arouse desires and emotions of the appropriate motivational type?' The point is often made by saying that the only reasons we should care about are 'internal' reasons, reasons that are (or could become) part of the system of desires and motivations of the person involved. If that does not happen in our case, there is something peculiar about calling these reasons for that person at all, or saying that a rational argument has established something that this person ought to believe or to do.[19]

[19] See Williams (1981) for one forceful argument along these lines.

But the question of what reasons can become 'internal' for a person seems to depend very much on what desires and emotions that person happens to have. And this seems to be something that rational argument can do little or nothing about. Even if the argument should convince our person to *believe* its conclusion (this argument says), it is hard to see how it could reform her desires. But unless this happens, nothing much has been accomplished.

This is a deep and complex issue in philosophy, and, indeed, in life. I cannot even fully state it here, much less resolve it, but what I want to suggest, as my second friendly amendment to Taylor's account, is that I think he needs to supplement his account of reason with a picture of the connection between argument and motivation, between reason and passion. And I suspect that if one probes this issue, one will find that the sceptic about reason is in the grip, here too, of a picture of reason that is the outcome of a relatively recent naturalistic understanding of the human being, a picture that could not withstand the scrutiny of our deepest beliefs concerning who we are.



There has been a tendency in philosophy—ever since Descartes analysed the passions in connection with his view of the mind-body split, dissociating them strongly from beliefs—to think of passions as brute feelings, more or less impervious

to reasoning, coming from altogether different parts of our nature. Most philosophers have long since given up the Cartesian picture of the mind-body split; but many still retain the associated picture of emotions as feelings that are distinct from and relatively impervious to reasoning. If one looks back to the correspondence of Descartes with Princess Elizabeth, that very astute and sceptical philosophical mind, one finds in her challenges (and in their mutual references to Stoic ideas) an older and also, I believe, more adequate view of the passions, one that was the dominant picture in the thought of most of the ancient Greek philosophers, and one that was extremely important to their picture of the ways in which philosophy can be practical, changing people and societies for the better.

This picture (which one can find in different forms in Aristotle and in the Epicureans, but above all in the Stoics) insists that emotions or passions [20]

[20] I am using 'emotions' and 'passions' interchangeably; both words are historically well entrenched, and both have, for many centuries, been used more or less interchangeably to designate a species of which the most important members are anger, fear, grief, pity, love and joy.

are highly discriminating evaluative responses, very closely connected to beliefs about what is valuable and what is not. Grief, for example, is intimately linked to the belief that some object or person, now lost, has profound importance; it is a recognition of that importance. Anger involves and rests upon a belief that one has been wronged in a more than trivial way. And so forth. What follows from this view is that a rational argument can powerfully influence a person's passions and motivations. If rational argument can show either that the supposed bad event (the death or insult or whatever) did not take place, or that the purported occasion for grief or anger is really not the sort of thing one ought to care very much about, then one can actually change a person's psychology in a more than intellectual way. The Stoics and Epicureans took on this task in many ways, showing people, for example, that the worldly goods and goods of reputation that are commonly the bases of anger should not be valued as people value them. Such arguments do not just operate on the surface of the mind. They conduct a searching scrutiny of the whole of the person's mental and emotional life. And this is so because emotions are not just animal urges, but fully human parts of our outlook on the world. [21]

[21] See Nussbaum (1987, 1989*a*, 1989*b*).

On this basis the Hellenistic philosophers built up a powerful picture of a practical philosophy, showing how good arguments about topics like death and the reasons for anger could really change the heart, and, through that, people's personal and social lives. They compared the philosopher to the doctor: through reasoning he or she treats and heals the soul. It seems to me that such an account of motivational change through argument would be an attractive and important addition to Taylor's picture. I believe that an account like this can be defended

end p.239

today, [22]

[22] Aspects of it are being powerfully defended in several fields, including philosophy, cognitive psychology, anthropology, psychoanalysis, and legal studies: see Nussbaum (1990) for discussion and references, and see especially Lutz (1988), A. Rorty (1980, 1988), and de Sousa (1987).

and that the defence will make a powerful contribution to the defence of practical reason.

If Taylor chooses to develop his picture of practical reasoning along these lines, he will need to move even further than he has already from the scientific model of reasoning. For good 'therapeutic' argument may wish to make use of techniques like story-telling and vivid exemplification—techniques that lie far indeed from the scientific model—in order to bring to light hidden judgements of importance and to

give a compelling picture of a life in which such judgements are absent. All this would fit in well with Taylor's emphasis in his paper on the 'biographical' nature of argument, and its function of bringing hidden things to light. Such further developments would yield arguments that could, I believe, be defended as Taylor defends his examples here: as examples of rational argument and epistemic progress.

**Bibliography**


Booth, W. (1988). *The Company We Keep: An Ethics of Fiction*. Berkeley and Los Angeles: University of California Press.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Derrida, J. (1976). *Of Grammatology*, trans. G. C. Spivak. Baltimore: Johns Hopkins University Press.

—— (1979). *Spurs: Nietzsche's Styles*, trans. B. Harlow. Chicago: University of Chicago Press.

de Sousa, R. (1987). *The Rationality of Emotions*. Cambridge, Mass.: MIT Press.

Fish, S. (1980). *Is There a Text in this Class? The Authority of Interpretive Communities*. Cambridge, Mass.: Harvard University Press.

—— (1982). 'Working on the Chain Gang: Interpretation in Law and Literature', *Texas Law Review*, 551.

—— (1985). 'Anti-Professionalism', *New Literary History*, 17.

Foucault, M. (1984). *The Use of Pleasure*. History of Sexuality, ii, trans. R. Hurley. New York: Pantheon.

Lutz, C. (1988). *Unnatural Emotions*. Chicago: University of Chicago Press.

Marglin, F.A., and Marglin, S. (1990). *Dominating Knowledge: Development, Culture, and Resistance*. Oxford: Clarendon Press.

Nussbaum, M. (1985). 'Sophistry about Conventions: A Reply to Stanley Fish', *New Literary History*, 17. Reprinted in Nussbaum (1990).

—— (1986). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.

—— (1987). 'The Stoics on the Extirpation of the Passions', *Apeiron*, 20, 129-77.


end p.240


—— (1988). 'Reading for Life', a review of Booth (1988), *Yale Journal of Law and the Humanities*, 1. Reprinted in Nussbaum (1990).

—— (1989*a*). 'Beyond Obsession and Disgust: Lucretius' Genealogy of Love', *Apeiron*, 22, 1-59.

—— (1989*b*). 'Mortal Immortals: Lucretius on Death and the Voice of Nature', *Philosophy and Phenomenological Research*, 50, 305-51.

—— (1990). *Love's Knowledge: Essays on Philosophy and Literature*. Oxford and New York: Clarendon Press and Oxford University Press.

—— and Sen, A. (1989). 'Internal Criticism and Indian Rationalist Traditions', in M. Krausz (ed.), *Relativism: Interpretation and Confrontation*. Notre Dame: University of Notre Dame Press.

Posner, R. A. (1988). *Law and Literature: A Misunderstood Relation*. Cambridge, Mass.: Harvard University Press.

Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.

—— (1987). *The Many Faces of Realism*. La Salle, Ill.: Open Court.

Rorty, A. (ed.) (1980). *Explaining Emotions*. Berkeley: University of California Press.

—— (1988). *Mind in Action. Essays in the Philosophy of Mind*. Boston: Beacon Press.

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
—— (1982). *Consequences of Pragmatism: Essays*, 1972-1980. Minneapolis: University of Minnesota Press.
Sen, A. (1980). 'Plural Utility', *Proceedings of the Aristotelian Society*, 80.
—— (1985). *Commodities and Capabilities*. Amsterdam: North-Holland.
—— (1987). *The Standard of Living*. *Tanner Lectures on Human Values 1985*, ed. G. Hawthorn. Cambridge: Cambridge University Press.
—— and Williams, B. (1987). *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.
Taylor, C. (1964). *The Explanation of Behaviour*. London: Routledge and Kegan Paul.
—— (1985*a*). *Human Agency and Language. Philosophical Papers I*. Cambridge: Cambridge University Press.
—— (1985*b*). *Philosophy and the Human Sciences. Philosophical Papers II*. Cambridge: Cambridge University Press.
—— (1989). *Sources of the Self: The Making of Modern Identity*. Cambridge, Mass., and London: Harvard University Press.
Vlastos, G. (1983). 'The Socratic Elenchus', *Oxford Studies in Ancient Philosophy*, i. Oxford and New York: Oxford University Press.
White, J. B. (1989). 'What Can a Lawyer Learn from Literature?', *Harvard Law Review*, 102, 2014-47.
Wiggins, D. (1987). *Needs, Values, Truth: Essays in the Philosophy of Value*. Oxford and New York: Basil Blackwell.
Williams, B. (1973). 'A Critique of Utilitarianism', in J. C. Smart and B. Williams, *Ultilitarianism For and Against*. Cambridge: Cambridge University Press.
—— (1981). 'Internal and External Reasons', in *Moral Luck: Philosophical Papers*, 1973-1980. Cambridge: Cambridge University Press.
—— (1985). *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.

end p.241

# Non-Relative Virtues: An Aristotelian **Approach**

show chapter abstract and keywords
hide chapter abstract and keywords

**Martha Nussbaum**

All Greeks used to go around armed with swords.
Thucydides, *History of the Peloponnesian War*
The customs of former times might be said to be too simple and barbaric. For Greeks used to go around armed with swords; and they used to buy wives from one another; and there are surely other ancient customs that are extremely stupid. (For example, in Cyme there is a law about homicide, that if a man prosecuting a charge can produce a certain number of witnesses from among his own relations, the defendant will automatically be convicted of murder.) In general, all human beings seek not the way of their ancestors, but the good.
Aristotle, *Politics*, 1268a39 ff.
One may also observe in one's travels to distant countries the feelings of recognition and affiliation that link every human being to every other human being.
Aristotle, *Nicomachean Ethics*, 1155a21-2

## [1]

The virtues are attracting increasing interest in contemporary philosophical debate. From many different sides one hears of a dissatisfaction with ethical theories that are remote from concrete human experience. Whether this remoteness results from the utilitarian's interest in arriving at a universal calculus of satisfactions or from a Kantian concern with universal principles of broad generality, in which the names of particular contexts, histories, and persons do not occur, remoteness is now being seen by an increasing number of moral philosophers as a defect in an approach to ethical questions. In the search for an alternative approach, the concept of virtue is playing a prominent role. So, too, is the work of Aristotle, the greatest defender of an ethical approach based on the concept of virtue. For Aristotle's work seems, appealingly, to combine rigour with concreteness, theoretical power with sensitivity to the actual circumstances

end p.242

of human life and choice in all their multiplicity, variety, and mutability.

But on one central point there is a striking divergence between Aristotle and contemporary virtue theory. To many current defenders of an ethical approach based on the virtues, the return to the virtues is connected with a turn towards relativism—towards, that is, the view that the only appropriate criteria of ethical goodness are local ones, internal to the traditions and practices of each local society or group that asks itself questions about the good. The rejection of general algorithms and abstract rules in favour of an account of the good life based on specific modes of virtuous action is taken, by writers as otherwise diverse as Alasdair MacIntyre, Bernard Williams, and Philippa Foot, [1]

[1] See MacIntyre (1981), and by contrast MacIntyre (1988); Foot (1978); Williams (1984, 1985); Walzer (1983, 1987).

to be connected with the abandonment of the project of rationally justifying a single norm of flourishing life for all human beings and a reliance, instead, on norms that are local both in origin and in application.

The position of all these writers, where relativism is concerned, is complex; none unequivocally endorses a relativist view. But all connect virtue ethics with a relativist denial that ethics, correctly understood, offers any transcultural norms, justifiable by reference to reasons of universal human validity, by reference to which we may appropriately criticize different local conceptions of the good. And all suggest that the insights we gain by pursuing ethical questions in the Aristotelian virtue-based way lend support to relativism.

For this reason it is easy for those who are interested in supporting the rational criticism of local traditions and in articulating an idea of ethical progress to feel that the ethics of virtue can give them little help. If the position of women, as established by local traditions in many parts of the world, is to be improved, if traditions of slave-holding and racial inequality, religious intolerance, aggressive and warlike conceptions of manliness, and unequal norms of material distribution are to be criticized in the name of practical reason, this criticizing (one might easily suppose) will have to be done from a Kantian or utilitarian viewpoint, not through the Aristotelian approach.

This is an odd result, as far as Aristotle is concerned. For it is obvious that he was not only the defender of an ethical theory based on the virtues, but also the defender of a single objective account of the human good, or human flourishing. This account is supposed to be objective in the sense that it is justifiable by reference to reasons that do not derive merely from local traditions and practices, but rather from features of humanness that lie beneath all local traditions and are there to be seen whether or

not they are in fact recognized in local traditions. And one of Aristotle's most obvious concerns was the criticism of existing moral traditions, in his own city and in others, as unjust or repressive, or in other ways incompatible with human flourishing. He uses his account of the virtues as a basis for this criticism of local traditions: prominently, for example, in Book II of the *Politics*, where he frequently argues against existing social forms by



pointing to ways in which they neglect or hinder the development at some important human virtue.[2]

[2] For examples of this, see Nussbaum (1988*a*).

Aristotle evidently believed that there is no incompatibility between basing an ethical theory on the virtues and defending the singleness and objectivity of the human good. Indeed, he seems to have believed that these two aims are mutually supportive.

Now the fact that Aristotle believed something does not make it true (though I have sometimes been accused of holding that position!). But it does, on the whole, make that something a plausible *candidate* for the truth, one deserving our most serious scrutiny. In this case, it would be odd indeed if he had connected two elements in ethical thought that are self-evidently incompatible, or in favour of whose connectedness and compatibility there is nothing interesting to be said. The purpose of this paper is to establish that Aristotle did indeed have an interesting way of connecting the virtues with a search for ethical objectivity and with the criticism of existing local norms, a way that deserves our serious consideration as we work on these questions. Having described the general shape of the Aristotelian approach, we can then begin to understand some of the objections that might be brought against such a non-relative account of the virtues, and to imagine how the Aristotelian could respond to those objections.

## [2]

The relativist, looking at different societies, is impressed by the variety and the apparent non-comparability in the lists of virtues she encounters. Examining the different lists, and observing the complex connections between each list and a concrete form of life and a concrete history, she may well feel that any list of virtues must be simply a reflection of local traditions and values, and that, virtues being (unlike Kantian principles or utilitarian algorithms) concrete and closely tied to forms of life, there can in fact be no list of virtues that will serve as normative for all these varied societies. It is not only that the specific forms of behaviour recommended in connection with the virtues differ greatly over time and place, it is also that the very areas that are singled out as spheres of virtue, and the manner in which they are individuated from other areas, vary so greatly. For someone who thinks this way, it is easy to feel that Aristotle's own list, despite its pretensions to universality and objectivity, must be similarly restricted, merely a reflection of one particular society's perceptions of salience and ways of distinguishing. At this point, relativist writers are likely to quote Aristotle's description of the 'great-souled' person, the *megalopsuchos*, which certainly contains many concrete local features and sounds very much like the portrait of a certain sort of Greek gentleman, in order to show that Aristotle's list is just as culture-bound as any other.[3]

[3] See, for example, Williams (1985: 34-6); Hampshire (1983: 150ff.).

But if we probe further into the way in which Aristotle in fact enumerates and

individuates the virtues, we begin to notice things that cast doubt upon the suggestion that he simply described what was admired in his own society. First of all, we notice that a rather large number of virtues and vices (vices especially) are nameless, and that, among the ones that are not nameless, a good many are given, by Aristotle's own account, names that are somewhat arbitrarily chosen by Aristotle, and do not perfectly fit the behaviour he is trying to describe.[4]

[4] For 'nameless' virtues and vice, see *EN* (*Nicomachean Ethics*) 1107b1-2, 1107b7-8, 1107b30-1, 1108a17, 1119a10-11, 1126b20, 1127a12, 1127a14; for recognition of the unsatisfactoriness of names given, see 1107b8, 1108a5-6, 1108a20 ff. The two categories are largely overlapping, on account of the general principle enunciated at 1108a16-19, that where there is no name a name should be given, satisfactory or not.

Of such modes of conduct he writes, 'Most of these are nameless, but we must try . . . to give them names in order to make our account clear and easy to follow' (*EN* 1108a 6-19). This does not sound like the procedure of someone who is simply studying local traditions and singling out the virtue-names that figure most prominently in those traditions.

What *is* going on becomes clearer when we examine the way in which he does, in fact, introduce his list. For he does so, in the *Nicomachean Ethics*[5]

[5] It should be noted that this emphasis on spheres of experience is not present in the *Eudemian Ethics*, which begins its discussion with a list of virtues and vices. This seems to me a sign that that treatise expresses a more primitive stage of Aristotle's thought on the virtues—whether earlier or not.

by a device whose very straight-forwardness and simplicity has caused it to escape the notice of most writers on this topic. What he does, in each case, is to isolate a sphere of human experience that figures in more or less any human life, and in which more or less any human being will have to make *some* choices rather than others, and act in *some* way rather than some other. The introductory chapter enumerating the virtues and vices begins with an enumeration of these spheres (*EN* II. 7); and each chapter on a virtue in the more detailed account that follows begins with 'Concerning *X* . . . ', or words to this effect, where *X* names a sphere of life with which all human beings regularly and more or less necessarily have dealings.[6]

[6] For statements with *peri* ('concerning') connecting virtues with spheres of life, see *EN* 1115a6-7, 1117a29-30, 1117b25 and 27, 1119b23, 1122a19, 1122b34, 1125b26, 1126b13—and *EN* II. 7 throughout. See also the related usages at 1126b11, 1127b32.

Aristotle then asks, what is it to choose and respond well within that sphere? And what is it to choose defectively? The 'thin account' of each virtue is that it is whatever being stably disposed to act appropriately in that sphere consists in. There may be, and usually are, various competing specifications of what acting well, in each case, in fact comes to. Aristotle goes on to defend in each case some concrete specification, producing, at the end, a full or 'thick' definition of the virtue.

Here are the most important spheres of experience recognized by Aristotle, along with the names of their corresponding virtues:[7]

[7] My list here inserts justice in a place of prominence. (In the *EN* it is treated separately, after all the other virtues, and the introductory list defers it for that later examination.) I have also added at the end of the list categories corresponding to the various intellectual virtues discussed in *EN* VI, and also to *phronesis*, or practical wisdom, discussed in *EN* VI as well. Otherwise the order and wording of my list closely follows II. 7, which gives the programme for the more detailed analyses from III. 5 to IV.

end p.245

| **Sphere** | **Virtue** |
|---|---|
| 1. Fear of important damages, esp. death | Courage |
| 2. Bodily appetites and their | Moderation |

| | |
|---|---|
| pleasures | |
| 3. Distribution of limited resources | Justice |
| 4. Management of one's personal property, where others are concerned | Generosity |
| 5. Management of personal property, where hospitality is concerned | Expansive hospitality |
| 6. Attitudes and actions with respect to one's own worth | Greatness of soul |
| 7. Attitude to slights and damages | Mildness of temper |
| 8. 'Association and living together and the fellowship of words and actions' | |
| *a*. Truthfulness in speech | Truthfulness |
| *b*. Social association of a playful kind | Easy grace (contrasted with coarseness, rudeness, insensitivity) |
| *c*. Social association more generally | Nameless, but a kind of friendliness (contrasted with irritability and grumpiness) |
| 9. Attitude to the good and ill fortune of others | Proper judgement (contrasted with enviousness, spitefulness, etc.) |
| 10. Intellectual life | The various intellectual virtues, such as perceptiveness, knowledge, etc. |
| 11. The planning of one's life and conduct | Practical wisdom |

There is, of course, much more to be said about this list, its specific members, and the names Aristotle chose for the virtue in each case, some of which are indeed culture-bound. What I want to insist on here, however, is the care with which Aristotle articulates his general approach, beginning from a characterization of a sphere of universal experience and choice, and introducing the virtue-name as the name (as yet undefined) of whatever it is to choose appropriately in

end p.246

that area of experience. On this approach, it does not seem possible to say, as the relativist wishes to, that a given society does not contain anything that corresponds to a given virtue. Nor does it seem to be an open question, in the case of a particular agent, whether a certain virtue should or should not be included in his or her life—except in the sense that she can always choose to pursue the corresponding deficiency instead. The point is that everyone makes some choices and acts somehow or other in these spheres: if not properly, then improperly. Everyone has *some* attitude, and corresponding behaviour, towards her own death; her bodily

appetites and their management; her property and its use; the distribution of social goods; telling the truth; being kind to others; cultivating a sense of play and delight, and so on. No matter where one lives one cannot escape these questions, so long as one is living a human life. But then this means that one's behaviour falls, willy-nilly, within the sphere of the Aristotelian virtue, in each case. If it is not appropriate, it is inappropriate; it cannot be off the map altogether. People will of course disagree about what the appropriate ways of acting and reacting in fact *are*. But in that case, as Aristotle has set things up, they are arguing about the same thing, and advancing competing specifications of the same virtue. The reference of the virtue term in each case is fixed by the sphere of experience—by what we shall from now on call the 'grounding experiences'. The thin or 'nominal' definition of the virtue will be, in each case, that it is whatever being disposed to choose and respond well consists in, in that sphere. The job of ethical theory will be to search for the best further specification corresponding to this nominal definition, and to produce a full definition.

## [3]

We have begun to introduce considerations from the philosophy of language. We can now make the direction of the Aristotelian account clearer by considering his own account of linguistic indicating (referring) and defining, which guides his treatment of both scientific and ethical terms, and of the idea of progress in both areas. [8]

[8] For a longer account of this, with references to the literature and to related philosophical discussions, see Nussbaum (1986*a*: ch. 8).

Aristotle's general picture is as follows. We begin with some experiences—not necessarily our own, but those of members of our linguistic community, broadly construed. [9]

[9] Aristotle does not worry about questions of translation in articulating this idea; for some worries about this, and an Aristotelian response, see below, sections 4 and 6.

On the basis of these experiences, a word enters the language of the group, indicating (referring to) whatever it is that is the content of those experiences. Aristotle gives the example of thunder. [10]

[10] *Posterior Analytics* II. 8, 93a21ff.; see Nussbaum (1986*a*: ch. 8).

People hear a noise in the clouds, and they then refer to it, using the word 'thunder'. At this point, it may be that nobody has any concrete account of the noise or any idea about what it really is. But the experience fixes a subject for further inquiry. From now on,

end p.247

we can refer to thunder, ask 'What is thunder?', and advance and assess competing theories. The thin or, we might say, 'nominal' definition of thunder is 'That noise in the clouds, whatever it is'. The competing explanatory theories are rival candidates for correct full or thick definition. So the explanatory story citing Zeus' activities in the clouds is a false account of the very same thing of which the best scientific explanation is a true account. There is just one debate here, with a single subject. So too, Aristotle suggests, with our ethical terms. Heraclitus, long before him, already had the essential idea, saying, 'They would not have known the name of justice, if these things did not take place.' [11]

[11] Heraclitus, fragment Diels-Kranz B23; see Nussbaum (1972).

'These things', our source for the fragment informs us, are experiences of injustice-presumably of harm, deprivation, inequality. These experiences fix the reference of the corresponding virtue word. Aristotle proceeds along similar lines. In the *Politics* he insists that only human beings, and not either animals or gods, will have our basic ethical terms and concepts (such as just and unjust, noble and base, good and bad),

because the beasts are unable to form the concepts, and the gods lack the experiences of limit and finitude that give a concept such as justice its points. [12]

[12] See *Politics* I. 2, 1253a1-18; that discussion does not explicitly deny virtues to the gods, but this denial is explicit in *EN* 1145a25-7 and 1178b10ff.

In the enumeration of the virtues in the *Nicomachean Ethics*, he carries the line of thought further, suggesting that the reference of the virtue terms is fixed by spheres of choice, frequently connected with our finitude and limitation, that we encounter in virtue of shared conditions of human existence. [13]

[13] Aristotle does not make the connection with his account of language explicit, but his project is one of defining the virtues, and we would expect him to keep his general view of defining in mind in this context. A similar idea about the virtues, and about the way in which a certain sort of experience can serve as a plausible basis for a non-relative account, is developed (without reference to Aristotle) in Sturgeon (1984).

The question about virtue usually arises in areas in which human choice is both non-optional and somewhat problematic. (Thus, he stresses, there is no virtue involving the regulation of listening to attractive sounds, or seeing pleasing sights.) Each family of virtue and vice or deficiency words attaches to some such sphere. And we can understand progress in ethics, like progress in scientific understanding, to be progress in finding the correct fuller specification of a virtue, isolated by its thin or nominal definition. This progress is aided by a perspicuous mapping of the sphere of the grounding experiences. When we understand more precisely what problems human beings encounter in their lives with one another, what circumstances they face in which choice of some sort is required, we will have a way of assessing competing responses to those problems, and we will begin to understand what it might be to act well in the face of them.

Aristotle's ethical and political writings provide many examples of how such progress (or, more generally, such a rational debate) might go. We find argument against Platonic asceticism, as the proper specification of moderation (appropriate

end p.248

choice and response *vis à vis* the bodily appetites), and in favour of a more generous role for appetitive activity in human life. We find argument against the intense concern for public status and reputation, and the consequent proneness to anger over slights, that was prevalent in Greek ideals of maleness and in Greek behaviour, together with a defence of a more limited and controlled expression of anger, as the proper specification of the virtue that Aristotle calls 'mildness of temper'. (Here Aristotle evinces some discomfort with the virtue term he has chosen, and he is right to do so, since it certainly loads the dice heavily in favour of his concrete specification and against the traditional one. [14]

[14] See *EN* 1107a5, where Aristotle says that the virtues and the corresponding person are 'pretty much nameless', and says, 'Let us call . . . ' when he introduces the names. See also 1125b29, 1126a3-4.

) And so on for all the virtues.

In an important section of *Politics* II, part of which forms one of the epigraphs to this paper, Aristotle defends the proposition that laws should be revisable and not fixed, by pointing to evidence that there is progress towards greater correctness in our ethical conceptions, as also in the arts and sciences. Greeks used to think that courage was a matter of waving swords around; now they have (the *Ethics* informs us) a more inward and a more civic and communally attuned understanding of proper behaviour towards the possibility of death. Women used to be regarded as property, to be bought and sold; now this would be thought barbaric. And in the case of justice as well we have, the *Politics* passage claims, advanced towards a more adequate

understanding of what is fair and appropriate. Aristotle gives the example of an existing homicide law that convicts the defendant automatically on the evidence of the prosecutor's relatives (whether they actually witnessed anything or not, apparently). This, Aristotle says, is clearly a stupid and unjust law; and yet it once seemed appropriate-and, to a tradition-bound community, must still be so. To hold tradition fixed is then to prevent ethical progress. What human beings want and seek is not conformity with the past, it is the good. So our systems of law should make it possible for them to progress beyond the past, when they have agreed that a change is good. (They should not, however, make change too easy, since it is no easy matter to see one's way to the good, and tradition is frequently a sounder guide than current fashion.)

In keeping with these ideas, the *Politics* as a whole presents the beliefs of the many different societies it investigates not as unrelated local norms, but as competing answers to questions about justice and courage (and so on) with which all societies are (being human) concerned, and in response to which they all try to find what is good. Aristotle's analysis of the virtues gives him an appropriate framework for these comparisons, which seem perfectly appropriate inquiries into the ways in which different societies have solved common human problems.

In the Aristotelian approach it is obviously of the first importance to distinguish two stages of the inquiry: the initial demarcation of the sphere of choice,

end p.249

of the 'grounding experiences' that fix the reference of the virtue term; and the ensuing more concrete inquiry into what the appropriate choice, in that sphere, *is*. Aristotle does not always do this carefully; and the language he has to work with is often not helpful to him. We do not have much difficulty with terms like 'moderation' and 'justice' and even 'courage', which seem vaguely normative, but relatively empty, so far, of concrete moral content. As the approach requires, they can serve as extension-fixing labels under which many competing specifications may be investigated. But we have already noticed the problem with 'mildness of temper', which seems to rule out by fiat a prominent contender for the appropriate disposition concerning anger. And much the same thing certainly seems to be true of the relativists' favourite target, *megalopsuchia*, which implies in its very name an attitude to one's own worth that is more Greek than universal. (A Christian, for example, will feel that the proper attitude to one's own worth requires an understanding of one's lowness, frailty, and sinfulness. The virtue of humility requires considering oneself *small*, not great.) What we ought to get at this point in the inquiry is a word for the proper attitude towards anger and offence, and for the proper attitude towards one's worth, that are more truly neutral among the competing specifications, referring only to the sphere of experience within which we wish to determine what is appropriate. Then we could regard the competing conceptions as rival accounts of one and the same thing, so that, for example, Christian humility would be a rival specification of the same virtue whose Greek specification is given in Aristotle's account of *megalopsuchia*, namely, the proper attitude towards the question of one's own worth. In fact, oddly enough, if one examines the evolution in the use of this word from Aristotle through the Stoics to the Christian fathers, one can see that this is more or less what happened, as 'greatness of soul' became associated, first, with the Stoic emphasis on the supremacy of virtue and the worthlessness of externals, including the body, and, through this, with the Christian denial of the body and of the worth of earthly life.[15]

[15] See Procope (forthcoming).

So even in this apparently unpromising case, history shows that the Aristotelian approach not only provided the materials for a single debate but actually succeeded in organizing such a debate, across enormous differences of both place and time. Here, then, is a sketch for an objective human morality based upon the idea of virtuous action-that is, of appropriate functioning in each human sphere. The Aristotelian claim is that, further developed, it will retain the grounding in actual human experiences that is the strong point of virtue ethics, while gaining the ability to criticize local and traditional moralities in the name of a more inclusive account of the circumstances of human life, and of the needs for human functioning that these circumstances call forth.

end p.250

## [4]

The proposal will encounter many objections. The concluding sections of this paper will present three of the most serious and will sketch the lines along which the Aristotelian might proceed in formulating a reply. To a great extent these objections were not imagined or confronted by Aristotle himself, but his position seems capable of confronting them.

The first objection concerns the relationship between singleness of problem and singleness of solution. Let us grant for the moment that the Aristotelian approach has succeeded in coherently isolating and describing areas of human experience and choice that form, so to speak, the *terrain* of the virtues, and in giving thin definitions of each of the virtues as whatever it is that choosing and responding well within that sphere consists in. Let us suppose that the approach succeeds in doing this in a way that embraces many times and places, bringing disparate cultures together into a single debate about the good human being and the good human life. Different cultural accounts of good choice within the sphere in question in each case are now seen not as untranslatably different, but as competing answers to a single general question about a set of shared human experiences. Still, it might be argued, what has been achieved is, at best, a single discourse or debate about virtue. It has not been shown that this debate will have, as Aristotle believes, a single answer. Indeed, it has not even been shown that the discourse we have set up will have the form of a *debate* at all-rather than a plurality of culturally specific narratives, each giving the thick definition of a virtue that corresponds to the experience and traditions of a particular group. There is an important disanalogy with the case of thunder, on which the Aristotelian so much relies in arguing that our questions will have a single answer. For in that case what is given in experience is the definiendum itself, so that experience establishes a rough extension, to which any good definition must respond. In the case of the virtues, things are more indirect. What is given in experience across groups is only the *ground* of virtuous action, the circumstances of life to which virtuous action is an appropriate response. Even if these grounding experiences are shared, that does not tell us that there will be a shared appropriate response.

In the case of thunder, furthermore, the conflicting theories are clearly put forward as competing candidates for the truth; the behaviour of those involved in the discourse about virtue suggests that they are indeed, as Aristotle says, searching 'not for the way of their ancestors, but for the good'. And it seems reasonable in that case for them to do so. It is far less clear, where the virtues are concerned (the objector continues), that a unified practical solution is either sought by the actual participants or a desideratum for them. The Aristotelian proposal makes it possible to conceive of

a way in which the virtues might be non-relative. It does not, by itself, answer the question of relativism.

The second objection goes deeper. For it questions the notion of spheres of

end p.251

shared human experience that lies at the heart of the Aristotelian approach. The approach, says this objector, seems to treat the experiences that ground the virtues as in some way primitive, given, and free from the cultural variation that we find in the plurality of normative conceptions of virtue. Ideas of proper courage may vary, but the fear of death is shared by all human beings. Ideas of moderation may vary, but the experiences of hunger, thirst, and sexual desire are (so the Aristotelian seems to claim) invariant. Normative conceptions introduce an element of cultural interpretation that is not present in the grounding experiences, which are, for that very reason, the Aristotelian's starting point.

But, the objector continues, such assumptions are naïve. They will not stand up either to our best account of experience or to a close examination of the ways in which these so-called grounding experiences are in fact differently constructed by different cultures. In general, first of all, our best accounts of the nature of experience, even perceptual experience, inform us that there is no such thing as an 'innocent eye' that receives an uninterpreted 'given'. Even sense-perception is interpretative, heavily influenced by belief, teaching, language, and in general by social and contextual features. There is a very real sense in which members of different societies do not see the same sun and stars, encounter the same plants and animals, hear the same thunder.

But if this seems to be true of human experience of nature, which was the allegedly unproblematic starting point for Aristotle's account of naming, it is all the more plainly true, the objector claims, in the area of the human good. Here it is only a very naïve and historically insensitive moral philosopher who would say that the experience of the fear of death, or of bodily appetites, is a human constant. Recent anthropological work on the social construction of the emotions, [16]

[16] See, for example, Harré (1986); Lutz (1988).

for example, has shown to what extent the experience of fear has learned and culturally variant elements. When we add that the object of the fear in which the Aristotelian takes an interest is death, which has been so variously interpreted and understood by human beings at different times and in different places the conclusion that the 'grounding experience' is an irreducible plurality of experiences, highly various and in each case deeply infused with cultural interpretation, becomes even more inescapable.

Nor is the case different with the apparently less complicated experience of the bodily appetites. Most philosophers who have written about the appetites have treated hunger, thirst, and sexual desire as human universals, stemming from our shared animal nature. Aristotle himself was already more sophisticated, since he insisted that the object of appetite is 'the apparent good' and that appetite is therefore something interpretative and selective, a kind of intentional awareness. [17]

[17] See Nussbaum (1978: notes on ch. 6), and Nussbaum (1986*a*: ch. 9).

But he does not seem to have reflected much about the ways in which historical and cultural differences could shape that awareness. The Hellenistic philosophers who immediately followed him did so reflect, arguing that the experience

end p.252

of sexual desire and of many forms of the desire for food and drink are, at least in part, social constructs, built up over time on the basis of a social teaching about

value; this is external to start with, but it enters so deeply into the perceptions of the individual that it actually forms and transforms the experience of desire.[18]

[18] A detailed study of the treatment of these ideas in the three major Hellenistic schools is presented in Nussbaum (forthcoming *b*); portions are published in Nussbaum (1986*b*, 1987*a*, 1989, 1990*a*); see also Nussbaum (1988*b*).

Let us take two Epicurean examples. People are taught that to be well fed they require luxurious fish and meat, that a simple vegetarian diet is not enough. Over time, the combination of teaching with habit produces an appetite for meat, shaping the individual's perceptions of the objects before him. Again, people are taught that what sexual relations are all about is a romantic union or fusion with an object who is seen as exalted in value, or even as perfect. Over time, this teaching shapes sexual behaviour and the experience of desire, so that sexual arousal itself responds to this culturally learned scenario.[19]

[19] The relevant texts are discussed in Nussbaum (forthcoming *b*); see also Nussbaum (1986*b*, 1989, 1990*a*).

.

This work of social criticism has recently been carried further by Michel Foucault, in his *History of Sexuality*.[20]

[20] Foucault (1984).

This work has certain gaps as a history of Greek thought on this topic, but it does succeed in establishing that the Greeks saw the problem of the appetites and their management in an extremely different way from that of twentieth-century Westerners. To summarize two salient conclusions of his complex argument: first, the Greeks did not single out the sexual appetite for special treatment; they treated it alongside hunger and thirst, as a drive that needed to be mastered and kept within bounds. Their central concern was with self-mastery, and they saw the appetites in the light of this concern. Furthermore, where the sexual appetite is concerned, they did not regard the gender of the partner as particularly important in assessing the moral value of the act. Nor did they as morally salient a stable disposition to prefer partners of one sex rather than the other. Instead, they focused on the general issue of activity and passivity, connecting it in complex ways with the issue of self-mastery.

Work like Foucault's — and there is a lot of it in various areas, some of it very good — shows very convincingly that the experience of bodily desire, and of the body itself, has elements that vary with cultural and historical change. The names that people call their desires and themselves as subjects of desire, the fabric of belief and discourse into which they integrate their ideas of desiring: all this influences, it is clear, not only their reflection about desire, but also their experience of desire itself. Thus, for example, it is naïve to treat our modern debates about homosexuality as continuations of the very same debate about sexual activity that went on in the Greek world.[21]

[21] See also Halperin (1990); Winkler (1990); Halperin, Winkler, and Zeitlin (1990).

In a very real sense there was no 'homosexual experience' in a culture that did not contain our emphasis on the gender of the partner, the subjectivity of inclination, and the permanence of

end p.253

appetitive disposition, nor our particular ways of problematizing certain forms of behaviour.

If we suppose that we can get underneath this variety and this constructive power of social discourse in at least one case — namely, with the universal experience of bodily pain as a bad thing — even here we find subtle arguments against us. For the

experience of pain seems to be embedded in a cultural discourse as surely as the closely related experiences of the appetites, and significant variations can be alleged here as well. The Stoics had already made this claim against the Aristotelian virtues. In order to establish that bodily pain is not bad by its very nature, but only by cultural tradition, the Stoics had to provide some explanation for the ubiquity of the belief that pain is bad and of the tendency to shun it. This explanation would have to show that the reaction is learned rather than natural, and to explain why, in the light of this fact, it is learned so widely. This they did by pointing to certain features in the very early treatment of infants. As soon as an infant is born, it cries. Adults, assuming that the crying is a response to its pain at the unaccustomed coldness and harshness of the place where it finds itself, hasten to comfort it. This behaviour, often repeated, teaches the infant to regard its pain as a bad thing — or, better, teaches it the concept of pain, which includes the notion of badness, and teaches it the forms of life its society shares concerning pain. It is all social teaching, they claimed, though this usually escapes our notice because of the early and non-linguistic nature of the teaching.[22]

[22] The evidence for this part of the Stoic view is discussed in Nussbaum (forthcoming *b*); for a general account of the Stoic account of the passions, see Nussbaum (1987*a*).

These and related arguments, the objector concludes, show that the Aristotelian idea that there can be a single, non-relative discourse about human experiences such as mortality or desire is a naïve one. There is no such bedrock of shared experience, and thus no single sphere of choice within which the virtue is the disposition to choose well. So the Aristotelian project cannot even get off the ground.

Now the Aristotelian confronts a third objector, who attacks from a rather different direction. Like the second, she charges that the Aristotelian has taken for a universal and necessary feature of human life an experience that is contingent on certain non-necessary historical conditions. Like the second, she argues that human experience is much more profoundly shaped by non-necessary social features than the Aristotelian has allowed. But her purpose is not simply, like the second objector's, to point to the great variety of ways in which the 'grounding experiences' corresponding to the virtues are actually understood and lived by human beings. It is more radical still. It is to point out that we could imagine a form of human life that does not contain these experiences — or some of them — at all, in any form. Thus the virtue that consists in acting well in that sphere need not be included in an account of the human good. In some cases, the experience may even be a sign of *bad* human life, and the corresponding virtue therefore no

end p.254

better than a form of non-ideal adaptation to a bad state of affairs. The really good human life, in such a case, would contain neither the grounding deficiency nor the remedial virtue.

This point is forcefully raised by some of Aristotle's own remarks about the virtue of generosity. One of his arguments against societies that eliminate private ownership is that they thereby do away with the opportunity for generous action, which requires having possessions of one's own to give to others.[23]

[23] *Politics* 1263b11ff.

This sort of remark is tailor-made for the objector, who will immediately say that generosity, if it really rests upon the experience of private possession, is a dubious candidate indeed for inclusion in a purportedly non-relative account of the human virtues. If it rests upon a grounding experience that is non-necessary and is capable of being evaluated in different ways, and of being either included or eliminated in

accordance with that evaluation, then it is not the universal the Aristotelian said it was.

Some objectors of the third kind will stop at this point, or use such observations to support the second objector's relativism. But in another prominent form this argument takes a non-relativist direction. It asks us to assess the grounding experiences against an account of human flourishing, produced in some independent manner. If we do so, the objector urges, we will discover that some of the experiences are remediable deficiencies. The objection to Aristotelian virtue ethics will then be that it limits our social aspirations, encouraging us to regard as permanent and necessary what we might in fact improve to the benefit of all human life. This is the direction in which the third objection to the virtues was pressed by Karl Marx, its most famous proponent.[24]

[24] For discussion of the relevant passages in Marx, see Lukes (1987). For an acute discussion of these issues I am indebted to an exchange between Alan Ryan and Stephen Lukes at the Oxford Philosophical Society, March 1987.

According to Marx's argument, a number of the leading bourgeois virtues are responses to defective relations of production. Bourgeois justice, generosity, etc., presuppose conditions and structures that are not ideal and that will be eliminated when communism is achieved. And it is not only the current *specification* of these virtues that will be superseded with the removal of the deficiency. It is the virtues themselves. It is in this sense that communism leads human beings beyond ethics. The Aristotelian is thus urged to inquire into the basic structures of human life with the daring of a radical political imagination. It is claimed that when she does so she will see that human life contains more possibilities than are dreamed of in her list of virtues.

## [5]

Each of these objections is profound. To answer any one of them adequately would require a treatise. But we can still do something at this point to map out



an Aristotelian response to each one, pointing the direction in which a fuller reply might go.

The first objector is right to insist on the distinction between singleness of framework and singleness of answer, and right, again, to stress that in constructing a debate about the virtues based on the demarcation of certain spheres of experience we have not yet answered any of the 'What is *X*?' questions that this debate will confront. We have not even said much about the structure of the debate itself, beyond its beginnings — about how it will both use and criticize traditional beliefs, how it will deal with conflicting beliefs, how it will move critically from the 'way of one's ancestors' to the 'good' — in short, about whose judgements it will trust. I have addressed some of these issues, again with reference to Aristotle, in two other papers;[25]

[25] Nussbaum (1986*a*; ch. 8), and Nussbaum and Sen (1989).

but much more remains to be done. At this point, however, we can make four observations to indicate how the Aristotelian might deal with some of the objector's concerns here. First, the Aristotelian position that I wish to defend need not insist, in every case, on a single answer to the request for a specification of a virtue. The answer might well turn out to be a disjunction. The process of comparative and critical debate will, I imagine, eliminate numerous contenders — for example, the view of justice that prevailed in Cyme. But what remains might well be a (probably small) plurality of acceptable accounts. These accounts may or may not be capable

of being subsumed under a single account of greater generality. Success in the eliminative task will still be no trivial accomplishment. If we should succeed in ruling out conceptions of the proper attitude to one's own human worth that are based on a notion of original sin, for example, this would be moral work of enormous significance, even if we got no further than that in specifying the positive account.
Second, the general answer to a 'What is *X*?' question in any sphere may well be susceptible of several or even of many concrete specifications, in connection with other local practices and local conditions. The normative account where friendship and hospitality are concerned, for example, is likely to be extremely general, admitting of many concrete 'fillings'. Friends in England will have different customs, where regular social visiting is concerned, from friends in ancient Athens. Yet both sets of customs can count as further specifications of a general account of friendship that mentions, for example, the Aristotelian criteria of mutual benefit and well-wishing, mutual enjoyment, mutual awareness, a shared conception of the good, and some form of 'living together'. [26]

[26] See Nussbaum (1986*a*: ch. 12).

Sometimes we may want to view such concrete accounts as optional alternative specifications, to be chosen by a society on the basis of reasons of ease and convenience. Sometimes, on the other hand, we may want to insist that a particular account gives the only legitimate specification of the virtue in question for that concrete context; in that case, the concrete account could be viewed as a part of a longer or fuller version

end p.256

of the single normative account. The decision between these two ways of regarding it will depend upon our assessment of its degree of non-arbitrariness for its context (both physical and historical), its relationship to other non-arbitrary features of the moral conception of that context, and so forth.
Third, whether we have one or several general accounts of a virtue, and whether this/these accounts do or do not admit of more concrete specifications relative to ongoing cultural contexts, the particular choices that the virtuous person, under this conception, makes will always be a matter of being keenly responsive to the local features of his or her concrete context. So in this respect, again, the instructions the Aristotelian will give to the person of virtue do not differ from part of what a relativist would recommend. The Aristotelian virtues involve a delicate balancing between general rules and a keen awareness of particulars, in which process, as Aristotle stresses, the perception of the particular takes priority. It takes priority in the sense that a good rule is a good summary of wise particular choices, and not a court of last resort. Like rules in medicine and navigation, ethical rules should be held open to modification in the light of new circumstances; and the good agent must therefore cultivate the ability to perceive and correctly describe his or her situation finely and truly, including in this perceptual grasp even those features of the situation that are not covered under the existing rule.
I have written a good deal elsewhere on this idea of the 'priority of the particular', exactly what it does and does not imply, in exactly what ways that particular perception is and is not prior to the general rule. Those who want clarification on this central topic will have to turn to those writings. [27]

[27] Nussbaum (1986*a*: ch. 10); Nussbaum (1985, 1987*b*).

What I want to stress here is that Aristotelian particularism is fully compatible with Aristotelian objectivity. The fact that a good and virtuous decision is context-sensitive does not imply that it is right only *relative to*, or *inside*, a limited context, any more

than the fact that a good navigational judgement is sensitive to particular weather conditions shows that it is correct only in a local or relational sense. It is right absolutely, objectively, anywhere in the human world, to attend to the particular features of one's context; and the person who so attends and who chooses accordingly is making, according to Aristotle, the humanly correct decision, period. If another situation should ever arise with all the same ethically relevant features, including contextual features, the same decision would again be absolutely right. [28]

[28] I believe, however, that some morally relevant features, in the Aristotelian view, may be features that are not, even in principle, replicable in another context. See Nussbaum (1986*a*: ch. 10) and Nussbaum (1985, 1990*b*).

It should be stressed that the value of contextual responsiveness and the value of getting it right are seen by the Aristotelian as mutually supportive here, rather than in tension. For the claim is that only when we have duly responded to the complexities of the context, seeing it for the very historical situation it is, will

end p.257

we have any hope of making the right decision. Short of that, the importation of plausible general values, however well intentioned, may do no good at all, and may actually make things worse. Nor, the Aristotelian argues, have we been sufficiently responsive to the context before us if we do not see the humanity in it: do not, that is, respond to the claims of human need, the strivings towards the good, the frustrations of human capability, that this situation displays to the reflective person. To study it with detached scientific interest, as an interesting set of local traditions, is not to respond sufficiently to the concrete situation it is; for whatever it is, it is concretely human.

An example from the development context will illustrate this mutual support. In *A Quiet Revolution*, an eloquent study of women's education in rural Bangladesh, [29]

[29] Chen (1986).

Martha Chen describes the efforts of a government development group, the Bangladesh Rural Advancement Committee, to increase the rate of female literacy in certain rural areas. The project began from a conviction that literacy is an important ingredient in the development of these women towards greater capability to live well. It was seen as closely linked with other important values, such as economic flourishing, autonomy, and self-respect. This conviction did not derive from the local traditions of the villages, where women had in fact little autonomy and no experience of education; it derived from the experiences and reflections of the development workers, who were themselves from many different backgrounds and two nationalities. (Chen herself is an American with a Ph.D. in Sanskrit.) The group as a whole lacked experience of the concrete ways of life of rural women, and thus had, as Chen says, 'no specific concepts or strategies' [30]

[30] Chen (1986: p. ix).

for working with them. In the first phase of the programme, then, the development workers went directly to the rural villages with their ideas of literacy and its importance, offering adult literacy materials borrowed from another national programme, and trying to motivate the women of the communities they entered to take them on.

But their lack of contextual knowledge made it impossible for them to succeed, in this first phase. Women found the borrowed literacy materials boring and irrelevant to their lives. They did not see how literacy would help them; even the accompanying vocational training was resisted, since it focused on skills for which there was little demand in that area. Thus failure led the agency to rethink their approach. On the

one hand, they never abandoned their basic conviction that literacy was important for these women; their conclusion, based on wide experience and on their picture of what the women's lives might be, still seemed sound. On the other hand, they recognized that far more attention to the lives and thoughts of the women involved would be necessary if they were going to come up with an understanding of what literacy might do and be for them. They began to substitute for the old approach a more participatory one, in which local co-operative groups brought together development workers with local women,

end p.258

whose experience and sense of life were regarded as crucial. This concept of the co-operative group led to a much more complex understanding of the situation, as the development workers grasped the network of relationships within which the women had to function and the specific dimensions of their poverty and constraints, and as the women grasped the alternative possibilities and began to define for themselves a set of aspirations for change. The result, which continues, has been a slow and complex evolution in the role of women in the villages. A visiting journalist wrote, some years later:

I saw the seeds of the quiet revolution starting in village women's lives. At the meeting houses BRAC has built, the wives, young and old, are learning to read and write. Forbidden from doing marketing, they now at least can keep the accounts . . . In one fishing village, the women have even become the bankers, saving over $2000 and lending it to their men to buy better equipment. It started in the simplest way—they collected a handful of rice a week from each family, stored it, and sold it in the market. About 50 villages from each area have thriving women's cooperatives, investing in new power-pumps or seed, and winning respect for their members. [31]

[31] Cited in Chen (1986: 4-5). Chen stresses that one important factor in this later success was that the group had no dogmatic adherence to an abstract theory of development, but had a flexible and situation-guided approach.

This is how the Aristotelian approach works—hanging on to a general (and open-ended) picture of human life, its needs and possibilities, but at every stage immersing itself in the concrete circumstances of history and culture. Chen's detailed narrative—which in its very style manifests a combination of Aristotelian commitment to the human good and Aristotelian contextual sensitivity—shows that the two elements go, and must go, together. If the development workers had approached these women as alien beings whose ways could not be compared with others and considered with a view to the human good, no change would have taken place—and the narrative convinces the reader that these changes have been good. On the other hand, general talk of education and self-respect did nothing at all until it came from within a concrete historical reality. Immersion made it possible to get the choice that was humanly right.

Thus the Aristotelian virtue-based morality can capture a great deal of what the relativist is after, and still make a claim to objectivity, in the sense we have described. In fact, we might say that the Aristotelian virtues do better than the relativist virtues in explaining what people are actually doing when they scrutinize the features of their context carefully, looking at both the shared and the non-shared features with an eye to what is best. For, as Aristotle says, people who do this are usually searching for the good, not just for the way of their ancestors. They are prepared to defend their decisions as good or right, and to think of those who advocate a different course as disagreeing about what is right, not just narrating a different tradition.

Finally, we should point out that the Aristotelian virtues, and the deliberations they guide, unlike some systems of moral rules, remain always open to revision

in the light of new circumstances and new evidence. In this way, again, they contain the flexibility to local conditions that the relativist would desire—but, again, without sacrificing objectivity. Sometimes the new circumstances may simply give rise to a new concrete specification of the virtue as previously defined; in some cases it may cause us to change our view about what the virtue itself is. All general accounts are held provisionally, as summaries of correct decisions and as guides to new ones. This flexibility, built into the Aristotelian procedure, will again help the Aristotelian account to answer the questions of the relativist, without relativism.

## [6]

We must now turn to the second objection. Here, I believe, is the really serious threat to the Aristotelian position. Past writers on virtue, including Aristotle himself, have lacked sensitivity to the ways in which different traditions of discourse, different conceptual schemes, articulate the world, and also to the profound connections between the structure of discourse and the structure of experience itself. Any contemporary defence of the Aristotelian position must display this sensitivity, responding somehow to the data that the relativist historian or anthropologist brings forward.

The Aristotelian should begin, it seems to me, by granting that with respect to any complex matter of deep human importance there is no 'innocent eye', no way of seeing the world that is entirely neutral and free of cultural shaping. The work of philosophers such as Putnam, Goodman, and Davidson [32]

[32] See Putnam (1979, 1981, 1988); Goodman (1968, 1978); Davidson (1984).

—following, one must point out, from the arguments of Kant and, I believe, from those of Aristotle himself [33]

[33] On his debt to Kant, see Putnam (1988); on Aristotle's relationship to 'internal realism', see Nussbaum (1986*a*: ch. 8).

—have shown convincingly that even where sense-perception is concerned, the human mind is an active and interpretative instrument, and that its interpretations are a function of its history and its concepts, as well as of its innate structure. The Aristotelian should also grant, it seems to me, that the nature of human world interpretations is holistic and that the criticism of them must, equally, be holistic. Conceptual schemes, like languages, hang together as whole structures, and we should realize, too, that a change in any single element is likely to have implications for the system as a whole.

But these two facts do not imply, as some relativists in literary theory and in anthropology tend to assume, that all world interpretations are equally valid and altogether non-comparable, that there are no good standards of assessment and 'anything goes'. The rejection of the idea of ethical truth as correspondence to an altogether uninterpreted reality does not imply that the whole idea of searching for the truth is an old-fashioned error. Certain ways in which people see the world can still be criticized exactly as Aristotle


criticized them: as stupid, pernicious, and false. The standards used in such criticisms must come from inside human life. (Frequently they will come from the society in question itself, from its own rationalist and critical traditions.) And the inquirer must attempt, prior to criticism, to develop an inclusive understanding of the conceptual scheme being criticized, seeing what motivates each of its parts and how they hang together. But there is so far no reason to think that the critic will not be able to reject the institution of slavery, or the homicide law of Cyme, as out of line

with the conception of virtue that emerges from reflection on the variety of different ways in which human cultures have had the experiences that ground the virtues. The grounding experiences will not, the Aristotelian should concede, provide precisely a single, language-neutral bedrock on which an account of virtue can be straightforwardly and unproblematically based. The description and assessment of the ways in which different cultures have constructed these experiences will become one of the central tasks of Aristotelian philosophical criticism. But the relativist has, so far, shown no reason why we could not, at the end of the day, say that certain ways of conceptualizing death are more in keeping with the totality of our evidence and the totality of our wishes for flourishing life than others; that certain ways of experiencing appetitive desire are for similar reasons more promising than others. Relativists tend, furthermore, to understate the amount of attunement, recognition, and overlap that actually obtains across cultures, particularly in the areas of the grounding experiences. The Aristotelian, in developing her conception in a culturally sensitive way, should insist, as Aristotle himself does, upon the evidence of such attunement and recognition. Despite the evident differences in the specific cultural shaping of the grounding experiences, we do recognize the experiences of people in other cultures as similar to our own. We do converse with them about matters of deep importance, understand them, allow ourselves to be moved by them. When we read Sophocles' *Antigone*, we see a good deal that seems strange to us; and we have not read the play well if we do not notice how far its conceptions of death, womanhood, and so on differ from our own. But it is still possible for us to be moved by the drama, to care about its people, to regard their debates as reflections upon virtue that speak to our own experience, and their choices as choices in spheres of conduct in which we too must choose. Again, when one sits down at a table with people from other parts of the world and debates with them concerning hunger, or just distribution, or in general the quality of human life, one does find, in spite of evident conceptual differences, that it is possible to proceed as if we were all talking about the same human problem; and it is usually only in a context in which one or more of the parties is intellectually committed to a theoretical relativist position that this discourse proves impossible to sustain. This sense of community and overlap seems to be especially strong in the areas that we have called the areas of the grounding experiences. And this, it seems, supports the Aristotelian claim that those

end p.261

experiences can be a good starting point for ethical debate.

Furthermore, it is necessary to stress that hardly any cultural group today is as focused upon its own internal traditions and as isolated from other cultures as the relativist argument presupposes. Cross-cultural communication and debate are ubiquitous facts of contemporary life, and our experience of cultural interaction indicates that in general the inhabitants of different conceptual schemes do tend to view their interaction in the Aristotelian and not the relativist way. A traditional society, confronted with new technologies and sciences, and the conceptions that go with them, does not in fact simply fail to understand them, or regard them as totally alien incursions upon a hermetically sealed way of life. Instead, it assesses the new item as a possible contributor to flourishing life, making it comprehensible to itself, and incorporating elements that promise to solve problems of flourishing. Examples of such assimilation, and the debate that surrounds it, [34]

[34] Abeysekera (1986).

suggest that the parties do in fact recognize common problems and that the traditional society is perfectly capable of viewing an external innovation as a device

to solve a problem that it shares with the innovating society. The village woman of Chen's narrative, for example, did not insist on remaining illiterate because they had always been so. Instead, they willingly entered into dialogue with the international group, viewing co-operative discussion as a resource towards a better life. The parties do in fact search for the good, not the way of their ancestors; only traditionalist anthropologists insist, nostalgically, on the absolute preservation of the ancestral.

And this is so even when cross-cultural discourse reveals a difference at the level of the conceptualization of the grounding experiences. Frequently the effect of work like Foucault's, which reminds us of the non-necessary and non-universal character of one's own ways of seeing in some such area, is precisely to prompt a critical debate in search of the human good. It is difficult, for example, to read Foucault's observations about the history of our sexual ideas without coming to feel that certain ways in which the Western contemporary debate on these matters has been organized, as a result of some combination of Christian moralism with nineteenth-century pseudoscience, are especially silly, arbitrary, and limiting, inimical to a human search for flourishing. Foucault's moving account of Greek culture, as he himself insists in a preface, [35]

[35] Foucault (1984: ii, preface).

provides not only a sign that someone once thought differently, but also evidence that it is possible for *us* to think differently. (Indeed, this was the whole purpose of genealogy as Nietzsche, Foucault's precursor here, introduced it: to destroy idols once deemed necessary, and to clear the way for new possibilities of creation.) Foucault announced that the purpose of his book was to 'free thought' so that it could think differently, imagining new and more fruitful possibilities. And close analysis of spheres of cultural

end p.262

discourse, which stresses cultural differences in the spheres of the grounding experiences, is being combined, increasingly, in current debates about sexuality and related matters, with a critique of existing social arrangements and attitudes, and an elaboration of new norms of human flourishing. There is no reason to think this combination incoherent. [36]

[36] This paragraph expands remarks made in a commentary on papers by D. Halperin and J. Winkler at the conference on 'Homosexuality in History and Culture' at Brown University, February 1987; Halperin's paper is now in Halperin (1990) and Winkler's in Halperin, Winkler, and Zeitlin (1990). The proposed combination of historically sensitive analysis with cultural criticism was forcefully developed, at the same conference, by Abelove (1987).

As we pursue these possibilities, the basic spheres of experience identified in the Aristotelian approach will no longer, we have said, be seen as spheres of *uninterpreted* experience. But we have also insisted that there is much family relatedness and much overlap among societies. And certain areas of relatively greater universality can be specified here, on which we should insist as we proceed to areas that are more varied in their cultural expression. Not without a sensitive awareness that we are speaking of something that is experienced differently in different contexts, we can none the less identify certain features of our common humanity, closely related to Aristotle's original list, from which our debate might proceed.

1. *Mortality*. No matter how death is understood, all human beings face it and (after a certain age) know that they face it. This fact shapes every aspect of more or less every human life.
2. *The body*. Prior to any concrete cultural shaping, we are born with human bodies,

1. *Mortality*. No matter how death is understood, all human beings face it and (after a certain age) know that they face it. This fact shapes every aspect of more or less every human life.
   whose possibilities and vulnerabilities do not as such belong to any culture rather than any other. Any given human being might have belonged to any culture. The experience of the body is culturally influenced; but the body itself, prior to such experience, provides limits and parameters that ensure a great deal of overlap in what is going to be experienced, where hunger, thirst, desire, and the five senses are concerned. It is all very well to point to the cultural component in these experiences. But when one spends time considering issues of hunger and scarcity, and in general of human misery, such differences appear relatively small and refined, and one cannot fail to acknowledge that 'there are no known ethnic differences in human physiology with respect to metabolism of nutrients. Africans and Asians do not burn their dietary calories or use their dietary protein any differently from Europeans and Americans. It follows then that dietary requirements cannot vary widely as between different races.'[37]
   [37] Gopalan (forthcoming).
   This and similar facts should surely be focal points for debate about appropriate human behaviour in this sphere. And by beginning with the body, rather than with the subjective experience of desire, we get, furthermore, an opportunity to criticize the situation of people who are so persistently deprived that their *desire* for good things has actually decreased. This is a further advantage of the Aristotelian approach, when contrasted

end p.263

   with approaches to choice that stop with subjective expressions of preference.
3. *Pleasure and pain*. In every culture, there is a conception of pain; and these conceptions, which overlap very largely with one another, can plausible be seen as grounded in universal and pre-cultural experience. The Stoic story of infant development is highly implausible; the negative response to bodily pain is surely primitive and universal, rather than learned and optional, however much its specific 'grammar' may be shaped by later learning.
4. *Cognitive capability*. Aristotle's famous claim that 'all human beings by nature reach out for understanding'[38]
   [38] *Metaphysics* I. 1.
   seems to stand up to the most refined anthropological analysis. It points to an element in our common humanity that is plausibly seen, again, as grounded independently of particular acculturation, however much it is later shaped by acculturation.
5. *Practical reason*. All human beings, whatever their culture, participate (or try to) in the planning and managing of their lives, asking and answering questions about how one should live and act. This capability expresses itself differently in different societies, but a being who altogether lacked it would not be likely to be acknowledged as a human being, in any culture.[39]
   [39] See Nussbaum (1988), where this Aristotelian view is compared with Marx's views on truly human functioning.
6. *Early infant development*. Prior to the greater part of specific cultural shaping, though perhaps not free from all shaping, are certain areas of human experience and development that are broadly shared and of great importance for the Aristotelian virtues: experiences of desire, pleasure, loss, one's own finitude, perhaps also of envy, grief, and gratitude. One may argue about the merits of one

with approaches to choice that stop with subjective expressions of preference. or another psychoanalytical account of infancy. But it seems difficult to deny that the work of Freud on infant desire and of Klein on grief, loss, and other more complex emotional attitudes has identified spheres of human experience that are to a large extent common to all humans, regardless of their particular society. All humans begin as hungry babies, perceiving their own helplessness, their alternating closeness to and distance from those on whom they depend, and so forth. Melanie Klein records a conversation with an anthropologist in which an event that at first looked (to Western eyes) bizarre was interpreted by Klein as the expression of a universal pattern of mourning. The anthropologist accepted her interpretation. [40]

[40] Klein (1984): 247-63.

7. *Affiliation*. Aristotle's claim that human beings as such feel a sense of fellowship with other human beings, and that we are by nature social animals, is an empirical claim; but it seems to be a sound one. However varied our specific conceptions of friendship and love are, there is a great point in seeing them as overlapping expressions of the same family of shared human needs and desires.
8. *Humour*. There is nothing more culturally varied than humour; and yet, as Aristotle insists, some space for humour and play seems to be a need of

end p.264

any human life. The human being is not called the 'laughing animal' for nothing; it is certainly one of our salient differences from almost all other animals, and (in some form or other) a shared feature, I somewhat boldly assert, of any life that is going to be counted as fully human.

This is just a list of suggestions, closely related to Aristotle's list of common experiences. One could subtract some of these items and/or add others. [41]

[41] See Nussbaum (1990*c*) for a slightly longer list, including discussion of our relationship to other species and to the world of nature. It is very interesting to notice that three other lists in this volume, prepared independently, contain almost the same items as this one: Dan Brock's list of basic human functions used in quality of life measures in medical ethics; Erik Allardt's enumeration of the functions observed by Finnish social scientists; and Robert Erikson's list of functions measured by the Swedish group. Only the last two may show mutual influence. So much independent convergence testifies to the ubiquity of these concerns and to their importance.

But it seems plausible to claim that in all these areas we have a basis for further work on the human good. We do not have a bedrock of completely uninterpreted 'given' data, but we do have nuclei of experience around which the constructions of different societies proceed. There is no Archimedean point here, no pure access to unsullied 'nature'—even, here, human nature—as it is in and of itself. There is just human life as it is lived. But in life as it is lived, we do find a family of experiences, clustering around certain focuses, which can provide reasonable starting points for cross-cultural reflection.

This paper forms part of a larger project. The role of the preliminary list proposed in this section can be better understood if I briefly set it in the context of this more comprehensive enterprise, showing its links with other arguments. In a paper entitled 'Nature, Function, and Capability: Aristotle on Political Distribution', [42]

[42] Nussbaum (1988*a*).

I discuss an Aristotelian conception of the proper function of government, according to which its task is to make available to each and every member of the community the basic necessary conditions of the capability to choose and live a fully good human life, with respect to each of the major human functions included in that fully good life. I examine sympathetically Aristotle's argument that, for this reason, the task

of government cannot be well performed, or its aims well understood, without an understanding of these functionings. A closely connected study, 'Aristotelian Social Democracy', [43]

[43] Nussbaum (1990*c*).

shows a way of moving from a general understanding of the circumstances and abilities of human beings (such as this list provides) to an account of the most important human functions that it will be government's job to make possible. It shows how this understanding of the human being and the political task can yield a conception of social democracy that is a plausible alternative to liberal conceptions. Meanwhile, in a third paper, 'Aristotle on Human Nature and the Foundations of Ethics', [44]

[44] Nussbaum (forthcoming *a*).

I focus on the special role of two of the human capabilities

end p.265

recognized in this list: affiliation (or sociability) and practical reason. I argue that these two play an architectonic role in human life, suffusing and also organizing all the other functions — which will count as truly human functions only in so far as they are done with some degree of guidance from both of these. Most of the paper is devoted to an examination of Aristotle's arguments for saying that these two elements are parts of 'human nature'. I argue that this is not an attempt to base human ethics on a neutral bedrock of scientific fact outside of human experience and interpretation. I claim that Aristotle seeks, instead, to discover, among the experiences of groups in many times and places, certain elements that are especially broadly and deeply shared. And I argue that the arguments justifying the claims of these two to be broad and deep in this way have a self-validating structure: that is, anyone who participates in the first place in the inquiry that supports them affirms, by that very fact, her own recognition of their salience. This is an important continuation of the project undertaken in this paper, since it shows exactly how Aristotle's 'foundation' for ethics can remain inside human history and self-interpretation, and yet still claim to be a foundation.

## [7]

The third objection raises, at bottom, a profound conceptual question: What is it to inquire about the *human* good? What circumstances of existence go to define what it is to live the life of a *human being*, and not some other life? Aristotle likes to point out that an inquiry into the human good cannot, on pain of incoherence, end up describing the good of some other being, say a god—a good that, on account of our circumstances, it is impossible for us to attain. [45]

[45] Cf. *EN* 1159*a* 10-12, 1166a 18-23.

What circumstances then? The virtues are defined relatively to certain problems and limitations, and also to certain endowments. Which ones are sufficiently central that their removal would make us into different beings, and open up a wholly new and different debate about the good? This question is itself part of the ethical debate we propose. For there is no way to answer it but to ask ourselves which elements of our experience seem to us so important that they count, for us, as part of who we are. I discuss Aristotle's attitude to this question elsewhere, and I shall simply summarize here. [46]

[46] Nussbaum (forthcoming *a*).

It seems clear, first of all, that our mortality is an essential feature of our circumstances as human beings. An immortal being would have such a different form

of life, and such different values and virtues, that it does not seem to make sense to regard that being as part of the same search for good. Essential, too, will be our dependence upon the world outside us: some sort of need for food, drink, the help of others. On the side of abilities, we would want to include cognitive

end p.266

functioning and the activity of practical reasoning as elements of any life that we would regard as human. Aristotle argues, plausibly, that we would want to include sociability as well, some sensitivity to the needs of and pleasure in the company of other beings similar to ourselves.

But it seems to me that the Marxist question remains, as a deep question about human forms of life and the search for the human good. For one certainly can imagine forms of human life that do not contain the holding of private property—not, therefore, those virtues that have to do with its proper management. And this means that it remains an open question whether these virtues ought to be regarded as virtues, and kept upon our list. Marx wished to go much further, arguing that communism would remove the need for justice, courage, and most of the bourgeois virtues. I think we might be sceptical here. Aristotle's general attitude to such transformations of life is to suggest that they usually have a tragic dimension. If we remove one sort of problem—say, by removing private property—we frequently do so by introducing another—say, the absence of a certain sort of freedom of choice, the freedom that makes it possible to do fine and generous actions for others. If things are complex even in the case of generosity, where we can easily imagine the transformation that removes the virtue, they are surely far more so in the cases of justice and courage. And we would need a far more detailed description than Marx ever gives us of the form of life under communism, before we could even begin to see whether it would in fact transform things where these virtues are concerned, and whether it would or would not introduce new problems and limitations in their place. In general it seems that all forms of life, including the imagined life of a god, contain boundaries and limits. [47]

[47] See Nussbaum (1986*b*: ch. 11).

All structures, even that of putative limitlessness, are closed to something, cut off from something—say, in that case, from the specific value and beauty inherent in the struggle against limitation. Thus it does not appear that we will so easily get beyond the virtues. Nor does it seem to be so clearly a good thing for human life that we should.

## [8]

The best conclusion to this sketch of an Aristotelian programme for virtue ethics was written by Aristotle himself, at the end of his discussion of human nature in *Nicomachean Ethics* I:

So much for our outline sketch for the good. For it looks as if we have to draw an outline first, and fill it in later. It would seem to be open to anyone to take things further and to articulate the good parts of the sketch. And time is a good discoverer or ally in such things. That's how the sciences have progressed as well: it is open to anyone to supply what is lacking. (*EN* 1098a20-6)

end p.267

**Bibliography**

Abelove, H. (1987). 'Is Gay History Possible?' Paper (unpublished) delivered at the Conference on Homosexuality in History and Culture, Brown University, Feb. 1987.

Abeysekera, C. (1986). Address (unpublished) to the WIDER Conference on Value and Technology, summer 1986.

Chen, M. (1986). *A Quiet Revolution: Women in Transition in Rural Bangladesh*. Cambridge, Mass.: Schenkman Publishing Company.
Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
Foot, P. (1978). *Virtues and Vices*. Berkeley: University of California Press.
Foucault, M. (1984). *Histoire de la sexualité*, ii, iii. Paris: Les Belles Lettres.
Goodman, N. (1968). *Languages of Art*. Indianapolis: Hackett.
—— (1978). *Ways of World-Making*. Indianapolis: Hackett.
Gopalan, C. (1992). 'Undernutrition: Measurement and Implications', In S. Osmani (ed.), *Nutrition and Poverty*. Oxford: Clarendon Press.
Halperin, D. (1990). *One Hundred Years of Homosexuality and Other Essays on Greek Love*. New York: Routledge, Chapman, and Hall.
——, Winkler, J., and Zeitlin, F. (1990). *Before Sexuality*. Princeton: Princeton University Press.
Hampshire, S. (1983). *Morality and Conflict*. Cambridge, Mass.: Harvard University Press.
Harré, R. (ed.) (1986). *The Social Construction of Emotions*. Oxford: Basil Blackwell.
Klein, M. (1984). 'Our Adult World and its Roots in Infancy', in *Envy, Gratitude and Other Works 1946-1963*. London: Hogarth Press, 247-63.
Lukes, S. (1987). *Marxism and Morality*. Oxford: Clarendon Press.
Lutz, C. (1988). *Unnatural Emotions*. Chicago: University of Chicago Press.
Macintyre, A. (1981). *After Virtue*. Notre Dame: Notre Dame University Press.
—— (1988). *Whose Justice? Which Rationality?* Notre Dame: Notre Dame University Press.
Nussbaum, M. (1972). '*Psuche* in Heraclitus', *Phronesis*, 17, 1-17, 153-70.
—— (1978). *Aristotle's De Motu Animalium*. Princeton: Princeton University Press.
—— (1985). 'The Discernment of Perception: An Aristotelian Model of Public and Private Rationality', *Proceedings of the Boston Area Colloquium for Ancient Philosophy*, 1, 151-201. Also in Nussbaum (1990).
—— (1986*a*). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.
—— (1986*b*). 'Therapeutic Arguments: Epicurus and Aristotle', in M. Schofield and G. Striker (eds.), *The Norms of Nature*. Cambridge: Cambridge University Press, 31-74.
—— (1987*a*). 'The Stoics on the Extirpation of the Passions', *Apeiron*, 20, 129-77.
—— (1987*b*). "Finely Aware and Richly Responsible": Literature and the Moral Imagination', in A. Cascardi (ed.), *Literature and the Question of Philosophy*. Baltimore: Johns Hopkins University Press, 169-91.
—— (1988*a*). 'Nature, Function, and Capability: Aristotle on Political Distribution', *Oxford Studies in Ancient Philosophy*, suppl. vol., 145-84.
—— (1988*b*). 'Narrative Emotions: Beckett's Genealogy of Love', *Ethics*, 98, 225-54.
end p.268
—— (1989). 'Beyond Obsession and Disgust: Lucretius' Genealogy of Love', *Apeiron*, 22, 1-59.
—— (1990*a*). 'Mortal Immortals: Lucretius on Death and the Voice of Nature', *Philosophy and Phenomenological Research*, 50, 305-51.
—— (1990*b*). *Love's Knowledge: Essays on Philosophy and Literature*. Oxford: Oxford University Press.
—— (1990*c*). 'Aristotelian Social Democracy', in R. B. Douglass, G. Mara, and H. Richardson (eds.). *Liberalism and the Good*. New York: Routledge, 203-52.

—— (forthcoming *a*). 'Aristotle on Human Nature and the Foundations of Ethics', in a volume on the philosophy of Bernard Williams, ed. J. Altham and R. Harrison. Cambridge: Cambridge University Press.
—— (forthcoming *b*). *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. The Martin Classical Lectures, 1986.
—— and Sen, A. (1989). 'Internal Criticism and Indian Rationalist Traditions', in M. Krausz (ed.), *Relativism*. Notre Dame: Notre Dame University Press.
Procope, J. (forthcoming). 'Hochherzigkeit', *Reallexikon für Antike und Christentum*, 14.
Putnam, H. (1979). *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
—— (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
—— (1988). *The Many Faces of Realism: The Carus Lectures*. La Salle, Ill. Open Court.
Sturgeon, N. (1984). Review of Foot, *Journal of Philosophy*, 81, 326-33.
Walzer, M. (1983). *Spheres of Justice*. Oxford: Basil Blackwell.
—— (1987). *Interpretation and Social Criticism*. Cambridge, Mass.: Harvard University Press.
Williams, B. (1984). 'Philosophy', in *The Legacy of Greece*, ed. M. I. Finley. Oxford: Oxford University Press.
—— (1985). *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.
Winkler, J. (1990). *The Constraints of Desire*. New York: Routledge, Chapman, and Hall.

end p.269

# Martha Nussbaum: Non-Relative Virtues: An Aristotelian Approach

show chapter abstract and keywords
hide chapter abstract and keywords

**Susan Hurley**

Does appeal to forms of life underwrite or undermine claims to objectivity? The question arises in various contexts and forms. The forms of life appealed to may be linguistic or other practices, or aspects of experience or thought. Objectivity may seem threatened in various ways: by relativism, given the variety of forms of life; by idealism, if a claim about the way things are is made to depend on forms of life that are ultimately down to us, not the world; by verificationism, if it is suggested that necessary conditions for our practices of determining something to be true are necessary conditions of its *being* true. On the other hand, some, such as Wittgenstein, have thought that the appeal to practices keeps such threats from arising, by engaging one's mind immediately with the world and other minds: if we are able to do certain things well, together, there is no further real issue as to how our private intentions can ever match up. Martha Nussbaum's paper considers whether appeal to the varying forms of life that ground talk of specific virtues tends to support relativism, and argues that it does not. She sketches forms of life that are distinctive of human beings, hedged in as we are in some sense between the gods and the beasts. In contrast with the gods, we have essentially in common certain basic reference-fixing experiences of the conditions of limited, mortal human (as opposed to godlike) life, which allow us to address the same question — how should one act with respect to *that* sort of situation? — even though we may disagree about

the best answer. In contrast with the beasts, we also have essentially in common certain distinctive capacities, for sociality and practical reason, our valued exercises of which may be appealed to in the course of ethical disagreement. Correctly understood, she argues, the role of these specific and distinctively human forms of life in ethics, and in Aristotelian ethics in particular, does not support relativism and does support objectivism.

I am in substantial accord with her conclusion, and indeed have myself offered different arguments for related conclusions, which I will not repeat here. [1]

[1] See my *Natural Reasons* (1989), esp. Part I.

Her responses to the first objection she considers to her position, concerning singleness of solution, seem to me correct. The second objection she considers concerns the lack of culturally neutral, uninterpreted common starting points in experience or practice; while I have certain reservations about her formulation of a response to this objection in terms of different conceptual schemes, I shall pass over those issues here. What I will focus on are certain issues arising out of suggestions she makes in responding to the third objection, concerning the possibility of human life that simply lacks

end p.270

certain forms of life altogether. These issues concern another way in which arguments that appeal to forms of life may seem to undermine claims to objectivity. Nussbaum suggests that a particular kind of self-validating argument is available in discussions of the value to be placed on certain of the features of our common humanity that she lists, such as sociality and practical reason. This is a kind of argument that gets its force not from any logical inconsistency in the position it argues against, but rather from a pragmatic inconsistency: some aspect of the form of life, the activity, practice, or mentality, of the sceptic has implications inconsistent with the content of his view. As we shall see below, such pragmatic inconsistencies often play a role in transcendental arguments — arguments concerning how certain forms of life are possible, what the necessary conditions for them are. Nussbaum suggests an example of what I am calling pragmatic inconsistency: the sceptic's very participation in discussion of the value of practical reason in the course of expressing his sceptical view about it in some sense affirms its value, and so is at odds with the sceptical content of his view. She further suggests that there may be cases in which the essential properties as a human being of the person holding a sceptical view would be negated by the living out of that view, so that such a life — say, the life of a god, or of a beast — could not be a life for that person.

I want to consider two questions about these suggestions. (1) What is the significance with respect to issues about objectivity of the fact that certain expressions of some propositions presuppose other inconsistent propositions? (2) How is our capacity for self-determination related to the way in which we may be constrained in the evaluative views we can defend by what we essentially are as human beings, our essential limitations and aspirations?

The first question arises in the debate about whether transcendental arguments undermine objectivity. They may seem to involve some kind of idealism or relativism or verificationism, by deriving conclusions about the world and/or other possible minds from features of our forms of life, so cutting the world and/or other minds down to the size of our minds. [2]

[2] See Nagel (1986: ch. 6); Williams (1981).

Ross Harrison has offered a rebuttal of the view that transcendental arguments must involve idealism. He points out that such arguments are of the valid inferential form:

*p*, necessarily if *p* then *q*; therefore *q*. The minor premiss *p* is often established pragmatically, by immersion in certain forms of life: experience, language, thought. Various celebrated if problematic arguments, which are not my focus here, support the conditional premisses of such inferences: arguments made by Kant, Strawson, Davidson, etc., concerning the necessary conditions of experience, language, content, thought. But we should notice, Harrison says, that for such inferences to be valid the minor premiss '*I*' need not be necessary. There might not, for example, have been any minds, or any language, at all. Far from being mind-dependent, the world might have been mindless,

end p.271

hence not the object of judgement or thought at all. Such an admission is hardly idealist in implication, nor does it tend to undermine a view of the world as objective. While there is clearly something right about it, Harrison's argument does not seem to go to the heart of the intuition that transcendental arguments somehow tend to undermine objectivity. To this end, we should distinguish between two ways in which the minor premiss of a transcendental argment gets established, even if not established as necessary. This is the premiss for which support may be pragmatic, conceded by something the sceptic does or is, his form of life. Such pragmatic support can take at least two forms. It can be either presuppositional, or demonstrative.

*Presuppositional* support derives from our presupposing that *p* in many of our activities and from the fact that forgoing such activities would be terribly costly. Consider, for example, the argument that the contents of the reactive attitudes presupposed by many of our relationships and practices imply, under certain conditions, blame and guilt, and that such attitudes, presupposed as they are in many of our valued ways of living, would be terribly costly to forgo. We have no good reason to forgo them, when we consider what the consequences would be in terms of the impoverishment of our lives. [3]

[3] See Strawson (1974).

Consider also the argument, suggested by Nussbaum, that participation in philosophical discussion of the value of practical reason presupposes that there is reason to go in for such discussion, which implies the value of practical reason. Or the argument put forward by Putnam that successful reference by someone to vats in asserting that he might be a brain in a vat presupposes that he is in the right sort of a causal connection with vats — not, that is, merely a generic causal connection to artificial skull — in which case he is not a brain in a vat. [4]

[4] Putnam (1979, 1981).

Harrison's point, applied to such arguments, would be that there is no necessity that we go in for reactive attitudes, or for philosophical discussion of the value of practical reason; nor does any necessity attach to reference to vats. We could simply cut ourselves off from other people, refrain from deliberation, or language, or thought about vats, and nothing would be presupposed; in some cases, the price of pragmatically consistent scepticism is silence, isolation, perhaps ever denaturing ourselves, and it may very well not be worth paying, but nevertheless we could pay it. If we were to, nothing could be demonstrated about blame, or the value of practical reason, or whether we are brains in vats. (Nussbaum seems to agree on this when she stresses, in another paper, that there is no hard external obstacle, only internal appeal to what we already value, to block certain sceptical views. [5]

[5] Nussbaum (forthcoming).

)

But it seems to me that this point is not enough to keep the threat to objectivity at bay. The worry that persists can be expressed thus: there is a problem about the relationship between the form of life and the truth of the proposition it presupposes. How can whether we choose to pay the terrible costs of for-going

end p.272

our presuppositions tell us anything about whether the sceptic is wrong? Surely, the objectivist will say, whether he is wrong cannot depend on what forms of life we go in for and hence on what we presuppose; such dependence would itself undermine objectivity. Of course, our withholding our presuppositions does not make the sceptic right; but how can we regard our presuppositions as making him wrong, without giving up objectivity? An objective world is not obliged to conform to our presuppositions. As Stroud points out with respect to conditional necessities to the effect that 'such-and-such is necessary if we are to think and speak as we now do', no theoretical justification is thereby given of our acceptance of the propositions at issue, since we could give up our present ways of thinking and speaking and adopt others—however unattractive a prospect that might be.[6]

[6] Stroud (1982: 126).

This leads me to the second, stronger kind of pragmatic support for the minor premiss *'p'—demonstrative support*. Perhaps it is only really here that transcendental arguments completely avoid the threat to objectivity. If so, it would be important to distinguish weaker, presuppositional varieties of transcendental argument from stronger, demonstrative ones. An indication of demonstrative support is that the heroic sceptical move of forgoing the form of life that supports the minor premiss is utterly pointless, self-defeating.[7]

[7] Compare Stroud's 'privileged class', in Stroud (1982: 127).

Suppose, for example, we have a sceptic about the existence of language. We point out to him that his expressions of his scepticism are self-defeating; his expression demonstrates the existence of a particular meaningful sentence, which necessarily implies that language exists. It would be utterly pointless for him now heroically to renounce language; indeed, it would be self-defeating in some-what the same way that it would be self-defeating for a sceptic about other minds to proceed in trying to win his argument by systematically killing off his fellows. There would be no point in renouncing language unless one had already conceded implicitly the existence of language; the whole point of renouncing it is to avoid having one's sceptical thesis refuted by it. In this sense, the sceptic's expression of his thesis does not merely presuppose something inconsistent with his thesis; it actually demonstrates the falsity of it. Harrison's point still holds: there is no necessity that there should have been language; the world might have been languageless. Nevertheless, the threat to objectivity involved when the world is held to conform to our presuppositions is absent. In these cases, our forms of life do not merely presuppose something inconsistent with the scepticism in question; rather, our forms of life demonstrate the truth of something inconsistent with scepticism. Other examples of the demonstrative variety of pragmatic support might be arguments that involve appeal to the existence of experience, or thought, as necessarily implying the existence of embodied, or world-situated, persons. Again, if the conditional premiss, 'necessarily if *p* then *q*', holds (which I am not discussing at all here),

end p.273

no pragmatic retreat route from the minor premiss '*p*' is available: refraining from experience, or thought—that is, deliberately doing away with it by, say, suicide—would be self-defeating as well as heroic on the part of a sceptic in a way that refraining from personal relationships, practical reason, or attempted reference to

vats would not be. Even though it is perfectly possible that there might not have been any experience, or thought, in the universe, doing away with one's own experience or thought concedes its existence; by contrast, refraining from the latter activities makes no concession about the reasons for going in for them or their success.

Now I return to the second question: I will try to apply the above points to arguments which appeal for pragmatic support to human nature, to what we essentially are. Suppose that there are various distinctive features, both negative and positive, limitations and capacities, that are essential to us as human beings. If so, then those that Nussbaum highlights in her reply to the third objection—for example, limitations owing to mortality and physiological requirements, the capacity for practical rationality, and sociality—may plausibly be claimed to be among the strongest candidates for such essential features. Certain kinds of lives, then, could not be lives for us as human beings, given our nature; they would rather be lives for gods, or asocial anthromorphs, or grazing animals, or shellfish. [8]

[8] I am drawing here also on Nussbaum (forthcoming).

Towards the end of 'Non-Relative Virtues', Nussbaum asks 'What circumstances of existence go to define what it is to live the life of a human being, and not some other life?' and comments: 'Aristotle likes to point out that an inquiry into the human good cannot, on pain of incoherence, end up describing the good of some other being, say a god, a good, that on account of our circumstances, it is impossible for us to obtain.' If someone is tempted to ask the question: But is it best, after all, to live the life of a human being, rather than one of these other kinds of life? The response cannot be simply that it is impossible for human beings to decide to live lives that lack these characteristics. If that were true, the discussion would lose much of its point. As Nussbaum emphasizes, [9]

[9] Ibid.

there is no hard external obstacle to or knock-down argument against a human being's deciding to live the life of a grazing animal or an asocial anthromorph, even though the cost of doing so would be a radical impoverishment of life. Nevertheless, as she suggests, the very raising of the question for thought or discussion reveals the committed participation of the person who raises it in practical reason, which is pragmatically inconsistent with denying there is reason to go in for it.

What kind of pragmatic inconsistency is there between participating in practical reason and denying its value? If someone were just to keep quiet, empty his mind, and graze, nothing would have been shown. By thus declining to participate in practical reason, he does not covertly participate in it or affirm its value; there is no self-defeatingness about such a heroic response. So, as I



suggested earlier, this pragmatic inconsistency seems to provide presuppositional rather than demonstrative support for the importance of practical reason. Is the presuppositional rather than demonstrative force of this pragmatic inconsistency a consequence of the particular context, namely, theories of human nature and scepticism about value? Or can we find a way to run a pragmatic argument in this context that provides the stronger, demonstrative kind of support? This would strengthen the claim that appeal to forms of life underwrites rather then undermines objectivity in this context.

I suggest we consider the idea put forward by various philosophers, that a distinctive property of persons is their capacity for higher-order attitudes. To have attitudes about one's own and others' attitudes is to be an interpreter, and in particular a self-interpreter; and self-interpretation that takes the form of working out one's own

higher-order desires, for example, in the face of conflicting values, can be regarded as a kind of self-determination.[10]

[10] I elaborate a view of this kind in Hurley (1989).

Consider the view, then, that a cluster of tightly related capacities—for higher-order attitudes, self-interpretation, and self-determination—is essential to persons; we are pragmatically committed to there being reason to exercise them, and hence to their value, through our regular exercise of them. Consider now the position of a sceptic about the value of these capacities and their exercise. Is it open to him to avoid the force of the pragmatic inconsistency of his position by opting out? I suggest not: the desire or intention not to have higher-order attitudes, even if successful (e.g. one deliberately undergoes brain surgery), is itself a higher-order attitude; a belief that one does not have higher-order beliefs (as opposed to a mere lack of higher-order beliefs) is self-defeating. The desire not to be a person, in these respects is itself the characteristic mark of a person, an exercise of the capacity of persons for self-determination, even if it results in lobotomy or suicide, or deliberate affirmation of or submersion of oneself in a culture that represses one's status as a person, say, because one is a member of some racial group, or religion, or because one is female. We find the kind of self-defeatingness here characteristic of the stronger, demonstrative kind of pragmatic argument. However, we do not seem to get all the way to full demonstrative support, even so: the sceptic's attempt to opt out of these forms of life does not demonstrate the truth of the premiss that he is sceptical about, namely, that there is reason to exercise the capacities in question, merely his practical commitment to it, which could still be unjustified. Perhaps in the evaluative realm this kind of middle ground between the presuppositional and demonstrative cases is the strongest kind of pragmatic support we can hope for; at any rate, I have not been able to think of an evaluative example of fully demonstrative support. Nevertheless, we can get something stronger than mere presuppositional support. Perhaps not all human beings have these capacities. But, if so, this fact can be accommodated in appropriate cases by saying that some human beings

end p.275

are not persons, rather than that they are not really human.[11]

[11] See Frankfurt (1971) on the use of the term 'wanton' here, to contrast with 'person'.

This slippage from talk of what is essential to us as human beings to talk of what is essential to us as persons has advantages; there is less of a strain in regarding some lacks as impugning someone's personhood than in regarding them as impugning his humanity. And this slippage allows more scope for that characteristic human activity of misguided efforts at self-determination, for example, emulation of the beasts, or the gods.

In conclusion: I have sketched the way Nussbaum's view about the relationship between relativism and the practices that ground the virtues, and especially her response to the third objection she considers, raise a more general issue, about the relationship between appeal to forms of life and objectivity. And I have pursued her suggestion of a certain pragmatic form of argument, from participation in forms of life to value, in order to consider further senses in which the appeal to forms of life may or may not threaten objectivity.

**Bibliography**

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Frankfurt, H. (1971). 'Freedom of the Will and the Concept of a Person', in G. Watson (ed.), *Free Will*. Oxford: Oxford University Press.

Harrison, R. (1982). 'Transcendental Arguments and Idealism', in G. Vesey (ed.), *Idealism Past and Present*, Royal Institute of Philosophy Lecture Series 13. Cambridge: Cambridge University Press.
Hurley S.L. (1989). *Natural Reasons: Personality and Polity*. New York: Oxford University Press.
Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
Nussbaum, M. (forthcoming). 'Aristotle on Human Nature and the Foundations of Ethics', in a volume on the philosophy of Bernard Williams, ed. J. Altham and R. Harrison. Cambridge: Cambridge University Press.
Putnam, H. (1979). *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
—— (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.
Strawson, P.F. (1959). *Individuals*. London: Methuen.
—— (1966). *The Bounds of Sense*. London: Methuen.
—— (1974). 'Freedom and Resentment', in *Freedom and Resentment and Other Essays*. London: Methuen.
Stroud, B. (1982). 'Transcendental Arguments', in R. Walker (ed.), *Kant on Pure Reason*. Oxford: Oxford University Press.
Williams, B. (1981). 'Wittgenstein and Idealism', in *Moral Luck*. Cambridge: Cambridge University Press.

end p.276

# Pluralism and the Standard of **Living**

show chapter abstract and keywords
hide chapter abstract and keywords

**Paul Seabright**

The proper province of a government's economic policy is the enhancement of the standard of living of its citizens. Such a remark, while true, is hardly informative. For in what does the standard of living consist? The intention of this paper is to sketch part of an answer to this question, involving some necessary (though not sufficient) conditions that characterize the concept. In particular, I want to examine how a notion of the standard of living, intended as a criterion of use to governments implementing policy, might be sensitive to a relatively pluralist theory of society. By a pluralist theory I understand an account of how society should be arranged that incorporates the possibility of multiple and non-trivially divergent views of the good life for individual human beings. The rider about non-trivial divergence is important: a theory of society is not pluralist merely because it assigns some citizens to be philosophers and others to be accountants, snake-charmers, or car-park attendants. A genuinely pluralist theory must count it a social good that there may exist multiple views of the individual good that are not subsumable under an encompassing theory of the individual good. This does not imply that a pluralist social theory must be neutral between all or even many theories of the individual good; for it to count as pluralist, it is enough for it to be compatible with the assertion of more than one theory of the individual good. I shall not argue for the correctness of such a point of view. In this paper I want to see instead whether it is possible for a government's economic policies to incorporate it if it is correct. Or does the very notion of an economic policy presuppose that a government has its own comprehensive theory of the individual good? Nevertheless, although I do not consider the correctness of pluralism as such, I shall argue that the possibility of a pluralist economic policy has important implications for the coherence of pluralism even in principle.

The argument proceeds in four stages. In the first I shall argue for the general propriety of allowing a moral or political theory to depend intrinsically on the practical limitations upon implementing its recommendations. By this I mean more than the fatuous point that a theory should not make impossible recommendations. I mean that our choice of a moral or political theory should not always be significantly influenced by our intuitive reactions to what it recommends in impossible or sufficiently unlikely circumstances. In the second stage I shall use this conclusion, along with a series of specific constraints on implementation,

end p.393

to help characterize the notion of the standard of living, giving some necessary conditions that any adequate standard must meet. In the third stage I shall draw analogies with other arguments in philosophy and economics that illuminate the connection between this notion and a pluralist theory of society. In the fourth stage I shall use the notion thus characterized to answer some specific questions about living standards and pluralism. These include the question of how the standard of living can incorporate values that are specific to certain cultures without thereby becoming relativist, and the question of whether economic policies need be intrinsically biased towards what is sometimes called commoditization, and to the market economy.

# 1 The Limitations on Moral Theories

It is a common practice among philosophers and social theorists to test the adequacy of a general theory by searching for possible counter-examples. Robustness against counter-example is definitive of a successful general theory, and on some views definitive of a theory, *tout court*. In ethics and social philosophy, where the data of our theories consist of our intuitive reactions to different situations, this search frequently takes the form of developing elaborately artificial situations designed to abstract away from the messy and confusing detail of every day. Thus we may situate moral dilemmas on Mars or behind a veil of ignorance or on tramlines stacked with Nobel Prizewinners, rather than in the Cabinet or in *EastEnders*[1]

[1] A soap opera on British television.

or in our own families. The process by which the intuitions in these abstract cases inform our more general judgements has been variously characterized, and John Rawls's notion of 'reflective equilibrium' represents one of the clearest and best known of such accounts.[2]

[2] Rawls, 1971, esp. 48-51.

While I have no intention of disputing the frequent usefulness and clarificatory value of such exercises,[3]

[3] Indeed, the exercise of the imagination is an essential component in our moral thinking.

I do want to argue that we can sometimes reasonably object to the result of such a thought experiment that the situation described is improbable or bizarre. The reason for this is not that I wish to reject the idea that a general theory must be general. 'Not *Fa*' is still and will always be conclusively incompatible with 'For all *x, Fx*'. The objection instead is to the idea that the judgements delivered by intuition in improbable cases must always amount to counter-examples.[4]

[4] It is of course true that our *initial* reactions to improbable situations may well change on further reflection. It is then trivially true that initial judgements do not always constitute counter-examples. The claim here is the stronger one that *settled* judgements of improbable situations do not always constitute counter-examples.

This line of argument makes best sense within a broadly naturalistic view of ethics. Our ethical intuitions are the product of a long process of both biological and cultural evolution. This evolution has been shaped in a major way by the

end p.394

need to cope with certain recurrent problems and difficulties faced by human beings in society. It has been shaped much less by very rare problems and difficulties that have affected very few people. And, at the risk of labouring the point, I should add that it has not been shaped at all by the need to cope with situations that have never happened or could never happen. It should not be surprising, therefore, if the intuitions we have concerning improbable or impossible situations turn out to be quixotic or mercurial. It is unwise to trust them very far.

At this point it is important to distinguish two conclusions that do not follow from this naturalistic line of reasoning, in order to protect from guilt by association the conclusion that does. First, from the fact that certain moral intuitions have arisen in us as a result of the need to cope with evolutionary contingencies it does not follow that these intuitions are morally defensible ones for us to have: aggression, vindictiveness, and jealousy all have comprehensible evolutionary roots. Instead, what the naturalistic argument illuminates is a question at the heart of the method of reflective equilibrium: do we have any reason to think that an equilibrium exists? Our intuitions can be considered as inputs to this reflective process. Certain requirements of consistency may have been imposed on those of our intuitions that have felt the evolutionary squeeze; others which have never done so may well be wildly inconsistent with what we believe in more familiar contexts. If we require a moral theory *both* to respect the requirements of reflective equilibrium *and* to be applicable to all logically possible situations, we may require the impossible.

A second unwarranted conclusion would be that this evolutionary reasoning implies an anti-objectivist or even just a non-cognitivist account of ethics. Many aspects of reality are too complex for many of reality's inhabitants to understand. The findings of modern physics are certainly beyond the comprehension of a dog, and we have no guarantee that the physical nature of reality will not turn out to be beyond our own comprehension too. To believe that we alone in the animal kingdom had reached the evolutionary plateau that would enable us to comprehend our own environment in its entirety would be a return to the image of ourselves as a chosen species, with no longer a God to be chosen by. So if likewise it turned out that we could not develop ethical theories to cover all possible contingencies without veering into contradictions, that would be a comment upon our moral capacities, not upon the existence or otherwise of objective moral truths.

The conclusion that does follow from this naturalistic argument is that, if we cannot be sure that there exists a moral theory that consistently satisfies our intuitions in all logically possible cases, we had better not throw out all candidates for a theory that happen to fail the test of our intuitions in certain improbable circumstances. We must find ways to choose among those theories that satisfy our intuitions in a sufficiently large number of cases. And here it seems to me inescapable to judge between cases not merely on grounds of number but also on grounds of importance. This is not the place to propose a detailed

end p.395

criterion, but such a criterion will certainly tend, other things being equal, to ascribe greater weight to our intuitions concerning cases that are more central to our everyday practices of judging, evaluating, and acting, than to those that are not. [5]

[5] This is not the same as a procedure that weights cases according to their probability of occurrence. Some cases might have a low probability of occurrence but still be sufficiently like actual cases to be a

valuable test of our intuitions. Others might be equally improbable, but so different in nature from actual cases as to be useless to us. Examples of the former occur when we ask questions like: 'If it were possible, by pressing a button, painlessly to kill an unknown and distant individual with no ties to family and friends, would it be wrong to do so?' An example of the latter might be the question: 'Would loyalty be so important a virtue if human beings reproduced asexually and did not age?'

This issue bears a close relation to an issue in the literature on social choice theory. That literature has been much concerned with the condition of unrestricted domain, which is a condition requiring any reasonable social welfare function or social decision function to apply to any logically possible ordering of individual preferences. To some it seems obvious that the question of what a defensible social decision function should be is independent of the particular decision problems it is applied to in any individual case: if a particular outcome is impossible, it is ruled out of the choice set as infeasible, and if a particular combination of preferences is impossible, it is ruled out as an input to the function. But at all events this infeasibility should not affect the choice of the original function. In other words, on this view you should first decide the moral theory and then apply it to some particular choices; there is no sense in allowing your choice of a moral theory to depend on what choices you will happen to face.

The problem with such a line of argument is that by imposing the condition of unrestricted domain along with other conditions (such as that the function respect transitivity), one effectively requires social preferences over feasible outcomes to be sensitive to what social preferences *would* be over outcomes that are improbable or physically impossible. This might not matter if we knew that social welfare functions compatible with all our intuitions existed. But the social choice literature has documented with depressing thoroughness the circumstances in which *no* social welfare function may exist and satisfy the restrictions that we might intuitively wish to place on it. It is clear that some of our intuitions somewhere may have to give. To insist upon retaining the condition of unrestricted domain is then to admit only those social welfare functions that fail to satisfy some of our other intuitions. The result is paradoxical: the price of requiring a theory to be compatible with all *possible* preferences over states consists in no longer requiring the theory to be compatible with even all *actual* preferences over theories.

The main reason, then, for not necessarily requiring a moral theory to satisfy our intuitions in improbable or impossible circumstances is that a theory compatible with all our intuitions may not exist. A subsidiary reason, familiar from considerations in the philosophy of language, is that it may be dubiously intelligible

end p.396

to ascribe to creatures capable of intentional action preferences and beliefs bizarrely different from our own. It is not always evident what may count as a genuine possibility, and certain situations may appear intelligible until we inquire closely just what circumstances would make us judge that such a possibility had indeed occurred.[6]

[6] For instance, it may seem logically possible that there should exist a creature all of whose beliefs were false, until one begins to reflect what circumstances could lead us to attribute entirely false beliefs to a language user. We would cease to call such a creature a language user long before that point had been reached.

This is a particularly powerful argument against the condition of unrestricted domain of preferences in the social choice literature. Here it underlines that in addition to the uncertain *reliability* of our intuitions in improbable or impossible circumstances, we must consider the uncertain *intelligibility* of those hypothetical circumstances themselves.

In the context of the present argument the upshot is this. A political theory, which is one kind of moral theory, deploys a variety of moral concepts in accordance with our intuitions about their application. One with which this paper will be much concerned is the concept of a contract, but there are of course many others, concepts like utility, the good life, self-respect, dignity, obligation, right, the citizen, and so on. I am arguing that it will often be legitimate to characterize a political theory in terms of its adequacy to these concepts as we actually have them, not as we might have had them. For example, it is not in itself an objection to retributive theories of punishment that under conceivable circumstances (those described by Nagel (1976) and Parfit (1976) in their discussions of brain bisection, for instance) the boundaries between persons might be uncertain and the question 'Who committed this crime?' might have no answer. Our theories of punishment have developed to deal with persons as they are. And as they are, persons are distinct entities characterized by physical continuity and a high degree of psychological continuity. If grotesque possibilities become actual we may, of course, need to develop our theories to cope with them. But unless there is some real chance of their doing so we would be unwise to allow our choice of a theory of punishment to be determined by their mere possibility.

There is nothing unusual or original in this form of argument. The emphasis in Rawls's recent work on the political, not metaphysical nature of justice as fairness [7]

[7] Rawls, 1985.

can best be understood as calling attention to the need to fashion justice to the needs of human communities that already have enough cohesion and common purpose to make justice possible at all. Using a theory of justice to speak to societies so fragmented that they cannot begin to co-operate is a futile exercise, one which therefore teaches us nothing about the kind of justice that more unified communities require. We can say more: if certain rules could never be implemented in any community that was recognizably like our own, then a theory of justice had better respect that constraint, and not seek to judge communities by their degree of resemblance to communities in which such rules are possible.

end p.397

## 2 Constraints and the Standard of Living

It was stated earlier that a government's economic policy should concern itself with the standard of living of its citizens. I want now to turn this around and claim that the way to understand the notion of the standard of living is as that aspect of individuals' well-being that falls within the proper sphere of society's concern. My perspective in making this claim will be broadly contractarian. [8]

[8] This does not imply that I think the contractarian perspective to be without problems. But the present argument is of strictly limited scope: if contractarianism is correct, and if pluralism is correct, then economic policy-making can be consistently pluralist.

Society is the institution of co-operation between individuals to their mutual benefit (it may not be only that but it is at least that). In particular, society makes possible the implementation of a framework—involving, for example, property rights and the organization of both production and consumption—without which individuals could not enjoy many of the activities that are central to their well-being. I am not here concerned with the difficult question as to whether an adequate theory of justice must ensure that all individuals are at least as well off in society as they would be on their own (the greatest problem for such a view being posed not by some of the very badly off, whose position is certainly unjust, but by able and law-abiding members of society who in the state of nature would be successful cocaine barons). [9]

[9] Roemer, 1985, is one important version of such a view.

It is enough for my purposes to observe that in taking part in the institutions of production, consumption, and exchange as they in fact exist, all individuals gain benefits from the co-operation that underlies these institutions. We can therefore reasonably ask what features of these institutions would have emerged from a process of agreement, or more formally of contract, between the individuals concerned, if society had indeed emerged from agreement instead of higgledy-piggledy from history.

Many such features will depend upon how precisely such a contract is specified (for instance, whether individuals are behind a veil of ignorance or aware of some features of their situation in life). For present purposes all that matters is that it is a contract. But—and here lies the connection with the first stage of argument of this paper—it is a contract that is recognizably like contracts made by real people in the real world. In the real world, for example, contracts are frequently made contingent on certain specified circumstances ('if the insured suffers a financial loss the insurer undertakes to indemnify him against such loss'). But the circumstances must always be, at least in principle, publicly verifiable. An insurance company may indemnify an individual against a loss of money, which is verifiable by recourse to the individual's bank, or against a loss of health, which is verifiable by recourse to his doctor. But it will not—it could not—indemnify him against a loss of utility, except in so far as a loss of utility might happen to be dependent on something that is publicly verifiable. So if we are to make sense of a contract at all, we must understand it as something

end p.398

public, and the contingencies on which it might depend as being public too. Members of society could and perhaps would contract to insure themselves against the hazards of birth and circumstance, but only against the publicly verifiable hazards. They would not contract to insure themselves against all inequalities in *utility*. A contract that did that would be no contract in any sense that we understand the term; it would be a kind of divine intervention.

Similarly, all contracts specify some kind of exchange, and no party can contract to offer something which it is not in his power to supply. This has two aspects. First, I cannot contract to give you something which I cannot ever deliver to anyone. So I cannot offer you my utility or my state of health or my relationships with my friends, though I can offer you my meditation techniques or my patent medicine or my well-thumbed copy of *How to Win Friends and Influence People*. Second, I cannot contract to give you something that must be given freely and without requiring a return if it is given at all. So I cannot make a contract promising you my respect or my love or my agreement with the conclusions of your latest book (I can promise to treat you with respect, but that is something different). And members of society contracting with each other to share the benefits of co-operation cannot make such contracts either. Perhaps this is unfortunate, for the kinds of good we cannot trade by contract include many of the most important elements in the flourishing of human beings. But the fact remains that these goods are outside the sphere of contracts as we know them or even conceivably could know them.

Indeed, from another point of view it is not unfortunate at all, for there would be something distinctly unattractive about even the attempt to think of these goods as within the contractual sphere. One can think of many kinds of motivation that are not only inappropriate for their intended effect but become in the process of conception ridiculous. Human beings need self-respect, and in order to attain self-respect they must have the respect of their fellow human beings. But it would be absurd if they sought to do so by contracting to have respect for each other. In like manner, there

can be something by turns comic and sinister about politicians' undertaking to create greater altruism or to improve the quality of family life, as distinct from ensuring that individuals have the means at their disposal to create these conditions themselves. Not only is this something they *cannot* deliver, but they ought not even to try.

In fact, the sphere of contracts consists more or less of rights over scarce [10]

[10] The rider about scarcity reflects the fact that for things that are in indefinite abundance there would be no gains from co-operation and no need to make contracts. Such things (oxygen in the atmosphere, for example) will not be components of the standard of living, though they may become so if the possibility of their scarcity arises (through air pollution, for example).

physical commodities (including rights of access to such things as library books and land), and of rights to the performance of services, these being publicly verifiable, intentional activities performed by or caused by other human beings. Social contracts must be contracts for the exchange of these commodities or services, contracts that are contingent on publicly verifiable circumstances. And

end p.399

the standard of living of citizens is represented by the command that they have over these commodities or services, given these circumstances. More precisely, one may say that the standard of living of individuals consists of those components of their well-being the enhancement of which would be the appropriate subject of a social contract between individuals wishing to share the benefits of social co-operation.

It will be evident by now how this account makes room for pluralism. There are a great many aspects of the good life for individuals the enhancement and distribution of which are simply not society's business. This is because, on the contractarian account I am outlining, nothing is society's business unless it could be the subject of an appropriate hypothetical social contract. Thus it is not the business of society at large whether people have happy marriages or believe in God, because these are not the kinds of thing people could contract to do. It may of course be the business of their friends or of their priests, individuals with whom their relationship is not contractual. And it will certainly be society's business whether people have the resources to make happy marriages possible (such as the ability for a couple to work and live in reasonable proximity to each other), and the ability to call on the state's resources to defend them from attacks on their worship of their own God in their own way. But their standard of living represents their command over these resources, and not the outcome that results from it. A pluralist social theory, which allows for the possibility of fundamentally different conceptions of the ends to which these resources may be put, will be superior to a theory which requires society to evaluate social outcomes in every detail, when many aspects of these outcomes are not ones society should be in the business of evaluating at all. The reason they are not is that the kind of world in which these aspects could be the subject of contract would be a world so unlike our own (peopled by individuals completely transparent to each other, and able to trade mental states with each other at will) as to be no fit point of comparison with the world we know.

In describing the standard of living as involving command over resources rather than the outcomes that result, I should make one point clear. The value of command over given resources may well be dependent on circumstances, provided these circumstances are publicly verifiable. Thus it may readily be granted that a disabled person has a lower living standard than a fully healthy person commanding the same resources, because the disability is a publicly verifiable circumstance. Amartya Sen has argued that this shows that the concept of the standard of living must be concerned with evaluating outcomes rather than control over resources, [11]

[11] He writes (Sen, 1985: 16) that the standard of living 'must be directly a matter of the life one leads rather than of the resources and means one has to lead a life'.

but I hope it is by now clear that, on the present account, it need show nothing of the kind. If standards of living were concerned with evaluating outcomes in their entirety, we should be obliged to state that whenever two individuals derive different levels of happiness from a given

end p.400

endowment of commodities, their standards of living are different. The present theory differs from Sen's in allowing only some reasons for the divergence of utility levels to count in standard of living comparisons—namely, those that are sufficiently publicly observable to be the basis of a contract. Other claims, such as that one individual may be of a cheerful disposition while another is somewhat morose, may be true without being relevant to the standard of living.

So far a fairly stark contrast has been drawn between those components of the good life for individuals that are, and those that are not, society's business. But it is evident on reflection that there exists a spectrum of possibilities rather than a sharp division. First of all, there is no sharp line between what is and what is not a publicly verifiable event, certain kinds of mild mental disability being a good example of a borderline case. There is nothing wrong with the idea that our conception of the standard of living might be sensitive to developments in diagnostic medicine, for example. But the emphasis on verifiability will tend to impose a burden of proof on borderline cases proposed for inclusion. Two individuals commanding the same resources will be held to have the same standard of living unless a convincing case can be made for their differing in some verifiable respect that might itself be the subject of a contingency clause in a social contract. So two things will need to be established: first, that the difference is verifiable and, second, that it would itself be relevant to the contract. Consider the case of differences in sex: these are (normally) readily verifiable. But are they such as to justify the claim that men and women with the same command over resources differ in standard of living, in a way that society might perhaps wish to correct? This will depend on the degree to which inequalities between the sexes consist of resource inequalities as opposed to inequalities in other respects, and the answer to this is likely to differ between societies and over time. If existing inequalities consisted primarily of unequal access to resources, there would be a less compelling case for adjusting living standard measures to reflect gender differences for given levels of command over resources.

A second kind of borderline problem concerns what kinds of service activity could be the subject of a contract. The problem is most perplexing within associations that are either not contractual or only partly contractual, like families. But here we may distinguish between problems of inclusion and problems of allocation. On the present account the time spent on housework, for example, even though it is frequently performed unpaid and without even a verbal contract, would count as a (negative) component of the standard of living, because it is the kind of activity that households need to perform and some reduction in which they would usually be willing to contract to achieve. Some problem of inclusion still arises because of the difficulty of distinguishing some such activities from those considered hobbies, where the activity is itself valuable. If, for example, I am a keen gardener (the example is hypothetical) I may not be willing to contract for anyone else to do my gardening for me, because it would then in some sense not be my garden any more. But in general there is no doubt that one significant component of poverty in many countries is the

end p.401

time required for unpaid household tasks, a component that the familiar measures of the standard of living routinely ignore.

Separate from the problem of inclusion is a problem of allocation that may arise *between* members of the household. Suppose, for example, that household tasks are shared very unequally. We would be inclined to say that the standard of living of individuals within the household is to that extent unequal. But suppose it were objected that relations within the household are not an appropriate subject for a contract, that there could not be a social contract aiming to equalize time spent on household tasks? This is a difficult area, but if the premiss of the objection were accepted we should have to withdraw the judgement of inequality within the household. Of course, that premiss may be disputed: one of the features of modern feminism has been precisely to draw attention to the tacitly contractual way in which many aspects of relations between the sexes may be understood. Certainly, the theory being advanced here is in no way undermined by the reflection that the concepts of a contract and of what is fit material for contractual exchange are themselves variable and evolving over time.

As the example of household work suggests, the limitation of the notion of the standard of living to the sphere of commodities and services, broadly conceived, does not restrict it to *marketed* commodities and services. In fact, it includes many services of a very general kind. So, for example, the extent to which members of a society enjoy freedom of speech and of association is certainly a part of their standard of living. This is because these freedoms are safeguarded by the resources of society in that people are protected from assaults upon their exercise of free speech or free association. Such freedoms (and the resources to protect them) they could and would contract with each other to guarantee (though not, of course, to an unlimited extent). But the extent to which they *use* these freedoms to say and do worthwhile things is not part of their standard of living, though it may be an important part of what gives value of their lives. Two societies that devote equal effort and resources to television do not differ in their standard of living merely because one society is innovative and creative while the other broadcasts rubbish. One cannot contract to have good taste.

Nevertheless, non-marketed resources do raise problems of a different kind, namely those of valuation. The valuation of marketed resources in measures of the standard of living is commonly made at their market price. How reasonable is this, and how should non-marketed resources be treated? The argument so far has confined itself to considering what aspects of individuals' lives should be counted as part of their standard of living. For all practical purposes, though, what is required is not a complete description of the standard of living of each individual in society, but some aggregation or summary of the information this contains. Two kinds of aggregation are needed: first, aggregation over the many different components of an individual's standard of living, in order to compare individuals with each other. Second, aggregation over the many different individuals in society for the purpose of comparing one society with another or with

end p.402

itself at different times or in different circumstances. The basic conceptual problems these forms of aggregation raise are very similar. All aggregation involves throwing away information. The best way to aggregate cannot be determined independently of the use to which the information is to be put. Different uses may require different methods of aggregation, and different degrees of aggregation: for example, for some purposes we may need to reduce the measure of the standard of living to a single

scalar number; for others we may be happy to reduce it to a vector so that we can make comparisons in several dimensions. There is no reason whatever to believe that there is a single optimal aggregate measure of the standard of living either of an individual or of a society.

This point has been obscured by the implicit utilitarianism of much of the economic literature on the standard of living, and the related literature on such questions as the measurement of inequality. It follows from the standard axioms of individual rational choice that individuals choose their consumption so as to equate the ratio of the marginal utility of each commodity to the marginal utility of money with the price of that commodity in money terms. A measure of the standard of living can be constructed adding quantities of commodities traded and using market prices as weights. This measure represents the individual's consumption, and the sum of these measures across all individuals is the national consumption. If—and it is a big if—utility in this sense represents the appropriate maximand for society, and if all consumers have the same marginal utility of money, then a government that acts to maximize the national consumption will thereby maximize the utility of marginal changes in the availability of goods and services to the population.

There are numerous and well-rehearsed problems with this line of reasoning, of which here are six:

1. One may deny that individual consumers maximise anything.
2. Even if they do, one may deny that what they maximize is utility in any interesting sense.
3. Even if it is, one may deny that society should be concerned only with utilities.
4. Even if it should, one may deny that society should be concerned only with the sum of utilities and not with their distribution.
5. Even if it should do that, one may deny that individuals have the same marginal utility of money.
6. Even if they do, one may deny that maximizing utility at the margin is equivalent to global maximization of utility.

The important point in this context is that only if it is true that society's goals can be appropriately represented by the maximization of some one quantity inherent in all commodities, will it follow that there is a single optimal way to aggregate the standards of living either of individuals or of societies.

This paper is concerned with the proper province of government policy, not with what the criteria of that policy should be. So the question of how to aggregate living standards, which requires answers to the latter problem, cannot be



addressed fully here. All the same, for many ordinary purposes market prices will provide a useful first approximation to a weighing system for those components of living standards for which markets exist. The reason for this is that if the standard of living comprises those commodities and services that can be exchanged by contract, one cultural indicator of their value will be the amounts of other commodities that citizens would require under the terms of some social contract in order to induce them to exchange. The justification for this need not be utilitarian (and typically will not be, since these amounts represent relative *marginal*, not average utilities). The justification might be directly contractarian: these amounts represent the relative weights that have been agreed as a result of the contractual process. And relative value by this criterion will be related to the amounts of other commodities that

citizens in fact require in exchange under the terms of actual contracts. The two will not always or even usually be the same, partly for reasons of market imperfection that are familiar in the economics literature (such as the presence of externalities), and partly because the terms under which individuals make actual contracts are not the terms that would be specified by hypothetical social contracts. Actual contracts reflect substantial inequalities in power and wealth that social contracts might well seek to redress. But here as before the nature of real contracts between real individuals represents an important starting point in the analysis of what social contracts might be.

Non-marketed resources naturally pose greater informational problems than do marketed resources, but the conceptual issues are very similar. Likewise, the valuation of rights such as the right to free speech depends upon what would be the terms of a social contract in which that right was guaranteed: how much of society's resources would be devoted to policing it in the optimal social contract? The more thoroughgoing the right the more expensive its defence, and the higher the valuation placed upon having that right. For those who blanch at the thought of 'valuing' rights in this fashion, I should repeat that the values thus derived represent purely a measure of the weight these rights have in that sphere of life that it is society's business to organize. Their overall importance in and contribution to the flourishing of individual citizens is something that society has no business to put a value upon.

## 3 Contractarianism and Pluralism

That there is a natural affinity between broadly contractarian views of the justification of social theories, and a pluralistic conception of the social good, has a well-established pedigree in the history of social thought. The connection has not always been made explicitly, but one place where it is so made is John Locke's *Letter Concerning Toleration*. On the basis of the claim that 'the commonwealth [is] a society of men constituted only for the procuring, preserving, and

end p.404

advancing their own civil interests', Locke argues that it can be no business of the state what religion its citizens believe in:

it appears not that God has ever given any such authority to one man over another, as to compel any one to his religion. Nor can any power be vested in the magistrate by the consent of the people, because no man can so far abandon the care of his own salvation as blindly to leave to the choice of any other, whether prince or subject, to prescribe to him what faith or worship he shall embrace. For no man could, if he would, conform his faith to the dictates of another. All the life and power of true religion consists in the inward and full persuasion of the mind; and faith is not faith without believing.

The irrelevance of contractual agreement to the procurement of spiritual well-being contrasts with the more material sphere:

For those things that are necessary to the comfortable support of our lives are not the spontaneous products of nature, nor do offer themselves fit and prepared for our use. This part therefore draws on another care, and necessarily gives another employment. But the depravity of mankind being such that they had rather injuriously prey upon the fruits of other men's labours than take pains to provide for themselves, the necessity of preserving men in the possession of what honest industry has already acquired, and also of preserving their liberty and strength, whereby they may acquire what they further want, obliges men to enter into society with one another, that by mutual assistance and joint force they may secure unto each other their properties, in the things that contribute to the comfort and happiness of this life,

leaving in the meanwhile to every man the care of his own eternal happiness, the attainment whereof can neither be facilitated by another man's industry, nor can the loss of it turn to another man's prejudice, nor the hope of it be forced from him by any external violence. [12]

[12] Locke, 1689: 128-9, 154. I am grateful to Tom Baldwin for directing me to this source.

Though Locke is not consistent in his demarcation between the contractual and the non-contractual spheres (at some points treating it as coextensive with the distinction between this world and the next, and elsewhere exempting atheists from the right to religious toleration), the idea that there exists such a distinction is central to Locke's argument, as it is to the argument advanced here.

More recently than Locke, Ronald Coase's approach to the analysis of externalities in production and consumption [13]

[13] Coase, 1960. There is a useful discussion by Farrell (1987).

illustrates another natural connection between contractualism and pluralism. Prior to Coase's work, it had been generally accepted that the appropriate response to an externality such as industrial pollution was to tax the polluting firm so as to bring private and social costs into equality at the margin. Coase pointed out that it was not evident that the polluting firm was the right agent to tax, because of the symmetrical character of externalities. For just as another firm downstream of the first might be said to suffer as a result of the actions of the polluting firm, so the polluting firm itself, if taxed, could be said to suffer as a result of the presence of another, a presence that made the pollution more costly in its effects (for the sake of the example, it is assumed that the damage due to the pollution is confined to



its effect on the downstream firm). The right solution, said Coase, was for the two agents to get together and bargain their way to an optimal outcome. It might be that this outcome would involve, not a reduction in the output of the polluting firm, but a (less costly) move by the other firm to another site where it would be unaffected by the first firm's activities.

One way to characterize Coase's argument is this: using taxes and subsidies to equalize marginal private and social costs will bring about local efficiency, in the sense that no reallocation of resources at the margin will make at least one agent better off without making any other worse off. But locally efficient points may not be globally efficient, [14]

[14] Indeed, there are particular reasons to fear this in the case of externalities. Dasgupta and Heal (1979: ch. 5) and Starrett (1972) discuss some of the reasons why.

because there may exist large resource reallocations (like moving whole factories) that dominate them. In order to ensure that globally efficient points are attained it is necessary for agents to bargain. The analogy with the present argument is that to make non-local comparisons of living standards (comparisons, that is, between individuals or societies that are very different in their values and ways of life), it may be necessary to consider explicitly what would be the terms of a bargain or contract between the different parties. For if we adopt the measures (such as measures of real income) that would be appropriate to local comparisons (within societies, perhaps), we may make incompatible judgements about their standards of living. [15]

[15] This point is considered explicitly in Christopher Bliss's paper in this volume.

Each society may have inferior living standards when viewed by the local criteria of the other. To compare them we must consider those terms of a bargain on which the diverse individuals or societies could agree.

## 4 Living Standards and Pluralism

Economic development has always provoked anxieties about conflicts of value, from fears about the provocation of decadence by riches to more subtle worries about the effect of development on the diversity of cultural traditions. In these matters, economics is often believed, with some justification, to represent philistinism incarnate. The account of living standards outlines here suggests that this need not always be so. First, consider the question whether measures of living standards can be sensitive to the different values of different cultures. In a simple sense they can certainly be so, because living standards have nothing to say about many important areas of human life, which individuals and associations are therefore free to order according to whatever values they hold dear. A more interesting question is whether those elements of individuals' lives that count as part of their standard of living, and the valuation of those elements, might themselves be sensitive to the different values of different cultures.

There is every reason why they should be, and the extent of this sensitivity will depend on the purpose for which the measures of living standards are

end p.406

developed. A measure of living standards developed for comparing countries may differ from a measure used in determining the policy of one particular country. This will be so for two kinds of reason. First, different conceptions of the good life will place different emphasis on the commodities and services necessary for their pursuit. [16]

[16] The commodities and services discussed here therefore represent a much broader spectrum than Rawls's primary goods, which are defined as 'things that every rational man is presumed to want' (Rawls, 1971: 62).

Second, different cultures vary to an extent in their views as to what is and what is not the proper subject of a contract. Anyone who has watched an American broadcast of televangelism will be aware that religion has a much more explicitly contractual character in some communities than in others. The point about a social contract is that it must rest on the agreement of all parties. Social contracts within reasonably homogeneous societies will therefore be able to take much more for granted than a social contract involving the citizens of the whole world.

There is no answer to the question 'Which social contract is the right one to make?' Considerations of justice affect relations within villages, within regions, within countries, within the world. To each of these questions a different social contract may be appropriate, and from each of these contractual exercises a different conception of the standard of living may emerge. The degree of divergence between these will depend in a fairly evident way on the degree to which parties to a social contract are required to abstract from the particularities of their own position. If they did not abstract at all the exercise would be futile, but there are grave doubts about the coherence of their abstracting from every value that gives individual character to their judgements—a point that has been familiar since Hegel's criticism of Kant. This is too large a question to explore here, but it may be observed that unless it is possible to delimit a set of values that must be held by all rational beings *qua* rational beings *and* that are sufficiently rich to characterize a social contract, then the outcome of a social contract will be sensitive to the specification of the parties involved. So, therefore, will a contractarian conception of the standard of living. This no more condemns such a conception to relativism than a road map is made relativist because the way to Tipperary depends on where you start from.

Finally, to what extent must governments' attempts to enhance the standard of living bias economic development towards the commoditization of society?

Commoditization is a difficult concept to understand and is used in a variety of senses. [17]

[17] Hart, 1985.

Most straightforwardly, it refers to the tendency of trade (either trade in general or monetized trade in particular) to penetrate social institutions. Existing measures of national income undoubtedly incorporate a bias towards the measurement of those goods and services that are in fact traded. In the theory advanced here such a bias is quite unwarranted, and involves essentially a confusion between the terms of a social contract and the terms of actual contracts. Nevertheless, it may still be objected that the emphasis in this theory



upon those components of the good life that *could be* the subject of contracts incorporates the same bias in a more subtle disguise.

There are two things to say about this objection. The first is that it sometimes incorporates the suggestion that a bias of this kind is in some way responsible for the widely observed tendency of human societies to mediate more and more of their consumption through the market as they become more prosperous. I do not think this is a suggestion to be taken seriously, at least not in this form. For one thing, it is not even true at all levels of development: for example, one may cite the growth of large corporations and their tendency to determine many aspects of their employees' lives either through fringe benefits or as an intrinsic part of their working conditions. Corporations can be viewed as institutions that supersede the market mechanism, as Ronald Coase long ago pointed out. [18]

[18] Coase, 1937. Recent work by Hart and Moore (1988) has developed this in the context of a formal model.

But even barring this exception, the spread of the market is due partly to certain general facts about human beings, about their tastes and about the technologies of production, that make the benefits from specialization in production tend dramatically to outweigh any benefits of specialization in consumption. This case must not be overstated, since it remains one of the most devastating elements of Marx's critique of capitalism that it has tended so to exaggerate specialization in production as to cramp and wither the human potential of all those who perform society's most routine and repetitive tasks. Serious as this criticism is, the causes of such a phenomenon are independent of whether the province of society's concern be restricted to commodities and services in the way I have argued here.

The second thing to say about this objection is that we should indeed be legitimately concerned about the development of our lives in all sorts of ways outside the broad sphere of commodities and services discussed here. If economic development were to lead to all social relations' becoming infected with the mentality of the market, that would be a matter for the gravest concern. We should indeed worry lest prosperity lead to what one might with some licence call the death of the soul (though by any standards the effect of poverty on the soul is of far more pressing concern in the world today). What is denied here is that governments may be legitimately concerned with such matters. It is hard to conceive that a theory which brought these concerns within the province of governments could retain any pretensions to pluralism at all.



**Bibliography**

Coase, Ronald (1937). 'The Nature of the Firm', *Economica*, 4, 386-405.

—— (1960). 'On Problem of Social Cost', *Journal of Law and Economics*, 3, 1-44.

Dasgupta, P., and Heal, G. (1979). *Economic Theory and Exhaustible Resources*, Cambridge: Cambridge University Press.

Farrell, Joseph (1987). 'Information and the Coase Theorem', *Journal of Economic Perspectives*, 1, 113-29.
Hart, Keith (1985). 'Commoditisation and the Standard of Living', in Sen, 1985.
Hart, Oliver, and Moore, John (1988). 'Property Rights and the Nature of the Firm', STICERD Discussion Paper, London: London School of Economics.
Locke, John (1689). *A Letter Concerning Toleration*. Oxford: Blackwell. 1956.
Nagel, Thomas (1976). 'Brain Bisection and the Unity of Consciousness', in Jonathan Glover (ed.), *The Philosophy of Mind*. Oxford: Oxford University Press.
Parfit, Derek (1976). 'Personal Identify', in Jonathan Glover (ed.), *The Philosophy of Mind*. Oxford: Oxford University Press.
Rawls, John (1971). *A Theory of Justice*. Oxford: Oxford University Press.
—— (1985). 'Justice as Fairness: Political not Metaphysical', *Philosophy and Public Affairs*, 14, 223-51.
Roemer, John (1985). 'A General Theory of Exploitation and Class', in Roemer (ed.), *Analytical Marxism*. Cambridge: Cambridge University Press.
Sen, A. K. (1985). *The Standard of Living* (The Tanner Lectures). Cambridge: Cambridge University Press.
Starrett, David (1972). 'Fundamental Nonconvexities in the Theory of Externalities', *Journal of Economic Theory*, 4, 180-99.

end p.409

# Paul Seabright: Pluralism and the Standard of Living

show chapter abstract and keywords
hide chapter abstract and keywords

**Derek Parfit**

## [I]

In his rich and stimulating paper, Paul Seabright claims that governments should take a narrow view of human well-being. The 'aspect of . . . well-being that falls within the proper sphere of society's concern'—or, for short, the standard of living—consists, he argues, of 'rights over scarce physical commodities' and 'rights to the performance of services'. Seabright admits that, on this narrow view, the standard of living fails to include 'many of the most important elements in the flourishing of human beings'. But these elements, he believes, 'are simply not society's business'. It would be quite illegitimate for a government, when deciding between public policies, to take such things into account.

Seabright's premisses are contractualist. 'Nothing is society's business', he writes, 'unless it could be the subject of an appropriate hypothetical social contract'. Since he does not defend this assumption, I shall not question it here. But I shall question Seabright's view about what *could* be covered by such a hypothetical agreement.

At times he suggests that contracts can cover only what we can either exchange, or do. Thus he writes, 'it is not the business of society at large whether people have happy marriages' since this is not the kind of thing that people could 'contract to do'. Nor can the standard of living include such things as the quality of the arts or television, since 'one cannot contract to have good taste'.

On this assumption, it would not be society's business whether people have good health. But Seabright agrees that health is part of the standard of living. He therefore turns, like Dworkin, to the notion of insurance. Though being healthy is not something that we could contract to do, we can insure against bad health.

We can insure against other things too. By appealing to the notion of insurance, a contractualist might reach a much broader view about the proper role of government.

If he imagines a sufficiently ideal insurance scheme, there would indeed be few limits.

Seabright rejects this line of thought. He insists that, as contractualists, we should not appeal to an Ideal Insurer. We should imagine only *feasible* insurance schemes; and these would be far more limited in scope. The chief restriction is that, to be feasible, an insurance scheme can cover only publicly verifiable events. This is why, though our health affects our standard of living, the happiness of our marriages does not. We could not reasonably expect to insure against marital disharmony or distress. Nor could we expect to insure against seeing rubbish on television.

end p.410

Must a contractualist accept this restriction? There are indeed good reasons, to do with verifiability and moral hazard, why we cannot insure against unhappiness. And, to be enforcible, contract law cannot turn on aesthetic judgements. But these points do not apply to many areas of public life. Suppose there are reasons to believe that certain kinds of housing provision, or certain details of family law, increase the incidence of depression in tower block dwellers, or mothers of young children. Why may not planners take such things into account? That one cannot insure against depression may be claimed to be irrelevant. Or consider the question of which buildings, or parts of the countryside, the government should protect. Judgements about beauty cannot be part of contract law. Why should this exclude them from planning decisions? Why may not the standards of verifiability depend on the context, and on the purposes in question?

One relevant difference is precisely that between public planning and private insurance. Thus, in the running of a health service, doctors may try to assess the effects of different treatments on the quality of people's lives. In making this assessment, they may try to gather evidence about such intangibles as comfort, dignity, or *joie de vivre*. Such evidence would fall far below the standards of verifiability that would be required by a private health insurance scheme. But this seems irrelevant.

The point is partly this. In the case of private insurance, what the claimant gains is a loss to the insurance company. Hence the scope for moral hazard; and hence the reasonable requirement that, for a claim to be legally enforcible, the relevant contingency be both well defined and easy to observe. But in areas like health administration, town planning, or education, the question is not whether some insurance company should be forced to bear a loss. The question is which of various policies is, on balance, best. With such a question, there is no need for the same degree of verifiability. In a choice between different policies, there is no special burden of proof which needs to be met. Why should we not here give some weight to *any* evidence which seems to bear on individual well-being? In Sen's words, is it not better to be vaguely right than precisely wrong?

Seabright would reply that this is to misunderstand his argument. He requires verifiability, not because he wants precision, but because of a more general view about the role in moral theory of imaginary cases. He claims that, to be relevant, an imaginary social contract must be '*recognizably like contracts made by real people in the real world*'. [1]

[1] In the printed version of his paper, Seabright qualifies this claim. There are, he admits, various differences between 'actual contracts' and 'hypothetical social contracts'. But even so 'the nature of real contracts . . . represents an important starting point in the analysis of what social contracts might be'. In this qualified form, Seabright's claim seems too weak to support his conclusions.

For his defence of this crucial claim, Seabright refers us to the first part on his paper. This attempts to show that 'our choice of a moral or political theory should not . . . be significantly influenced by our intuitive reactions to what

end p.411

it recommends in impossible or sufficiently unlikely circumstances'. [2]

[2] In the printed version of his paper, Seabright inserts the word 'always'. But this makes his claim too weak for his purposes.

One of Seabright's points is this. We should not reject a theory simply because it is counter-intuitive in some cases; for all we know, this may be true of *every* theory. This point seems to me correct. But it does not show that we should ignore our intuitions about imaginary cases. It shows only that we should not reject one theory unless we have another theory which seems better.

Seabright makes two other claims. He points out that, if an imagined case is sufficiently bizarre, it may be doubtfully intelligible. Even if we can understand the case, we may fail to realize what it would involve. These are indeed good grounds for doubting our intuitive reactions. But this point applies only to some imaginary cases. It does not cover cases which depart from reality in understandable ways. As Seabright notes, by considering such cases, we can 'abstract away from the messy and confusing detail' of ordinary life. Compare the use, in science, of artificial tests, or impossible thought experiments. If we can clearly imagine what would be involved in such cases, we need some different ground for doubting our intuitive reactions.

Seabright's third point does apply to all imaginary cases. He suggests that, since 'our ethical intuitions are the product of a long process of both biological and cultural evolution', we should expect them to be trustworthy only in the kinds of case in which they arose. These must all be cases which could actually occur.

This claim seems to me too strong. If a case is impossible only because we have imagined away various complicating features, there seems no reason to distrust our reactions to those features that remain. Moreover, as Seabright remarks, we should expect these kinds of evolution to distort our intuitions in various ways. Thus selective pressure favours partiality, tribalism, and aggression. Similarly, when we think about actual cases, we may be influenced by an awareness that some moral claim would threaten our own privileged position. Thinking about non-actual cases may help us to rise above these distorting influences. [3]

[3] In his discussion of these claims, Seabright's only example concerns punishment and personal identity. He writes, 'it is not an objection to retributive theories of punishment that under conceivable circumstances (those described by Nagel and Parfit . . . ) the boundaries between persons might be uncertain and the question "who committed this crime?" might have no answer.' 'Our theories of punishment have developed to deal with persons as they are', not as they might be in such bizarre cases. This misunderstands my aim in appealing to such cases. This was to show that we have certain false beliefs about the nature of personal identity, beliefs which apply also to actual cases, or to persons as they are. If retributive theories rest on such false beliefs, this *is* an objection.

Let us now return to Seabright's main argument. How do his claims about imaginary cases support that argument?

They provide, I believe, little support. If we doubt Seabright's argument, this is not because, in some impossible imagined case, we find his moral view counter-intuitive. Since this is not the issue, it is irrelevant whether, and when,

end p.412

we should distrust our intuitions. The question is whether, in a contractualist theory, our imagined social contract *must* be as verifiable as an actual contract. There may be ways to defend this claim; but the issues raised are different.

Consider, for example, Rawls's imagined 'original position'. This is intended as an artificial model to help us to work out the implications of certain assumptions about

moral reasoning. If we reject this model, our objection cannot be that, when we consider such a case, we should not trust our moral intuitions. Rawls's argument does not appeal to these intuitions.

There is a further problem. Seabright allows the contractualist model to be, in various ways, unrealistic. Thus he does not exclude a Rawlsian veil of ignorance. The parties to the social contract can be assumed to know nothing about themselves. They are quite unaware of their own aims, abilities, ideals, attitudes to risk, and every other individual feature. An imagined contract between such people is not, in *this* respect, 'recognizably like contracts made by real people in the real world'. Why must it be, in other respects, just like them? Seabright objects that, if we allow the contract to cover unverifiable features—such as the happiness of our marriages—we are imagining individuals who are 'completely transparent to each other'. But he allows us to imagine individuals who are completely opaque to themselves.

Seabright might reply that, even if the contract can be *made* in an unreal world, it must be imagined to *apply* in the real world. This reply is not, I think, sufficient. We should indeed, within a contractualist approach, prefer principles that would be easy to apply. It is in part for this reason that Rawls states his Second Principle in terms, not of well-being, but of primary goods.

Much of Seabright's argument could be recast along these lines. But such an argument could not, I think, yield his conclusion. It could not show that, in every choice between social or economic policies, governments should consider only those facts that could have entered into private contracts or insurance schemes. Such an argument would have to admit that, in different contexts, different degrees of verifiability would be appropriate. Claims about injustice, like claims about rights, need to be simple, and, as far as possible, verifiable. But the grounds for thinking this apply with much less force to a vast range of policy decisions. Besides the examples I have mentioned—health administration, family law, town planning, and the protection of the countryside—there are countless others. If we insist that, in all such cases, we would admit as relevant only verifiable facts—if we echo Mr Gradgrind—we shall not get good decisions. Practical considerations count *against* this narrow view. [4]

[4] He gives another argument in passing. Some elements of a person's well-being are not society's business because their value essentially depends on their being the activities or achievements of this person. This is why it would be both comic and sinister for a government to undertake to 'create greater altruism or improve the quality of family life. . . . Not only is this something they *cannot* deliver, but they ought not even to try.'

These claims are plausible, but I believe they exclude little. Though governments cannot *directly* improve the quality of our family life, they can help us to achieve this ourselves, and thus *indirectly* promote this element in our well-being. They can try to make marriages happier by altering family law, promoting creches and flexible work arrangements, and subsidizing marriage guidance counselling. All of these fall within the 'services' which Seabright's formula allows, since one can contract to give or to receive such services.

To exclude such services, Seabright would have to claim that these elements in well-being would lose their value if they were indirectly assisted in this way. On this view, the quality of family life would not really count as higher if its improvement came from state-provided marriage guidance: nor would there be value in greater altruism if this came from state-provided moral education. In the case of virtue, Kant made such a claim; but I doubt that Seabright would agree.

end p.413

## [II]

I turn to a more particular question raised by Seabright's approach. On his account, the standard of living consists of our command over various goods and services. How should we assess their value?

Seabright at times suggests that the value of goods and services depends on the cost of their provision. Thus he writes: 'the valuation of rights such as the right to free speech depends upon what would be the terms of a social contract in which that right was guaranteed: how much of society's resources would have to be devoted to policing it? The more thoroughgoing the right the more expensive its defence, and the higher the valuation placed upon having that right.'
This seems the wrong test. The value of a right cannot be measured by asking how many resources would have to be devoted to policing it. The right to free speech would then have little value, since the government can 'police' this right simply by ceasing to prosecute people for what they say.
Seabright applies this test to other kinds of goods. Thus he writes: 'Two societies that devote equal effort and resources to television do not differ in their standard of living merely because one society is innovative and creative while the other broadcasts rubbish.' On this view, it is the cost of services, not their effects of quality, which is the measure of their value. But cost seems never to be the measure of value. If it were, we could not *waste* resources.
Later, however, Seabright writes: 'In describing the standard of living as involving command over resources rather than the outcomes that result, I should make one point clear. The value of command over given resources may well be dependent on circumstances . . . Thus it may readily be granted that a disabled person has a lower living standard than a fully healthy person commanding the same resources.' Seabright here concedes that we should assess someone's standard of living, not in terms of the cost of the goods and services which this person commands, but in terms of their value to this person. If you and I have the same resources, but I am disabled, I have a lower standard of living.
This seems to contradict Seabright's earlier claim. If two communities spend equal amounts on television, but one broadcasts rubbish, the value of having television may be, for those in the second community, less. And, if two communities spend as much on education, but one has much worse educated children, the value of their education may be less. On this, which seems the better view, governments should assess their policies in much broader terms.

end p.414

After granting that disablement lowers one's standard of living, Seabright continues, 'Amartya Sen has argued that this shows that the concept of the standard of living must be concerned with evaluating outcomes rather than control over resources, but I hope it is now clear that, on the present account, it need show nothing of the kind.' This is not clear to me. If we believe that a disabled person is worse off than a healthy person with the same resources, Sen seems right to claim that, in our assessment of the standard of living, we look not only at people's control over resources, but also at one aspect of the outcome, namely, what people can *do* with these resources. If people are disabled, their resources enable them to do less. [5]

[5] Seabright might reply that, on this approach, we are still *evaluating their control over resources*. But this would make his point trivial. Any view, even the purest hedonism, could be stated in these terms.

Seabright continues: 'The present theory differs from Sen's in allowing only some reasons for the divergence of utility levels to count in standard of living comparisons—namely, those that are sufficiently publicly observable to be the basis of a contract.' But this seems inaccurate on three counts. Sen does not allow *all* such reasons to count in the standard of living. What he does count—capabilities and functionings—he does not assess in terms of their effects on utility levels. And both

capabilities and functionings *are* publicly verifiable. As before, Seabright's premisses seem to allow a broader conclusion.

## [III]

I have not yet mentioned another element in Seabright's view: his appeal to pluralism. This I find enigmatic.

'By a pluralist theory', Seabright writes, 'I understand an account of how society should be arranged that incorporates the possibility of multiple and non-trivially divergent views of the good life for individual human beings.'

How might a theory 'incorporate' this 'possibility'? Is it enough for the theory to admit that there might *be* such divergent views? This would be a very weak constraint. It is hard to imagine a theory which would deny this.

Perhaps by the 'possibility' of divergent views, Seabright means, not that such views might *exist*, but that they might be *correct*. This suggests two readings of his definition.

On one reading, a social theory is pluralist if it admits that divergent views might *all* be correct. But this cannot be right. When Seabright calls these views 'non-trivially divergent', he clearly means that they are incompatible. At most one of them could be correct.[6]

[6] Such views, he claims, are 'not subsumable under an encompassing theory of the good'. (We might replace 'correct' by 'reasonable'. Perhaps, to be pluralist, a theory must admit that there are divergent views which could all be reasonably held. But this constraint also seems too trivial.)

On the second reading, a theory is pluralist if it admits that there are divergent



views *any one of which* might be correct. This suggests that, to be pluralist, a theory must be *neutral* between these views. And this would be most simply true if this theory does not itself contain any view about the good. Thus Seabright asks 'whether it is possible for a government's economic policies to incorporate [pluralism]. Or does the very notion of an economic policy presuppose that the government has its own comprehensive theory of the individual good?' This seems to imply that, to be pluralist, a theory must not contain its own view about the good.

But this seems not to be what Seabright means. He also writes: 'This does not imply that a pluralist social theory must be neutral between all or even many theories of the individual good; for it to count as pluralist, it is enough for it to be compatible with the assertion of more than one theory of the individual good.' If a pluralist theory, rather than being neutral between all views about the good, is incompatible with many of these views, it must itself contain some view about the good.

Seabright seems to have in mind a theory which is both (*a*) incompatible with some views about the goods and (*b*) compatible with at least two views that are incompatible with each other. I can imagine ways in which (*a*) and (*b*) could both be true. But I am too unclear what Seabright means to take these comments further.[7]

[7] Seabright also claims that, to be pluralist, a theory 'must count it a social good that there may exist multiple views of the individual good'. If we delete 'may', this seems to define pluralists as those who welcome the holding by different people of incompatible views about the good.

There are various reasons why a theory might welcome such diversity in people's views. It may hold, with John Stuart Mill, that rivalry between these views will help each view to develop, and make it more intelligently and sincerely held. But from this form of pluralism there seems no argument to Seabright's conclusion on the proper sphere of government. A theory could be in this sense pluralist while allowing a government's decisions to reflect, at any time, the views of the good which are then most widely held. This theory's attitude to competing views would then be like a democrat's attitude to competing political parties.