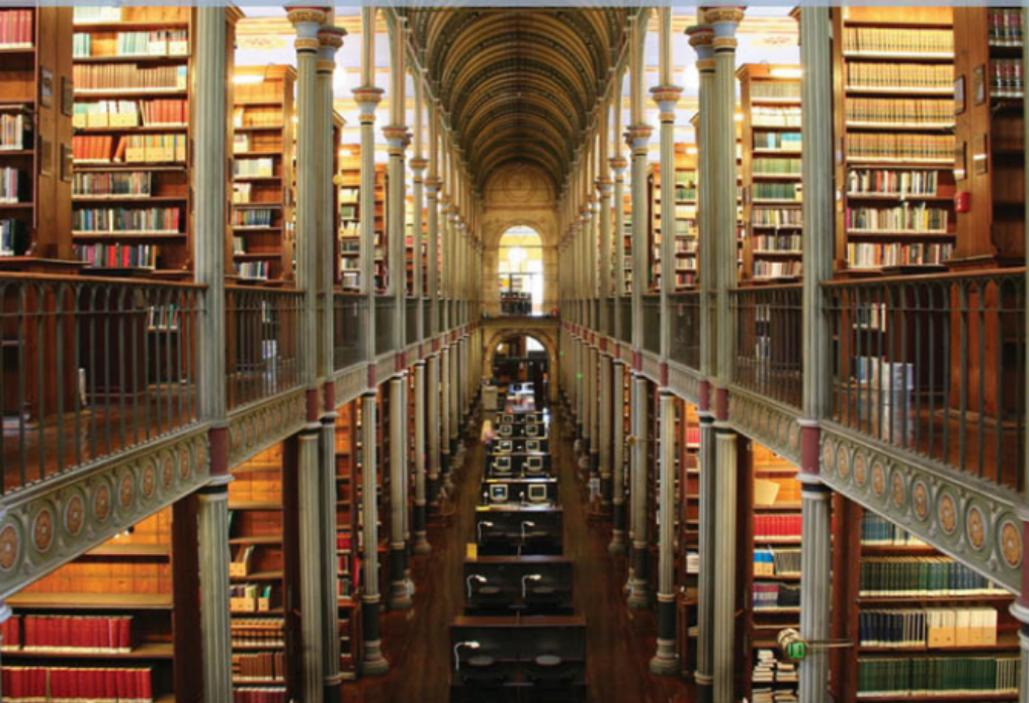


OXFORD



Samuel Scheffler

The Rejection of Consequentialism

A Philosophical Investigation of the Considerations Underlying Rival
Moral Conceptions

The Rejection of Consequentialism

A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions

Samuel Scheffler

REVISED EDITION
CLARENDON PRESS · OXFORD

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States by

Oxford University Press Inc., New York

Except where otherwise stated, © Samuel Scheffler 1982

The moral rights of the authors have been asserted

Database right Oxford University Press (maker)

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

ISBN 0-19-823510-0 (Hbk.)

ISBN 0-19-823511-9 (Pbk.)

*For
Israel Scheffler
and
Rosalind Z. Scheffler*

Preface

According to an ancient if occasionally unfashionable view, the subject matter of moral philosophy is organized in the first instance around the question of how people ought to live their lives. That is certainly how I conceive of the subject, and as a consequence it has sometimes seemed to me that only a fool or a fanatic could seriously think himself 'professionally' competent to express and defend views in this area. Despite these scruples, which perhaps represent my better judgement, I am submitting the work that follows for the reader's consideration. In mitigation, I can only say that if the subject matter of moral philosophy is vast and daunting, as it is, and if the complexity and power of the experiences that typically prompt moral reflection sometimes make the theorist's abstractions seem hollow and glib, as they do, it is also true that the question which animates the subject as I conceive it is vivid and gripping and demands our attention, even if all too often we acknowledge the demand only by contriving to ignore it.

This book grows out of a dissertation which I submitted for the Ph.D. at Princeton in 1977. But my interest in the topics it deals with is as longstanding as my interest in philosophy itself. The first philosophy course I took as an undergraduate at Harvard was a course on ethics taught by Roderick Firth. At the time, I found myself strongly drawn to the deontological views, though not the epistemological intuitionism, of W. D. Ross, and utilitarianism I found thoroughly abhorrent. Rejecting Ross's own intuitionism, I began to worry about how a deontological view might be defended. My worries have only increased since that time, as the reader of this book will discover, and have led me in directions that have sometimes surprised and dismayed me. This book charts the current state of my thinking.

My debts are many and large, and I will mention only the greatest of them. A memorable seminar taught by Thomas Nagel during my first year as a graduate student at Princeton convinced me that I wanted to undertake the project which eventually became my dissertation. As a teacher and as a dissertation adviser, Nagel was everything a student could hope for, and more. My debt to him is greater than any brief acknowledgement could begin to suggest.

In the course of writing my dissertation, I was greatly helped by long and frequent discussions with John Campbell, Peter Railton, and T. M. Scanlon. I am particularly indebted to Railton for helping me to appreciate for the first time the real force and power of a consequentialist view.

After the dissertation was completed, Railton and Gilbert Harman made fundamental criticisms which caused me to rethink the entire project.

When I began to write again two years later in Berkeley, I benefited greatly from the opportunity to discuss my work in seminars, and I am grateful to the students and colleagues who attended those seminars. More generally, I would like to thank the members of the Berkeley philosophy department for providing such a congenial and supportive environment in which to work. I am particularly grateful to Janet Broughton, Thompson Clarke, Linda Foy, and Barry Stroud.

I received extremely helpful written comments on the penultimate draft of my manuscript from Campbell and Railton, and also from Shelly Kagan. I am sure that they will not be fully satisfied with my attempts to meet their objections, but I am equally sure that the book has been much improved as a result of my attempts to respond to the points they raised. I am grateful to them for taking my views seriously, and for taking the time to express their reservations and make suggestions for improvement.

My debt to Derek Parfit is, quite simply, extraordinary. His written comments on several drafts were extremely helpful and led to major improvements at points far too numerous to acknowledge individually. Indeed, his written comments on the penultimate draft were almost as long as that draft itself, and I blush to think of the errors the book would have contained but

for his detailed and sensitive criticism. For all of this, and also for his support and encouragement, I am very deeply grateful. In fact, I suspect that only those who have received comparable assistance from him, and they are by no means few in number, will be able to fully appreciate the real extent of my debt.

Since this work does grow out of my Ph.D. dissertation, I would like to take this opportunity to thank the Danforth Foundation for giving me fellowship support when I was a graduate student. Most of the final version of the book was completed during a period of leave which was extended beyond the normal length of time by a Humanities Research Fellowship from the University of California, Berkeley, and I would like to thank the University for making possible that period of free time.

Finally, I want to express my appreciation to Kathryn G. Dreith, who prepared the book's index, and whose assistance at the proofreading stage was invaluable.

SAMUEL SCHEFFLER

Berkeley, California

December, 1981

Preface to the Revised Edition

This new edition of *The Rejection of Consequentialism* comprises the original text along with three additional essays I have written in the years since the book's initial appearance. All three of these essays have been previously published, and two of the three—'Agent-Centred Restrictions, Rationality, and the Virtues' and 'Deontology and the Agent: A Reply to Jonathan Bennett'—are reprinted here without revision. The third essay—'Prerogatives without Restrictions'—has been revised since its original publication and appears here for the first time in its present form.

The issues that are dealt with in this book have been widely discussed in the years since it was first published, and many

people have made valuable contributions to the discussion. I have not attempted to survey all of these contributions or to update the book in the light of everything that has been said. I have responded to some points in the essays included here, but I have done so only when I felt that I had something specific to add, by way of clarification or supplementation, to my earlier arguments.

In 'Agent-Centred Restrictions, Rationality, and the Virtues' I attempt further to elucidate the paradoxical character of agent-centred restrictions. In response to Philippa Foot's suggestion that the appearance of paradox may derive from a distinctively consequentialist assumption, which she takes to be illicit, that it always makes sense to talk about the relative value of overall states of affairs, I argue that the puzzle stems from something deeper. Agent-centred restrictions seem paradoxical, I maintain, because they appear to clash with a powerful conception of rationality that has strong roots in our everyday thought, and which is not unique to consequentialism. In 'Deontology and the Agent' I attempt, in response to questions raised by Jonathan Bennett, to clarify my use of the term 'agent-centred' and its function in my argument. In the course of this discussion, I consider the relationship between agent-centredness on the one hand and the distinction between doing and allowing on the other, and assess the respective roles played by these notions in the characterization of deontology. Finally, in 'Prerogatives without Restrictions', I suggest a modification in the design of an agent-centred prerogative to avoid an objection raised by Shelly Kagan and others. Although all three of these essays respond to specific writings by other philosophers, each attempts further to develop certain of the book's themes and arguments, and each can, I hope, stand on its own.

That I have not tried to update the book in a more general or systematic way is due largely to the nature of its aims. Although the book assigns an important role to 'hybrid theories' in the development of its overall argument, its primary aim is not to advance a new philosophical theory. Instead, the book grew out of my efforts to understand how certain basic features of common-sense moral thought might best be motivated and explained, and its primary aspiration is to focus attention on

certain issues that need to be addressed if we are to satisfy ourselves that either common-sense morality or consequentialism is fully defensible.

The book's central contention is that there are some features of common-sense moral thought that are much easier to explain than others. In particular, it is much easier to identify a comprehensible rationale for what I call an 'agent-centred prerogative' than for what I call 'agent-centred restrictions'. The most important function of hybrid theories, in the context of the book's overall argument, is to make this contention vivid and to help raise some basic questions about its implications. More specifically, hybrid theories help to frame two challenges—one to common-sense morality and one to consequentialism. The challenge to common-sense morality is to dispel the impression that agent-centred restrictions are puzzling or paradoxical, either by identifying a plausible rationale for them or by showing that the demand for such a rationale is misplaced. The challenge to consequentialism is better to defend its insistence on construing morality in purely impersonal terms, given the apparent availability of a rationale for at least one agent-centred moral principle, and given that the rationale for that principle seems no less comprehensible or compelling than the rationale for consequentialism itself.

The primary aim of this book is to articulate these twin challenges and to emphasize the importance of coming to terms with them. If the book retains a measure of interest, it is, I believe, because these challenges continue to strike us as pertinent, and because we are not yet persuaded that either one of them has fully been met.

SAMUEL SCHEFFLER

Berkeley, California

November, 1993

Contents

CHAPTER ONE The Project and Its Motivation	1
CHAPTER TWO Outline of a New Theory of Normative Ethics	14
CHAPTER THREE The Independence and Distinctness of the Personal Point of View	41
CHAPTER FOUR The Defence of Agent-Centred Restrictions: Intuitions in Search of a Foundation	80
CHAPTER FIVE The Project Reconsidered	115
APPENDICES Agent-Centred Restrictions, Rationality, and the Virtues	133
Deontology and the Agent: A Reply to Jonathan Bennett	152
Prerogatives Without Restrictions	167
INDEX	193

This page intentionally left blank

1 The Project and Its Motivation

As John Rawls has written, ‘The two main concepts of ethics are those of the right and the good . . . The structure of an ethical theory is . . . largely determined by how it defines and connects these two basic notions.’¹ Among ethical theories, those that I call ‘act-consequentialist’ may be characterized roughly as follows. Such theories first specify some principle for ranking overall states of affairs from best to worst from an impersonal point of view. In other words, the rankings generated by the designated principle are not agent-relative; they do not vary from person to person, depending on what one’s particular situation is. For they do not embody judgements about which overall states of affairs are best for particular individuals, but rather judgements about which states of affairs are best, all things considered, from an impartial standpoint. After giving some principle for generating such rankings, act-consequentialists then require that each agent in all cases act in such a way as to produce the highest-ranked state of affairs that he is in a position to produce.² Different act-consequentialist

¹ *A Theory of Justice* (Harvard University Press, 1971), p. 24.

² Obviously, this formulation is oversimplified. Most such theories do not require the agent to act in such a way as to produce the best actual state of affairs that is available. Rather they require that the agent perform the available act that has the highest expected value, where the expected value of an act is a function of each of its various possible outcomes and of their probabilities of occurrence. I will, however, continue to use the oversimplified formulation because it highlights the features of consequentialism that are relevant to this discussion. For my purposes, nothing is lost by avoiding the more complicated and more accurate formulation. It should also be emphasized that when I speak of the act-consequentialist as requiring agents to produce the best overall outcomes or states of affairs, I do not mean that the act-consequentialist divides what happens into the act and the outcome, and evaluates only the latter with his overall ranking principle. Rather, the act itself is initially evaluated as part of the overall outcome or state of affairs. The act-consequentialist first ranks overall outcomes, which are understood, in this broad way, to include the acts necessary to produce them, and then directs the agent to produce the best available outcome so construed.

theories incorporate different conceptions of the overall good: that is, different principles for ranking overall states of affairs from best to worst. But all such theories share the same conception of the right which requires each agent in all cases to produce the best available outcome overall.

Act-consequentialism is not the only kind of consequentialism; other variants include rule-consequentialism and motive-consequentialism. These views typically differ somewhat from act-consequentialism in what they require of agents,³ though they share with act-consequentialism the feature of ranking overall states of affairs impersonally, and the general idea that the best states of affairs are *somehow* to be promoted. I will not be discussing these other variants of consequentialism in this book. Although I believe that my main lines of argument could be modified to cover them, the only kind of consequentialism that I will actually consider in the book is act-consequentialism. Since this is so, and since the term ‘act-consequentialism’ is cumbersome, I will, beginning with the next paragraph and throughout the rest of the book, use the terms ‘consequentialism’ and ‘consequentialist’ to mean ‘act-consequentialism’ and ‘act-consequentialist’, except where I explicitly state otherwise. This is purely an abbreviatory device, adopted for the sake of simplicity and ease of exposition; it is not intended to suggest either that my discussion encompasses all of the various forms of consequentialism, or that act-consequentialism is the only legitimate form of consequentialism.

In contrast to consequentialist conceptions, standard deontological views maintain that it is sometimes wrong to do what will produce the best available outcome overall. In other words, these views incorporate what I shall call ‘agent-centred restrictions’: restrictions on action which have the effect of denying that there is any non-agent-relative principle for ranking overall states of affairs from best to worst such that it is

³ For some complications, see David Lyons, *Forms and Limits of Utilitarianism* (Oxford University Press, 1965).

always permissible to produce the best available state of affairs so characterized.

Classical utilitarianism, which ranks states of affairs according to the amount of total satisfaction they contain, is the most familiar consequentialist view.⁴ But classical utilitarianism is widely thought to be too crude a theory. Although its defenders point with approval to its simplicity, critics charge that this simplicity is achieved at too high a cost. They argue that utilitarianism relies on implausible assumptions about human motivation, incorporates a strained and superficial view of the human good, and ignores a host of important considerations about justice, fairness, and the character of human agency. More generally, they accuse utilitarianism of relentless insensitivity to the nature of a person, and suggest that it has forfeited any serious claim to account for the complex and varied considerations that intrude on the moral life, and which give rise to the severest tests of our decency. Indeed, utilitarianism has gained a reputation for moral clumsiness that is unparalleled among ethical theories. Bernard Williams, writing that ‘the simple-mindedness of utilitarianism disqualifies it totally’, suggests that ‘[t]he day cannot be too far off in which we hear no more of it’.⁵

And yet that day refuses to come; we continue to hear a great

⁴ To avoid confusion, there are a number of points about this characterization of the view that I call ‘classical utilitarianism’ which should be noted at the outset. First, I intend it to be understood that whenever I speak of this view as ranking states of affairs according to the amount of total or aggregate satisfaction they contain, I mean, of course, total *net* satisfaction (that is, total satisfaction minus dissatisfaction). Second, classical utilitarianism as I understand it is a view that has hedonistic and non-hedonistic variants. Thus in characterizing the view I have deliberately used the term ‘satisfaction’, which can be understood either hedonistically, as referring to a kind of feeling, or non-hedonistically, as referring to the satisfaction of people’s preferences, whatever the preferences may be for. In the course of this book, I will distinguish between the hedonistic and non-hedonistic variants of classical utilitarianism only when the distinction is relevant to the topic under discussion. Whenever I do not make the distinction explicitly, the reader is to understand what I say as applying to both the hedonistic and non-hedonistic variants. Finally, the term ‘utilitarianism’ is used in the philosophical literature in connection with a wide range of moral views, from rule-utilitarianism to the principle that the right act maximizes average utility to the so-called ‘ideal utilitarianism’ associated with G. E. Moore. But I will use the term exclusively to refer to the classical act-utilitarianism described in the text.

⁵ J. J. C. Smart and Bernard Williams, *Utilitarianism For and Against* (Cambridge University Press, 1973), p. 150.

deal about utilitarianism. Cynics may suppose that the explanation for this lies in the philosopher's penchant for keeping half-dead horses just barely alive so that he can continue to beat them with a moderately clear conscience. My diagnosis is different: I believe that utilitarianism refuses to fade from the scene in large part because, as the most familiar consequentialist theory, it is the major recognized normative theory incorporating the deeply plausible-sounding feature that one may always do what would lead to the best available outcome overall. Despite all of utilitarianism's faults (including, no doubt, its misidentification of the best outcomes), its incorporation of this one plausible feature is in my opinion responsible for its persistence. Moral conceptions that include agent-centred restrictions, of course, reject this feature. Although a full characterization and discussion of these restrictions will not be presented until Chapter Four, they have already been identified as restrictions on action which have the effect of denying that there is *any* non-agent-relative principle for ranking overall states of affairs such that it is always permissible to produce the best available state of affairs so construed. If an adequate theoretical rationale for such restrictions could be identified, it would provide a reason for rejecting utilitarianism's deeply plausible-sounding feature. But although, as I will indicate in Chapter Four, it is easy to think of cases in which agent-centred restrictions seem intuitively appropriate, and although such restrictions constitute the heart of most familiar deontological moral conceptions, it is, as I will also indicate in Chapter Four, surprisingly difficult to find persuasive hypotheses in the literature as to what their underlying theoretical rationale might be.

Faced with what I take to be serious difficulties in the attempt to provide an adequate rationale for agent-centred restrictions, and faced with the plausibility of the idea that it is always permissible to do what would have the best outcome overall, I wish in this book to reconsider the rejection of consequentialism. What I will do, more specifically, is to undertake a comparative examination of two different kinds of non-consequentialist moral conceptions. The standard deontological theories I call 'fully agent-centred' conceptions constitute the more familiar of these two kinds. By virtue of including agent-centred

restrictions, these conceptions deny that there is any non-agent-relative principle for ranking overall states of affairs from best to worst such that it is always permissible to produce the best available state of affairs so characterized. And in addition to including such restrictions, these conceptions also deny that one must do what would have the best outcome overall on all of those occasions when the restrictions do *not* forbid it. In other words, fully agent-centred conceptions maintain that, given any impersonal principle for ranking overall states of affairs from best to worst, there will be some circumstances in which one is not permitted to produce the best available state of affairs, and still other circumstances in which one is permitted but not required to do so.

Non-consequentialist conceptions of the second kind I will consider are much less familiar. Indeed, I am unaware of any previous discussion of them in the literature. These ‘hybrid’ conceptions, as I refer to them, depart from consequentialism through their incorporation of something I call an ‘agent-centred prerogative’, which has the effect of denying that one is always required to produce the best overall states of affairs, and which is thus in some form a feature of fully agent-centred conceptions as well.⁶ At the same time, however, hybrid conceptions are akin to consequentialist conceptions in their rejection of agent-centred restrictions: that is, in their acceptance of the idea that it is always permissible to do what would produce the best overall state of affairs. In other words, hybrid conceptions are like fully agent-centred conceptions and unlike consequentialist conceptions in maintaining that one need not always do what would produce the best outcome; but they are like consequentialist conceptions and unlike fully agent-centred conceptions in accepting the plausible-sounding idea that one *may* always do what would produce the best outcome.

The agent-centred prerogative, as I will argue, is responsive

⁶ That is, since fully agent-centred conceptions do, as I have said, deny that one is required to do what would have the best overall outcome on all of those occasions when the agent-centred restrictions do not forbid it, they in effect include an agent-centred prerogative of some form, although not necessarily of the very same form as I describe in Chapter Two of this book, and although the term ‘agent-centred prerogative’ is of course my own. For further discussion of this and related points, see the last paragraph of Chapter Two and footnote ⁷ of Chapter Four.

⁷ This passage calls to mind a kind of moral conception of which I have not yet taken note. The non-consequentialist conceptions I have been discussing depart from consequentialism either by incorporating an agent-centred prerogative alone, or an agent-centred prerogative *and* agent-centred restrictions. Non-consequentialist conceptions of the second type, which I have been calling ‘fully-agent centred’ conceptions, include two kinds of permissions not to produce the best states of affairs: those conferred by the prerogative and those entailed by the restrictions. As this passage suggests, however, one can imagine moral conceptions that did include agent-centred restrictions but that did not include any permissions not to produce the best states of affairs *except* those entailed by the restrictions.

to certain important anti-consequentialist intuitions. To this extent it is on a par with agent-centred restrictions. But, as I will also argue, there is a significant asymmetry between the two agent-centred features; that is, it is much easier to identify a plausible theoretical foundation for the former than it is for the latter. Thus an agent-centred prerogative can be motivated and defended not merely by showing that it has intuitive appeal in certain cases, but also by demonstrating that there is a plausible principled rationale which underlies it. To the extent that this rationale is compelling, hybrid conceptions may, by virtue of incorporating such a prerogative, seem more attractive than consequentialist conceptions. And at the same time, the fact that it is indeed possible to provide a persuasive theoretical rationale for one departure from consequentialism will make the difficulties in providing such a rationale for agent-centred restrictions seem all the more striking, whatever the intuitive appeal of such restrictions. The upshot is that hybrid theories, intermediate between consequentialist and fully agent-centred conceptions and less familiar than either, may in the end seem preferable to both. At the very least, they will emerge as a serious alternative. Or so I hope to show.

My argument will be somewhat circuitous. I will start out, in the remainder of this chapter, by reviewing two important and influential objections to utilitarianism, each of which gives voice to a kind of intuitive uneasiness about utilitarianism which many people feel. I will try to explain how these two objections are related to each other, but I will not consider the merits of the objections or possible replies to them right away. Instead, in Chapter Two, I will sketch the outline of one particular hybrid conception which is clearly capable of accommodating both objections. Its inclusion of an agent-centred prerogative enables the view to accommodate one of the objections, and its inclusion of what I call a 'distribution-sensitive' conception of the overall good enables it to accommodate the other. Once the outline of this hybrid theory is before us, I will turn in Chapter Three to an evaluation of the two objections to utilitarianism which prompted the construction of the hybrid view. I will argue that even if the objections as formulated in the recent literature and discussed here can be answered

by the utilitarian, there is nevertheless a deep, principled rationale underlying each of the two features that enable the hybrid theory sketched in Chapter Two to meet those objections. Thus even if the two objections as they stand can be rebutted, the intuitive uneasiness to which each gives voice can be provided with a more defensible principled foundation. That it may not be possible to say the same for agent-centred restrictions and the intuitive objections to which *they* respond I will argue in Chapter Four. As I have already indicated, I will try to show that although such restrictions have considerable intuitive appeal, it is surprisingly difficult to find persuasive hypotheses as to what their underlying rationale might be. This difficulty is all the more troubling in light of the fact that it *is* possible to identify such a rationale for an agent-centred prerogative, and may be seen by some as a reason for preferring a hybrid conception such as the one outlined in Chapter Two to a fully agent-centred conception. In Chapter Five I will reconsider the structure of the investigation undertaken in the first four chapters, and further examine the implications of that investigation.

Let me begin, then, by reviewing the two objections to utilitarianism to which I have alluded, and by explaining how these two criticisms, one dealing with personal integrity and the other with distributive justice, are related to each other. Williams argues that utilitarianism erodes the integrity of individuals by virtue of its strong doctrine of 'negative responsibility'. On this doctrine, one is as responsible for outcomes one fails to prevent as for outcomes one directly brings about, even when a crucial component of the outcome one fails to prevent consists of someone *else's* doing something. Thus, for example, one must abandon the projects one cares most about and strive to thwart the evil projects of others, any time doing so will avert a worse overall state of affairs. As a result of this doctrine, Williams claims, the link between one's deepest commitments and concerns, on the one hand, and one's actions, on the other hand, is thoroughly and systematically severed. Williams poses the issue as follows: 'how can a man, as a utilitarian agent, come to regard as one satisfaction among others, and a dispensable one, a project or attitude round which he has built his life, just

because someone else's projects have so structured the causal scene that that is how the utilitarian sum comes out?' To require that he do this, Williams writes, 'is to alienate him in a real sense from his actions and the source of his action in his own convictions. It is to make him into a channel between the input of everyone's projects, including his own, and an output of optimific decision; but this is to neglect the extent to which *his* actions and *his* decisions have to be seen as the actions and decisions which flow from the projects and attitudes with which he is most closely identified. It is thus, in the most literal sense, an attack on his integrity.'⁷

Despite the suggestiveness of Williams's remarks, his formulation leaves obscure precisely what feature of utilitarianism is supposed to alienate the agent from his actions and hence to undermine his integrity.⁸ One natural way to read him is as maintaining that utilitarianism alienates an agent from his actions by making the permissibility of the agent's devoting energy to his projects and commitments dependent on the state of the world viewed from an impersonal standpoint. If, through no fault of the agent's, things get bad enough from the impersonal standpoint, his projects become dispensable. But if it is *this* feature of utilitarianism that attacks the agent's integrity, it is doubtful that any theory but complete egoism could avoid doing so. For virtually *any* moral theory will make the permissibility of pursuing one's own projects depend at least in part on the state of the world from an impersonal standpoint. Virtually any moral view will hold that if things get bad enough from the impersonal standpoint, the agent's projects become dispensable. Different moral views do of course differ on the question of how bad things have to get from the impersonal standpoint before the agent is required to abandon his projects. But if the objection from integrity is interpreted as an objection to the in-principle dispensability of the agent's projects, then it must be regarded as a criticism of almost all non-egoistic theories,

⁷ Smart and Williams, pp. 116–17.

⁸ This point is also made by Nancy Davis in 'Utilitarianism and Responsibility', *Ratio* 22 (1980): 15–35, and by John Harris in 'Williams on Negative Responsibility and Integrity', *Philosophical Quarterly* 24 (1974): 265–73.

and not as an objection to which utilitarianism is distinctively vulnerable.

I believe, however, that the objection can be reconstrued in such a way that it could not be directed at all non-egoistic theories. It should be seen as arising not in response to utilitarianism's insistence on the in-principle dispensability of the agent's projects, but rather in response to the discrepancy between the way in which concerns and commitments are *naturally* generated from a person's point of view quite independently of the weight of those concerns in an impersonal ranking of overall states of affairs, and the way in which utilitarianism requires the agent to treat the concerns generated from his point of view as altogether dependent for their *moral* significance on their weight in such a ranking. In other words, utilitarianism incorporates a conception of the right which requires each agent in all cases to produce the best available outcome overall. It requires the agent to pursue his projects, commitments, and personal relationships whenever and to the extent that doing so would have the best overall outcome impersonally judged, and to neglect or abandon them whenever and to the extent that *that* would have the best overall outcome impersonally judged. Utilitarianism thus requires the agent to allocate energy and attention to the projects and people he cares most about *in strict proportion* to the value from an impersonal standpoint of his doing so, even though people typically acquire and care about their commitments quite independently of, and out of proportion to, the value that their having and caring about them is assigned in an impersonal ranking of overall states of affairs. It is *this* feature of utilitarianism which may be thought to alienate the agent 'from his actions and the source of his action in his own convictions', and thereby to undermine his integrity. So construed, the objection based on integrity could not be directed against those non-egoistic theories that do not share the utilitarian conception of the right, for they lack the feature that is now being said to generate the objection. While holding that there are circumstances in which people are required to turn their attention away from their own personal concerns, these theories do not require that agents devote energy and attention to their projects and commitments *in strict proportion* to

the value from an impersonal standpoint of their doing so.

In his discussion of integrity, Williams explicitly focuses on instances when utilitarianism requires an agent to abandon his projects because 'someone else's projects have so structured the causal scene that that is how the utilitarian sum comes out'. Now while it may seem particularly intrusive that the permissibility of pursuing one's own projects and commitments should be so dependent on the projects and commitments of other people, the objection based on integrity as I am construing it does not arise specifically in response to such cases. It arises in response to the utilitarian conception of the right which requires each agent in all cases to produce the best available outcome overall. It urges that this conception of the right holds the agent's ability to permissibly pursue his own projects and plans unacceptably hostage to the state of the world viewed from an impersonal standpoint, regardless of the extent to which that state has been shaped by human activity specifically. Thus although the objection as I am construing it cannot be directed against all non-egoistic theories, it is nevertheless in this respect more general than Williams's formulation suggests. And it applies not only to utilitarianism but to every consequentialist theory. For every consequentialist theory incorporates the same conception of the right. Thus every such theory must answer the charge of alienating the agent from his actions and the source of his action in his own convictions, and hence of undermining his integrity.

What is the relation between this objection to consequentialist theories and the objection to utilitarianism based on distributive justice? Because it is concerned to maximize *total aggregate* satisfaction or utility, classical utilitarianism demands that we channel resources to the relatively well-off whenever that will lead to the required maximization. As a result, it is alleged, classical utilitarianism will frequently require us to ignore the misery of a few people and concentrate instead on increasing the pleasures of the many simply in order to maximize aggregate satisfaction. And, it is said, this is morally unacceptable. This objection to utilitarianism is a familiar one, and it may not be immediately evident how it is connected to the criticism based on integrity. The connection can be made