beyond consequentialism

paul hurley

OXFORD

Beyond Consequentialism

Consequentialism, the theory that morality requires us to promote the best overall outcome, is the default alternative in contemporary moral philosophy, and is highly influential in public discourses beyond academic philosophy. Paul Hurley argues that current discussions of the challenge of consequentialism tend to overlook a fundamental challenge to consequentialism. The standard consequentialist account of the content of morality, he argues, cannot be reconciled to the authoritativeness of moral standards for rational agents. If rational agents typically have decisive reasons to do what morality requires, then consequentialism cannot be the correct account of moral standards.

Hurley builds upon this challenge to argue that the consequentialist case for grounding the impartial evaluation of actions in the impartial evaluation of outcomes is built upon a set of subtle and mutually reinforcing mistakes. Through exposing these mistakes and misperceptions, he undermines consequentialist arguments against alternative approaches that recognise a distinct conception of impartiality appropriate to the evaluation of actions. A moral theory that recognizes a fundamental role for such a distinct conception of impartiality can account for the rational authority of moral standards, but it does so, Hurley argues, by taking morality beyond consequentialism in both its standard and non-standard forms.

Paul Hurley is Professor of Philosophy at Claremont McKenna College, California.

This page intentionally left blank

Beyond Consequentialism

Paul Hurley

OXFORD
UNIVERSITY PRESS

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 Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

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

© in this volume Paul Hurley 2009

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2009
First published in paperback 2011

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 the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by Laserwords Private Limited, Chennai, India
Printed in the United Kingdom by
Lightning Source UK Ltd., Milton Keynes

ISBN 978–0–19–955930–5 (Hbk)
ISBN 978–0–19–969843–1 (Pbk)

For
My Wife, Mary Braet
and
My Mother, Joy Hurley

This page intentionally left blank

Acknowledgments

I am grateful for Kurt Baier's advice to develop the implications of my earlier work for the debate between consequentialists and their critics—he saw the outlines of this book long before I did. A special debt of gratitude is owed as well to David Gauthier, Jean Hampton, and Shelly Kagan for their invaluable role in the early development of many of the arguments presented here. The writings of T. M. Scanlon and Samuel Scheffler have also served as an ongoing source of insight and inspiration. I have benefited from discussions of this material with Stephen Darwall, Fred Schueler, Mark LeBar, Elizabeth Radcliffe, Peter Graham, John Fischer, Corey Brettschneider, Ted Hinchman, Voula Tsounas, Andrews Reath, Alan Nelson, Pierre Keller, Derek Parfit, Michael Meyer, Aaron James, Margaret Gilbert, Hilary Bok, Geoff Sayre-McCord, Amy Kind, Suzanne Obdrzalek, Stephen Davis, Noell Birondo, Peter Kung, Peter Thielke, Ann Davis, Emma Sabean, Michael Green, and Masahiro Yamada. I am particulary grateful for extensive written and oral comments on substantial portions of the manuscript from Gary Watson, Charles Young, Alex Rajczi, Dion Scott-Kakures, and Mark Lance, and owe an even deeper debt of gratitude to Arthur Ripstein, David Cummiskey, and Rivka Weinberg, who provided invaluable comments and criticisms on the entire manuscript at various crucial stages in its development. I have received helpful feedback on parts of this material from audiences at UC Riverside, Georgetown University, Ohio University, Occidental College, UC Santa Barbara, The University of New Mexico, UC Irvine, and the Claremont Colleges Works in Progress Group. I have also presented portions of this material to many of my classes, and have benefited greatly from many of these discussions. Particular thanks are due to the members of the Philosophy Department at Santa Clara University for inviting me to spend an extremely productive term as a visiting Fagothey Professor.

Additional debts of gratitude are owed to Peter Momtchiloff and Catherine Berry of Oxford University Press for their expertise in overseeing the publication of this book, to Celina Rosas for timely technical assistsance, to Kim Allen for copy editing, to Andrew Hawkey and Cory Davia for

proofreading, and to my sons Paul and Joe, who help me keep my feet on the ground when my head is in the clouds. Gratitude is also due to the journal *Ethics* for granting me permission to use revised portions of my essay "Does Consequentialism Make Too Many Demands, or None at All?" (Volume 116, July (2006), pp. 680–706).

Contents

1. Introduction 1
2. The Challenge to Consequentialism: A Troubling Normative Triad 10
3. The Demandingness Objection: Too Demanding, or Not Demanding at All? 35
4. Harnessing Williams to Sharpen the Challenge to Consequentialism 62
5. Deflating the Challenge *of* Consequentialism 106
6. From Impersonality to Interpersonality: Alternative Conceptions of Impartiality 139
7. Impartial Evaluation and Rational Authority 177
8. Generalizing to Other Forms of Consequentialism 226

References 264
Index 271

This page intentionally left blank

1

Introduction

Morality requires agents to perform the act that promotes the best overall state of affairs—it never permits us to bring about a worse state of affairs when a better one is available. To endorse this claim is to be committed to the paradigmatic version of consequentialism, upon which an action is morally right just in case its performance leads to the best state of affairs.[1] Such moral theories, and variations upon them, can with some plausibility claim a status as the default alternative in contemporary moral philosophy.[2] All roads to a systematic theoretical understanding of our moral practices can seem to lead down slippery slopes to consequentialism. Other approaches can appear upon reflection to smack of post hoc intuition mongering,

[1] This formulation is taken from Tim Scanlon (1998, p. 81). There are almost as many formulations of even the standard form of act consequentialism as there are consequentialists, nor are the differences in these formulations trivial. (See, for example, Shelly Kagan's alternative formulation (1989, p. 8).) But these differences are not relevant to the arguments that follow.

[2] These variations upon standard act consequentialism are usefully categorized by Shelly Kagan, who makes a helpful distinction between foundational consequentialism, which holds that "the ultimate justification for . . . genuine normative factors . . . lies in an appeal to overall good," and factoral consequentialism, which holds that "the goodness of overall consequences is the one and only normative factor with intrinsic moral significance" (1998, p. 213). Kagan's "factoral" consequentialism is sometimes characterized as well as "normative" consequentialism. One could in theory be a foundational consequentialist but not a factoral consequentialist, holding, for example, that although ultimate justification of normative factors appeals to overall states of affairs, such an appeal justifies normative factors other than the goodness of overall states of affairs. One could also in theory be a factoral/normative but not a foundational consequentialist, holding, for example, that although the appeal to the goodness of overall states of affairs is the only relevant normative factor, the ultimate ground for adopting this factor is not itself provided by overall states of affairs. Several recent accounts that reject foundational/philosophical consequentialism but defend normative/factoral consequentialism will be taken up in Section 8.5. Kagan also distinguishes versions of consequentialism by what sorts of things they hold to "be the primary focus of our moral evaluations" (1998, p. 214). Standard alternatives include actions, motives, and rules. Standard act consequentialism takes the proper focal point of direct evaluation to be actions, and embraces both factoral (normative) and foundational (philosophical) consequentialism with respect to the focal point of action. Variations upon such standard act consequentialism that shift the focal point from actions to rules and motives will be taken up in Section 8.4.

or to run afoul of Occam's razor. Beyond philosophy consequentialism is rarely mentioned but widely used. Its pervasive deployment in spheres such as economics, public policy, and jurisprudence is one of the more striking developments of the last century and a half. In public policy, for example, it is now commonplace to rank policies in terms of the better or worse consequences that will result overall from their implementation, often measured in monetary terms of benefits vs. costs.[3] In economics the rational course of action by an agent is taken to be the one that maximizes that agent's own welfare, utility, or preference satisfaction, but the moral course of action is often taken to be that of maximizing overall social welfare, utility, or preference satisfaction.[4] In many areas of contemporary jurisprudence, particularly in the United States, the right strategy is taken to be that which maximizes overall benefit. Markets are taken to be the most effective tools for implementing this consequentialist strategy, hence the role of the laws and the courts is taken to be that of mimicking the market (hence maximizing benefit) in areas in which markets (due to externalities, etc.) fail.[5] Even most of its advocates readily concede that the theory has a host of counter-intuitive implications and conflicts with many of our deeply held moral judgments. Yet efforts to supply such judgments with an underlying rationale can seem to lead ineluctably away from such intuitions and particular judgments and back towards consequentialism.[6]

Much recent work in ethics has consisted of efforts either to mitigate the counter-intuitiveness of generally consequentialist approaches to ethics, or to strike against the fundamental theoretical challenge that consequentialism is taken to provide to considered moral judgments and alternative moral theories. It is the thesis of this book that these discussions of the challenge *of* consequentialism tend to overlook a fundamental challenge *to* consequentialism, an unresolved tension between the theory and many of its most fundamental presuppositions. My project is to demonstrate that the traditional considerations that are taken to drive the challenge *of* consequentialism collapse in the face of this challenge *to* consequentialism. Many others have raised objections to consequentialism, but it

[3] See Elizabeth Anderson's discussion of cost–benefit analysis (1993, ch. 9), as well as Amartya Sen (2001) and Henry Richardson (2001).

[4] See, for example, Amartya Sen's characterization of a version of act utilitarianism (1985, p. 175).

[5] See, for example, Richard Posner's account (1981).

[6] For representative discussions of the pressure towards consequentialism that results from theoretical reflection see T. M. Scanlon (1982, pp. 103–28), and Shelly Kagan (1989, chs 1 and 2).

is often open to the consequentialist to respond that such criticisms beg certain of the crucial questions at issue. Critics argue that consequentialism cannot take persons or rights seriously, for example, but consequentialists respond that it is precisely their view that treats each and every person with full and equal seriousness. If rights are to be taken seriously, isn't the right approach to maximize the extent to which they are upheld overall?[7] My approach, by contrast, identifies certain tensions within the consequentialist approach itself. An appreciation of the nature of these tensions grounds the articulation of a fundamental challenge to the theory from within. Plausible steps for meeting this challenge, I will argue, lead us naturally beyond consequentialism; indeed, lead us to certain distinctly non-consequentialist commitments.

The fundamental challenge to consequentialism can be introduced by way of two claims. These claims are typically endorsed not only by advocates of consequentialist moral theories, but by defenders of the standard Aristotelian, Hobbesian, and Kantian alternatives to consequentialism as well. The first claim is that there are some acts that morality prohibits, and others that it requires of us. The second is that we should do what morality requires; we typically have decisive reasons to act in accordance with such moral requirements and prohibitions.[8] The standard alternatives to consequentialism are theories both of the standards set by morality and of the decisive reasons that we have to conform to such standards. In particular, each is a theory of the relationship between reasons and rightness (or "moral" virtues) upon which we have decisive reasons to do what the correct standards for right or virtuous action require of us and to avoid doing what they prohibit. None of these traditional approaches challenges these claims that morality establishes contentful standards of conduct (that morality is contentful) and that we have decisive reasons to do what such standards require (that morality is in this sense rationally authoritative).

[7] For classic examples of this sort of criticism and this sort of response see Robert Nozick's criticism (1974, ch. 3) and Samuel Scheffler's response (1982, ch. 4).

[8] I am following Derek Parfit in characterizing reasons to act as "decisive" when such reasons "are stronger than our reasons to act in any of the other possible ways" (2008, ch. 1, sec. 1, quoted with permission of the author). I will also follow Parfit in characterizing reasons as "sufficient" to do something "when these reasons are not weaker than, or outweighed by, our reasons to do anything else." (ibid., ch. 1, sec. 1) Such formulations allow for the possibility that in cases in which we do not have decisive reasons to do any one particular thing, we might have sufficient reasons to do many different things. This spare taxonomy of reasons will suffice for the arguments that are developed in this and the following three chapters, although it will be elaborated in key respects in Chapter 5.

Rather, they are attempts to provide theories of the relationship between what we have reasons to do and what it is right or virtuous to do that supply rationales for both the content that they take moral standards to have and the rational authority of such standards.

Consequentialism is often presented as an alternative in kind to such theories. But this is misleading. Like advocates of these other approaches, standard act consequentialists often appeal to both the claim that morality is contentful and the claim that morality is authoritative. But, unlike these alternatives, consequentialist moral theories are not in the first instance theories of the relationship between reasons to act and right actions. They are instead theories of the relationship between right actions and good overall states of affairs, upon which an action is morally right just in case its performance leads to the best state of affairs. The traditional alternatives are fundamentally theories of the relationship between reasons to act and right (or virtuous) action; consequentialism is fundamentally a theory of the relationship between right action and good states of affairs. Consequentialism thus provide a rationale for the content of morality, but such a rationale can be embraced without taking up any position at all concerning the authoritativeness of such moral standards. Unlike the alternatives, such a theory can be embraced by someone who rejects the claim that moral standards are authoritative. Do the ends justify the means? The consequentialist answer can be deceptively nuanced: their moral standards for right action are only satisfied if an agent acts to promote the best overall state of affairs, but this consequentialist theory of the content of moral standards is agnostic as to whether agents ever have even sufficient (much less decisive) reasons to do what is morally right.

This contrast between consequentialism and the other traditional approaches can be brought into focus with the example of Carl the card-carrying consequentialist. Let us assume that Carl accepts one from among the accounts of practical reason that can be and frequently are espoused by consequentialists, for example that the rational agent has decisive reasons to pursue her own happiness, or the maximal satisfaction of her preferences, or her own well-being, or the effective satisfaction of her plans, projects, and commitments. As a card-carrying consequentialist, he also accepts that the morally right action for him to perform is the action that brings about the best overall state of affairs, for example that maximizes overall happiness or maximizes the overall satisfaction of

people's projects and commitments. Carl further accepts the plausible claim that what furthers his own happiness or individual preference satisfaction or the satisfaction of his plans and projects often diverges from what maximizes overall happiness or aggregate preference satisfaction. Carl, while continuing to be a card-carrying consequentialist, draws the obvious conclusion from these commitments: he clearly has decisive reasons *not* to do the right thing in such routine cases. Carl is crystal clear about the action on his part that would be necessary to conform to what he, as a card-carrying consequentialist, recognizes as the correct moral standards. He is equally clear that he routinely has decisive reasons to do what such standards identify as wrong, and should not perform the right action in such cases. Because Harry the Hobbesian and Kate the Kantian espouse theories of the relationship between reasons to act and right actions, theories that purport to provide a rationale for both claims, such a result—that agents routinely have decisive reasons not to do what their theories identify as right—would constitute a devastating objection to their theories. By contrast, Carl can proceed blithely on as a card-carrying consequentialist while embracing such a result. Indeed, he can cite whatever grounds he takes there to be for the truth of his theory of the content of moral standards as providing grounds for rejection of the claim that morality is rationally authoritative. Carl might even conclude that the only reason the authoritativeness of morality has seemed plausible is that we have not recognized the full implications of the truth of consequentialism, and the stark contrast that in fact routinely obtains between what is morally right and wrong (properly, consequentially understood) and what we have decisive or sufficient reasons to do or not to do.[9] This consequence that

[9] Carl may adopt a strategy of indirection, holding that the best strategies for achieving his happiness and overall happiness do not aim at these directly, but involve the identification of certain rules of thumb, or perhaps even certain hard and fast rules that he believes will, if followed, lead to his greatest happiness and the greatest happiness overall. (For an insightful discussion of various forms of indirect consequentialism see Eric Wiland (2007, esp. sec. 2).) But Carl will be clear that what he has sufficient or decisive reasons to do will be determined by following the strategies for achieving his own happiness, not those for achieving the greatest overall happiness. In particular, he will have no reason even to take the overall happiness maximizing strategies into account unless doing so can somehow be shown to be relevant to the strategies for securing his own happiness. Absent some implausible congruence between two such sets of strategies, our more indirect Carl will persist in taking himself often to have decisive reasons to do what he recognizes to be morally wrong. Indirect act consequentialist strategies typically distinguish the context of justification from the context of deliberation. But Carl is clear that he is not entitled to presuppose the rational authority of his moral standards, hence that the context of justification diverges from the context of moral evaluation. With this divergence, indirection seems

is blithely accepted by Carl would of course be rejected by virtually all actual consequentialists. To surrender the rational authority of their moral standards as the price for the vindication of their account of the content of such standards would be for them a pyrrhic victory, losing the war as a cost of winning one battle. I will argue in what follows that consequentialism loses both the battle and the war.

This disanalogy between consequentialism and the traditional alternatives receives remarkably little direct attention. In part this is because consequentialists typically invoke the claim that morality is authoritative even when they do not attempt to provide a supplementary rationale for it. Their critics often do not challenge this appeal to the rational authority of moral standards because they share it. It is what they do not share, the consequentialist theory of what moral standards in fact require and prohibit, that becomes the focus of attention. Peter Singer's famous argument for a demanding moral obligation to help the global poor, for example, moves directly from the claim that such action is required to conform to consequentialist moral standards to the claim that we ought to do it: if consequentialist morality requires us to spend ourselves down to our own subsistence to prevent absolute poverty elsewhere, and he believes that it does, then he takes it to be clear, drawing upon the authoritativeness of morality, that we ought to do what morality requires.[10] He offers a consequentialist rationale for the conviction that we are morally required to bring about the best overall consequences, but no rationale is offered for the authoritativeness of moral requirements thus understood. Singer himself concedes that given his account of practical reason there are significant obstacles to providing any such rationale for the authoritativeness of consequentialist moral standards,[11] effectively conceding that he is not clear

unlikely to establish the rational authority of moral requirements. Such indirect variations upon standard act consequentialism will be taken up in more detail in Chapter 8.

[10] Singer (1979, pp. 230–1). Singer also draws upon an alleged analogy between his extremely exacting moral requirement to help the global poor, which does not intuitively provide us with decisive reasons, and a moral requirement to rescue a drowning child, which does intuitively provide us with decisive reasons. Once the illicit nature of the move by the consequentialist from moral requirement to decisive reasons becomes apparent, this problematic analogy must be called upon to carry the entire burden of the argument. But once the analogy is separated from any consequentialist claim to the rational authority of moral standards, Singer's analogy threatens to cut in the other direction for the consequentialist, suggesting that rational agents do not have decisive reasons to do what consequentialist morality requires in either case.

[11] Ibid., pp. 320–2.

how he can avoid something along the lines of Carl's argument. Carl will readily accept Singer's argument that it is morally right to spend down to subsistence and wrong to pursue any other course of action, but he will reject any claim that he ought to pursue such a course of action; indeed, he will take himself to have decisive reasons not to do so. Singer's argument fails without an appeal to authoritativeness underwriting the move from the claim that spending ourselves down to subsistence is morally right to the claim that we ought to do so.

The fact that a theory of the relationship between right actions and good overall states of affairs does not without augmentation provide a rationale for the authoritativeness of its moral standards is particularly noteworthy because even many consequentialists grant that we often seem to have sufficient reasons *not* to do what their exacting moral standards require. Such consequentialists allow that if it is right as they claim to bring about the best overall state of affairs, then we often seem to have sufficient reasons not to do what is right, and sometimes seem to have decisive reasons not to do what is right. Indeed, I will demonstrate in what follows that various accounts of practical reason presupposed by standard act consequentialists share certain features that provide obstacles to any reconciliation between the consequentialist theory of the content of moral standards and any commitment to the authority of such moral standards. Taken together, the theories of practical reasons and of moral standards advocated by standard act consequentialists not only do not provide a rationale for the claim that morality is authoritative, they provide a rationale for rejecting the authoritativeness of consequentialist moral standards. Such consequentialists are under tremendous pressure to follow Carl, and abandon the second claim. A dilemma threatens. If theirs is the right account of the content of morality, then it seems that morality cannot have the rational authority that even they commonly take it to have. Consequentialist moral standards are vindicated only by marginalizing the role of morality in practical reason and deliberation. If, however, morality is authoritative, then consequentialism cannot be the correct account of what morality requires and prohibits. Arguments such as Singer's, this suggests, are attempts to eat one's cake while having it too: they appeal to the authoritativeness of morality in the implementation of their theory of the content of morality, but acceptance of their theories of morality and practical reason undermines the legitimacy of just such an appeal. The result is a deep practical schizophrenia.

The problem is not just that consequentialism requires us to be moral saints,[12] it is that it presupposes that we are rationally required to be moral sinners even as it morally requires us to be moral saints.

Over the course of the next three chapters I will clarify the nature of the obstacles to acceptance of these two claims by standard act consequentialists. Armed with this more nuanced understanding of the inability of standard act consequentialists to account for both the content and the authoritativeness of their moral standards, I will turn in Chapter 5 to an examination of the traditional rationales that have been offered in defense of their account of the content of moral standards. I will show that many of these arguments turn on a misappropriation of certain deep intuitions concerning the authoritativeness of moral judgments. The intuition that it is always right to do the best that one can do in the circumstances, what one has decisive reasons to do, is hijacked by the consequentialist and redeployed as an alleged intuition that it is always right to do what brings about the best overall state of affairs, an intuition that appears to support the consequentialist's rationale for his theory of the content of morality. With the illicit nature of this appropriation exposed, many of the traditional rationales offered in support of the consequentialist theory of the content of morality lose much if not all of their apparent force.

One central argument for a consequentialist theory of the content of moral standards, an argument appealing to impartiality, avoids this criticism. It is the project of Chapters 6 and 7 to address this impartiality argument in support of the consequentialist theory of the content of moral standards, and the project of Chapter 8 to demonstrate that the arguments developed in the first seven chapters against standard act consequentialism extend to the most common consequentialist variations upon this standard form of the theory. The argument from impartiality begins with the widely held assumption that morality requires impartiality—the adoption of a standpoint of equal concern for all persons. It asserts that the conception of impartiality deployed in the consequentialist's assessment of better and worse overall *states of affairs* is the fundamental conception of impartiality for moral evaluation of *actions* as well. I will argue in Chapters 6 and 7 that there are profound obstacles to taking the consequentialist's conception of impartiality to be the fundamental moral conception for the evaluation of actions. I will also

[12] See Susan Wolf (1982) for the argument that consequentialism requires agents to be moral saints.

demonstrate that one natural response to the inadequacy of consequentialist rationales for the authoritativeness of morality, a response explored by Samuel Scheffler, suggests an alternative conception of impartiality, a conception appropriate in the first instance to the impartial evaluation of actions rather than states of affairs. This alternative conception of impartiality, a conception that plays a central role in many alternatives to consequentialism, avoids the obstacles that confront attempts to take the consequentialist's impersonal conception of impartiality to be the fundamental conception for the moral evaluation of actions. Moreover, an approach that takes this alternative conception of impartiality to be fundamental to morality avoids the obstacles that confront consequentialist theories to establishing the rational authority of moral requirements.

Such an approach, in short, holds out the prospect, closed to the consequentialist, of being incorporated into an account that can provide a plausible rationale for both claims. Non-foundational forms of consequentialism of the sort formulated by David Cummiskey, Brad Hooker, and Derek Parfit, accounts that attempt to ground recognizably consequentialist rules or principles in non-consequentialist foundations, might seem to be resistant to such arguments. Much of Chapter 8 will be spent demonstrating that the arguments developed in Chapters 6 and 7 against standard act consequentialism have purchase upon these non-foundational variants as well.

To summarize: the next three chapters will demonstrate that attempts to establish the authoritativeness of consequentialist moral standards are confronted by profound and debilitating obstacles that provide a fundamental challenge to the plausibility of the theory; Chapters 5–8 will demonstrate that an exploration of the source of these difficulties takes us beyond consequentialism.

2

The Challenge to Consequentialism: A Troubling Normative Triad

2.1 Challenges of and to Consequentialism

In this and the next several chapters my focus will be upon the paradigmatic form of consequentialism, standard act consequentialism.[1] The standard act consequentialist directly evaluates acts (not rules or motives) as right just in case they bring about the state of affairs that is evaluated as best overall from an impersonal standpoint.[2] Moreover, for such a standard act consequentialist the rationale for the adoption of a consequentialist normative principle is itself consequentialist, based in the appeal to the

[1] More specifically, and deploying a set of distinctions that will be clarified along the way, the standard form of consequentialism that is our initial focus is direct (not indirect), foundational (as well as normative/factoral), impersonal/agent-neutral (not evaluator-relative), act (not, e.g. rule or motive) consequentialism. The challenge that will be developed to standard act consequentialism in these initial chapters will be extended to other versions of consequentialism in Chapter 8.

[2] Here I adopt the standard characterization of the standpoint from which states of affairs are impartially evaluated as better and worse overall as the impersonal standpoint. In what follows I will characterize as "impersonal" and "agent-neutral" reasons that are based in the appeal to the impersonal value of states of affairs. Non-impersonal reasons, reasons that are not based in the impersonal value of states of affairs, will be characterized in what follows as "agent-relative" reasons. Thus, if certain reasons are impartial, general, and universalizable, but not based in the impersonal value of states of affairs, these reasons will be agent-relative reasons in this sense. The agent-relative/agent-neutral distinction is drawn in many different ways in the philosophical literature (for a different way of drawing the distinction, see Christine Korsgaard's "The Reasons We Can Share," in (1996*b*, sec. II)). The reasons I characterize in what follows as agent-neutral are identified as agent-neutral on all of these proposals for drawing the distinction, but many of these distinctions identify certain classes of non-impersonal reasons in addition as agent-neutral. It is in an effort to avoid the confusion that results from appeal to these many different ways of drawing the distinction that I will restrict the term "agent-neutral" in what follows to impersonal reasons, and identify all non-impersonal reasons as "agent-relative".

impersonal value of states of affairs.[3] One reason for this focus on standard act consequentialism is that many of the rationales that apparently provide support for consequentialism support standard act consequentialism, and do not readily extend to many consequentialist variations.[4] Such variants upon standard act consequentialism, moreover, are often beset by distinctive problems that have been pointed out with brutal effectiveness, often by other consequentialists.[5] Another reason is that this proliferation of variants has made consequentialism itself into something of a moving target. It becomes unclear in many cases what is being advocated, and even in some cases whether there is any meaningful sense in which the view defended is still consequentialist. But the most fundamental reason for this initial focus on standard act consequentialism is that despite the many sources of dissatisfaction with this form of the theory, sources reflected in myriad departures from it, the deepest challenges to the standard form of the theory have not yet been clearly articulated. My project in the next several chapters is to articulate these deeper challenges. I will argue in Chapter 8 that these deeper challenges that emerge in our consideration of standard act consequentialism generalize to these other forms of consequentialism. Moreover, I will argue that the tools for moral theory that come into view through developing these deeper challenges to act consequentialism provide the framework for moving beyond consequentialism in any form to an approach that recognizes the promotion of overall states of affairs as playing an important but circumscribed role within a comprehensive moral theory.

We can begin to clarify the nature of the challenge to consequentialism by locating the theory within a triad of claims. The first is the claim that morality is authoritative, that agents typically have decisive reasons to do—are in this sense rationally required to do—what they are morally required to do, and rationally required not to do what they are morally prohibited from doing. Some version of such a claim that rational agents have decisive reasons to do what morality requires, at least typically, and decisive reasons to avoid doing what morality prohibits, is often recognized

[3] See David Cummiskey (1996, pp. 7–11), Shelly Kagan (1998, pp. 213–14), T. M. Scanlon (1982, pp. 103–28), and Brad Hooker (2000*b*, pp. 222–38), for discussions of the distinction between normative/factoral and foundational/philosophical consequentialism.

[4] Many of these rationales are taken up in Chapter 5.

[5] See, for example, the criticisms put forward by Shelly Kagan in his essay "Evaluative Focal Points" (2000) and parts 2 and 3 of Tim Mulgan's *The Demands of Consequentialism* (2001).

as having deep intuitive appeal: if something is morally prohibited, then you shouldn't do it. The utilitarian sympathizer Henry Sidgwick begins *The Methods of Ethics* with this common thought "that wrong conduct is essentially irrational," and the contemporary consequentialist sympathizer Shelly Kagan suggests that one can bracket inquiry into the question as to whether it is ever rational to act immorally because "most of us take ourselves to be quite prepared to act as we morally should, if only we can determine just what it is that morality requires of us."[6] J. J. C. Smart argues that agents are morally required to maximize overall happiness, and takes agents to have decisive reasons to do what such consequentialist morality requires, asserting that "the only reason for performing an action A rather than an alternative action B is that doing A will make mankind . . . happier than will doing B."[7] Non-consequentialists are often more emphatic in their endorsement and defense of some such claim. Stephen Darwall, for example, maintains that such rational authority "captures part of what is felt to be special about moral requirements," and Jay Wallace maintains that "an agent who has serious interest in morality ought to regard compliance with moral principles as a fixed constraint on what may be counted as a good life."[8] I will demonstrate in the next chapter that many of the central debates between consequentialists and their critics presuppose that moral requirements have such rational authority. In particular, pervasive concerns that consequentialism is in particular need of

[6] See Henry Sidgwick (1982, p. 23) and Kagan (1998, p. 11). See also Kagan's discussion of moral requirement and morally decisive reasons (1998, pp. 66–7). For an invocation of the claim by a contractualist critic of consequentialism, see Scanlon (1998, p. 1). A similar claim is identified by Samuel Scheffler as the "claim of overridingness" (1992*a*, pp. 25–8), by David Brink as the "supremacy thesis" (1997, p. 255), and by Jay Wallace as the "optimality thesis," upon which "someone has most reason to comply with moral demands" (2006, p. 130). Wallace identifies commitment to such a claim as a marker for a "practical conception of ethics" (ibid., pp. 73–5).

[7] Smart (1973, p. 30). I have already shown that although arguments such as Peter Singer's do not explicitly appeal to such a claim, they do implicitly rely upon it. His argument moves from the claim that if the consequentialist moral requirement to assist the global poor is not outweighed by anything of comparable "moral significance," then we ought to do what morality requires. If the claim were only that such an action is required by morality, without any claim about whether or not we have any reason to do what morality requires, then Singer's argument would result only in a set of standards, and not in any rational demands upon agents that they conform to such standards. The argument has the apparent force that it does precisely because it implicitly draws upon the claim that a decisive moral requirement typically provides a rational agent with a decisive reason—with what they ought, simpliciter, to do. If Singer explicitly allowed that on his theory of reasons there is no basis for any presumption that agents rationally ought to perform the action identified by his moral standards as right, it would be clear that his argument fails to establish its intended result.

[8] Darwall (2006 p. 26); Wallace (2006, p. 134).

defense against charges that it is too demanding, confining, and alienating become utterly mysterious unless it is assumed that agents typically have decisive reasons to conform their actions to what is morally prohibited or required.

The second claim is that some practical reasons are fundamentally non-impersonal. We have reasons to pursue courses of action that are not based in the impartial evaluation of states of affairs, reasons that are sometimes sufficient to pursue courses of action that will not bring about the best overall consequences. Such a claim that some reasons are fundamentally non-impersonal, which I will refer to as the non-impersonality of practical reason, is readily endorsed by virtually all alternative theories of the authoritativeness and contentfulness of morality, and is a feature of virtually every theory of practical reason endorsed by consequentialists. Mill inquires into the "sanctions" to which the principle of utility is susceptible precisely because he recognizes that rational agents as he understands them often have good fundamental reasons to pursue their own happiness at the expense of overall happiness, hence that an account must be provided of what is necessary in order to alter these agents and their circumstances such that they typically do have decisive reasons to act to bring about the best overall states of affairs. Absent such circumstances, Mill allows, agents will not have decisive reasons to do what consequentialist moral standards require.[9] Peter Singer endorses a Humean account of practical reason, upon which "reason in action applies only to means, not ends. The ends must be given by our wants and desires."[10] Because these desires are frequently to perform actions that do not result in the best overall states of affairs, the reasons to which they give rise similarly count in favor of actions other than those that would bring about the best overall state of affairs. He agrees with Sidgwick that my concern for my own existence, for example, is fundamentally different from my concern for yours, that this difference is reflected in fundamental non-impersonal reasons, and that for the consequentialist this fact "stands in the way of attempts to show that to act rationally is to act ethically."[11] Samuel Scheffler points out that human beings are "moved by their attachment to particular people, their loyalty to causes, their pursuit of goals, their respect for principles, and their delight in forms of activity . . . the constitution and expression of distinct human

[9] But see Millgram (2000). [10] Singer (1993, p. 320). [11] Ibid., p. 322.

selves could not take the form that it does if human beings did not treat such a wide and diverse range of considerations as providing them with reasons for action."[12] At a minimum, accounts of practical reason allow what Derek Parfit characterizes as "stronger reasons to care about our own well-being and the well-being of those we love." Even granting that we also have distinctive impersonal reasons to maximize overall well-being, it seems clear that such non-impersonal and impersonal reasons will often conflict, and that in such cases "we can ask what, all things considered, we have most reason to do."[13] When considerations of our own well-being and the well-being of our loved ones are sufficiently compelling, we will presumably not have decisive reasons to perform the action that best promotes overall well-being, and will have sufficient reasons to act otherwise.

My formulation of this claim adopts the standard characterization of the standpoint from which states of affairs are impartially evaluated as better and worse overall as the impersonal standpoint, and characterizes as impersonal reasons the agent-neutral reasons that appeal to such an impersonal standpoint.[14] On virtually every account of practical reason and deliberation put forward either by consequentialists or their critics, agents have non-impersonal reasons to act, reasons that are not themselves grounded, even indirectly, in appeals to the impersonal value of states of affairs. Advocates of any such theory are committed to the fundamental non-impersonality of practical reason. One form of this claim that some practical reasons are fundamentally non-impersonal is a claim that the plans, projects, interests, and commitments of agents have rational significance independent of whatever rational significance they have in the determination of the

[12] Scheffler (2004, p. 237). [13] Both quotations are from of Parfit (2004, p. 355).

[14] I will follow Shelly Kagan in relying, at least initially, on our intuitive grasp of this distinction, although much will be said by way of refining this intuitive grasp in subsequent chapters. As I suggested in the preceding chapter, agent-neutral reasons will be understood in what follows as reasons based in the appeal to the impersonal value of states of affairs. Because the impersonal value of a state of affairs "gives every agent the same reason to promote it" (Scanlon 1998, p. 80), it provides a reason that is neutral among agents—an agent-neutral reason. Non-impersonal reasons, by contrast, will be understood as any reasons that are not based in the appeal to the impersonal value of states of affairs. All such non-impersonal reasons will be characterized as agent-relative or agent-centered rather than agent-neutral reasons. Such agent-relative reasons may be of varying kinds, including partial reasons to promote states of affairs that are better for the agent in question and impartial reasons to act that are not reasons to promote overall states of affairs. From the impersonal standpoint, any such agent-relative reason to perform or not to perform some action, whether partial or impartial, will appear as a special reason for some particular agent "that does not apply in the same way to others" (ibid., p. 83).

best overall state of affairs, independent significance that manifests itself at the most fundamental level of practical reason.[15] Reasons with such independent rational significance sometimes provide agents with sufficient reasons to act in ways that fail to bring about the best overall state of affairs.

The third claim asserts the consequentialist account of the content of morality, that an action is morally right just in case its performance leads to the best overall state of affairs. Such an act (or acts, in the case of ties) is not merely morally permitted but morally required, and all other acts are morally prohibited. This third member of the triad, the consequentialist account of moral content, is a claim concerning the relationship of distinctively moral evaluations (right/wrong, morally required/morally prohibited) of actions to evaluations of overall states of affairs (better/worse); the second, the non-impersonality of practical reason, is a claim concerning the relationship of certain evaluations of actions and reasons for action (decisive/sufficient/better/worse) to evaluations of states of affairs impersonally considered (better/worse); the first, the authoritativeness of morality, is a claim concerning the relationship between moral evaluations of actions as right/wrong and morally required/morally prohibited and evaluations of actions and reasons for action as sufficient or decisive, better or worse.

These three claims, (1) RAMS: the rational authoritativeness of moral standards (agents have decisive reasons to do what they are morally required to do), (2) NIR: the non-impersonality of practical reason (agents have some fundamentally non-impersonal reasons that sometimes provide them with sufficient reasons not to bring about the best overall state of affairs, and (3) CMS: the consequentialist theory of moral standards, are in considerable tension with each other; acceptance of any two appears to generate obstacles to endorsing the third. If agents have decisive reasons to do what they are morally required to do (RAMS), and they are morally required to bring about the best overall state of affairs (CMS), then they have decisive reasons to bring about the best overall state of affairs. But NIR assures us that whether or not morality is entirely impersonal, reason is not. Agents

[15] Samuel Scheffler argues that an agent's own concerns and commitments have moral significance independent of whatever significance they have "in an impersonal ranking of overall states of affairs" (1982, p. 9). I am adopting Scheffler's familiar terminology to make the analogous point about the independent rational significance of such concerns and commitments.

often have sufficient reasons not to bring about the best state of affairs. Similarly, if agents are morally required to bring about the best overall state of affairs (CMS), and they have non-impersonal reasons that are sometimes sufficient to act in ways that do not bring about the best overall state of affairs (NIR), they do not have decisive reasons in such cases to do what they are morally required to do. But this is difficult to reconcile with RAMS, the claim that agents have decisive reasons to do what they are morally required to do. Finally, if agents have decisive reasons to do what they are morally required to do, and to avoid actions that are morally prohibited (RAMS), but they often have sufficient reasons not to bring about the best state of affairs (NIR), then agents cannot be morally required to bring about the best overall state of affairs and morally prohibited from any other course of action. But consequentialism (CMS) holds that this is precisely what morality does require of agents. Peter Railton has captured the tension quite nicely. If the consequentialist's central claim (CMS) is embraced, he argues, a dilemma threatens. Given a plausible account of reason and reasonable expectations of the sort that involves NIR, we must remove "obligation so far from reasonable expectation that we no longer expect most people in our society to come close to carrying out their obligations," thereby giving up on anything in the neighborhood of RAMS and marginalizing morality in the process. If we insist on holding on to RAMS as well as CMS, however, the result will be an account which tells us that we can reasonably expect people to do things that NIR assures us no one could reasonably be expected to do, an account upon which "most people will be amazed at what is expected of them."[16]

It has been suggested by some consequentialists that although moral requirement is impersonal and rational requirement is not, the first member of our triad (RAMS) can nonetheless be satisfied because there either is, or can be, an almost constant congruence between the action required to bring about the best overall state of affairs and decisive reasons to act, for example between the action that will promote my happiness and the action that will promote overall happiness, or between the action that will maximize the satisfaction of my preferences and the action that will maximize overall preference satisfaction. But since Sidgwick such claims

[16] These quotations are both from Peter Railton's essay "How Thinking About Character and Utilitarianism Might Lead to Rethinking the Character of Utilitarianism" (2003, p. 238).

of congruence have been recognized, even by most consequentialists themselves, as highly dubious.[17]

Alternatively, many contemporary consequentialists have been led by the tension among the three members of this triad to abandon standard act consequentialism (CMS) in favor of indirect, non-foundational, and other forms of consequentialism. Although such variants bring a host of new difficulties in their wake, they are taken to hold out the prospect of closing the yawning gap that confronts standard act consequentialism between what agents have sufficient or decisive reasons to do and what act consequentialist morality requires. But I will present a series of arguments in Chapters 4–7 that undermine the plausibility of an appeal to a foundational role for the consequentialist's impersonal conception of impartiality in moral evaluation. These arguments demonstrating that the impersonal conception of impartiality cannot play such a foundational role will also allow us to extend the arguments against standard act consequentialism to indirect variants and variants that focus upon rules rather than acts. They will also reveal the structural features of an alternative conception of impartiality, an interpersonal conception, that can plausibly be called upon to play just such a foundational moral role. I will demonstrate in Chapter 8 that the availability of such a plausible alternative conception of impartiality presents significant challenges to arguments by Brad Hooker, David Cummiskey, and Derek Parfit that consequentialist norms can be grounded in non-consequentialist foundations.

2.2 Elaboration of the Normative Triad: Preview of Chapter 3

I have up to this point provided only the briefest sketch of the tensions among these claims that will be explored and developed throughout the

[17] I suggest in the next chapter that Mill himself is plausibly understood as pursuing this congruence strategy. See Sidgwick (1981, pp. 499–501) for his rejection of this aspect of Mill's argument. For a contemporary discussion of such prospects for congruence, see Scheffler (1992). See, in particular, his arguments that any morality that is to aspire even to a "potential congruence" (ibid., p. 4) with the point of view of the individual agent must be a moderate morality that is difficult to reconcile with consequentialism. Even the potential for congruence requires that "the most demanding moral theories are mistaken" (ibid., p. 6). If consequentialist standards lack rational authority, it becomes unclear what reasons there are to establish such a congruence if it is lacking, or to maintain such a congruence if it has somehow been established.

next several chapters. But even this brief sketch might be taken to show that no triad of claims is necessary to generate a challenge to consequentialism because NIR and CMS alone prove sufficient. The two claims put together suggest that agents often have sufficient reasons to ignore what morality requires and do what morality prohibits. If morality prohibits some course of action, however, isn't it the case that we have decisive reasons not to perform that action—that we are rationally required, not just morally required, not to perform it? As T. M. Scanlon and countless others have pointed out, "the fact that a certain action would be morally wrong seems to provide a powerful reason not to do it, one that is, at least normally, decisive against any competing considerations."[18] NIR, however, suggests that we often fail to have decisive reasons to perform the action that consequentialist moral standards identify as right.

It is one of the striking features of consequentialist moral theory, however, that it not only invites, but in some of its formulations appears to presuppose a radically different understanding of moral requirement and prohibition, rightness and wrongness, an understanding which is agnostic with respect to any commitment to RAMS. As we saw in the first section, this is because consequentialism, unlike, for example, traditional Hobbesian and Kantian alternatives, is not a theory of the rational authority of its moral standards.[19] These alternatives, by contrast, are theories of the rational authority of a certain set of moral standards. Hobbesian moral theories, for example, identify moral standards as rational constraints that rational persons typically have decisive reasons to adopt.[20] Their accounts of the rational authority of rightness are constitutive components of their accounts of standards for right action themselves. A similar point can be made about traditional Kantian alternatives.

Consequentialism, by contrast, is a theory of the relationship not of rightness of actions to goodness (sufficiency and decisiveness) of reasons for acting, but of the rightness of actions to the goodness of overall states of affairs. It can be, and frequently has been, augmented with accounts of the rational authority, or lack thereof, of such standards

[18] See Scanlon (1998, p. 1). Although some version of this claim is commonly presupposed or explicitly invoked as a platitude, it has been challenged. See Joshua Gert's *Brute Rationality* (2004) for the articulation of an account that challenges this claim.

[19] This point will be developed and clarified in the next chapter.

[20] See Hobbes's *Leviathan* (1994). For a contemporary Hobbesian approach, see David Gauthier (1986).